GLUTEN, FROM PLANT TO PLATE: IMPLICATIONS FOR PEOPLE WITH CELIAC DISEASE

EDITED BY: Michelle Lisa Colgrave, Katharina Anne Scherf, Melanie Downs and Alberto Caminero

PUBLISHED IN: Frontiers in Nutrition and Frontiers in Plant Science

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88966-939-4
DOI 10.3389/978-2-88966-939-4

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

GLUTEN, FROM PLANT TO PLATE: IMPLICATIONS FOR PEOPLE WITH CELIAC DISEASE

Topic Editors:
Michelle Lisa Colgrave, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia
Katharina Anne Scherf, Karlsruhe Institute of Technology (KIT), Germany
Melanie Downs, University of Nebraska-Lincoln, United States
Alberto Caminero, McMaster University, Canada

Citation: Colgrave, M. L., Scherf, K. A., Downs, M., Caminero, A., eds. (2021). Gluten, from Plant to Plate: Implications for People with Celiac Disease. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-939-4

Table of Contents

frontiers in Nutrition

EDITORIAL
published: 06 May 2021
doi: 10.3389/fnut.2021.680418

Editorial: Gluten, From Plant to Plate: Implications for People With Celiac Disease

Michelle Lisa Colgrave[1]*, *Katharina Anne Scherf*[2], *Melanie Downs*[3] and *Alberto Caminero*[4]

[1] Commonwealth Scientific and Industrial Research Organisation (CSIRO) Agriculture & Food, Brisbane, QLD, Australia, [2] Department of Bioactive and Functional Food Chemistry, Institute of Applied Biosciences, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, [3] University of Nebraska-Lincoln, Food Science and Technology Department, Lincoln, NE, United States, [4] Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

Keywords: gluten, celiac disease, cereal breeding, wheat, food processing

Editorial on the Research Topic

Gluten, From Plant to Plate: Implications for People With Celiac Disease

Gluten is the collective name for a class of proteins found in wheat, rye, and barley. Eating gluten triggers an autoimmune reaction in the ∼70 million people globally affected by celiac disease (CD), which causes the gut to react to gluten with intestinal inflammation and epithelial cell damage. In addition, wheat proteins may trigger respiratory, skin or food allergies and non-celiac gluten/wheat sensitivity (NCGS). Recently, more and more evidence has been emerging to support an increasing prevalence of gluten-related disorders in the population. This increase in prevalence has been too quick to be explained by genetic drift, pointing toward a change in environmental exposures as risk modifiers.

Gluten-free (GF) foods are now commonplace, offering consumers greater choice and availability. While many of these foods are made from non-gluten-containing grains, contamination of these inherently gluten-free products can occur during harvest, transport, or processing. Moreover, these foods are expensive and may be nutritionally inferior to gluten-containing products. The differences in nutritional properties of GF foods has led to research on ways to remove or reduce gluten from wheat and barley to provide new fiber, mineral, and vitamin options for those who must avoid gluten. This has led to research in classical plant breeding and the use of gene technology. An alternative approach to producing celiac-safe foods is via processing, wherein processes, such as separation, filtration, and/or application of enzymes, aim to remove gluten from gluten-containing ingredients. With many of these processed products entering the market, questions remain over the safety of these products and controversy over a suitable test to determine the gluten content remains. The question why an increasing number of people are affected by gluten-related disorders also needs to be answered. While improved diagnostics and awareness may partly explain the rise, further factors such as the use of vital wheat gluten in many food products, changes in wheat processing and in wheat protein composition may be responsible. In addition, it has been proposed that other environmental factors such as introduction of gluten to infant diets, breastfeeding patterns, alterations in the gut microbiota and infections could also dictate the development of gluten-related disorders.

This e-book is a compilation of 15 research and/or review papers written by 70 authors. This Special Research Topic comprises contributions from leading experts in the fields of plant breeding, food processing, clinical immunology, and gluten analysis to share their latest findings and help improve the quality and safety of foods for CD patients and other gluten-related disorders.

As critical overviews of the field, Shewry reviews wheat gluten, focusing on functional properties, and its role in triggering coeliac disease and gluten-related disorders. Wieser et al. provide a

Edited and reviewed by:
Dietrich Knorr, Technical University of Berlin, Germany

***Correspondence:**
Michelle Lisa Colgrave
michelle.colgrave@csiro.au

Specialty section:
This article was submitted to Nutrition and Food Science Technology, a section of the journal Frontiers in Nutrition

Received: *14 March 2021*
Accepted: *06 April 2021*
Published: *06 May 2021*

Citation:
Colgrave ML, Scherf KA, Downs M and Caminero A (2021) Editorial: Gluten, From Plant to Plate: Implications for People With Celiac Disease. Front. Nutr. 8:680418. doi: 10.3389/fnut.2021.680418

balanced review on the benefits of wheat consumption contrasted with the adverse effects for individuals suffering from wheat-related disorders.

Starting with the plant, Tanner et al. examined the accumulation of hordein storage proteins in developing barley grains using a combination of enzyme-linked immunosorbent assay (ELISA), western blot and liquid chromatography tandem mass spectrometry (LC-MS/MS) demonstrating maximum protein accumulation late in grain development. Shifting to wheat, Altenbach, Chang, Yu et al. described the genetic transformation of bread wheat to reduce the omega-1,2 gliadins, the wheat gluten proteins that present immunodominant epitopes relevant in celiac disease. Subsequently, the same group silence a subset of the alpha-gliadins and demonstrate reduced reactivities of antibodies (IgG and IgA) from a celiac disease patient cohort. Marín-Sanz et al. explore the impact of temperature and nitrogen availability of grain filling in bread wheat focusing specifically on the gliadin and glutenin protein fractions.

Shifting to gluten analysis, Panda and Garber review the use of antibody-based methods for accuracy in the quantitation of gluten in fermented or hydrolyzed foods and the inherent challenges due to the lack of appropriate reference materials and variable proteolysis. Next Alves et al. review the primary proteomic approaches used in the identification and quantitation of gluten peptides related to CD-activity and gluten-related allergies. In a complementary study, Daly et al. describe an update to the GluPro database, that provides a solid foundation for proteomic analysis of gluten proteins from gluten containing cereals. This database will enable identification of peptide markers for use in new gluten quantitation methods based on coeliac toxic motifs present in all relevant cereal species. To this end, Lexhaller et al. characterize and quantify cereal-specific gluten protein types by LC-MS/MS, allowing known wheat allergens and celiac disease-active peptides to be identified and laying the foundation for development of reference materials. Subsequently, five wheat cultivars were assessed by Schall et al. for their use as reference materials wherein their protein content, protein composition and responses to different ELISA methods were evaluated.

Osorio et al. examine the ability to detoxify gluten proteins using "glutenases" and employing site-directed mutagenesis aimed at the glutamine specific endoprotease from barley (EP-B2), and a prolyl endopeptidase from *Flavobacterium meningosepticum* (Fm-PEP).

From a clinical perspective, Pinto-Sanchez and Bai review the current strategies for follow up of patients with celiac disease, describing new tools for monitoring adherence to the gluten-free diet which could alter patient treatment. The isolation and purification of oat avenin for clinical trials aiming to establish the safety of oats in the diets of those with CD is reported by Tanner et al. Lastly, the Prolamin Working Group provide recommendations regarding clinical, analytical and legal aspects of CD, identifying those areas that require future multidisciplinary collaborative efforts.

AUTHOR CONTRIBUTIONS

MC prepared the initial draft concept of the Special Topic. MC, KS, MD, and AC refined the topic, compiled a list of potential contributing authors, and were associate editors for all submitted manuscripts. All authors contributed to the article and approved the submitted version.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 09 May 2019
doi: 10.3389/fpls.2019.00580

Elimination of Omega-1,2 Gliadins From Bread Wheat (*Triticum aestivum*) Flour: Effects on Immunogenic Potential and End-Use Quality

Susan B. Altenbach[1], Han-Chang Chang[1], Xuechen B. Yu[2,3], Bradford W. Seabourn[4], Peter H. Green[3,5] and Armin Alaedini[2,3,5,6]**

[1] Western Regional Research Center, United States Department of Agriculture-Agricultural Research Service, Albany, CA, United States, [2] Department of Medicine, Columbia University, New York, NY, United States, [3] Institute of Human Nutrition, Columbia University, New York, NY, United States, [4] Hard Winter Wheat Quality Laboratory, Center for Grain and Animal Health Research, United States Department of Agriculture-Agricultural Research Service, Manhattan, KS, United States, [5] Celiac Disease Center, Columbia University, New York, NY, United States, [6] Department of Medicine, New York Medical College, Valhalla, NY, United States

OPEN ACCESS

Edited by:
Michelle Lisa Colgrave, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia

Reviewed by:
Wujun Ma, Murdoch University, Australia
Craig F. Morris, Wheat Health, Genetics, and Quality Research (USDA-ARS), United States
Francisco Barro, Instituto de Agricultura Sostenible (IAS), Spain

***Correspondence:**
Susan B. Altenbach
susan.altenbach@ars.usda.gov
Armin Alaedini
aa819@columbia.edu

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: 21 February 2019
Accepted: 18 April 2019
Published: 09 May 2019

Citation:
Altenbach SB, Chang H-C, Yu XB, Seabourn BW, Green PH and Alaedini A (2019) Elimination of Omega-1,2 Gliadins From Bread Wheat (Triticum aestivum) Flour: Effects on Immunogenic Potential and End-Use Quality. Front. Plant Sci. 10:580. doi: 10.3389/fpls.2019.00580

The omega-1,2 gliadins are a group of wheat gluten proteins that contain immunodominant epitopes for celiac disease (CD) and also have been associated with food allergies. To reduce the levels of these proteins in the flour, bread wheat (*Triticum aestivum* cv. Butte 86) was genetically transformed with an RNA interference plasmid that targeted a 141 bp region at the 5′ end of an omega-1,2 gliadin gene. Flour proteins from two transgenic lines were analyzed in detail by quantitative two-dimensional gel electrophoresis and tandem mass spectrometry. In one line, the omega-1,2 gliadins were missing with few other changes in the proteome. In the other line, striking changes in the proteome were observed and nearly all gliadins and low molecular weight glutenin subunits (LMW-GS) were absent. High molecular weight glutenin subunits (HMW-GS) increased in this line and those that showed the largest increases had molecular weights slightly less than those in the non-transgenic, possibly due to post-translational processing. In addition, there were increases in non-gluten proteins such as triticins, purinins, globulins, serpins, and alpha-amylase/protease inhibitors. Reactivity of flour proteins with serum IgG and IgA antibodies from a cohort of CD patients was reduced significantly in both transgenic lines. Both mixing time and tolerance were improved in the line without omega-1,2 gliadins while mixing properties were diminished in the line missing most gluten proteins. The data suggest that biotechnology approaches may be used to create wheat lines with reduced immunogenic potential in the context of gluten sensitivity without compromising end-use quality.

Keywords: celiac disease, wheat allergy, gliadins, gluten proteins, gluten-related disorders, proteomics, wheat flour quality

Abbreviations: 2-DE, two-dimensional gel electrophoresis; AAI, alpha amylase/protease inhibitors; CD, celiac disease; HMW-GS, high molecular weight glutenin subunits; LMW-GS, low molecular weight glutenin subunits; MS/MS, tandem mass spectrometry; RNAi, RNA interference.

INTRODUCTION

Wheat is a major food crop grown throughout the world that is used in a wide range of different food products because of the unique viscoelastic properties of the flour. These properties are conferred by the gluten proteins, a complex group of proteins that account for about 70–80% of the total flour protein and are unusual in that they contain large regions of repetitive sequences with high proportions of glutamine and proline. The gluten proteins are divided into two major groups referred to as gliadins and glutenins. The gliadins are present in the flour as monomers and consist of alpha, gamma, delta and omega types, each with distinct structures, N-terminal sequences and repetitive motifs. These proteins confer extensibility to wheat flour dough. In comparison, the glutenins are present as large insoluble polymers made up of two types of proteins that are linked by disulfide bonds. These proteins are referred to as HMW-GS and LMW-GS. The glutenin polymers contribute elasticity to wheat flour dough. Ultimately, the composition of the gluten proteins in the flour, determined by both the genetics of the plant and the growth environment, is critical for end-use quality.

Each gluten protein group is encoded by many similar genes and there is considerable allelic variation among different wheat cultivars. Most bread wheat cultivars contain only six HMW-GS genes, while the numbers of gliadin and LMW-GS genes are much higher. Only recently with the availability of a high-quality wheat genome sequence from the reference wheat Chinese Spring (Zimin et al., 2017; International Wheat Genome Sequencing Consortium [IWGSC], 2018) has it been possible to determine accurately the complexity of these gene families in a single wheat cultivar. Indeed, a complete set of genes assembled and annotated by Huo et al. (2018a,b) from Chinese Spring included 102 genes of which 47 were alpha gliadins, 14 were gamma gliadins, five were delta gliadins, 19 were omega gliadins and 17 were LMW-GS. Of these, 26 alpha, 11 gamma, two delta, five omega gliadin and 10 LMW-GS encoded full-length proteins, while the remaining genes were either partial sequences or pseudogenes.

While the complexity of the wheat gluten proteins and their genes makes wheat research challenging, this is compounded by the fact that some of the same proteins that determine the commercial value of the flour also trigger human health conditions, including CD and IgE-mediated food allergies (Scherf et al., 2016, for review). T-cell epitopes that are relevant for CD have been identified in all of the major gluten protein groups. A list compiled by Sollid et al. (2012) includes five distinct epitopes from alpha gliadins, ten from gamma gliadins, two from omega gliadins, two from LMW-GS and one from HMW-GS. Of the gluten proteins, alpha gliadins have been thought to harbor some of the most important epitopes and a protease-resistant 33-mer peptide found in some alpha gliadins has been shown to be particularly toxic (Shan et al., 2002). Interestingly, the 33-mer peptide consists of six overlapping CD epitopes. The numbers of epitopes in individual gluten proteins also can vary considerably within a cultivar. For example, of the 26 alpha gliadins from Chinese Spring, only one protein contains the 33-mer peptide while nine proteins, all encoded by the B genome, do not contain any of the previously described CD epitopes. The rest contain from one to nine CD epitopes (Huo et al., 2018a). Certain omega gliadins also are immunodominant in CD. In fact, Tye-Din et al. (2010) reported that epitopes found in the omega-1,2 gliadins have a level of immunogenicity similar to the 33-mer peptide.

Food allergies to wheat also are complex. In a survey of 60 patients, Battais et al. (2005a) demonstrated that sera from allergy patients reacted with gluten proteins in all of the major groups and that the observed reactivity correlated with both the age and the symptoms of the patient. Using overlapping synthetic peptides, they identified IgE binding epitopes in alpha, gamma and omega gliadins and found that epitopes in food allergy were different from those in CD (Battais et al., 2005b).

A better understanding of the relationships between specific gluten proteins and their contributions to human health conditions and end-use functional properties is important for efforts to develop wheat that will be less likely to trigger immunogenic responses or better tolerated by patients with CD and food allergies. If the most highly immunogenic proteins could be eliminated from wheat flour without jeopardizing the functional properties of the flour, the introduction of that wheat into the marketplace may make it possible to reduce the numbers of people that become sensitized to wheat in the future. Alternately, making wheat flour safe for patients who already have CD or food allergy would require that all immunogenic proteins be eliminated or substantially reduced in the flour. This would include gluten proteins in all of the major groups and would likely impact the functional properties of the flour unless only those proteins within each group that contain harmful epitopes are identified and targeted.

This study focuses on the omega-1,2 gliadins, a subgroup of omega gliadins that are highly immunogenic. Omega gliadins are unusual even among gluten proteins in that they consist almost entirely of repetitive motifs with only short regions of unique sequence at their N- and C-terminal ends. The two types of omega gliadins, referred to as omega-1,2 gliadins and omega-5 gliadins, differ in N-terminal sequences and repetitive motifs. The omega-1,2 gliadins begin with ARE, ARQ, or KEL and contain the repetitive motif PQQPFP, while the omega-5 gliadins usually begin with the N-terminal sequence SRL and contain multiple copies of FPQQQ and QQIPQQ. The omega-1,2 gliadins are important in CD and in allergy patients that show a reaction to hydrolyzed wheat proteins (HWPs) in food products and cosmetics (Denery-Papini et al., 2012), while the omega-5 gliadins are the major sensitizing allergens in wheat-dependent exercise-induced anaphylaxis (WDEIA), a serious food allergy that occurs in sensitized individuals when the ingestion of wheat is followed by physical exercise (Morita et al., 2003).

In previous studies, we used RNAi to reduce the levels of omega-5 gliadins in wheat flour, resulting in transgenic plants with reduced IgE reactivity to sera from WDEIA patients without adverse effects on flour end-use quality (Altenbach and Allen, 2011; Altenbach et al., 2014b, 2015). The transgenic plants also had more stable protein compositions when produced under different levels of post-anthesis fertilizer. In this study, our goal was to design an RNAi construct that would target only the omega-1,2 gliadins in hopes of reducing the levels of immunodominant CD epitopes in wheat flour.

MATERIALS AND METHODS

Plant Material

The United States hard red spring wheat *Triticum aestivum* 'Butte 86' was grown in a greenhouse with daytime/nighttime temperatures of 24/17°C as described previously (Altenbach et al., 2003). Plants were watered by drip irrigation with 0.6 g/l of Peters Professional 20-20-20 water-soluble fertilizer (Scotts-Sierra Horticultural Products Company, Marysville, OH, United States).

RNA Interference Construct and Transformation of Plants

A 141 bp fragment from the 5′ end of an omega-1,2 gliadin gene was selected as the trigger for the RNAi construct. This fragment was amplified from 20 DPA endosperm RNA using primers QF18 and QR18 described in Altenbach and Kothari (2007), inserted in opposite orientations on either side of a 146 bp intron from a wheat starch synthase gene, then placed under the regulatory control of the HMW-GS Dy10 promoter and the HMW-GS Dx5 terminator as described in Altenbach and Allen (2011). The final construct was verified by DNA sequencing. Transformation of wheat plants with the construct and the plasmid pAHC25 that facilitates selection of transgenic plants with phosphinothricin (Christensen and Quail, 1996) was as described in detail in Altenbach and Allen (2011). Identification of putative transgenic plants by PCR analysis and initial screening of grain proteins from transgenic lines by SDS-PAGE were described previously (Altenbach and Allen, 2011). Homozygous lines were selected for transgenic plants in which the omega-1,2 gliadins were specifically eliminated from the grain without significant changes on other gluten proteins or where omega-1,2 gliadins as well as other gliadins and LMW-GS were eliminated from the grain.

Protein Extraction and Analysis by Two-Dimensional Gel Electrophoresis (2-DE)

Grain from selected lines was pulverized into a fine powder and sifted sequentially through #25, 35, and 60 mesh screens. Total proteins were extracted from the resulting flour with SDS buffer (2% SDS, 10% glycerol, 50 mM DTT, 40 mM Tris-Cl, pH 6.8) and quantified using a modified Lowry assay as described in Dupont et al. (2011). Three separate extractions of flour were each analyzed three times by 2-DE as described in detail previously (Dupont et al., 2011). Gels were digitized using a calibrated scanner and analyzed using Progenesis SameSpots Version 5.0 (TotalLab, Ltd., Newcastle upon Tyne, United Kingdom). Identifications of individual protein spots in the Butte 86 non-transgenic line were reported in Dupont et al. (2011). Individual spots in transgenic lines were deemed to show significant changes from the non-transgenic if they had ANOVA p-values < 0.02 and had changes in average normalized spot volumes that were greater than 20%.

Identification of Proteins in 2-DE Spots by Mass Spectrometry

The identities of proteins in selected 2-DE spots were confirmed by MS/MS. Protein spots #1–6 from Butte 86, #1–3 from transgenic line SA-30-118a-5 and #7–15 from transgenic line SA-30-118b-3 were excised from triplicate 2-D gels, placed in 96-well plates and digested individually with chymotrypsin, thermolysin, or trypsin using a DigestPro (Intavis, Koeln, Germany). Protein spots #1–6 from SA-30-118b-3 were digested with only trypsin. The resulting samples were then analyzed using an Orbitrap Elite mass spectrometer (Thermo Scientific, San Jose, CA, United States) as described in detail in Vensel et al. (2014). For analysis of spectral data, two search engines, Mascot[1] and XTandem![2], were used to interrogate a database of 125,400 sequences that included Triticeae sequences downloaded from NCBI on 06-18-2018, gluten protein sequences from Chinese Spring reported by Huo et al. (2018a,b), Butte 86 sequences from Dupont et al. (2011) and Altenbach et al. (2011), Xioayan 81 gliadin sequences from Wang et al. (2017), and common MS contaminant sequences contained in the common Repository of Adventitious Proteins (cRAP)[3]. Data from the two searches and the three enzyme digestions were compiled and further validated using Scaffold version 4.7.5[4] with a protein threshold of 99%, peptide threshold of 95% with 20 ppm error, and a minimum of 4 peptides. The decoy false discovery rate (FDR) in the analysis was 0%.

Patients

Serum samples were from 20 CD patients with elevated levels of IgG antibody to gluten [14 female, 17 white race, mean (SD) age 40.7 (18.5) years] and 20 CD patients with elevated levels of IgA antibody to gluten [13 female, 19 white race, mean (SD) age 46.0 (17.3) years]. Positivity for IgG or IgA antibody reactivity to gluten was determined as described previously (Samaroo et al., 2010). All cases of CD were positive for antibody reactivity to transglutaminase 2 (the most sensitive and specific serologic marker of CD), determined as previously described (Lau et al., 2013). In addition, all patients were biopsy-proven, diagnosed with CD according to previously described criteria (Alaedini and Green, 2005), and on a gluten-containing diet. Serum samples were obtained under institutional review board-approved protocols at Columbia University. This study was approved by the Institutional Review Board of Columbia University Medical Center. All serum samples were maintained at −80°C to maintain stability.

Assessment of Immune Reactivity by ELISA and 2-D Immunoblotting

Serum IgG and IgA antibody reactivities to gluten were measured separately by enzyme-linked immunosorbent assay (ELISA) as described previously (Moeller et al., 2014; Uhde et al., 2016),

[1] www.matrixscience.com
[2] https://www.thegpm.org/TANDEM/
[3] ftp://ftp.thegpm.org/fasta/cRAP/
[4] http://www.proteomesoftware.com/

with some modifications. Gluten proteins were extracted from the non-transgenic Butte 86 and transgenic lines as described before (Huebener et al., 2015). A 2 mg/mL stock solution of the gluten extract in 70% ethanol was prepared. Wells of 96-well Maxisorp round-bottom polystyrene plates (Nunc, Roskilde, Denmark) were coated with 50 μL/well of a 0.01 mg/mL solution of protein extract in 0.1 M carbonate buffer (pH 9.6) or left uncoated to serve as controls. After incubation at 37°C for 1 h, all wells were washed and blocked by incubation with 1% bovine serum albumin (BSA) in PBS containing 0.05% Tween-20 (PBST) for 1.5 h at room temperature. Serum samples were diluted at 1:200 for IgA and at 1:800 for IgG measurement, added at 50 μL/well in duplicates, and incubated for 1 h. Each plate contained a positive control sample from a patient with biopsy-proven CD and elevated IgG and IgA antibodies to gluten. After washing, the wells were incubated with HRP-conjugated anti-human IgG (GE Healthcare, Piscataway, NJ, United States) or IgA (MP Biomedicals, Santa Ana, CA, United States) secondary antibodies for 50 min. The plates were washed and 50 μL of developing solution, containing 27 mM citric acid, 50 mM Na_2HPO_4, 5.5 mM *o*-phenylenediamine, and 0.01% H_2O_2 (pH 5), was added to each well. After incubating the plates at room temperature for 20 min, absorbance was measured at 450 nm. All serum samples were tested in duplicate. Absorbance values were corrected for non-specific binding by subtraction of the mean absorbance of the associated uncoated wells. The corrected values were normalized according to the mean value of the positive control duplicate on each plate.

Two-dimensional immunoblotting was also used to assess reactivity. Following 2-DE as described above, proteins were transferred onto nitrocellulose membranes using the iBlot Dry Blotting System (Life Technologies, Carlsbad, CA, United States). The membranes were incubated for 1 h in a blocking solution made of 5% milk and 0.5% BSA in a solution of Tris-buffered saline containing 0.05% Tween-20 (TBST). Incubation with patient serum specimens (1:2000 for IgA and 1:4000 for IgG determination in dilution buffer containing 10% blocking solution and 10% fetal bovine serum in TBST) was done for 1 h. Serum samples from representative patients with elevated IgA and/or IgG antibody reactivity to gluten were included. HRP-conjugated anti-human IgA and IgG were used as secondary antibodies. Detection of bound antibodies was by the ECL system (Millipore, Billerica, MA, United States) and the FluorChem M imaging system (ProteinSimple, San Jose, CA, United States). Following immunodetection, bound antibodies were removed from the nitrocellulose membranes with Restore Western blot stripping buffer (Thermo Scientific, Rockford, IL, United States) and the membrane proteins were visualized using colloidal gold stain (Bio-Rad, Hercules, CA, United States). Each immunoblot was aligned to its corresponding colloidal gold-stained membrane using the SameSpots software (version 5.0) (TotalLab Ltd., Newcastle upon Tyne, United Kingdom).

Analysis of Flour End-Use Quality

End-use functionality tests were conducted at the USDA-ARS-HWWQL (Manhattan, KS, United States) using methods approved by American Association of Cereal Chemists International (AACCI, 1961, 1985, 1988, 1995) that are routinely used for assessment of wheat breeding lines. Wheat was converted to straight grade flour using a Quadramat Senior experimental flour mill following AACCI Method 26-10.02. Flour protein and moisture contents (14%mb) were determined by near-infrared reflectance (NIR) using AACCI method 39-11.01; mixing properties were determined on 10 g (14%mb) flour samples using a Mixograph (TMCO, National Mfg., Lincoln, NE, United States) according to AACCI Method 54-40.02; and SDS sedimentation tests were conducted in adherence to AACCI Method 56-60.01. Averages and standard deviations from triplicate samples were calculated for the non-transgenic and SA-30-118a-5 transgenic lines.

RESULTS

Selection of a Target Region for the RNAi Construct

The sequences of two omega-1,2 gliadins, omega-D1 and omega-D2, whose genes were identified from Chinese Spring by Huo et al. (2018b) are shown in **Figure 1A**. These proteins contain ∼70% proline + glutamine and have central regions that contain 20 QQPFP and either 45 or 53 PQQ motifs, respectively. These motifs are also found in other gliadins and LMW-GS and are present within the sequences of several characterized CD epitopes (**Supplementary File 1**). For example, 21 of 26 alpha gliadins, all delta gliadins and three of 10 LMW-GS from Chinese Spring contain a QQPFP motif while the 11 gamma gliadins contain from two to nine copies of the sequence. Multiple copies of the PQQ motif are also found in all other gliadins and LMW-GS from Chinese Spring. To design a RNAi construct specific for omega-1,2 gliadin genes, a 141 bp region that included 34 bp of the 5′ untranslated region as well as the portion of the gene encoding the signal peptide and N-terminal region of the protein was selected as the trigger sequence (**Figure 1B**). The specificity and potential off-target effects of the construct were assessed by comparing the trigger sequence to the genomic regions containing gluten protein genes from chromosomes 1A, 1B, 1D, 6A, 6B, and 6D from Chinese Spring (NCBI Accessions MG560140, MG560141, MG560142, MH338176, MH338181, and MH338193, respectively) (Huo et al., 2018a,b). Of the full-length omega-1,2 gliadin genes from Chinese Spring, the target region had 30 and 110 bp regions of identity with the omega-D1 gene and 20, 33, and 44 bp regions of identity with the identical omega-D2 and -D3 genes. In addition, there were multiple regions of identity that ranged from 20 to 90 bp with pseudogenes omega-A1, -A2, -A3, and -A4. Of the other expressed gluten protein genes described by Huo et al. (2018a,b), only seven contained regions of identity greater than 20 bp. These were the omega-5 gliadin genes, omega-B3 and -B6, containing 34 bp regions of identity, and LMW-GS genes, LMW-B2, -B3, -D1, -D6, and -D8, containing 23 bp regions of identity. Identities were within the portions of the genes encoding signal peptides.

A

omega-D1

MKTFLIFVLLAMAMNIATAARQLN PSNKELQS PQQSFSHQQQPFPQQPY PQQPYPSQQPYPS QQPFP TPQPQF PQQS
QQPFTQ PQQPTPLQ PQQPFPQQPQQPQQPFPQ PQQPFP WQ PQQPFPQ TQQSFPLQ PQQPFPQQPQQPFP QPQLQF PQ
Q PEQII PQQPQQ PFLLES QQPFPQQPQQPFPQ PQQLIPMQ PQQPFPQQSQQS QQPFPG PQQ LFPELQQPI PQQPQQP
FPLQ PQQPFPQQS QQPFPQQPQQ PCPLQ PQQPFPQQPQQPFPQQPQQPFPLQ PQQPFPLR PQQ PFSQQ PQQ SQQSFP
QPQ PQQPQQ PSILQ PQQ PFLQ PQQQLSQQLEQTISQQ PQQPFPQQ PHQ PQQ PY PQQQPYGSSLTSIDGQ

omega-D2

MKTFLIFVLLAMAMKIATAARELNPSNKELQS PQQSFSHQQQPFPQQPY PQQPYPSQQPYPS QQPFP TPQQQF PQQS
QQPFTQ PQQPTPLQ PQQPFPQQPQQPQQPFPQ PQQPFP WQ PQQPFPQ TQQSFPLQ PQQPFPQQPQQPFP QPQLPF PQ
Q SEQII PQQPQQ PFPLQ PQQPFPQQPQQPFPQ PQQ PIPVQ PQQSFPQQSQQSQQPFAQ PQQLFPELQQPI PQQPQQP
FPLQ PQQPFPQQPQQPFPQQPQQSFPQQPQQPFPQQPQQPFPQQPQQPFPQQPQQPFPLR PQQPFPQQPQQSQQSFP
QPQ PQQPQQ PSILQ PQQ PL PQQPQQ PFQQ PQQQLSQQPEQTISQQ PQQPFPQQ PHQ PQQ PY PQQQPYGSSLTSIGGQ

AAGGCAAGCAAGCAGTAGTAACCACAAATCCAACATGAAGACCTTCCTCATCTTTGTCCTCCTTGCCATG
M K T F L I F V L L A M

B

GCGATGAACATCGCCACTGCTGCTAGGCAGCTAAACCCTAGCAACAAAGAGCTACAATCACCTCAACAATC
A M N I A T A A R Q L N P S N K E L Q S P Q Q

FIGURE 1 | Design of the target region for the RNAi construct. **(A)** Sequences of two omega-1,2 gliadins from Chinese Spring with PQQ and QQPFP motifs shown in red. The signal peptide is underlined. **(B)** Nucleotide sequence of the 141 bp target region used in the RNAi construct and the portion of the protein encoded by the target region. In the DNA sequence, the initiation codon is enclosed in a box. In the protein sequence, the signal peptide is underlined.

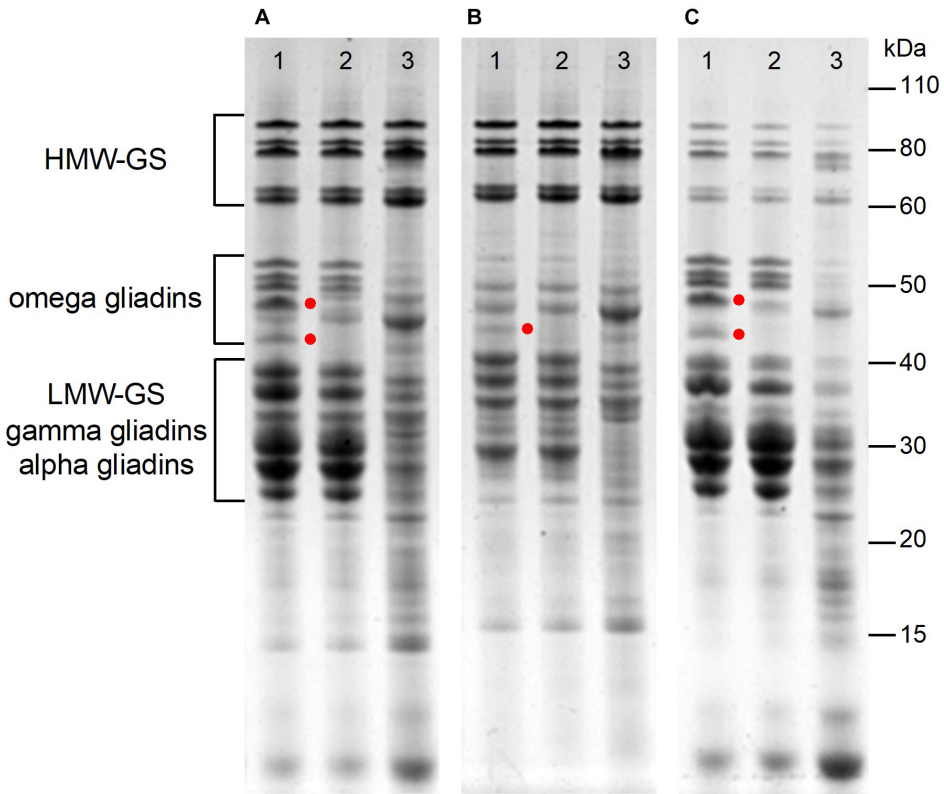


FIGURE 2 | Analysis of total protein **(A)**, glutenin **(B)**, and gliadin **(C)** in flour of non-transgenic (lane 1), and transgenic lines 118a-5 (lane 2) and 118b-3 (lanes 3). Red dots highlight bands that are present in the non-transgenic and absent in 118a-5.

Analysis of Flour Proteins From Transgenic Lines

Following transformation of Butte 86 plants with the RNAi construct, transgenic plants were identified that contained grain with altered protein profiles. **Figure 2** shows total protein, glutenin, and gliadin fractions from the grain of two transgenic lines that were selected for detailed analysis. In line SA-30-118a-5, referred to as 118a-5, several bands between 40 and

50 kDa were missing in the total protein, glutenin and gliadin fractions (**Figures 2A–C**, lane 2), but most other gluten proteins were not affected. In contrast, line SA-30-118b-3, referred to as 118b-3, showed notable changes in all protein fractions (**Figures 2A–C**, lane 3).

Analysis of total proteins from the non-transgenic control and 118a-5 by 2-DE is shown in **Figure 3**. Major differences between the two lines are highlighted in the boxes in **Figures 3A,B**. Although the positions of these spots in 2-DE were consistent with omega-1,2 gliadins identified by Dupont et al. (2011), the identifications were confirmed by MS/MS in the non-transgenic line (**Supplementary File 2** and **Figure 3C**). Spots 1 and 2 contained omega-1,2 gliadin BAN29067 as well as protein disulfide isomerase and beta-amylase. BAN29067 is identical to omega-D2 and -D3 from Chinese Spring except that it is missing the signal peptide. Spot 3 was identified as omega-1,2 gliadins BAN29067 and CAR82265 while spot 4 was identified as CAR82265. CAR82265 is the same as omega-D1 from Chinese Spring except for an extra amino acid at the N-terminus. This omega-1,2 gliadin contains a single cysteine residue. Spots 5 and 6 contained omega-1,2 gliadin AKB95614. ABK95614 is similar to the protein that would be encoded by pseudogene omega-A1 from Chinese Spring if the stop codon midway through the coding region was removed (**Supplementary File 2**). These spots were absent in 118a-5. Several minor spots from the transgenic line that appeared in the omega-1,2 gliadin region of the gel were also identified. Spot 1 in **Figure 3D** contained protein disulfide isomerase and beta-amylase while spots 2 and 3 were identified as beta-amylases. Quantitative 2-DE analyses revealed that there were few changes in the levels of individual proteins other than the omega-1,2 gliadins in the transgenic line (**Table 1** and **Supplementary File 3**). Gluten proteins accounted for 73.9% of the total normalized spot volume in the non-transgenic lines and 71.2% in the transgenic line. Gliadins accounted for 40.3% of the spot volume in the non-transgenic and 36.4% in 118a-5 while the proportions of glutenins were similar in both lines (**Supplementary File 3**).

In contrast, changes in the protein profile of transgenic line 118b-3 relative to the non-transgenic were quite dramatic (**Figure 4**). In addition to the omega-1,2 gliadins shown in the black box, nearly all gliadins and LMW-GS were suppressed (**Figure 4B**). A number of other proteins showed obvious increases in this line and were identified by MS/MS. Spots 1-4 in **Figure 4B** were identified as triticins, proteins similar to 11S storage proteins from dicots that have a large subunit and a small subunit cleaved from a larger precursor. Spots 1 and 2 in **Figure 4B** correspond to the large subunit encoded by the 5′ portions of the genes while spots 3 and 4 correspond to the small subunit encoded by the 3′ portions of the genes (**Supplementary File 4**). Spot 5 was identified as a purinin and spot 6 was identified as the endogenous alpha amylase/subtilisin inhibitor referred to as WASI. In addition, a number of spots that were either minor or undetectable in the HMW-GS region of the non-transgenic line increased in 118b-3 (**Figures 4B**, dashed box; 4D). Spots 7 and 8 in **Figure 4D** were identified by MS/MS as HMW-GS Ax2*,

TABLE 1 | Changes in the amounts of different classes of flour proteins in transgenic lines relative to the non-transgenic.

	% Change	
	118a-5	**118b-3**
Alpha gliadins	5.8	−58.9
Gamma gliadins	−7.1	−73.1
Omega-1,2 gliadins	−68.6	−74.6
Omega-5 gliadins	−10.1	−85.1
HMW-GS	7.6	27.1
LMW-GS	−6.6	−61.0
Purinins	16.5	150.9
Farinins	9.5	60.2
Triticins	11.8	108.1
Globulins	0.6	77.8
Serpins	11.3	51.4
AAI	8.8	90.5
Other non-gluten proteins	−1.4	58.9
Total gliadins	−12.7	−70.1
Total glutenins	0.0	−17.7
Total gluten proteins	−6.9	−45.7
Total non-gluten proteins	6.9	80.0

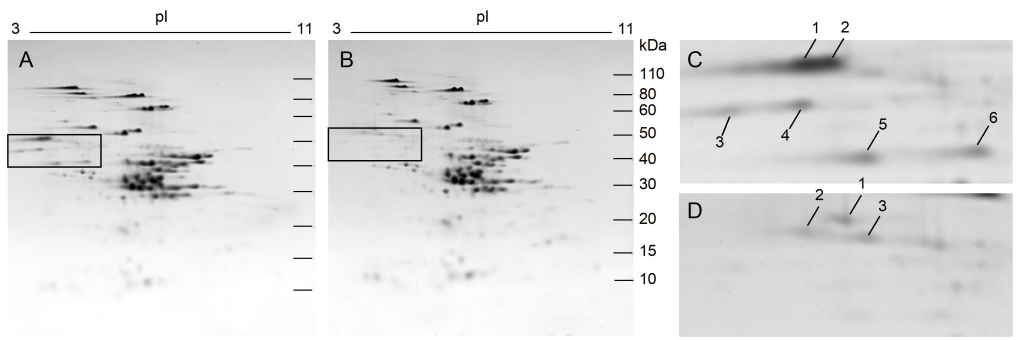


FIGURE 3 | 2-DE analysis of total flour proteins from non-transgenic **(A)** and transgenic plant 118a-5 **(B)**. The black boxes in **(A,B)** show the positions of the omega-1,2 gliadins. Regions of the gels containing omega-1,2 gliadins are enlarged in **(C)** (non-transgenic) and **(D)** (transgenic). Spots that are numbered were identified by MS/MS.

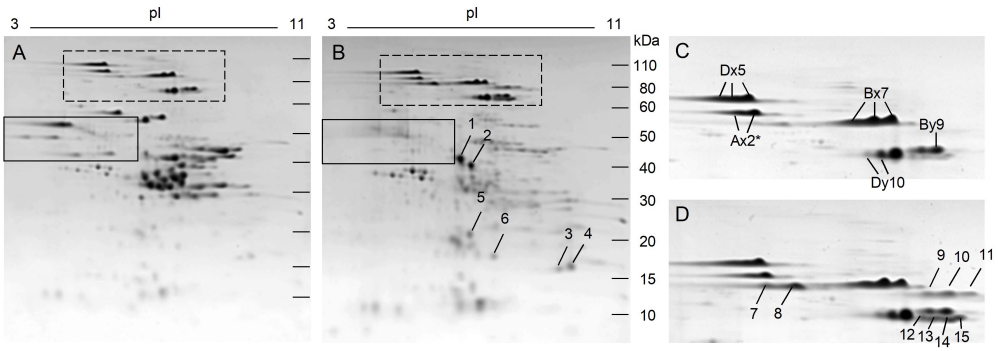


FIGURE 4 | 2-DE analysis of total flour proteins from non-transgenic **(A)** and transgenic plant 118b-3 **(B)**. The solid black boxes in **(A,B)** show the positions of the omega-1,2 gliadins while the dashed boxes show the positions of the HMW-GS. Regions of the gels containing HMW-GS are enlarged in **(C)** (non-transgenic) and **(D)** (transgenic). Arrows in **(B,D)** point to spots that increased in the transgenic line and were identified by MS/MS.

spots 10 and 11 as HMW-GS Bx7 and spots 12–15 as HMW-GS By9, consistent with identifications obtained from Butte 86 in a previous report by Dupont et al. (2011) (**Supplementary File 4**). Another transgenic line, SA-30-131b-5, showed a similar 2-DE profile. From the quantitative 2-DE analysis, HMW-GS accounted for 17% of the total protein in the non-transgenic, but 25% of the total protein in 118b-3 (**Supplementary File 5**). Interestingly, most of the increase was due to HMW-GS spots 7–15. All other gluten proteins accounted for 57.3% of the total protein in the non-transgenic, but only 21.6% of the protein in the transgenic (**Supplementary File 5**). The decline in total gluten protein in the transgenic was compensated by an increase in the non-gluten proteins from 25.7 to 53.5% of the total protein. Indeed, there were significant increases in most proteins within the major groups of non-gluten proteins, including triticins (108.1%), purinins (150.9%), farinins (60.2%), globulins (77.8%), serpins (51.4%), and AAI (90.5%) (**Table 1**).

Immunogenic Potential of Transgenic Lines

Reactivities of serum IgG and IgA antibodies from biopsy-proven CD patients toward gluten proteins were examined in the non-transgenic and transgenic wheat lines. Levels of detected IgG and IgA antibodies were highly diminished in the transgenic lines when compared to the non-transgenic line as determined by ELISA ($p < 0.0001$ for all comparisons) (**Figure 5**). All patients in the study had lower IgG and IgA reactivities to 118b-3 than to 118a-5, although differences were small for many patients. The molecular specificity of the reduction in CD antibody binding to gluten proteins was further examined by two-dimensional immunoblotting (**Figure 6**). For the representative patient shown in **Figure 6A**, IgG serum antibodies reacted with omega-1,2 gliadins, alpha and gamma gliadins, LMW-GS and serpins. Reactivity to omega-1,2 gliadins was eliminated in 118a-5 while reactivity to all gluten proteins was eliminated in 118b-3. IgA serum antibodies from the patient shown in **Figure 6D** showed the greatest reactivity to the omega-1,2 gliadins. This reactivity was eliminated in both 118a-5 and 118b-3.

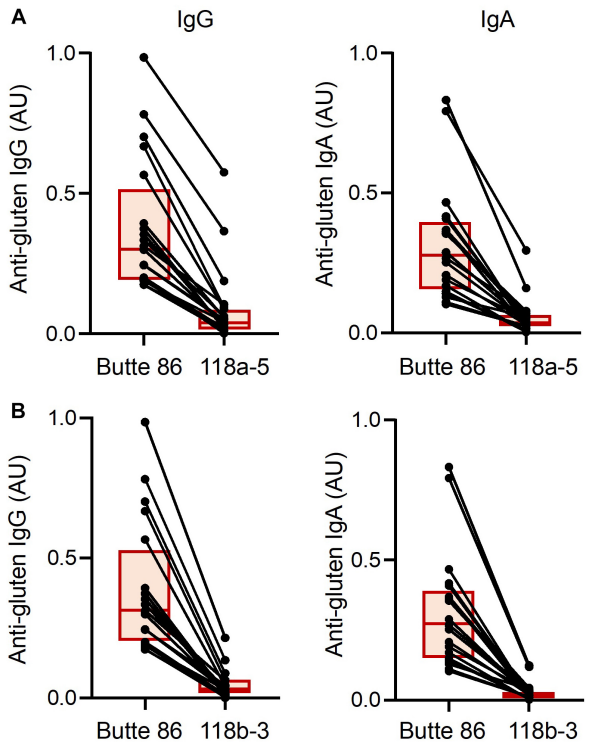


FIGURE 5 | Levels of IgG and IgA antibody reactivity for each of the 20 anti-gluten IgG-positive and 20 anti-gluten IgA-positive celiac disease patients toward gluten proteins from non-transgenic and transgenic plants 118a-5 **(A)** and 118b-3 **(B)**. Each individual is represented by a dot and the two points corresponding to the same individual are connected by a line. Each box indicates the 25th–75th percentiles of distribution, with the horizontal line inside the box representing the median.

End-Use Quality of Flour From Transgenic Lines

Non-transgenic and transgenic lines were grown in the greenhouse in sufficient quantities for end-use quality testing. The resulting grain from all lines had a vitreous appearance.

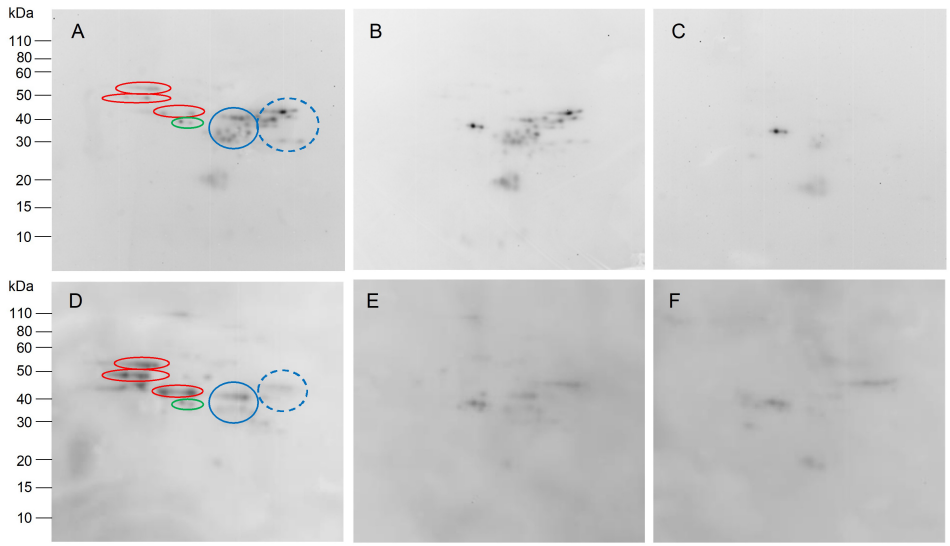


FIGURE 6 | Immunoblots showing IgG **(A–C)** and IgA **(D–F)** antibody reactivity in two representative celiac disease patients toward two-dimensionally separated total flour proteins from non-transgenic **(A,D)** and transgenic plants 118a-5 **(B,E)** and 118b-3 **(C,F)**. Omega-1,2 gliadins in **(A,D)** are shown in red ovals. Alpha and gamma gliadins are shown in blue circles, LMW-GS in blue dashed circles and serpins in green ovals.

Average kernel weight of 118a-5 was similar to that of the non-transgenic, 46.2 ± 1.4 mg for 118a-5 vs. 49.3 ± 2.6 mg for the non-transgenic. Grain protein and flour protein percentages also were similar for the two lines (**Table 2**). However, there were notable differences in the 10 g mixogram curves for each of the two lines (**Figures 7A,B**). Mix time was increased from 2.5 min for flour from the non-transgenic to 5.8 min for flour from 118a-5. In addition, mixing tolerance increased from 2 in the non-transgenic to 6 in 118a-5. There was also an increase in the SDS sedimentation volume from 62.8 ml in the non-transgenic to 66.1 ml in 118a-5 (**Table 2**).

In comparison, the average kernel weight was reduced ∼24% for 118b-3 and grain and flour protein percentages were 17.8 and 12.9% less, respectively, in 118b-3 than in the non-transgenic. The mixogram curve from 118b-3 flour was essentially flat, making it difficult to determine accurate mix times and tolerances (**Figure 7C**). The SDS sedimentation volume of 33.6 ml also was ∼46% less than the non-transgenic.

DISCUSSION

With the exception of short regions at the N- and C-termini, the omega-1,2 gliadins consist entirely of repetitive sequences that are also found in other gliadins and some LMW-GS. Careful selection of a 141 bp trigger from the 5′ region of the omega-1,2 gliadin gene for the RNAi construct made it possible to silence only the genes of interest in transgenic line 118a-5. Surprisingly, transformation with the same construct also resulted in transgenic lines in which nearly all gliadin and LMW-GS genes were suppressed. Currently, full-length sequences of only 13 alpha gliadin, 9 gamma gliadin and 5 LMW-GS genes from cv. Butte 86 are available (Altenbach et al., 2010a,b; Dupont et al., 2011). However, comparison of the RNAi target with a complete set of gluten protein gene sequences from Chinese Spring revealed that only omega-5 gliadin and a few LMW-GS genes had regions of identity with the trigger that were greater than 20 bp. Nonetheless, alpha and gamma gliadins were down-regulated effectively by the construct in some plants. As reviewed by Senthil-Kumar and Mysore (2011), there are many other factors that could contribute to off-target silencing, including the size of the trigger, the region of the gene that it targets and the specificity of the promoter that is used for expression. In addition, the site of integration in the genome and the copy number of the insertion can influence off-target effects.

Compensatory effects on the proteome were also noted in 118b-3 but not in 118a-5. Little is known about how the wheat grain compensates for alterations in different groups of storage

TABLE 2 | End-use quality data from non-transgenic and transgenic lines.

	Grain protein (%)	Flour protein[1] (%)	Mix time (min)	Mix Tolerance[2]	SDS sedimentation volume (ml)
Butte 86[3]	20.5 (0.8)	18.3 (0.5)	2.5 (0.1)	2	62.8 (0.45)
118a-5[3]	20.2 (0.9)	17.8 (0.8)	5.8 (0.1)	6	66.1 (0.87)

[1]*Based on 14% moisture.* [2]*Recorded on a 0–6 scale with 6 having the greatest tolerance.* [3]*Averages and (standard deviations) from flour samples from three biological replicates are reported.*

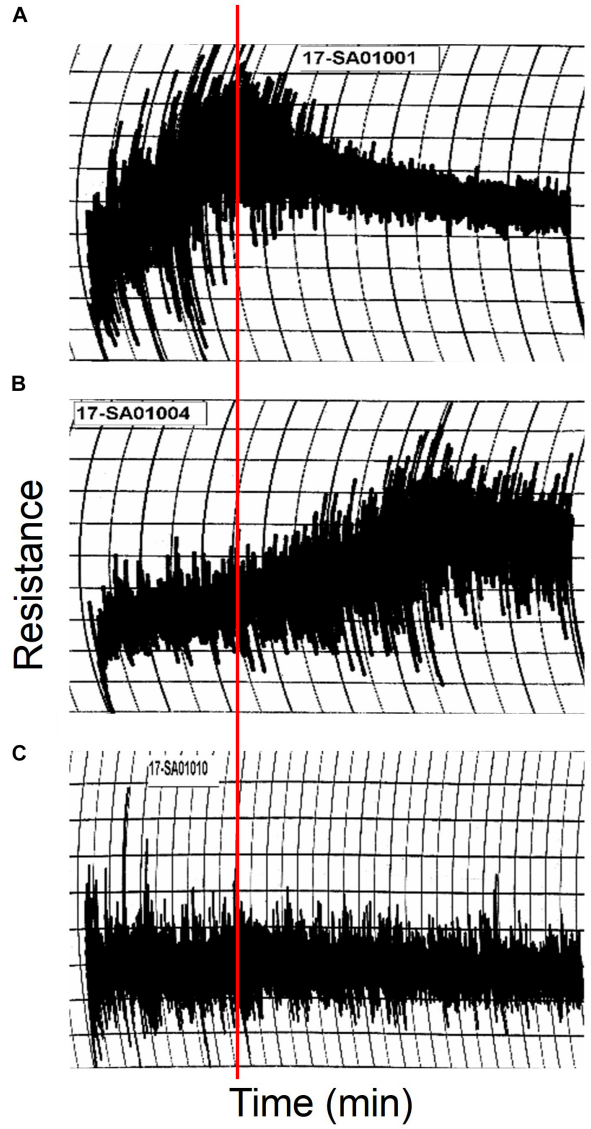


FIGURE 7 | 10 g mixogram curves produced using flour from non-transgenic **(A)** and transgenic lines 118a-5 **(B)** and 118b-3 **(C)**. The red line shows the position of the peak mixing time in the non-transgenic.

proteins. The omega-1,2 gliadins encompass only about 6.6% of the total flour protein in Butte 86. In 118a-5, it is likely that any compensation for the loss of these proteins was spread over the entire proteome. In contrast, gliadins and LMW-GS account for ∼58% of the total flour protein in Butte 86 and the large decreases in these proteins in 118b-3 were compensated by obvious increases in HMW-GS as well as in many groups of non-gluten proteins including triticins, serpins, purinins, globulins, and AAI. Among the HMW-GS, it is particularly curious that several protein spots that migrate faster than the major Ax2*, By9, and Bx7 spots in the second dimension of 2-DE were more prominent in the transgenic line. Differences between the apparent molecular weights of these protein spots and the major HMW-GS spots in the non-transgenic suggest that these spots could result from post-translational modifications. Recently, Nunes-Miranda et al. (2017) presented evidence that y-type HMW-GS encoded by the B genome are subject to proteolytic processing at the C-terminus by an asparaginyl endopeptidase. This processing event likely occurs at one of two asparagine residues 36 and 42 amino acids upstream of the C-terminus of the protein, resulting in the removal of a cysteine residue that may be involved in the formation of the glutenin polymer. Thus, the processed HMW-GS potentially could influence glutenin polymer size and end-use quality. Certain omega-1,2 gliadins and some LMW-GS have also been shown to undergo post-translational cleavage at the N-terminus by an asparaginyl endopeptidase (Dupont et al., 2004; Egidi et al., 2014). In the absence of the omega-1,2 gliadins and LMW-GS in 118b-3, it is possible that HMW-GS By9 may be a preferred substrate for the enzyme. While this explanation seems plausible, it is not supported by the MS/MS data that identified a number of peptides within 36 amino acids of the C-termini of the proteins in spots 12-15 (**Figure 4D** and **Supplementary File 4**). It is also notable that HMW-GS Ax2* and Bx7 do not contain any asparagine. However, asparaginyl endopeptidases are reported to cleave at aspartate residues with a lower efficiency and these residues are found in both proteins (Müntz et al., 2002). But, as with By9, C-terminal peptides (in spots 7 and 11) or peptides close to the C-termini (in spots 8 and 10) were identified by MS/MS.

Off-target and compensatory effects due to RNAi silencing of wheat gluten protein genes have also been observed in a number of studies. In RNAi experiments targeting the omega-5 gliadins, Altenbach et al. (2014a) reported differential effects on the proteome, although it was possible to identify transgenic plants in which only the omega-5 gliadins were reduced. Barro et al. (2015) evaluated seven different combinations of RNAi constructs with the goal of creating a wheat line that was devoid of CD epitopes and observed a wide variety of effects. Their constructs included target regions derived from both repetitive and non-repetitive portions of alpha, gamma and omega gliadin and LMW-GS genes. In some constructs, target regions from multiple genes were combined. Additionally, plants were sometimes transformed with more than one RNAi construct. In plants transformed with an RNAi construct targeting a 169 bp region corresponding to the Q-rich domain of a gamma gliadin, gamma gliadins decreased but were compensated by increases in omega gliadins (Gil-Humanes et al., 2010). Likewise, when a 377 bp region corresponding to the repetitive domain from an alpha gliadin was used as the target, alpha gliadins decreased while omega gliadins and HMW-GS increased. In one construct, a 132 bp region that encoded part of the signal peptide, N-terminal region and the repetitive domain of a LMW-GS was combined with a 173 bp region encoding a repetitive region from an omega gliadin. In the two resulting transgenic lines, LMW-GS decreased, omega, gamma and alpha gliadins were partially decreased and HMW-GS increased. In another construct, a 170 bp region that corresponded to the first non-repetitive domain of an alpha gliadin was combined with

a 191 bp region that corresponded to part of the signal peptide, N-terminal region and repetitive region of an omega gliadin. Both were from the most conserved regions of the genes. In the three lines reported, omega and gamma gliadins decreased, alpha gliadins were partially decreased, HMW-GS increased and LMW-GS were somewhat increased. Other combinations of RNAi plasmids resulted in transgenic plants with decreased CD toxicity. In plants transformed with two RNAi constructs, one targeting alpha gliadins and the other targeting both LMW-GS and omega gliadins, all gliadins and LMW-GS decreased while HMW-GS increased. These results are similar to those obtained the current study in 118b-3 using a construct that contained target sequences from only an omega-1,2 gliadin. Barro et al. (2015) did not evaluate constructs containing sequences derived only from omega gliadins. Additionally, they evaluated effects on the proteome by HPLC rather than 2-DE so it is not known whether increases in HMW-GS observed in their lines might reflect post-translational processing.

With regard to immunogenicity, the analyses with antibodies from CD patients demonstrated a highly significant reduction in binding to the transgenic lines when compared to the non-transgenic wheat. The data suggest that removal of the omega-1,2 gliadins is likely to have a considerable effect in eliminating the pathogenic sequences present in the studied Butte 86 wheat cultivar. Whether a similarly significant reduction in T cell reactivity to the transgenic lines would be observed remains to be seen in further analyses. However, considering the fact that the omega-1,2 gliadins contain known T cell epitopes and that most T cell epitopes are located within B cell epitope sequences of gluten proteins, it is highly likely that T cell reactivity to the transgenic lines would also be eliminated or diminished to a great extent.

With regard to flour quality, removal of the omega-1,2 gliadins from the flour resulted in flour with increased mix times and tolerances. There was also a small increase in SDS sedimentation volume in the absence of these proteins. Taken together, this suggests that the omega-1,2 gliadins have a negative effect on flour mixing properties. It should also be noted that the omega-1,2 gliadins in spots 3 and 4 in **Figure 3C** have been shown previously to be preferentially accumulated in small glutenin polymer fractions (Vensel et al., 2014). The omega-1,2 gliadins identified in these spots by MS/MS contain a single cysteine residue and likely function as chain terminators of the polymer. These proteins constitute about 20% of the omega-1,2 gliadins in Butte 86 and would be expected to reduce the size of the polymers and decrease dough strength. In similar studies using RNAi, omega-5 gliadins were also shown to have a negative effect on end-use quality. However, the omega-5 gliadins in Butte 86 do not contain any cysteine and are not associated with glutenin polymers (Altenbach et al., 2014b).

The mixing properties of flour from 118b-3 were diminished as evidenced by the flat mixing curve and the reduction in SDS sedimentation volumes. Perhaps this is not surprising given the absence of most gliadins and LMW-GS. Flat mixing curves were also observed by Gil-Humanes et al. (2014b) using a 35 g mixograph for transgenic lines with reduced levels of gliadins and, in some cases, LMW-GS. Many of their lines also showed reduced SDS sedimentation values. Nonetheless, Gil-Humanes et al. (2014a) demonstrated that the flour could be used to produce a reduced-gliadin bread that, while not optimal, was at least more acceptable than some of the current gluten-free bread options. Further studies are necessary to determine whether flour from 118b-3 could be used to produce a similar product. Alternately, flour from 118b-3 may prove useful as a base flour to test the effects of individual gluten protein components on flour functional properties. For example, specific types of LMW-GS or chain-terminating gliadins might be added in mixing studies to evaluate their roles on glutenin polymer formation. Additionally, different types and amounts of gliadins could be added to determine how the balance of gliadins and glutenins affects end-use quality. It may also be interesting to examine disulfide linkages between HMW-GS in transgenic line 118b-3 since most other gluten proteins are absent in this line.

While these studies provide new information about the role of specific gluten proteins in flour end-use quality and human health, an important question is whether these or similar lines could be of commercial value. Thus far, consumer attitudes have prevented the release of transgenic wheat in the marketplace both in the United States and abroad. Genome editing offers an alternative approach to RNAi and allows similar changes to be made to the flour proteome with the important advantage that the resulting plants are not considered transgenic and could be incorporated into conventional wheat breeding programs, at least in the United States. In some of the first genome editing studies, Sánchez-Léon et al. (2018) demonstrated success in using two single guide RNAs (sgRNAs) with homology to conserved regions of alpha gliadin genes to reduce both the levels of alpha gliadins and the immunoreactivity of the flour. Their study targeted the majority of alpha gliadin genes. However, the fact that only 20 bp of identical sequence are required for sgRNAs suggests that it should be possible to target individual gluten protein genes or small subgroups of genes provided that detailed sequence knowledge about the complete set of gluten protein genes in the wheat cultivar being modified is available. Given the complexity of the gluten proteins, the road ahead is not easy. The results nonetheless suggest that biotechnology approaches can be used in the future to improve the healthfulness of wheat, while maintaining or even improving its end-use qualities.

DATA AVAILABILITY

This manuscript contains previously unpublished data. The name of the repository and accession number are not available.

AUTHOR CONTRIBUTIONS

SA designed the study, analyzed the data, and wrote the manuscript. AA contributed to study design, data analysis, and writing of the manuscript. BS was responsible for end-use quality testing and interpretation of results. H-CC conducted genetic transformation experiments, analyses of flour proteins by 2-DE

and MS/MS analyses. XY was responsible for immunoblotting and ELISA experiments and interpretation of results. PG was responsible for subject recruitment and clinical characterization of patients. All authors contributed to editing of the manuscript and approved the manuscript.

FUNDING

This work was supported by United States Department of Agriculture, Agricultural Research Service CRIS 2030-21430-014-00D (to SA). Additional support was provided by the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant Number UL1 TR000040 (to AA).

ACKNOWLEDGMENTS

The authors thank Paul Allen and Julian DeLellis-Mitch for excellent technical support and Annamaria Simon-Buss, Leslie Harden, and Melissa Erickson-Beltran for expert advice and assistance with mass spectrometry work. Mention of a specific product by the USDA does not constitute an endorsement and does not imply a recommendation over other suitable products.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.00580/full#supplementary-material

REFERENCES

AACCI (1961). *Approved Methods of Analysis,* 11th Edn. Method 56-60.01. Sedimentation Test for Flour. St. Paul, MN: AACC International.

AACCI (1985). *Approved Methods of Analysis,* 11th Edn. Method 39-11.01. Near-Infrared Reflectance Method for Protein Determination in Wheat Flour. St. Paul, MN: AACC International.

AACCI (1988). *Approved Methods of Analysis,* 11th Edn. Method 26-10.02. Experimental Milling: Introduction, Equipment, Sample Preparation, and Tempering. St. Paul, MN: AACC International.

AACCI (1995). *Approved Methods of Analysis,* 11th Edn. Method 54-40.02. Mixograph Method. St. Paul, MN: AACC International.

Alaedini, A., and Green, P. H. (2005). Narrative review: celiac disease: understanding a complex autoimmune disorder. *Ann. Intern. Med.* 142, 289–298.

Altenbach, S. B., and Allen, P. V. (2011). Transformation of the US bread wheat 'butte 86' and silencing of omega-5 gliadin genes. *GM Crops* 2, 67–74. doi: 10.4161/gmcr.2.1.15884

Altenbach, S. B., Dupont, F. M., Kothari, K. M., Chan, R., Johnson, E. L., and Lieu, D. (2003). Temperature, water and fertilizer influence the timing of key events during grain development in a US spring wheat. *J. Cereal Sci.* 37, 9–20. doi: 10.1006/jcrs.2002.0483

Altenbach, S. B., and Kothari, K. M. (2007). Omega gliadin genes expressed in *Triticum aestivum* cv Butte 86: effects of post-anthesis fertilizer on transcript accumulation during grain development. *J. Cereal Sci.* 46, 169–177. doi: 10.1016/j.jcs.2007.02.001

Altenbach, S. B., Tanaka, C. K., and Allen, P. V. (2014a). Quantitative proteomic analysis of wheat grain proteins reveals differential effects of silencing of omega-5 gliadin genes in transgenic lines. *J. Cereal Sci.* 59, 118–125.

Altenbach, S. B., Tanaka, C. K., and Seabourn, B. W. (2014b). Silencing of omega-5 gliadins in transgenic wheat eliminates a major source of environmental variability and improves dough mixing properties of flour. *BMC Plant Biol.* 14:393. doi: 10.1186/s12870-014-0393-1

Altenbach, S. B., Tanaka, C. K., Pineau, F., Lupi, R., Drouet, M., Beaudouin, E., et al. (2015). Assessment of the allergenic potential of transgenic wheat (*Triticum aestivum*) with reduced levels of ω5-gliadins, the major sensitizing allergen in wheat-dependent exercise-induced anaphylaxis. *J. Agric. Food Chem.* 63, 9323–9332. doi: 10.1021/acs.jafc.5b03557

Altenbach, S. B., Vensel, W. H., and Dupont, F. M. (2010a). 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.* 10:7. doi: 10.1186/1471-2229-10-7

Altenbach, S. B., Vensel, W. H., and Dupont, F. M. (2010b). Integration of transcriptomic and proteomic data from a single wheat cultivar provides new tools for understanding the roles of individual alpha gliadin proteins in flour quality and celiac disease. *J. Cereal Sci.* 52, 143–151. doi: 10.1016/j.jcs.2010.04.006

Altenbach, S. B., Vensel, W. H., and Dupont, F. M. (2011). The spectrum of low molecular weight alpha-amylase/protease inhibitor genes expressed in the US bread wheat cultivar Butte 86. *BMC Res. Notes* 4:242. doi: 10.1186/1756-0500-4-242

Barro, F., Iehisa, J. C. M., Giménez, M. J., García-Molina, M. D., Ozuna, C. V., Comino, I., et al. (2015). Targeting of prolamins by RNAi in bread wheat: effectiveness of seven silencing-fragment combinations for obtaining lines devoid of coeliac disease epitopes from highly immunogenic gliadins. *Plant Biotech. J.* 14, 986–996. doi: 10.1111/pbi.12455

Battais, F., Courcoux, P., Popineau, Y., Kanny, G., Moneret-Vautrin, D. A., and Denery-Papini, S. (2005a). Food allergy to wheat: differences in immunoglobulin E-binding proteins as a function of age or symptoms. *J. Cereal Sci.* 42, 109–117. doi: 10.1016/j.jcs.2005.01.004

Battais, F., Mothes, T., Moneret-Vautrin, D. A., Pineau, F., Kanny, G., Popineau, Y., et al. (2005b). Identification of IgE-binding epitopes on gliadins for patients with food allergy to wheat. *Allergy* 60, 815–821. doi: 10.1111/j.1398-9995.2005.00795.x

Christensen, A. H., and Quail, P. H. (1996). Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Res.* 5, 213–218. doi: 10.1007/bf01969712

Denery-Papini, S., Bodinier, M., Larré, C., Brossard, C., Pineau, F., Triballeau, S., et al. (2012). Allergy to deamidated gluten in patients tolerant to wheat: specific epitopes linked to deamidation. *Allergy* 67, 1023–1032. doi: 10.1111/j.1398-9995.2012.02860.x

Dupont, F. M., Vensel, W., Encarnacao, T., Chan, R., and Kasarda, D. D. (2004). Similarities of omega gliadins from *Triticum urartu* to those encoded on chromosome 1A of hexaploid wheat and evidence for their post-translational processing. *Theor. Appl. Genet.* 108, 1299–1308. doi: 10.1007/s00122-003-1565-9

Dupont, F. M., Vensel, W. H., Tanaka, C. K., Hurkman, W. J., and Altenbach, S. B. (2011). Deciphering the complexities of the wheat flour proteome using quantitative two-dimensional electrophoresis, three proteases and tandem mass spectrometry. *Proteome Sci.* 9:10. doi: 10.1186/1477-5956-9-10

Egidi, E., Sestili, F., Janni, M., D'Ovidio, R., Lafiandra, D., Ceriotti, A., et al. (2014). An asparagine residue at the N-terminus affects the maturation process of low molecular weight glutenin subunits of wheat endosperm. *BMC Plant Biol.* 14:64. doi: 10.1186/1471-2229-14-64

Gil-Humanes, J., Pistón, F., Altamirano-Fortoul, R., Real, A., Comino, I., Sousa, C., et al. (2014a). Reduced-gliadin wheatbread: an alternative to the gluten-free diet for consumers suffering gluten-related pathologies. *PLoS One* 9:e90898. doi: 10.1371/journal.pone.0090898

Gil-Humanes, J., Pistón, F., Barro, F., and Rosell, C. M. (2014b). Shutdown of celiac disease-related gliadin epitopes in bread wheat by RNAi provides flours with increased stability and better tolerance to over-mixing. *PLoS One.* 9:e91931. doi: 10.1371/journal.pone.0091931

Gil-Humanes, J., Pistón, F., Tollefsen, S., Sollid, L. M., and Barro, F. (2010). Effective shutdown in the expression of celiac disease-related wheat gliadin

T-cell epitopes by RNA interference. *Proc. Natl. Acad. Sci. U.S.A.* 107, 17023–17028. doi: 10.1073/pnas.1007773107

Huebener, S., Tanaka, C. K., Uhde, M., Zone, J. J., Vensel, W. H., Kasarda, D. D., et al. (2015). Specific nongluten proteins of wheat are novel target antigens in celiac disease humoral response. *J. Proteome Res.* 14, 503–511. doi: 10.1021/pr500809b

Huo, N., Zhang, S., Zhu, T., Dong, L., Mohr, T., Hu, T., et al. (2018a). Gene duplication and evolution dynamics in the homeologous regions harboring multiple prolamin and resistance gene families in hexaploid wheat. *Front. Plant Sci.* 9:673. doi: 10.3389/fpls.2018.00673

Huo, N., Zhu, T., Altenbach, S., Dong, L., Wang, Y., Mohr, T., et al. (2018b). Dynamic evolution of α-gliadin prolamin gene family in homeologous genomes of hexaploid wheat. *Sci. Rep.* 8:5181. doi: 10.1038/s41598-018-23570-5

International Wheat Genome Sequencing Consortium [IWGSC] (2018). Shifting the limits in wheat research and breeding through a fully annotated and anchored reference genome sequence. *Science* 361:eaar7191.

Lau, N. M., Green, P. H. R., Taylor, A. K., Hellberg, D., Ajamian, M., Tan, C. Z., et al. (2013). Markers of celiac disease and luten sensitivity in children with autism. *PLoS One* 8:e66155. doi: 10.1371/journal.pone.0066155

Moeller, S., Canetta, P. A., Taylor, A. K., Arguelles-Grande, C., Snyder, H., Green, P. H., et al. (2014). Lack of serologic evidence to link IgA nephropathy with celiac disease or immune reactivity to gluten. *PLoS One* 9:e94677. doi: 10.1371/journal.pone.0094677

Morita, E., Matsuo, H., Mihara, S., Morimoto, K., Savage, A. W. J., and Tatham, A. S. (2003). Fast ω-gliadin is a major allergen in wheat-dependent exercise-induced anaphylaxis. *J. Derm Sci.* 33, 99–104. doi: 10.1016/s0923-1811(03)00156-7

Müntz, K., Blattner, F. R., and Shutov, A. D. (2002). Legumains – a family of asparagine-specific cysteine endopeptidases involved in propolypeptide processing and protein breakdown in plants. *J. Plant Physiol.* 159, 1281–1293. doi: 10.1078/0176-1617-00853

Nunes-Miranda, J. D., Bancel, E., Viala, D., Chambon, C., Capelo, J. L., Branlard, G., et al. (2017). Wheat glutenin: the "tail" of the 1By protein subunits. *J. Proteomics* 169, 136–142. doi: 10.1016/j.jprot.2017.05.019

Samaroo, D., Dickerson, F., Kasarda, D. D., Green, P. H. R., Briani, C., Yolken, R. H., et al. (2010). Novel immune response to gluten in individuals with schizophrenia. *Schizophr. Res.* 118, 248–255. doi: 10.1016/j.schres.2009.08.009

Sánchez-Léon, S., Gil-Humanes, J., Ozuna, C. V., Giménez, M. J., Sousa, C., Voytas, D., et al. (2018). Low-gluten, nontransgenic wheat engineered with CRISPR/Cas9. *Plant Biotech. J.* 16, 902–910. doi: 10.1111/pbi.12837

Scherf, K. A., Koehler, P., and Wieser, H. (2016). Gluten and wheat sensitivities – An overview. *J. Cereal Sci.* 67, 2–11. doi: 10.1016/j.jcs.2015.07.008

Senthil-Kumar, M., and Mysore, K. S. (2011). 744, 13–25. doi: 10.1007/978-1-61779-123-9_2

Shan, L., Molberg, O., Parrot, I., Hausch, F., Filia, F., Gray, G. M., et al. (2002). Structural basis for gluten intolerance in celiac sprue. *Science* 297, 2275–2279. doi: 10.1126/science.1074129

Sollid, L. M., Qiao, S.-W., Anderson, R. P., Gianfrani, C., and Koning, F. (2012). Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenet* 64, 455–460. doi: 10.1007/s00251-012-0599-z

Tye-Din, J. A., Stewart, J. A., Dromey, J. A., Beissbarth, T., van Heel, D. A., Tatham, A., et al. (2010). Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Sci. Transl. Med.* 2:41ra51. doi: 10.1126/scitranslmed.3001012

Uhde, M., Ajamian, M., Caio, G., DeGiorgio, R., Indart, A., Green, P. H., et al. (2016). Intestinal cell damage and systemic immune activation in individuals reporting sensitivity to wheat in the absence of coeliac disease. *Gut* 65, 1930–1937. doi: 10.1136/gutjnl-2016-311964

Vensel, W. H., Tanaka, C. K., and Altenbach, S. B. (2014). Protein composition of wheat gluten polymer fractions determined by quantitative two-dimensional gel electrophoresis and tandem mass spectrometry. *Proteome Sci.* 12:8. doi: 10.1186/1477-5956-12-8

Wang, D. W., Li, D., Wang, J., Zhao, Y., Wang, Z., Yue, G., et al. (2017). Genome-wide analysis of complex wheat gliadins, the dominant carriers of celiac disease epitopes. *Sci. Rep.* 7:44609. doi: 10.1038/srep44609

Zimin, A. V., Puiu, D., Hall, R., Kingan, S., Clavijo, B. J., and Salzberg, S. L. (2017). The first near-complete assembly of the hexaploid bread wheat genome *Triticum aestivum*. *GigaScience* 6, 1–7. doi: 10.1093/gigascience/gix097

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 16 May 2019
doi: 10.3389/fpls.2019.00649

Hordein Accumulation in Developing Barley Grains

Gregory J. Tanner[1†], Michelle L. Colgrave[2,3†], Malcolm J. Blundell[4], Crispin A. Howitt[4] and Antony Bacic[1,5]*

[1] School of Biosciences, University of Melbourne, Melbourne, VIC, Australia, [2] Commonwealth Scientific and Industrial Research Organisation Agriculture and Food, St Lucia, QLD, Australia, [3] School of Science, Edith Cowan University, Joondalup, WA, Australia, [4] Commonwealth Scientific and Industrial Research Organisation Agriculture and Food, Canberra, ACT, Australia, [5] La Trobe Institute for Agriculture and Food, La Trobe University, Bundoora, VIC, Australia

The temporal pattern of accumulation of hordein storage proteins in developing barley grains was studied by enzyme-linked immunosorbent assay (ELISA), western blot and liquid chromatography tandem mass spectrometry (LC-MS/MS). Hordein accumulation was compared to the pattern seen for two abundant control proteins, serpin Z4 (an early accumulator) and lipid transferase protein (LTP1, a late accumulator). Hordeins were detected from 6 days post-anthesis (DPA) and peaked at 30 DPA. Changes in fresh weight indicate that desiccation begins at 20 DPA and by 37 DPA fresh weight had decreased by 35%. ELISA analysis of hordein content, expressed on a protein basis, increased to a maximum at 30 DPA followed by a 17% decrease by 37 DPA. The accumulation of 39 tryptic and 29 chymotryptic hordein peptides representing all classes of hordein was studied by LC-MS/MS. Most peptides increased to a maximum at 30 DPA, and either remained at the maximum or did not decrease significantly. Only five tryptic peptides, members of the related B1- and γ1-hordeins decreased significantly by 21–51% at 37 DPA. Thus, the concentration of some specific peptides was reduced while remaining members of the same family were not affected. The N-terminal signal region was removed by proteolysis during co-translation. In addition to a suite of previously characterized hordeins, two novel barley B-hordein isoforms mapping to wheat low molecular weight glutenins (LMW-GS-like B-hordeins), and two avenin-like proteins (ALPs) sharing homology with wheat ALPs, were identified. These identified isoforms have not previously been mapped in the barley genome. Cereal storage proteins provide significant nutritional content for human consumption and seed germination. In barley, the bulk of the storage proteins comprise the hordein family and the final hordein concentration affects the quality of baked and brewed products. It is therefore important to study the accumulation of hordeins as this knowledge may assist plant breeding for improved health outcomes (by minimizing triggering of detrimental immune responses), nutrition and food processing properties.

Keywords: hordeins, accumulation, developing barley grain, multiple reaction monitoring mass spectrometry, gluten

OPEN ACCESS

Edited by:
Stefanie Wienkoop,
University of Vienna, Austria

Reviewed by:
Pasquale Ferranti,
University of Naples Federico II, Italy
Zhaojun Wei,
Hefei University of Technology, China

****Correspondence:***
Gregory J. Tanner
gregory.tanner@unimelb.edu.au;
read_right@grapevine.com.au

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Proteomics, a section of the journal Frontiers in Plant Science

Received: *15 December 2018*
Accepted: *30 April 2019*
Published: *16 May 2019*

Citation:
Tanner GJ, Colgrave ML, Blundell MJ, Howitt CA and Bacic A (2019) Hordein Accumulation in Developing Barley Grains. Front. Plant Sci. 10:649. doi: 10.3389/fpls.2019.00649

INTRODUCTION

Prolamins, are rich in proline and glutamine residues, and this is a collective name given to the alcohol-soluble, water-insoluble storage proteins that exist in wheat gluten (gliadin and glutenin), barley (hordein), rye (secalins), and oat (avenin). The barley hordeins consist of four closely related protein families that are categorized by molecular weight: the D-hordeins, a 105 kDa protein family coded for by a single gene with up to five post-translationally modified isoforms (Gu et al., 2003); the C-hordeins which are 55 and 65 kDa sulfur-poor proteins, coded for by 20–30 genes (Shewry et al., 1985a,b); the B-hordeins which are a group of sulfur-rich proteins running at 50 kDa, coded for by at least 13 genes, with at least two protein sub-families, the B1- and B3-hordeins (Kreis et al., 1983; Anderson, 2013); and the sulfur-rich gamma (γ)-hordeins comprising at least three isoforms of 35–45 kDa. The four hordein protein families are coded for by Hor-1 (C-hordeins), Hor-2 (B-hordeins), Hor-3 (D-hordeins), and Hor-4 (γ-hordeins) loci, located on barley chromosome 1H. The B-hordeins account for 70–90% of total hordeins; the C-hordeins form 10–30% of the hordein fraction; the gamma-hordeins and the D-hordeins are minor components accounting for 1–2, and 2–4%, respectively, of the hordeins (Shewry et al., 1985b).

The hordeins are amongst the triggers of coeliac disease (CD), a well characterized T-cell mediated disorder suffered by approximately 1% of most populations (Fasano et al., 2003; Lebwohl et al., 2018). In CD the immune system mounts an inappropriate reaction to particular peptide sequences in dietary gluten, reacting as if the gluten molecules were an invading microorganism (Anderson et al., 2000). Life-long avoidance of gluten remains the only treatment option for coeliacs. Other adverse reactions to gluten also exist including gluten intolerance (Biesiekierski et al., 2011, 2013; Choung et al., 2017), which affects approximately 10% of the population (Golley et al., 2015), and gluten allergy, a serious, rapid, IgE mediated allergy impacting 1% of the population (Snegaroff et al., 2006).

Hordeins accumulate in the starchy endosperm cells of developing barley grains, during grain filling (Roustan et al., 2018). Hordein synthesis proceeds linearly from approximately 10 to 30 days post-anthesis (DPA) (Brandt, 1976; Sorensen et al., 1989). Hordeins are coordinately expressed on the polyribosomes of the rough endoplasmic reticulum (RER) of starchy endosperm cells, and ∼20 amino acid N-terminal transit peptides are removed during co-translational processing and transport into the ER lumen (Cameron-Mills et al., 1978). Hordeins are then transported from the ER lumen to the central vacuole during maturation (Rechinger et al., 1993; Cameron-Mills et al., 1994; Ibl et al., 2014). The pathway taken by hordeins, from protein body to vacuole is not clear. It has been proposed that hordeins either pass through or by-pass the Golgi network (Bechtel et al., 1991; Muntz, 1998; Gomord et al., 1999; Bethke and Jones, 2000; Pereira et al., 2014).

Gluten proteins have traditionally been measured using ELISA, however, great care must be taken to match the calibration standard with the protein that is being measured (Tanner et al., 2013a). While ELISA may be suitable for total gluten determination in either unprocessed grain or raw ingredients, multiple reaction monitoring mass spectrometry (MRM-MS) is capable of identifying and quantifying individual gluten proteins (Tanner et al., 2013b). MRM-MS is a method where prototypic peptides representing all hordein families are detected and quantified (Colgrave et al., 2012, 2014, 2015, 2016b). In this study we have used both ELISA and MRM-MS to study the pattern of hordein accumulation during grain development to further our understanding of hordein function *in planta*. A fundamental understanding of hordein accumulation in the grain is required for plant breeding applications that aim to either improve the nutritional status of barley or reduce the coeliac reactivity, but may also apply to food and beverage processing.

MATERIALS AND METHODS

Plant Material

Barley cv Sloop was obtained from the Australian Grains Genebank, Department of Environment and Primary Industries (DEPI) Horsham, and germinated in a 50/50 (v/v) mixture of soil (Debco seed raising mix, Tyabb Victoria) and perlite, three plants per 20 cm pot, and grown at constant temperature of 19–24°C, under ambient light with 12 h daylight extension provided by 1500 W halogen lights at 400 μE for approximately 8 weeks until flowering. Plants were watered with a balanced nutrient solution (250 mL per pot of 2.5 g/L Aquasol, Yates Australia Padstow) once every week. Anthesis was determined by daily inspection and was taken as the first day that the anthers in the middle of the head dehisced. In practice this was when the head was about half extended from the flag leaf. Under these conditions, barley cv Sloop flowers first in the middle two grains and then the flowering spreads up and down the head over several days.

Antibodies

Rabbit polyclonal anti-peptide antibodies to LTP1 (lab designation V6177) and serpin Z4 (V6175) were produced by Genscript (Piscataway, United States) from antigenic peptides identified within LTP1 (P07597.1; D_{33}LHNQAQSSGDRQT$_{46}$) and serpin Z4 (P06293.2; R_{258}LSTEPEFIENHIP$_{271}$) respectively, as described in the Supplementary Material of Tanner et al. (2016). Anti-hordein MAbs (lab designation B4 and 23-3) were raised against C-TQQQLQQEQVGQ and C-SFLRPHISQQNS, respectively as in Tanner et al. (2016).

Total Protein Determination of Grains

Four replicates of two grains were harvested and snap frozen in liquid nitrogen (LN_2) on the indicated DPA and stored at −80°C until required. Grains were quickly weighed without thawing to determine fresh weight, and then ground in a mortar and pestle under LN_2, to an ice powder. Either 1 mL (6, 8, 10 DPA) or 2 mL (15, 20, 30, 37 DPA) of extraction buffer containing 8 M urea, 1% (w/v) DTT, 20 mM triethylamine-HCL (termed "Urea/DTT") and 1/1000 dilution of Sigma plant protease inhibitor (all adjusted to pH 6) was added. The mixture was ground as an ice slurry and allowed to thaw and centrifuged at 15,000 g for 5 min. The supernatants were aliquoted and frozen in LN_2 and could be

thawed and refrozen repeatedly without losing antigenicity or SDS-PAGE performance (Tanner et al., 2013a). Samples were reserved for liquid chromatography tandem mass spectrometry (LC-MS/MS), western blot and ELISA analysis. Total protein was determined by dye-binding (Bradford, 1976).

Enzyme-Linked Immunosorbent Analysis (ELISA)

Extracts from grains at the indicated DPA were diluted with ELISA Systems sample diluent and added to ELISA wells (ELISA Systems, Brisbane, Australia). ELISA plates were processed according to manufacturers' instructions. Urea/DTT extracts were diluted 1/1000 with PBST and an appropriate aliquot (50 μL of 6, 8, 10 DPA; or 10 μL of 15, 20, 25, 30, 37 DPA) diluted to 100 μL and calibrated against standard curve of 10–75 ng total hordein extracted from barley cv Sloop and expressed as total hordein (mg/g fresh weight). Total hordein was prepared by adding 20 mg flour into a bead beater to which 20 mg glass beads (0.1 mm, Edwards Instruments, Sydney) plus a stainless steel 10 mm ball bearing in 0.5 mL of 50% (v/v) isopropyl alcohol (IPA), plus 1% (w/v) DTT (IPA/DTT) was added and then extracted for 30 s at a frequency of 1/30 s^{-1}, centrifuged at 15,000 g for 5 min, and the process repeated and supernatants pooled. The protein content of the IPA/DTT supernatant was measured (Bradford, 1976) and an aliquot containing 1 mg of hordein was freeze dried and re-dissolved in 1 mL Urea/DTT to yield a 1.0 mg/mL total hordein standard.

SDS-PAGE

The required aliquot of protein solution was diluted with at least one volume of 6 M urea, 2% (w/v) SDS, 1% (w/v) DTT, 62.5 mM Tris–HCl (pH 6.8), 0.2% (w/v) Bromophenol Blue (termed "Urea/SDS") at room temperature and loaded on NuPAGE Bis-Tris 4–12% 1 mm gels (Thermo Fisher Scientific) and run at 200 V for 60 min. Protein bands were calibrated against pre-stained protein standards which were in turn calibrated against unstained protein standards (Invitrogen).

Western Blot

For each primary antibody, duplicate gels were run, with 2 μg of protein per lane, taken from the first and second replicate extractions respectively. Gels were rinsed in distilled water and blotted to nitrocellulose membrane using iBLOT2 (Thermo Fisher Scientific) using 20 V for 1 min; 23 V for 4 min; 25 V for 2 min. The blotted membranes were blocked overnight at 4°C in 1% (w/v) Tween 20, 5% (w/v) Diploma skimmed milk powder in PBST, rinsed and exposed to primary and secondary antibodies as follows. For detection of hordeins, Sigma polyclonal anti-gliadin-HRP was added at a ratio of 1:1000 with incubation for 1 h, washed with PBST for 3 × 10 min and imaged in Amersham ECL reagent. This commercial antibody has previously been shown to detect all gluten and hordein families (Colgrave et al., 2015). For detection of LTP1, anti-LTP (V6177) was added at a ratio of 1:1000 with incubation for 1 h, followed by 3 × 10 min PBST washes, then secondary Amersham donkey anti-rabbit-HRP added at a ratio of 1:1000 with incubation for 1 h, followed by 3 × 10 min PBST washes, and addition of ECL reagent as described above. For serpin detection, anti-serpin Z4 (V6175) was added at a ratio of 1:1000 with incubation for 1 h, followed by 3 × 10 min PBST washes, followed by secondary Amersham donkey anti-rabbit-HRP added at a ratio of 1:1000 with incubation for 1 h, followed by 3 × 10 min PBST washes, and addition of ECL reagent as described above. Representative blots are shown.

Hordein Mass Spectrometry

Protein Digestion

Protein extracts (100 μg protein, n = 4) were digested as previously described (Colgrave et al., 2017). In brief, protein was applied to a 10 kDa molecular weight cut-off filter (Millipore, Australia), and washed with two 200 μL of 8 M urea, 100 mM Tris–HCl (pH 8.5) with centrifugation (20,800×g, 10 min). For cysteine alkylation, 100 μL of 100 mM iodoacetamide in 8 M urea, 100 mM Tris–HCl was added and incubated at ambient temperature in the dark for 30 min. The filters were centrifuged (20,800×g, 10 min) to remove excess iodoacetamide and washed with two 200 μL volumes of 8 M urea, 100 mM Tris–HCl. The buffer was exchanged using 100 mM ammonium bicarbonate (pH 8.0) by two consecutive wash/centrifugation steps. Sequencing grade porcine trypsin or bovine chymotrypsin (Sigma-Aldrich, Australia) at a concentration of 250 μg/mL in 100 mM ammonium bicarbonate with 1 mM $CaCl_2$ (200 μL) was added to the protein on the 10 kDa filters and incubated for 16 h at 37°C in a wet chamber. The filters were transferred to fresh centrifuge tubes and the filtrate (digested peptides) were collected by centrifugation (20,800×g, 10 min). The filters were washed with 200 μL of 100 mM ammonium bicarbonate and the filtrates were combined and lyophilized. The tryptic peptides were resuspended in 100 μL of 1% formic acid and stored at 4°C until analysis.

Global Proteomic Profiling

Gluten-enriched fractions (5 μL; corresponding to 5 μg extracted protein) were analyzed as described previously (Colgrave et al., 2014) with chromatographic separation using a nano HPLC system (Shimadzu Scientific, Rydalmere, Australia) directly coupled to a TripleTOF 5600 MS (SCIEX, Redwood City, CA, United States). ProteinPilot™ 5.0 software (SCIEX) with the Paragon Algorithm (Shilov et al., 2007) was used for protein identification. Tandem mass spectrometry data collected in this study was searched against the Poaceae subset of the Uniprot database (version 2018/08; 1,693,876 sequences). The search parameters were defined as iodoacetamide modified for cysteine alkylation and either trypsin or chymotrypsin as the digestion enzyme. Modifications and cleavages were defined previously (Colgrave et al., 2014). The database search results were manually curated to yield the protein identifications (**Supplementary Tables S1**, **S2**) using a 1% global false discovery rate (FDR) determined by the in-built FDR tool within ProteinPilot software (Tang et al., 2008).

Identification of Prototypic Peptides

Peptide summaries generated by ProteinPilot (**Supplementary File S1**) were used to select peptides that yielded intense peaks and were fully tryptic or chymotryptic, i.e., no unusual or missed cleavages. The peptides were subjected to BLAST searching using the Uniprot BLASTp server limited to the taxonomy Poaceae (**Supplementary Table S3**). MRM transitions were determined for each peptide where the precursor ion (Q1) *m/z* and the fragment ion (Q3) *m/z* values were determined from the data collected in the discovery experiments. The peptides identified in the discovery experiments in this study were added to existing MRM methods as previously described (Colgrave et al., 2016b). Three transitions were used per peptide, with 39 tryptic and 29 chymotryptic peptides measured, wherein the three MRM transitions were required to co-elute and the peak areas of all three were summed (**Supplementary Table S4**).

Targeted MS

Reduced and alkylated tryptic peptides (5 μL, corresponding to 5 μg extracted protein) were chromatographically separated on a Shimadzu Nexera UHPLC and analyzed on a 6500 QTRAP mass spectrometer (SCIEX) as described previously (Colgrave et al., 2014). Quantification was achieved using scheduled MRM scanning experiments using a 40 s detection window for each MRM transition and a 0.3 s cycle time. Peaks were integrated using MultiQuant v3.0 (SCIEX) wherein all three transitions were required to co-elute at the same retention time (RT, min) with a signal-to-noise (S/N) > 3 for detection and a S/N > 5 for quantification. Graphs were generated in Graphpad Prism v6.

Phylogenetic Analysis

In total, 22 hordein-like protein sequences identified in the present study were aligned by MUSCLE[1], and subsequently phylogenetic analysis was performed in MEGA X software (Kumar et al., 2018), using the neighbor-joining method (Saitou and Nei, 1987).

RESULTS

Temporal Accumulation of Hordeins

The changes during barley grain development can be observed as the increase in the fresh weight of the grains over time, from a low at 6 DPA to a maximum at 20 DPA and then followed by a decrease of 35% at 37 DPA after which the grains are almost entirely desiccated (**Figure 1A**).

Hordein Determination by ELISA

The accumulation of hordeins was measured by ELISA, calibrated against an appropriate standard consisting of total hordein purified from barley cv Sloop. Measurement by ELISA shows a steady accumulation in hordein level from near zero levels at 6 DPA to a maximum at 30 DPA. The trends are similar when expressed on either a fresh weight basis or as a proportion of total protein. However, as the seed desiccates beyond this stage, the hordein level as a percentage of total protein decreases by 17% (**Figure 1B**). The level of hordein on a fresh weight basis decreases more significantly at 37 DPA (by 49%) compared to the peak value observed at 30 DPA (**Figure 1B**).

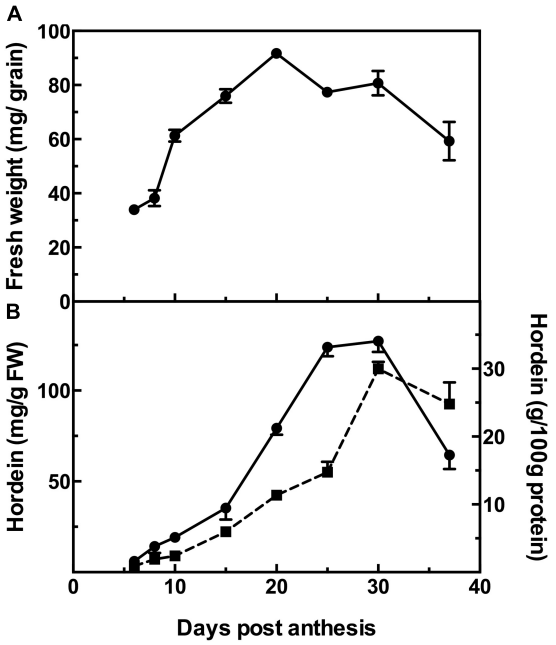


FIGURE 1 | Fresh weight of developing barley grain **(A)**; and hordein content by ELISA **(B)**, calibrated against barley cv Sloop hordein (mg/g fresh weight, FW: ●, left *y*-axis), or as g/100 g protein (■, right *y*-axis). Mean (n = 4) ± SE are shown; no error bars are shown when SE < symbol size.

Characterizing the Gluten in the Developing Endosperm

Grains at different stages of development were sampled, frozen and extracted as described above and subsequently subjected to cysteine modification (reduction and alkylation) and digestion using either trypsin or chymotrypsin. The resultant peptide solutions were analyzed by LC-MS/MS and the spectral datasets were searched against the Poaceae subset of the Uniprot-KB database. The protein and peptide summaries were curated to yield the gluten components. The protein identifications in barley cv Sloop (**Figure 2**) include a suite of B-hordeins, three C-hordeins, a single D-hordein and five γ-hordeins. Additionally, two avenin-like proteins (ALPs: F2EGD5 and M0VEH1) sharing homology with the C-terminal region of the γ-hordeins were identified. These ALPs contain a high proportion of glutamine (Gln; 21–25%) as observed in other gluten families (∼30%), but a lower proportion (∼7%) of proline (Pro) compared to ∼14% present in other hordein families. The ALPs are typically smaller (∼17 kDa) than the γ-hordeins (∼32–33 kDa). Notably, two of the isoforms identified within the B-hordein family included peptides mapping to wheat protein accessions (R9YTM4 and B9VUV5). The issue with protein inference (mapping peptides to proteins) is exacerbated in

[1]http://www.ebi.ac.uk/Tools/msa/muscle/

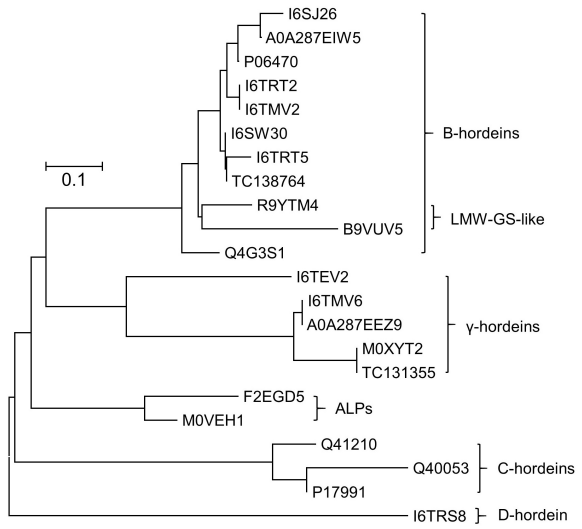


FIGURE 2 | Neighbor-joining phylogenetic tree showing the evolutionary relationship between hordein-like proteins present in *H. vulgare*. All sequences were retrieved from Uniprot database, except two proteins with a "TC" prefix which came from the TIGR database. Sequences were aligned using MUSCLE, and the analyzed in MEGAX (Kumar et al., 2018). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances that were used to infer the phylogenetic tree (scale bar, 0.1 amino acid substitutions per site). Two avenin-like proteins (ALPs) that share homology with the γ-hordeins were identified. Two novel B-hordein isoforms were identified based on peptides that map to wheat low molecular weight glutenins (LMW-GS), demonstrating that despite being absent in the barley genome, they were present in the barley lines studied.

the absence of complete genomes leading to the identification of orthologous proteins from related species as has occurred here. While the sequence coverage achieved for these protein isoforms was not complete, the peptides identified included amino acid substitutions that were absent in the barley genome. The peptides mapped to a central region of the B-hordeins that is commonly identified by proteomic studies employing trypsin (Colgrave et al., 2017). In total six peptides mapping to this region, starting with the conserved sequence "VFLQQQC," were identified proving the existence of at least six B-hordeins in cv Sloop. These novel proteins were termed low molecular weight-glutenin-subunit-like B-hordeins (LMW-GS B-hordein). Likewise, one B-hordein (TC138764) and one γ-hordein (TC131355) were identified using translated protein sequences from the TIGR database (Ouyang and Buell, 2004). Since both Golgi and non-Golgi transport routes of hordeins have been proposed the peptide spectral data was investigated for evidence of glycosylation. Both D-hordein and several γ-hordeins contain N-glycosylation consensus sequences (Asn-Xxx-Ser/Thr). Despite the predictions, no evidence of glycosylation was found indicating it was unlikely that the hordeins were transported to the vacuole via the Golgi bodies.

Western Blot

The effect of maturity on the hordein content of the developing endosperm was firstly examined by western blotting. The accumulation of the abundant grain proteins, the serpins and lipid transfer protein 1 (LTP1), was also studied. These proteins were intended as controls, and in immuno-localisation experiments described in a subsequent publication. The serpins were detected as three isoforms on western blots and the apparent MW calculated relative to protein standards: serpin Z7 (47.3 kDa), native serpin Z4 (44.3 kDa), and as hydrolyzed serpin ZX (minus the 4 kDa active loop, 41.6 kDa) (Roberts et al., 2003). These measured MW correspond to serpin Z7 (Q434392), serpin Z4 (P06293) and serpin ZX (Q40066) with theoretical MW of 42.7, 43.3, and 42.9 kDa, respectively. Serpin Z4 was detectable from 6 DPA, but unlike hordeins the bulk of serpin Z4 accumulation occurred only after mid-development, beyond 15 DPA and reached maximum level at 30 DPA (**Figure 3A**). Serpin Z4 reaches a maximum at 37 DPA, and does not decrease as some hordeins do, in the later stages of maturity. The remaining control protein LTP1 accumulated late in seed development (**Figure 3B**) and was not detectable until 15 DPA and accumulated to a maximum at about 30 DPA. Examination of these same proteins using data-dependent LC-MS (data not shown) revealed the same pattern of protein expression with LTP1 (UniProt: P07597) accumulating after 15 days with a 4-fold increase from 25 to 30 DPA. Serpin Z4 (UniProt: P06293) showed a gradual increase from 6 to 30 DPA. Hordeins accumulated early in seed development and were detectable on western blots by 6 DPA (**Figure 3C**) and the intensity of protein bands increased to a maximum at about 25–30 DPA. The different hordein families, the B-, C-, D-, and γ-hordeins appeared to accumulate synchronously. The ALP, and LMW-GS B-hordein are expected to run on an SDS-PAGE at 17 kDa and 34–40 kDa, respectively, but were not obvious in western blots.

Mass Spectrometric Analysis

The accumulation of individual hordein isoforms was examined by LC-MRM-MS/MS, using peptide markers representing each of the hordein families (Colgrave et al., 2016b). Most, but not all hordein, prototypic peptides measured by MS rise to a peak at 30 DPA and decreased slightly by 37 DPA. There were some exceptions and it appears that sequence-specific post-translational processing reduced the concentration of some, but not all, isoforms even within the same hordein family.

Several identified proteins do not appear in the barley protein databases but had high amino acid sequence similarity to hordeins (**Figure 2**). The genetic inter-relationships were revealed by phylogenetic analysis that shows the ALPs are more closely related to the γ-hordeins while the isoforms mapping to wheat LMW-GS proteins were closely related to the barley B-hordein family (**Figure 2**). These novel proteins generally followed the same trend observed with the other hordeins.

The peptides identified in this study and in previous analyses of barley cv Sloop (Colgrave et al., 2016a) were mapped to the predicted DNA sequences to firstly generate the protein sequence coverage and secondly to confirm that the N-terminal transit peptides were removed from all hordeins (**Supplementary Figure S1**). Transit peptides were invariably 19 amino acids long for the B-hordeins, ending in amino acids TIA except the two ALPs, which ended with AVA and VQS (**Table 1**).

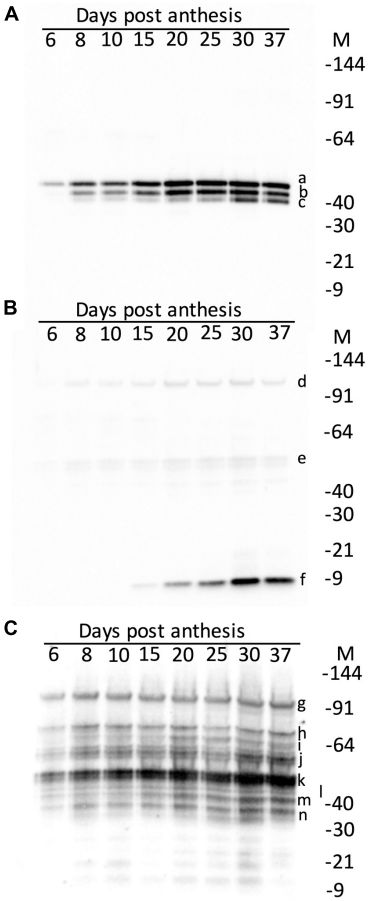


FIGURE 3 | Western blots of developing endosperm showing the effect of maturity (days post-anthesis, DPA) on: **(A)** serpins showing serpin Z7 (**a**, 47.3 kDa), serpin Z4 (**b**, 44.3 kDa) and serpin Z4 minus 4 kDa active loop (**c**, 41.6 kDa); **(B)** LTP1 showing trace artefactual 10-mer (**d**, 97.0 kDa), and 5-mer (**e**, 48.5 kDa) and mature LTP (**f**, 9.7 kDa); and **(C)** hordeins showing D-hordein (**g**, 93.9 kDa), C-hordeins (**h**, 70.5; **i**, 63.7; and **j**, 55.6 kDa), B-hordein (**k**, 47.8 kDa) partly obscuring γ1-hordein (**l**, 45.0 kDa), γ2-hordein (**m**, 40.0 kDa), γ3-hordein (**n**, 38.0 kDa). The relative molecular weights (given in parentheses, in kDa) were determined by calibration against Invitrogen pre-stained standards, which were in turn calibrated against Invitrogen unstained standards.

The C- and D-hordein transit peptides were either 20 or 21 residues long, respectively, ending in TTA. The γ-1-hordein transit peptides were all 19 residues long ending in ATS, while the single γ-3 hordein transit peptide was 15 residues long, ending with ATA. The sequences of three proteins (TC138764, Q40020, and P06471) did not start with MK residues and did not code for transit peptides. It is likely that these sequences represent incomplete protein sequences (protein fragments) rather than processed full-length proteins.

The protein sequence coverage of hordeins as determined by comparison of LC-MS/MS from this study together with data previously generated, was nearly complete in every case (**Supplementary Figure S1**) allowing unequivocal identity of hordein isoforms in barley cv Sloop.

TABLE 1 | Hordein transit peptide sequences.

Hordein family	Uniprot accession	Amino acid sequence
ALP	F2EGD5| M0VKM6	MKTMLILALIAFAATSAVA
B1-hordein	I6SJ22	MKTFLIFALLAIAATSTIA
B3-hordein	I6SJ26	MKTFLIFALLAIVATSTIA
LMW-GS-like B-hordein	B9VUV5| R9YTM4	MKTFLVFALLAVAATSAIA
C-hordein	Q41210| Q40053	MKTFLTFVLLAMVMSIVTTA
D-hordein	I6TRS8	MAKRLVLFVAVIVALVALTTA
Gamma-1-hordein	I6TMV6	MKILIILTILAMATTFATS
Gamma-1-hordein	M0XYT2	MKILIILIILAMATSFATS
Gamma-3-hordein	I6TEV2	MKIFLLFSLLGVATA

A selection of peptide markers that represented each class of barley gluten were monitored in the grains across the development timeline (DPA). In general, the peptides specific to the different hordein isoforms followed a similar pattern to the total hordein level determined by ELISA, increasing gradually to a maximum at day 30 followed by a slight decrease by 37 DPA. **Figure 4** shows the tryptic peptide markers that represent the ALPs, B-hordeins, D-hordein and γ-hordeins. The C-hordeins contain few trypsin cleavage sites and are not well represented after trypsin digestion (Colgrave et al., 2017) and so the accumulation of C-hordeins was followed using chymotryptic peptides (**Figure 5**).

The D-hordeins increase to a maximum at 37 DPA (**Figure 4D**) and the avenin-like proteins, B1/B3-hordeins, and γ-1-hordeins increase to a maximum at 25–30 DPA and then decrease by approximately 10% by 37 DPA (**Figures 4A,B,C,E**). The chymotryptic peptides follow a similar pattern (**Figures 5A,B,D,E**). However, γ-3-hordein increases gradually to a maximum at 37 DPA and does not decrease when monitoring either tryptic or chymotryptic peptides (**Figures 4F**, **5F**, respectively). Similarly, the C-hordeins accumulate to a maximum at 37 DPA and do not decrease (**Figure 5C**).

In the tryptic peptide data one of the ALPs (F2EGD5) increased to a significant maximum at 30 DPA (**Supplementary Figure S2A**), whereas another ALP isoform (M0VEH1) showed low level expression that did not significantly increase after 8 DPA (**Supplementary Figure S2B**). Generally, most but not all peptides derived from the same protein behave in the same manner. An exception is shown for the B1-hordeins, wherein two peptides, I6TRT2-T1 and -T3, both increased to a maximum value at 37 DPA, whereas the other two peptides in this protein (I6TRT2-T2 and -T4) decreased significantly after 25 DPA (**Supplementary Figure S3**). The B-hordein peptides monitored are not unique to a single B-hordein isoform (**Supplementary Table S3**) and as such the different patterns of expression may be the result of cumulatively monitoring more than one protein isoform wherein the isoforms have different protein expression profiles. Good agreement is noted between the two peptides from Q4G3S1 with a non-significant decrease in the level seen beyond 25 DPA (**Supplementary Figures S3F,G**). Three of the four B3-hordein peptide markers that map to I6SW30 increased to a maximum at 30 DPA with a subtle, but not significant decrease to 37 DPA (**Supplementary Figure S4**). The remaining peptide

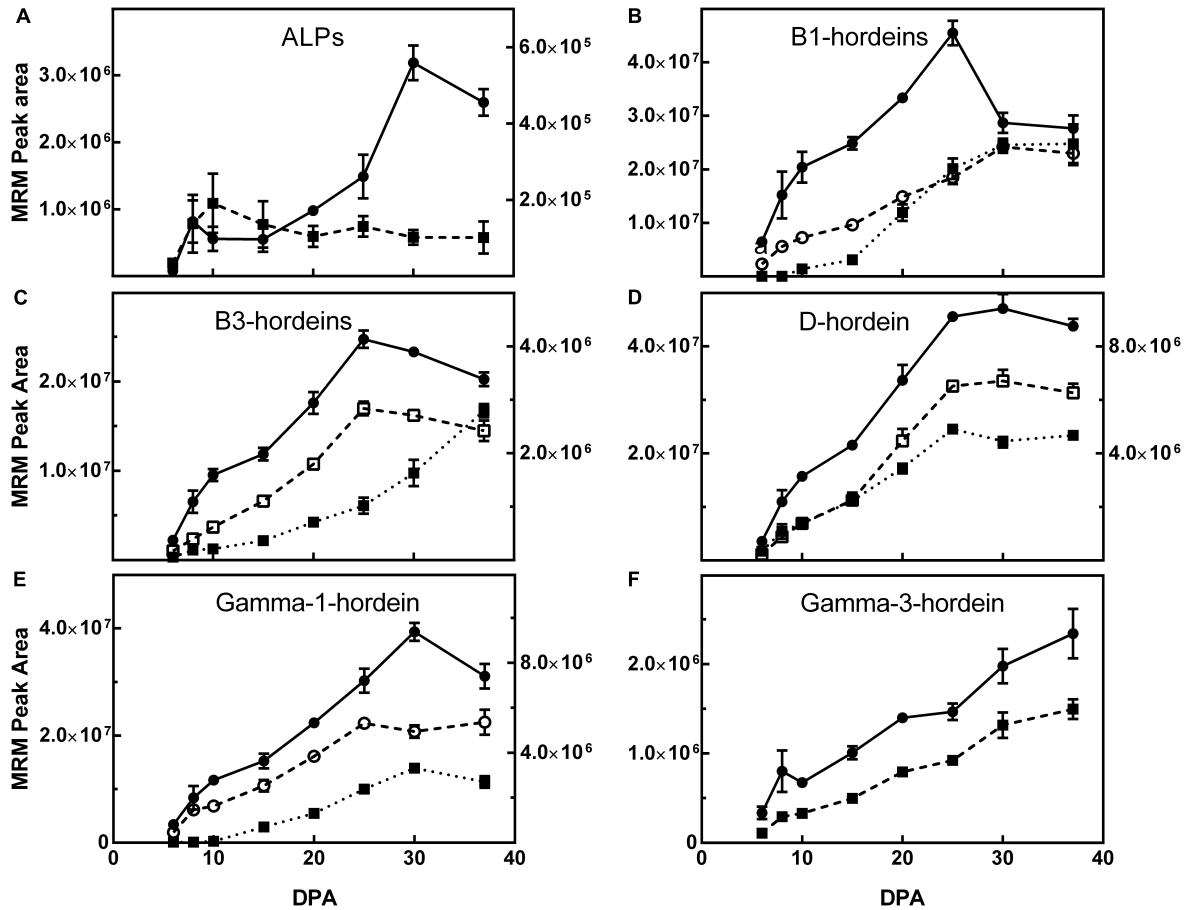


FIGURE 4 | The effect of maturity (days post-anthesis, DPA) on the accumulation of representative hordein family-specific peptides following trypsin cleavage. The mean multiple reaction monitoring (MRM) peak area ± SE (n = 3) from 5 μg extracted protein is shown for peptides mapping to: **(A)** avenin-like proteins (ALPs) F2EGD5-T1 (●, left y-axis) and M0VEH1-T1 (■, right y-axis); **(B)** B1-hordeins I6TRT2-T1 (○),I6TRT2-2 (●), I6TRT2-3 (■), and; **(C)** B3-hordeins Q4G3S1-1 (□, left y-axis), Q4G3S1-2 (●, left y-axis), and I6SJ26-T1 (■, right y-axis); **(D)** D-hordein I6TRS8-T1 (●, left y-axis), I6TRS8-T4 (□, right y-axis), and I6TRS8-T5 (■, right y-axis); **(E)** γ1-hordein I6TMV6-T1 (●, left-y axis), TC131355-T1 (■, right y-axis), I6TMV6-T5 (○, left y-axis); and **(F)** γ3-hordein I6TEV2-T1 (●), I6TEV2-T2 (■). The sequence of proteins and peptide markers are presented in **Supplementary Table S4**. For clarity, symbols of significance are not shown, but where points differ by 2 × SE, they are significantly different. Detailed statistical comparisons are shown in **Supplementary Figures**.

(**Supplementary Figure S4A**) which is notably common to I6SJ26 (I6SJ26-T1) increased to a significant maximum at 37 DPA in agreement with a second peptide mapping to the same isoform (I6SJ26-T2), and the third peptide I6SJ26-T3 decreased beyond 30 DPA. The two peptides mapping to low molecular weight glutenin subunit (LMW-GS: B9VUV5 and R9YTM4) isoforms from wheat, that shared homology with the B-hordeins, show different patterns of protein expression with B9VUV5 reaching a peak at 25 DPA (like B-hordein Q4G3S1) whereas R9YTM4 reached a significant peak at 37 DPA (**Supplementary Figure S5**).

The six D-hordein peptides all increased reaching a maximum between 25 and 37 DPA (**Supplementary Figure S6**). The variation noted between 25 and 37 DPA was not significant. The five γ-1-hordein (I6TMV6) peptides all reached a maximum at 30 DPA and decrease slightly to 37 DPA (**Supplementary Figure S7**). The remaining γ-1-hordein peptides, including the four M0XYT2 peptides and the single TC131355 peptide, increase to a maximum at 37 DPA (**Supplementary Figure S8**), whereas the single A0A287EEZ9 peptide peaked at 30 DPA. γ -3-Hordein specific peptides I6TEV2-T1 and I6TEV2-T2 both increased to a maximum at 37 DPA (**Supplementary Figure S9**). All C-hordein specific chymotryptic peptides increased to a significant maximum at either day 30 or 37 DPA (**Supplementary Figure S10**).

The small, but statistically significant different decreases in protein expression at 37 DPA indicate that sequence-specific differences occur in the accumulation of hordein proteins. Similar patterns are seen in a detailed comparison of the chymotryptic peptides (**Figure 5**).

DISCUSSION

It is assumed that extraction of different proteins by a vigorous solvent such as urea/DTT is similar. That many extracted hordeins measured by LC-MS/MS increase to a maximum at

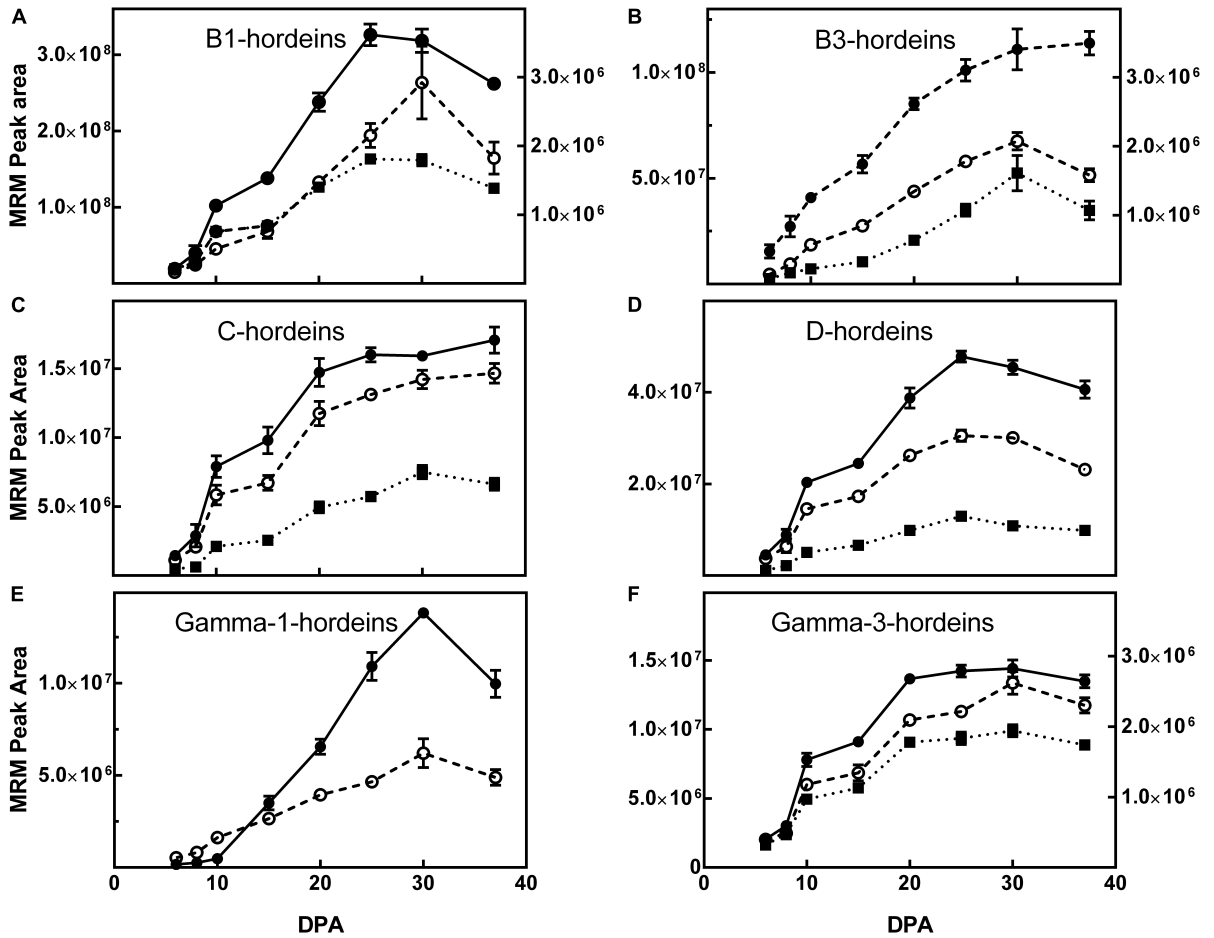


FIGURE 5 | The effect of maturity (days post-anthesis, DPA) on the accumulation of representative hordein family-specific peptides following chymotrypsin cleavage. The mean multiple reaction monitoring (MRM) peak area ± SE (n = 3) from 5 μg extracted protein is shown for peptides mapping to: **(A)** B1-hordeins P06470-C1 (○, right y-axis), P06470-C2 (■, left y-axis), and P06470-C3 (●, right y-axis); **(B)**: B3-hordeins I6SW30-C1 (●, left y-axis), I6SJ26-C1 ■, right y-axis), I6SJ26-C2 (○, left y-axis); **(C)** C-hordeins Q40053-C1 (○), Q40053-C3 (■) and Q41210-C4 (●); **(D)** D-hordein I6TRS8-C1 (●), I6TRS8-C5 (■) and I6TRS8-C6 (○); **(E)** γ1-hordein TC131355-C1 (○), γ1-hordein I6TMV6-C1 (●); **(F)** γ3-hordein I6TEV2-C1 (○, left y-axis), γ3 I6TEV2-C2 (●, right y-axis), and γ3 I6TEV2-C3 (■, right y-axis). The sequence of the proteins and peptide markers are shown in **Supplementary Table S4**. Where the SE is less than symbol size it is not shown. For clarity, symbols of significance are not shown but where points differ by 2 × SE they are significantly different. Detailed statistical comparisons are shown in **Supplementary Figures**.

37 DPA and do not decrease confirms that this assumption is correct. This is shown by the minor hordeins, the D-hordeins (**Figure 4D**), and the γ-3-hordein (**Figure 4F**). The C-hordeins, represent about 30% of total hordein and also increase to a maximum (**Figure 5C**). This confirms that the observed decrease in hordein level from 30 to 37 DPA measured by ELISA (**Figure 1B**) was genuine and not due to a failure to extract protein from the partially desiccated seeds.

Results of ELISA analysis of hordein content, expressed on a protein basis, increased to a maximum at 30 DPA followed by a 17% decrease by 37 DPA. The ELISA Systems kit used the Skerritt antibody, which is selective and detects D- and C-hordeins at least 50× more sensitively than γ- and B-hordeins with half-maximal signals given by 57, 84, 3640, and 19400 ppb hordein, respectively (Tanner et al., 2013a), however, it is unlikely that a change in relative composition is responsible for the decrease observed by ELISA. LC-MS/MS results of 39 tryptic and 29 chymotryptic hordein peptides showed most peptides increased from 6 DPA to a maximum at 30 DPA, often followed by a slight decrease to 37 DPA. ANOVA analysis showed these small decreases were not statistically significant, however, taken together they most likely account for the observed decrease at 37 DPA measured by ELISA.

Some hordeins behave differently compared to other members of the same protein family – either accumulating earlier or decreasing before other family members. This implies fine differential regulation of the expression of hordein genes, and is consistent with a recent study that showed that different classes of hordein transcripts had slightly different expression patterns in developing endosperm (Vinje et al., 2019). Although the reason for this is not apparent it has been observed for protein storage genes in other cereals. For example, a novel family of gliadin genes localized to the wheat group 1 chromosomes (1A, 1B, 1D), and with homology to hordeins was significantly upregulated by nitrogen levels during grain development (Wan et al., 2013).

Prolamin gene expression is tissue specific and developmentally regulated, and also sensitive to nitrogen and sulfur nutrition of the grain (Shewry and Halford, 2002).

The biological function of hordeins is unclear, nitrogen storage and involvement in protein trafficking have been suggested (Rechinger et al., 1993; Cameron-Mills et al., 1994). They account for a significant proportion of seed protein, up to 55% (Schalk et al., 2017), and are mobilized during germination. It has been suggested they provide a source of nitrogen for the germinating seed. However, in ULG 2.0, a hordein double null line obtained by combining Risø 56 (no B-hordein) and Risø 1508 (no C-hordeins), the B- and C-hordeins do not accumulate, with only 5% of the total hordein remaining (Tanner et al., 2010) yet this line does not suffer impaired germination compared to wild type cv Sloop (Tanner and Howitt, 2007). Hordeins may play a role in regulating disease resistance as the hordeins have a distant relationship to the amylase inhibitors and serpins of barley, but again ULG 2.0 flour does not preferentially support the growth of bacterial pathogens compared to wild type cv Sloop (Tanner and Howitt, 2007).

The approach used here of analyzing specific proteins by LC-MRM-MS/MS is highly sensitive and selective and hence is applicable to the rapid selection of elite lines. For example, it has been used to produce barley lines carrying significantly reduced hordein content in segregating populations of hordein triple null lines (Tanner et al., 2016). The method may be generalized to the rapid analysis and selection of lines with increased or decreased expression of any protein of agricultural significance.

CONCLUSION

Total hordein content detected by ELISA and LC-MS/MS, increased co-ordinately from 6 DPA to a peak level at 30 DPA. Five peptides, members of the B1- and γ1-hordeins decreased significantly by 37 DPA. The majority of hordein peptides including the remaining B1-, B3-, γ-1-, γ-3-, C and D-hordeins increase to a maximum and then either remain high or do not decrease significantly. Hordein accumulation was compared to two other abundant proteins which also accumulate during grain development, LTP1 (a late accumulator, detectable by 15 DPA) and serpin Z4 (an early accumulator, detectable by 6 DPA). In all cases the N-terminal transit peptide, coded for by the hordein genes was not observed in the mature proteins, confirming that these sequences were removed during transit into the RER. Small, but statistically significant, differences in the pattern of accumulation at 37 DPA indicating some sequence-specific differences occur in the accumulation of B-hordein proteins. Similar patterns were observed in a detailed comparison of the protein expression using chymotryptic peptides. Two novel barley B-hordein isoforms were detected mapping to the wheat LMW-GS proteins. From the LC-MS/MS data we can conclude that the pattern of accumulation of these proteins was similar to the bulk of the hordeins. The lack of evidence of hordein N-glycosylation indicated that it was unlikely that the hordeins were transported to the vacuole via the Golgi bodies.

AUTHOR CONTRIBUTIONS

GT and MC carried out the experimental work. All authors wrote the manuscript and contributed to the manuscript revision, read and approved the submitted version.

ACKNOWLEDGMENTS

The authors thank Mr. Alun Jones of the Molecular and Cellular Proteomics Mass Spectrometry Facility at the University of Queensland for access to some of the mass spectrometers used in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.00649/full#supplementary-material

REFERENCES

Anderson, O. D. (2013). The B-hordein prolamin family of barley. *Genome* 56, 179–185. doi: 10.1139/gen-2013-0016

Anderson, R. P., Degano, P., Godkin, A. J., Jewell, D. P., and Hill, A. V. S. (2000). In vivo antigen challange in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope. *Nat. Med.* 6, 337–342. doi: 10.1038/73200

Bechtel, D. B., Wilson, J. D., and Shewry, P. R. (1991). Immunocytochemical localization of the wheat storage protein triticin in developing endosperm tissue. *Cereal Chem.* 68, 573–577.

Bethke, P. C., and Jones, R. L. (2000). Vacuoles and prevacuolar compartments. *Curr. Opin. Plant Biol.* 3, 469–475. doi: 10.1016/s1369-5266(00)00115-1

Biesiekierski, J. R., Muir, J. G., and Gibson, P. R. (2013). Is gluten a cause of gastrointestinal symptoms in people without celiac disease? *Curr. Allergy Asthma Rep.* 13, 631–638. doi: 10.1007/s11882-013-0386-4

Biesiekierski, J. R., Newnham, E. D., Irving, P. M., Barrett, J. S., Haines, M., Doecke, J. D., et al. (2011). Gluten causes gastrointestinal symptoms in subjects without celiac disease: a double-blind randomized placebo-controlled trial. *Am. J. Gastroenterol.* 106, 508–514. doi: 10.1038/ajg.2010.487

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principle of dye binding. *Anal. Chem.* 72, 248–254. doi: 10.1006/abio.1976.9999

Brandt, A. (1976). Endosperm protein formation during kernel development of wild type and a high-lysine barley mutant. *Cereal Chem.* 53, 890–901.

Cameron-Mills, V., Ingversen, J., and Brandt, A. (1978). Transfer of in vitro synthesized barley endosperm proteins into the lumen of the endoplasmic reticulum. *Carlsb. Res. Commun.* 43, 91–102. doi: 10.1007/bf02906506

Cameron-Mills, V., Rechinger, K. B., and Simpson, D. J. (1994). The intracellular trafficking of hordein polypeptides in barley endosperm cells. *J. Cell. Biochem.* 104. doi: 10.1007/bf02906506

Choung, R. S., Larson, S. A., Khaleghi, S., Rubio-Tapia, A., Ovsyannikova, I. G., King, K. S., et al. (2017). Prevalence and morbidity of undiagnosed celiac disease from a community-based study. *Gastroenterology* 152, 830.e5–839.e5. doi: 10.1053/j.gastro.2016.11.043

Colgrave, M. L., Byrne, K., Blundell, M., Heidelberger, S., Lane, C. S., Tanner, G. J., et al. (2016a). Comparing multiple reaction monitoring and sequential window acquisition of all theoretical mass spectra for the relative quantification of barley gluten in selectively bred barley lines. *Anal. Chem.* 88, 9127–9135. doi: 10.1021/acs.analchem.6b02108

Colgrave, M. L., Byrne, K., Blundell, M., and Howitt, C. A. (2016b). Identification of barley-specific peptide markers that persist in processed foods and are capable of detecting barley contamination by LC-MS/MS. *J. Proteo.* 147, 169–176. doi: 10.1016/j.jprot.2016.03.045

Colgrave, M. L., Byrne, K., and Howitt, C. A. (2017). Food for thought: selecting the right enzyme for the digestion of gluten. *Food Chem.* 234, 389–397. doi: 10.1016/j.foodchem.2017.05.008

Colgrave, M. L., Goswami, H., Blundell, M., Howitt, C. A., and Tanner, G. J. (2014). Using mass spectrometry to detect hydrolysed gluten in beer that is responsible for false negatives by ELISA. *J. Chromatogr. A* 1370, 105–114. doi: 10.1016/j.chroma.2014.10.033

Colgrave, M. L., Goswami, H., Byrne, K., Blundell, M., Howitt, C. A., and Tanner, G. J. (2015). Proteomic profiling of 16 cereal grains and the application of targeted proteomics to detect wheat contamination. *J. Proteo. Res.* 14, 2659–2668. doi: 10.1021/acs.jproteome.5b00187

Colgrave, M. L., Goswami, H., Howitt, C. A., and Tanner, G. J. (2012). What is in a beer? proteomic characterization and relative quantification of hordein (gluten) in beer. *J. Proteo. Res.* 11, 386–396. doi: 10.1021/pr2008434

Fasano, A., Berti, I., Gerarduzzi, T., Not, T., Colletti, R. B., Drago, S., et al. (2003). Prevalence of celiac disease in at-risk and not-at-risk groups in the United States - a large multicenter study. *Arch. Inter. Med.* 163, 286–292.

Golley, S., Corsini, N., Topping, D., Morell, M., and Mohr, P. (2015). Motivations for avoiding wheat consumption in Australia: results from a population survey. *Public Health Nutr.* 18, 490–499. doi: 10.1017/S1368980014000652

Gomord, V., Wee, E., and Faye, L. (1999). Protein retention and localization in the endoplasmic reticulum and the Golgi apparatus. *Biochimie* 81, 607–618. doi: 10.1016/s0300-9084(99)80118-7

Gu, Y. Q., Anderson, O. D., Londeore, C. F., Kong, X. Y., Chibbar, R. N., and Lazo, G. R. (2003). Structural organization of the barley D-Hordein locus in comparison with its orthologous regions of wheat genomes. *Genome* 46, 1084–1097. doi: 10.1139/g03-071

Ibl, V., Kapusi, E., Arcalis, E., Kawagoe, Y., and Stoger, E. (2014). Fusion, rupture, and degeneration: the fate of in vivo-labelled PSVs in developing barley endosperm. *J. Exp. Bot.* 65, 3249–3261. doi: 10.1093/jxb/eru175

Kreis, M., Rahman, S., Forde, B. G., Pywell, J., Shewry, P. R., and Miflin, B. J. (1983). Sub-families of hordein messenger-rna encoded at the Hor-2 locus of barley. *Mol. Gen. Genet.* 191, 194–200. doi: 10.1007/bf00334813

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Lebwohl, B., Sanders, D. S., and Green, P. H. R. (2018). Coeliac disease. *Lancet* 391, 70–81.

Muntz, K. (1998). Deposition of storage proteins. *Plant Mol. Biol.* 38, 77–99. doi: 10.1007/978-94-011-5298-3_4

Ouyang, S., and Buell, C. R. (2004). The TIGR plant repeat databases: a collective resource for the identification of repetitive sequences in plants. *Nucleic Acids Res.* 32, D360–D363.

Pereira, C., Pereira, S., and Pissarra, J. (2014). Delivering of proteins to the plant vacuole-an update. *Int. J. Mol. Sci.* 15, 7611–7623. doi: 10.3390/ijms15057611

Rechinger, K. B., Simpson, D. J., Svendsen, I., and Cameron-Mills, V. (1993). A role for gamma3 hordein in the transport and targeting of prolamin polypeptides to the vacuole of developing barley endosperm. *Plant J.* 4, 841–853. doi: 10.1046/j.1365-313x.1993.04050841.x

Roberts, T. H., Marttila, S., Rasmussen, S. K., and Hejgaard, J. (2003). Differential gene expression for suicide-substrate serine proteinase inhibitors (serpins) in vegetative and grain tissues of barley. *J. Exp. Bot.* 54, 2251–2263. doi: 10.1093/jxb/erg248

Roustan, V., Roustan, P. J., Weidinger, M., Reipert, S., Kapusi, E., Shabrangy, A., et al. (2018). Microscopic and proteomic analysis of dissected developing barley endosperm layers reveals the starchy endosperm as prominent storage tissue for er-derived hordeins alongside the accumulation of barley protein disulfide isomerase (HvPDIL1-1). *Front. Plant Sci.* 9:1248. doi: 10.3389/fpls.2018.01248

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Schalk, K., Lexhaller, B., Koehler, P., and Scherf, K. A. (2017). Isolation and characterization of gluten protein types from wheat, rye, barley and oats for use as reference materials. *PLoS One* 12:e0172819. doi: 10.1371/journal.pone.0172819

Shewry, P. R., Bunce, N. A., Kreis, M., and Forde, B. G. (1985a). Polymorphism at the Hor 1 locus of barley (*Hordeum vulgare* L.). *Biochem. Genet.* 23, 391–404. doi: 10.1007/bf00499082

Shewry, P. R., Kreis, M., Parmar, S., Lew, E. J. L., and Kasarda, D. D. (1985b). Identification of γ-type hordeins in barley. *FEBS Lett.* 190, 61–64. doi: 10.1016/0014-5793(85)80427-0

Shewry, P. R., and Halford, N. G. (2002). Cereal seed storage proteins: structures, properties and role in grain utilization. *J. Exp. Bot.* 53, 947–958. doi: 10.1093/jexbot/53.370.947

Shilov, I. V., Seymour, S. L., Patel, A. A., Loboda, A., Tang, W. H., Keating, S. P., et al. (2007). The paragon algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol. Cell. Proteo.* 6, 1638–1655. doi: 10.1074/mcp.t600050-mcp200

Snegaroff, J., Bouchez-Mahiout, I., Pecquet, C., Branlard, G., and Lauriere, M. (2006). Study of ige antigenic relationships in hyper sensitivity to hydrolyzed wheat proteins and wheat-dependent exercise-induced anaphylaxis. *Int. Arch. Allergy Immunol.* 139, 201–208. doi: 10.1159/000091165

Sorensen, M. B., Cameronmills, V., and Brandt, A. (1989). Transcriptional and post-transcriptional regulation of gene-expression in developing barley endosperm. *Mol. General Genet.* 217, 195–201. doi: 10.1046/j.1365-313X.1998.00249.x

Tang, W. H., Shilov, I. V., and Seymour, S. L. (2008). Nonlinear fitting method for determining local false discovery rates from decoy database searches. *J. Proteo. Res.* 7, 3661–3667. doi: 10.1021/pr070492f

Tanner, G. J., Blundell, M. J., Colgrave, M. L., and Howitt, C. A. (2013a). Quantification of Hordeins by ELISA: the correct standard makes a magnitude of difference. *PLoS One* 8:e56456. doi: 10.1371/journal.pone.0056456

Tanner, G. J., Colgrave, M. L., Blundell, M. J., Goswami, H. P., and Howitt, C. A. (2013b). Measuring Hordein (Gluten) in beer - a comparison of elisa and mass spectrometry. *PLoS One* 8:e56452. doi: 10.1371/journal.pone.0056452

Tanner, G. J., Blundell, M. J., Colgrave, M. L., and Howitt, C. A. (2016). Creation of the first ultra-low gluten barley (*Hordeum vulgare* L.) for coeliac and gluten-intolerant populations. *Plant Biotechnol. J.* 14, 1139–1150. doi: 10.1111/pbi.12482

Tanner, G. J., and Howitt, C. A. (2007). *Barley with Low Levels of Hordein*. U.S. Patent No WO2009021285. Melbourne: FB RICE & CO.

Tanner, G. J., Howitt, C. A., Forrester, R. I., Campbell, P. M., Tye-Din, J. A., and Anderson, R. P. (2010). Dissecting the T-cell response to hordeins in coeliac disease can develop barley with reduced immunotoxicity. *Aliment. Pharmacol. Therapeu.* 32, 1184–1191. doi: 10.1111/j.1365-2036.2010.04452.x

Vinje, M. A., Walling, J. G., Henson, C. A., and Duke, S. H. (2019). Comparative gene expression analysis of the β-amylase and hordein gene families in the developing barley grain. *Gene* 693, 127–136. doi: 10.1016/j.gene.2018.12.041

Wan, Y., Shewry, P. R., and Hawkesford, M. J. (2013). A novel family of gamma-gliadin genes are highly regulated by nitrogen supply in developing wheat grain. *J. Exp. Bot.* 64, 161–168. doi: 10.1093/jxb/ers318

Conflict of Interest Statement: GT, CH, and MC are authors on patents related to gluten reduction by plant breeding.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Nutrition

REVIEW
published: 28 June 2019
doi: 10.3389/fnut.2019.00097

Detection and Quantitation of Gluten in Fermented-Hydrolyzed Foods by Antibody-Based Methods: Challenges, Progress, and a Potential Path Forward

Rakhi Panda and Eric A. E. Garber*

Division of Bioanalytical Chemistry, Office of Regulatory Science, Center for Food Safety and Applied Nutrition, United States Food and Drug Administration, College Park, MD, United States

Celiac disease (CD) affects ∼1 in 141 individuals in the United States, requiring adherence to a strict gluten-free diet. The Codex Standard and the European Commission states that gluten level of gluten-free foods must not exceed 20 ppm. The FDA requires food bearing the labeling claim "gluten-free" to contain <20 ppm gluten. Accurate quantitation of gluten in fermented-hydrolyzed foods by antibody-based methods is a challenge due to the lack of appropriate reference materials and variable proteolysis. The recent uses of proteases (e.g., proline endopeptidases or PEP) to hydrolyze immunopathogenic sequences of gluten proteins further complicates the quantitation of immunopathogenic gluten. The commercially available antibody-based methods routinely used to detect and quantitate gluten are not able to distinguish between different hydrolytic patterns arising from differences in fermentation processes. This is a severe limitation that makes accurate quantitation and, ultimately, a detailed evaluation of any potential health risk associated with consuming the food difficult. Utilizing gluten-specific antibodies, a recently developed multiplex-competitive ELISA along with western blot analysis provides a potential path forward in this direction. These complimentary antibody-based technologies provide insight into the extent of proteolysis resulting from various fermentation processes and have the potential to aid in the selection of appropriate hydrolytic calibration standards, leading to accurate gluten quantitation in fermented-hydrolyzed foods.

Keywords: gluten, fermentation, quantitation, competitive ELISA, hydrolysis, peptides

OPEN ACCESS

Edited by:
Michelle Lisa Colgrave, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia

Reviewed by:
Crispin Alexander Howitt, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia
Venkatesan Arul, Pondicherry University, India

****Correspondence:***
Rakhi Panda
rakhi.panda@fda.hhs.gov

Specialty section:
This article was submitted to Nutrition and Food Science Technology, a section of the journal Frontiers in Nutrition

Received: *06 May 2019*
Accepted: *13 June 2019*
Published: *28 June 2019*

Citation:
Panda R and Garber EAE (2019) Detection and Quantitation of Gluten in Fermented-Hydrolyzed Foods by Antibody-Based Methods: Challenges, Progress, and a Potential Path Forward. Front. Nutr. 6:97. doi: 10.3389/fnut.2019.00097

INTRODUCTION

Celiac disease (CD) is an immune mediated enteropathy triggered by the interaction of the prolamin and glutelin fractions of proteins from wheat, barley, and rye with the intestinal mucosa of sensitive individuals (1). Upon ingestion, proteases in the gastrointestinal tract degrade gluten proteins into peptides, which undergoes deamidation by transglutaminase. Subsequently, these peptides interact with human leukocyte antigen (HLA)-DQ2 or -DQ8 molecules evoking a T cell response, resulting in inflammation in the small intestine (2, 3). Gluten can be fractionated into

alcohol soluble prolamins and the alcohol insoluble glutelins. The wheat prolamins, gliadins, are monomeric proteins with molecular weight ranging from 30 to 50 kDa and can be classified into α/β, γ, and ω-type. The wheat glutelins, glutenins, can be divided into high molecular weight (HMW) glutenins with molecular weights of 66–88 kDa, and low molecular weight (LMW) glutenins with molecular weights falling in the range of the gliadin proteins, ∼32–45 kDa (4, 5). A typical feature of gluten T cell stimulating peptides is their high proline content. Proline constitutes 12–17% of gluten. The abundance of proline residues in gluten makes them highly resistant to complete proteolytic degradation in the human gastrointestinal track (6, 7).

Approximately 1 in 141 people in the US are affected by CD and adherence to a strict gluten-free diet is the only option to prevent inflammatory symptoms in sensitive individuals (8, 9). In 2013, the FDA issued a regulation defining and allowing the use of the term gluten-free for food that "does not contain an ingredient that is a gluten-containing grain (e.g., spelt wheat); an ingredient that is derived from a gluten-containing grain and that has not been processed to remove gluten (e.g., wheat flour); or an ingredient that is derived from a gluten-containing grain and that has been processed to remove gluten (e.g., wheat starch), if the use of that ingredient results in the presence of 20 parts per million (ppm) or more gluten in the food [i.e., 20 milligrams (mg) or more gluten per kilogram (kg) of food]; or inherently does not contain gluten; and that any unavoidable presence of gluten in the food is below 20 ppm gluten (i.e., below 20 mg gluten per kg of food)." It was further "recognized that some food matrices, such as fermented or hydrolyzed foods, may lack currently available scientifically valid methods that can be used to accurately determine if these foods contain ≥20 ppm gluten" (10). Recognizing the unique problems associated with the accurate detection and quantitation of gluten in fermented foods, a regulation regarding the use of gluten-free label for fermented, hydrolyzed, and distilled foods was proposed in 2015 (11).

Several qualitative and quantitative analytical methods are used for the detection and quantitation of gluten in foods. The strengths and limitations of each method have been summarized in **Table 1** (12–15). Enzyme-linked immunosorbent assays (ELISAs) are currently the most popular method used to detect and quantitate gluten in foods. Most commercial ELISAs for gluten quantitation employ monoclonal antibodies such as Skerritt, R5 and G12. A polyclonal antibody against gluten proteins is also available from the Morinaga Institutes of Biological Sciences, Inc., (MIoBS). The Skerritt antibody was raised against wheat gliadin and has been shown to recognize the HMW glutenins (16–18). The R5 antibody was raised against rye secalin and strongly binds to the QQPFP, QQQFP, LQPFP, and QLPFP epitopes in α-/β-, ω-, and γ-gliadins (19, 20). The G12 antibody was produced against a synthetic 33-mer (LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF) of α2-gliadin, believed to invoke immunopathogenicity and the antibody recognizes the QPQLPY epitope of the peptide (21, 22). Recently, a novel monoclonal antibody that recognizes deamidated gliadin was generated by Pi Bioscientific Inc. (23, 24). Detection and quantitation of intact gluten has been routinely performed using sandwich ELISAs (16–19, 21, 25–29). There are questions related to accuracy of the results with respect to antibody specificity, extraction procedure, lack of suitable reference materials as well as of scientific data to support the underlying assumptions for calculating the gluten content, that has been extensively reviewed in several previous publications (9, 30–32).

TABLE 1 | Common analytical techniques used for detection of gluten in foods.

Common gluten detection techniques	Strengths	Limitations
Sandwich ELISA	- Commercially available - Specific - Sensitive - Robust - Quantitative analysis of intact gluten is possible	- Not suitable for quantitation of fermented-hydrolyzed gluten - Lack of certified reference materials limit the accuracy of the results
Competitive ELISA	- Commercially available - Appropriate for fermented-hydrolyzed gluten	- Usually less sensitive and robust compared to sandwich ELISA - Appropriate calibrant is needed for accurate analysis results
Immunosensors/ Dipsticks/Lateral flow devices (LFDs)	- User friendly - Rapid analysis - Useful for on-site analysis - Commercially available	- Usually qualitative or semi-quantitative
Western blots	- Separates and detects gluten proteins according to their size - Can be used as a confirmatory technique for ELISA	- Less sensitive compared to ELISAs - Not commercially available - Requires expertise - Usually qualitative/semi-quantitative
Mass spectrometry	- Highly sensitive - Can directly detect proteins/peptides that are not detected by immunological techniques - Quantitative analysis is possible	- Expensive equipments - Requires expertise - Similar to the ELISAs need certified reference materials for accurate quantitation - Depends on publicly available databases of wheat and barley proteins, which in most cases are incomplete or are poorly curated
DNA-based methods	- Stable analyte - DNA is more efficiently extracted compared to proteins - Can be used as a highly sensitive screening method for the presence of gluten containing cereals - Quantitative analysis is possible using quantitative real-time PCR (Q-PCR)	- Unsuitable for highly processed or fermented-hydrolyzed foods
Aptamer-based assays	- New generation methods - Highly sensitive	- Extensive validation studies are lacking in different food matrices

The reliable detection and accurate quantitation of gluten in fermented-hydrolyzed foods is another challenge that warrant further discussion. This review will discuss the challenges involved in the detection and quantitation of fermented-hydrolyzed gluten by antibody-based methods and a potential path forward in overcoming the challenges. Although significant progresses have been made by using mass spectrometry-based methods in this direction, this review will only discuss mass spectrometry-based methods briefly and will particularly focus on antibody-based methods.

DIFFERENT FERMENTATION PROCESSES AND GLUTEN PROTEIN/PEPTIDE PROFILE DIFFERENCES

Cereal-based fermented-hydrolyzed foods can be classified into different categories depending on the grain source, type of fermenting organism, and differences in the fermentation process. Wheat, rye and barley are commonly used in fermented-hydrolyzed foods such as beers, soy sauces, vinegars, and sourdough breads.

Beers

Beer is the most widely consumed alcoholic beverage made from celiac-toxic cereals, mainly barley, and wheat. During mashing, malting, and fermentation, the gluten is proteolyzed by enzymes, resulting in the formation of peptides. Gluten-derived peptides tend to remain in the final beer product and often contain immunopathogenic sequences (33). Studies have detected peptide fragments from the putative immunotoxic 33-mer of α2-gliadin in several wheat and barley beers produced by different manufacturing processes, indicating the resistance of this peptide to proteolytic cleavage during the production of beers (17, 34, 35). The susceptibility of different gluten proteins to proteolysis during fermentation varies, thereby generating a very diverse range of peptides. A study by Colgrave et al. (36), indicated that B-hordein and D-hordein are more susceptible to hydrolysis compared to γ-3 hordeins. In recent years, several mass spectrometry studies have detected and characterized gluten proteins/peptides in both wheat and barley beers (17, 34–43).

Soy Sauces

Soy sauces are popular fermented foods that are commonly used to impart flavor. Soy sauce is produced in a two-stage fermentation process of soybean and wheat. First, koji (a mold-covered mixture of soybeans and wheat) is generated, which is mixed with salt water to form moromi. The moromi is allowed to age for several months, during which fermentation is catalyzed by lactic acid bacteria and yeast (44). Several studies have indicated the absence of any gluten-derived peptides in soy-based sauces using ELISAs or serum IgE binding studies (45–47), which is consistent with the extensive proteolysis that occurs during soy sauce fermentation. Although soy sauces produced by classical fermentation may lack the presence of gluten derived proteins/peptides, any changes to the fermentation process, or ingredients used may alter the extent and type of proteolysis and, possibly, the immunopathogenicity. A study by Hefle et al. (48) indicated that some soy sauces contained 10–30% residual activity by means of RAST inhibition assays using sera from soy-allergic subjects (48). A recent western blot study indicated the presence of gluten-derived proteins/peptides in several soy-based sauces. Intact gluten was detected in a teriyaki sauce and gluten-derived peptides were detected in one soy sauce and two Worcestershire sauces. The exact quantity of gluten in these products could not be ascertained from the immunoblot data; however, the detection of gluten-derived proteinaceous materials in these products indicate the potential for immunopathgenicity (49).

Vinegars

Malt vinegars are produced by fermentation of cereals containing gluten, mostly barley and wheat. Vinegars made from distilled ethanol, are generally produced from non-gluten-containing raw material such as corn, beet, or sugar cane, but in some cases also gluten-containing cereals. The raw materials are typically processed in a manner that avoids the presence of any non-volatile compounds (e.g., gluten) from the finished product. However, exceptions to this occur when the distillation process is poorly performed. Thus, it is not uncommon to observe gluten peptides in some vinegars (44). Gluten peptides have been detected in vinegars both in western blot as well as mass spectrometry studies (43, 49, 50). Further, immunopathogenic epitopes in the HMW glutenin peptides derived from a malt vinegar have been reported. However, it is unclear whether the amount of glutenin present is sufficient to pose a health risk for celiac patients (50).

Sourdough Breads

Sourdough is a mixture of flour (usually wheat and/or rye), water, and other ingredients that are fermented by naturally occurring lactic acid bacteria and yeasts. The potential of sourdough lactic acid bacteria as a source of proteolytic enzymes has also been investigated recently. Although primary proteolysis during sourdough fermentation is exerted by wheat or rye endogenous enzymes that are activated by the low pH, studies have shown that certain strains of lactic acid bacteria used in sourdough fermentation can produce peptidases that can proteolytically cleave the gliadin fraction of wheat gluten under certain conditions (51–54). However, as was observed in the production of beers, the glutenin fraction of gluten has been shown to be more resistant to microbial proteolysis, so sourdough breads can still pose a potential health risk for those with celiac disease (53, 54). Further, a study has shown that lactic acid fermentation of wheat flour does not degrade gluten sufficiently enough to decrease available transglutaminase 2 binding sites on α2-gliadin and, therefore, doesn't prevent the interaction of enzyme transglutaminase 2 with gluten, indicating another source of potential immunopathogenicity (55).

Protein/Peptide Profile Differences

Quantitation of gluten in fermented-hydrolyzed foods poses a challenge due to lack of methods that can recognize the highly variable proteolytic peptide patterns that vary between

fermentation processes, as well as due to the lack of suitable hydrolytic calibrants. This is further complicated by the lack of clinical information correlating peptide content with biological activity. Further, it is unknown how to interpret the immunopathogenicity based on the amount or profile of gluten protein/peptides being detected in several different fermented-hydrolyzed foods. The regulatory threshold of 20 ppm intact gluten was based on studies examining the immunopathogenicity of intact gluten. Whether the biological activity is the same for gluten peptides that are produced during fermentation is unknown (29, 56–59).

The protein/peptide profile generated during the fermentation of different foods is dependent on numerous parameters. These include the ingredients used, time, temperature, and fermenting organisms. A slight change in these parameters can lead to wide variations in the protein/peptide profile. As such, it is impossible to generalize the profile for the different fermentation processes. The protein/peptide profiles of different fermented foods were examined using a recently developed multiplex-competitive ELISA. The ELISA utilized HRP (Horseradish peroxidase)-conjugated gluten specific antibodies (G12, R5, 2D4, MIoBS, and Skerritt) from nine gluten ELISA test kits. The antibodies were utilized in a competitive ELISA format by multiplexing the nine gluten specific antibodies into a single assay plate as described previously (56). **Figure 1** shows the apparent gluten concentration values obtained for six different fermented-hydrolyzed food categories using the multiplex-competitive ELISA. Included in the analysis were barley beers, wheat beers, a model sorghum beer brewed with 200 ppm gluten (added prior to fermentation) and brewed in the presence and absence of a PEP (Brewers Clarex), sourdough breads, soy-based sauces (soy sauces, teriyaki sauces, and Worcestershire sauces), and vinegars. Since, the antibodies used in the multiplex-competitive ELISA displays different specificities (gliadin, glutenin, and deamidated gliadin), the profiles reflect the antigenic differences arising due to the different manufacturing processes. As illustrated in **Figure 1**, the protein/peptide profile as recognized by the different gluten specific antibodies varied among the different categories of fermented-hydrolyzed foods. For example, comparing the wheat and the barley beers, the apparent gluten concentration values of the wheat beers using all the nine antibodies were higher than the barley beers. Higher gluten content has been observed in wheat beers compared to barley beers in several previous studies (35, 38, 41, 60, 61). Further, by western blot, higher level of immunoreactive peptides have been identified in wheat beers compared to barley beers (35). Another interesting difference that was observed between the profiles of the wheat beers and the barley beers, using the multiplex-competitive ELISA, was the higher apparent gluten concentration values using the two G12 antibodies (a and b) in wheat beers compared to barley beers. This observation is consistent with a previous study, which showed high level of 33-mer equivalent peptides (specifically recognized by the G12 antibodies) in wheat beers compared to the barley beers (34). Further, the wheat beers, the model sorghum beers brewed with 200 ppm gluten, and the sourdough breads resulted in a comparatively high apparent gluten concentration by the Skerritt antibody (i and j), indicating the possible abundance of glutenin proteinaceous materials. However, this was not the case with the soy-based sauces and vinegars, which instead resulted in comparatively high apparent gluten concentration values with both the Neogen antibodies (e and f) and the Microbiologique gluten antibody (g), indicating a relatively higher abundance of gliadin, and deamidated gliadin. These results indicate that protein/peptide profile differences exist among various fermentation processes. Further, the recognition of the protein/peptide profile differences, as achieved by the multiplex-competitive ELISA, is not possible if a single gluten-specific antibody is used in an assay for the detection of gluten, which is usually the case with the commercially available ELISA kits. This limits the utility of the commercial ELISAs in accurately quantitating gluten in several different types of fermented-hydrolyzed foods. The recognition of the differences in the proteolytic patterns among the different fermentation processes by a gluten detection assay is essential for the selection of appropriate calibration standards of comparable digestion and similar peptide composition, leading to accurate quantitation of gluten in different categories of fermented-hydrolyzed foods.

PROLINE ENDOPEPTIDASES (PEP) TO REDUCE IMMUNOPATHOGENIC GLUTEN CONTENT

Several proteases [PEP derived from *Aspergillus niger* (AN-PEP), *Sphingomonas capsulate*, EP-B2 (cysteine endoprotease from germinating barley), ALV003 (mixture of cysteine endoprotease and PEP), and Pseudolysin (lasB)] have been recently used to enzymatically hydrolyze gluten proteins in an attempt to prevent proliferative responses in gluten specific T cells (58, 62–70). The *Aspergillus niger* derived PEP (AN-PEP) and the ALV003 have been evaluated in clinical trials for their effectiveness in mitigating gluten-induced immune responses in celiac patients (71, 72). PEP is a serine protease which proteolyzes the peptide bonds at the carboxyl end of prolines. The use of AN-PEP in hydrolyzing gluten present in wheat starch, wheat bran, and a non-alcoholic cereal-based beverage has been reported (73, 74). In the manufacture of beer, AN-PEP, commercially available as Brewers Clarex, has been frequently used to prevent chill-haze formation that involves hydrophobic interaction of polyphenols with proline-rich proteins in beer. This enzyme has an optimum pH around 4.5, making it suitable for use during fermentation to brew beer (75). There are several conflicting reports on the ability of AN-PEP to sufficiently proteolyze gluten and eliminate any immunopathogenicity. A study by Guardum and Bamforth indicated that addition of PEP during brewing process reduced the prolamin contents of beers (76). A mass spectrometric study also reported that AN-PEP was effective in eliminating all known immunopathogenic gluten epitopes during beer production (77). However, not all potentially immunopathogenic sequences were monitored in the study. A third study indicated that PEP could destroy gluten T-cell epitopes (64). In contrast, several recent studies utilizing mass spectrometry, ELISA, and western blot analysis indicated that PEP didn't completely degrade all gluten

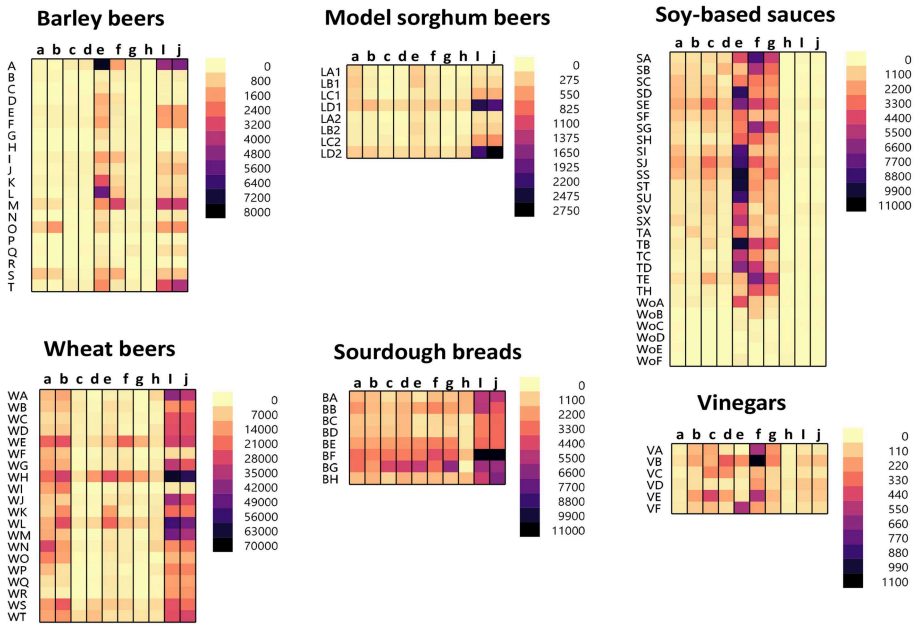


FIGURE 1 | Apparent gluten concentration (μg/mL) profiles of different categories of fermented-hydrolyzed foods (20 barley beers, 20 wheat beers, 8 model sorghum beers containing 200 μg/mL gluten brewed in the presence and absence of PEP, 8 sourdough breads, 27 soy-based sauces, and 6 vinegars) as analyzed by the multiplex-competitive ELISA utilizing gluten specific antibodies from the AgraQuant Gluten G12 (a), GlutenTox ELISA Competitive G12 (b), RIDASCREEN Gliadin (c), RIDASCREEN Gliadin Competitive (d), Veratox for Gliadin, cat # 8480 (e), Neogen Veratox for Gliadin R5 (f), Microbiologique Gluten (g), Morinaga Institute of Biological Sciences, Inc. (MIoBS) Gliadin (h), and AllerTek Gluten (25 (i) and 10 μg/mL (j), respectively coating antigen concentration) ELISA kits (56). For soy-based sauces, SA-SX represents soy sauces, TA-TH represents teriyaki sauces, and WoA-WoF represents Worcestershire sauces. For the model sorghum beers, LA-LC represent 200 ppm gluten containing beers brewed in the presence of different concentrations of PEP (25, 4, and 0.5 mL/31 gallon of wort, respectively) and LD represents 200 ppm gluten containing beer brewed in the absence of PEP. LA1-LD1 and LA2-LD2 represents two different replicate brews.

proteins and gluten proteins/peptides remain in the final beer produced by addition of PEP. Specifically, the HMW glutenin were resistant to the action of PEP during beer production (17, 39, 78). In addition, beer treated with PEP has been shown to cause a humoral response toward IgA or IgG antibodies, derived from the sera of 3 celiac disease-active patients, but there was no response from normal control subjects ($n = 31$, control group: $n = 29$), indicating that beers treated with PEP are still immunogenic (79). In another mass spectrometry study, gluten peptides that contained sequences associated with celiac disease were detected in a model wheat containing sorghum beer brewed in the presence of PEP. Included among the peptides detected were the LQLQPFPQPQLPY peptide, which is the beginning of the immunopathogenic 33-mer, and hydrolyzed HMW glutenin peptides containing immunogenic sequences (39).

We analyzed six different commercial gluten-reduced beers (brewed in the presence of PEP to reduce their gluten content) using the multiplex-competitive ELISA and western blots (49, 56). The apparent gluten concentration measured by the multiplex competitive ELISA was high for all the gluten-reduced beers with at least one gluten specific antibody (**Figure 2A**). Specifically, the Skerritt antibody and the two Neogen Varatox antibodies resulted in high apparent gluten concentrations with multiple gluten-reduced beers (**Figure 2A**). Although R5 antibodies from two other ELISA kits, RIDASCREEN gliadin (c) and RIDASCREEN gliadin competitive (d), were used in the multiplex-competitive ELISA, the apparent gluten concentration values with those antibodies were much lower compared to the Neogen Veratox R5 antibody (f). Differences in the sensitivity displayed by the same antibody derived from different test kits can be easily explained by differences in HRP conjugation resulting in higher catalytic activity. This leads to the question on a more complex issue of why differences in the performance of the same antibody in two different ELISA test kits arise. It could be due to differences in the handling of the antibody, such as in the coating of the microtiter plates and the chemistry associated with HRP conjugation altering the binding properties (affinity) toward the target analyte. More complex differences, that may alter the performance relative to defined calibration standards may arise from changes to the binding conditions, including the coating of the microtiter plates to block non-specific interactions. Lastly, all quantitative analyses are dependent on the calibration standards employed. In as much as there are no universally recognized gluten standards that are employed by all test kit manufacturers, it is possible that two kits employing identically prepared antibody reagents may differently calculate gluten content (9, 31, 80, 81).

In the western blot, 3 gluten-reduced beers, RA (Lane 5), RC (Lane 9), and RD (Lane 11), resulted in bands with the Veratox for Gliadin, cat # 8,480 detector antibody (**Figure 2B**). Bands at 17 kDa in beers RC (Lane 9) and RD (Lane 11) represent gluten-derived peptides, whereas binding observed to multiple protein

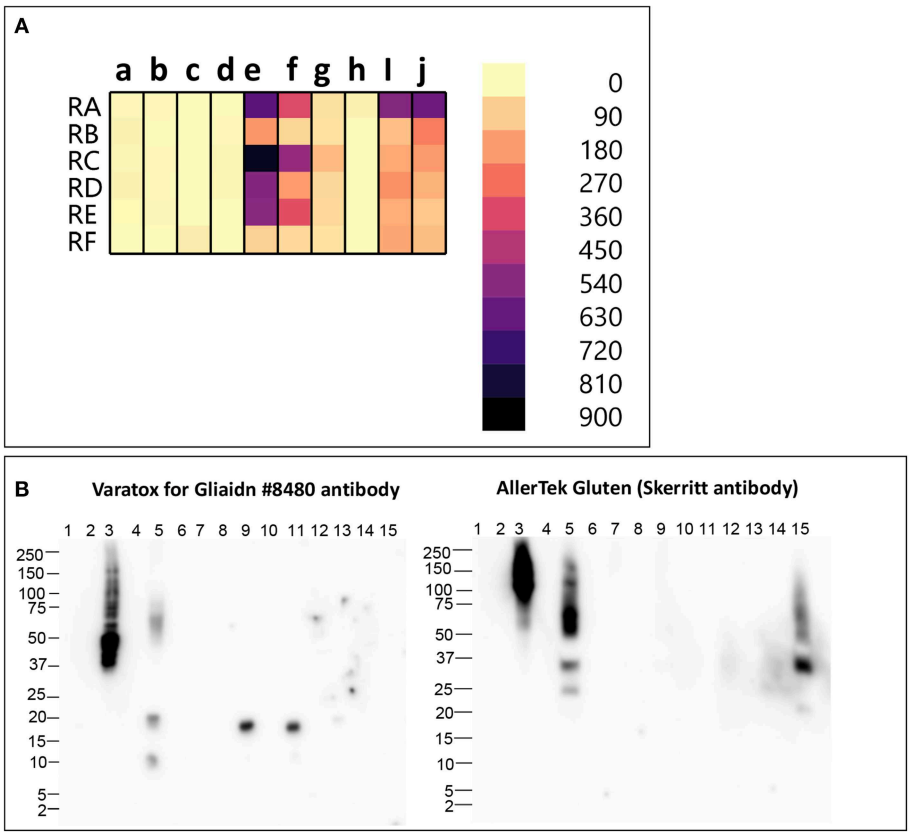


FIGURE 2 | (A) Apparent gluten concentration (μg/mL) profiles of gluten -reduced barley beers (RA, RB, RC, RD, RE, RF), obtained by the multiplex-competitive ELISA utilizing gluten specific antibodies from the AgraQuant Gluten G12 (a), GlutenTox ELISA Competitive G12 (b), RIDASCREEN Gliadin (c), RIDASCREEN Gliadin Competitive (d), Veratox for Gliadin, cat # 8480 (e), Neogen Veratox for Gliadin R5 (f), Microbiologique Gluten (g), Morinaga Institute of Biological Sciences, Inc. (MIoBS) Gliadin (h), and AllerTek Gluten (25 (i) and 10 μg/mL (j) coating antigen concentration) ELISA kits (56). **(B)** Western blot binding signal of gluten-reduced barley beers (RA, RB, RC, RD, RE, RF), using the detector antibodies of Veratox for Gliadin, cat # 8,480 and AllerTek Gluten ELISA kits. Lane information for western blot: Lane 1- Molecular weight marker, Lane 2- Empty, Lane 3- 2.5 μg/mL intact gluten standard, Lane 4- Empty, Lane 5- beer RA, Lane 6- Empty, Lane 7- beer RB, Lane 8- Empty, Lane 9- beer RC, Lane 10- Empty, Lane 11- beer RD, Lane 12- Empty, Lane 13- beer RE, Lane 14- Empty, Lane 14- beer RF (49).

bands at 10, 20, and 50–75 kDa in beer RA (Lane 2) indicate both intact gluten and gluten-derived peptides (**Figure 2B**). Beer RA is produced by a different manufacturer than beers RC and RD. Therefore, the differences in band pattern observed can be attributed to the differences in the manufacturing processes employed by the two companies. Nevertheless, gluten proteins/peptides remain in the final products, confirming the findings of previous mass spectrometry studies (36, 39). With the Skerritt antibody, beers RA (Lane 5) and RF (Lane 15) yielded multiple bands (20–150 kDa) both at higher and lower MW range (**Figure 2B**). Binding to Skerritt antibody indicates the presence of HMW glutenin (D Hordein) epitopes in gluten-reduced beers and again confirms the results of previous studies (17, 39, 78). The presence of HMW glutenin in gluten-reduced beers may not get accurately detected by gluten detection assays targeting only gliadin proteins. Further, studies have indicated that glutenin proteins can develop toxic response in celiac patients. Therefore, consumption of gluten-reduced beers may pose a potential concern for individuals with CD (17, 39, 82–84).

COMMERCIALLY AVAILABLE ELISA METHODS ARE NOT ACCURATE FOR FERMENTED/HYDROLYZED GLUTEN

Table 2 lists the various commercial ELISAs that are routinely used for the detection and quantitation of gluten in foods. ELISAs in both sandwich and competitive format are available. Sandwich ELISAs require two epitopes and therefore cannot detect short peptides lacking two antibody binding sites. However, celiac disease requires only a single immunopathogenic element, thereby making it possible for sandwich ELISAs to miss toxic gluten-derived peptides in fermented-hydrolyzed foods (17, 29). In contrast, competitive ELISAs recognize a single epitope and may be more effective in detecting immunopathogenic peptides derived from gluten in fermented-hydrolyzed foods. Competitive ELISAs based on R5 (RIDASCREEN® Gliadin Competitive) and G12 (GlutenTox® Competitive) monoclonal antibodies are marketed for detection and quantitation of gluten in fermented-hydrolyzed foods.

TABLE 2 | Manufacturer's specified properties of commercially available gluten ELISA test kits.

ELISA kits	Manufacturer	Target	Capture	Detector	LOD[a]	LOQ[a]	Upper limit[a]
AgraQuant ELISA Gluten G12	Romer Labs	QPQLPY	G12 monoclonal	G12 monoclonal	2	4	200
GlutenTox ELISA Competitive	Biomedal Diagnostics	QPQLPY	Gliadin	G12 monoclonal	–	3	48
RIDASCREEN Gliadin Sandwich	R-Biopharm, AG	QQPFP, QQQFP, LQPFP, QLPFP	R5 monoclonal	R5 monoclonal	3	5	80
RIDASCREEN Gliadin Competitive	R-Biopharm, AG	QQPFP, QQQFP, LQPFP, QLPFP	Gliadin	R5 monoclonal	2.6	10	270
Veratox for Gliadin, 8480	Neogen Corp.	Gluten	USDA monoclonal[b]	USDA monoclonal[b]	–	5	50
Veratox for Gliadin R5	Neogen Corp.	QQPFP, QQQFP, LQPFP, QLPFP	R5 monoclonal	R5 monoclonal	–	5	80
AllergenControl ^{TM} Gluten Sandwich	Microbiologique Inc.	Gliadin	2D4[c]	2D4[c]	-	2.5	80
Wheat Protein ELISA (MIoBS)	Morinaga Institute of Biological Sciences, Inc.	Gliadin	Polyclonal	Polyclonal	0.24	0.25	16
AllerTek Gluten	ELISA Technologies, Inc.	HMW[d] glutenin	Skerritt monoclonal	Skerritt monoclonal	–	5	80
GlutenTox ELISA Sandwich	Biomedal Diagnostics	QPQLPY	A1 monoclonal	A1 monoclonal	–	0.6	10

[a] *Expressed as mg/kg (ppm) gluten.*
[b] *Gluten specific monoclonal antibody developed and licensed by the U.S. Department of Agriculture (USDA).*
[c] *Deamidated gliadin specific antibody.*
[d] *HMW, High molecular weight.*

The RIDASCREEN® R5 competitive ELISA utilizes a mixture of pepsin-trypsin digested prolamin fractions from wheat, rye, and barley as the calibrator for quantitation purposes. Though awarded First Action by the Association of Official Analytic Chemists Official Methods of Analysis (AOAC OMA) (59, 85), the validation of this method was based on the detection of the reference material spiked into various foods and the AOAC OMA specifically states that depending on the fermentation conditions and resulting proteolysis, the validation performed may not be scientifically valid. It is critical that the calibration standards reflect the peptides produced by the proteolysis and the appropriate amount of residual intact gluten. Further, to be representative of real-world samples, the analyte must be incurred prior to processing (86, 87).

In a mass spectrometry study, we evaluated the potential of the hydrolyzed wheat prolamin (HWP, used as a calibrant in the R5 competitive ELISA) as a calibrant for the detection of gluten in a model sorghum beer containing 200 mg/L added gluten, brewed with or without the addition of PEP (17). By mass spectrometry, 274 unique gluten peptides were detected in HWP. However, only 4 peptides were represented in the peptide profile of a 200 mg/L gluten containing beer brewed without PEP and 1 was represented in that of the PEP containing beer (**Figure 3**). These disparities in the peptide profiles between HWP and the beers reflects the unsuitability of HWP as a calibrant for accurate quantitation of gluten in these beers. Although specific types of beers were brewed in the study, variability in fermentation conditions (time, temperature, pH) would likely result in a peptide profile not compatible with using HWP as a calibrant for accurate gluten quantitation.

Another limitation of the R5 competitive ELISAs is the use of gliadin as the calibrant without the inclusion of the glutenin fraction of gluten. Though the gliadin fraction of gluten is mainly responsible for exacerbating celiac disease, glutenin proteins have also been shown to stimulate celiac small intestinal T cells and can induce a toxic response in patients (83, 84). In another study, Tye-Din et al. (82) identified gluten T-cell stimulatory peptides that resembled the HLA-DQ8-restricted epitope present in HMW glutenin (82). Studies have shown that

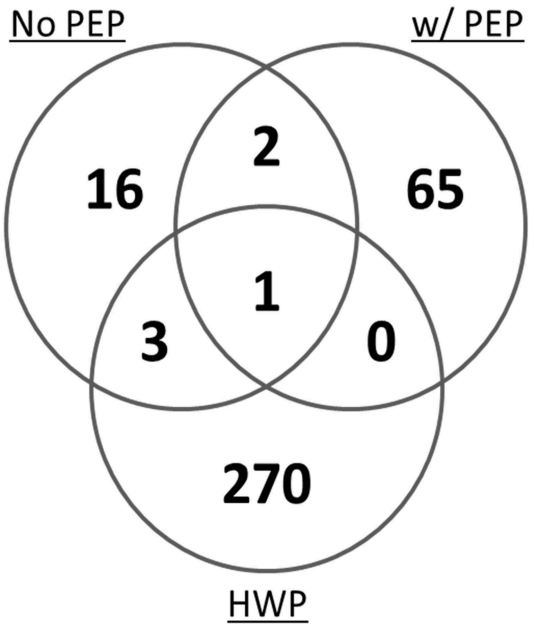


FIGURE 3 | Venn diagram of the total number of unique gluten peptides identified by mass spectrometry in a model sorghum beer containing 200 μg/mL gluten, brewed in the presence and absence of PEP, comparing it to that of the HWP. The samples were analyzed in triplicate and only peptides that were identified in at least two of the three injections were included (17).

peptides derived from HMW glutenins (known as D hordein in barley) are present in beers (17, 39). The presence of HMW glutenin-derived peptides have also been reported in sourdough breads and vinegars (50, 53, 54). Therefore, calibration standards based on gliadin proteins are likely not be suitable for accurate quantitation of gluten in fermented-hydrolyzed products.

A G12 antibody based competitive ELISA is also available for the detection of gluten. Although, the G12 antibody was raised against a prominent immunogenic gluten peptide, it may not recognize all known potential immunogenic-sequence-containing gluten peptides. This is complicated by the fact that not all immunogenic sequences associated with CD is known due to incomplete understanding of the pathogenesis of CD (21, 22, 88, 89). Studies have indicated that in fermented beverages such as beer, the reactivity of the G12 antibody to peptides correlates with potential celiac immunotoxicity. T cell epitopes in beer have been recognized with the highest affinity by the G12 antibody (34, 90). Further, the G12 antibody has been shown to be more efficient at immunocapturing the T-cell active peptides from a barley beer and a hydrolyzed gliadin from wheat compared to the R5 antibody (91). This indicates that the G12-based ELISA may be more suitable for the analysis of fermented-hydrolyzed gluten compared to the R5-based ELISA. However, no information is available on the calibrant used in the G12 competitive ELISA and validation studies have not been performed to establish the reliability of the = ELISA nor its ability to accurately quantitate gluten in fermented-hydrolyzed foods.

RECENT PROGRESS AND A POTENTIAL PATH FORWARD

Competitive ELISAs based on the G12 and R5 antibodies cannot distinguish between the protein/peptide profile pattern of different fermented foods, and while they target gliadins, they do not accurately detect glutenins, which also contain immunopathogenic sequences (17, 39, 83, 84). Mass spectrometry has the ability to differentiate the peptide profile differences among different fermentation processes (35, 36), thereby providing for a potential alternative to immunodiagnostic methods in developing suitable calibration standards for accurate quantitation of gluten in fermented-hydrolyzed foods. Semi-quantitation by mass spectrometry is possible by comparing mass areas measured from food samples against the appropriate calibration curve obtained by measuring mass areas of standard prolamin solutions (40). Recently, targeted approaches such as multiple reaction monitoring (MRM) mass spectrometry have been used for relative quantitation of gluten-derived tryptic peptides in fermented beverages such as beers (36, 92). In this method, multiple peptides are monitored per protein after trypsin digestion to compare the abundance of the original protein in different samples. MRM mass spectrometry combined with the use of synthetic peptide standards has been utilized for quantitation of six potentially immunopathogenic wheat gluten peptides in a range of native flours, processed products, sauces and beverages, including a light beer, and a vinegar (43). However, there are several limitations associated with using mass spectrometry as a quantitative method for routine analysis of gluten in fermented-hydrolyzed foods. A major limitation is dependence on identification of all the immunopathogenic sequences associated with celiac disease, however, such is not the case (88, 89). Further, similar to ELISA, accurate quantitation by mass spectrometry also requires suitable calibration standards, whereby the peptide content can be related to the regulatory threshold of 20 ppm intact gluten. Also, publicly available databases of plant protein sequences are incomplete, in particular for wheat and barley gluten proteins, further limiting the utility of mass spectrometry (36, 39).

Recognition of protein/peptide profile differences among different fermentation processes is the first step toward selection of appropriate calibration standards and, eventually, to the development of a method that can accurately quantitate gluten in fermented-hydrolyzed foods. Accurate quantitation requires that the calibration standard be identical to the protein/peptide profile of the fermented-hydrolyzed foods. This means encompassing all the gluten components (gliadin and glutenins, or any modifications resulting from fermentation, such as deamidated gliadin) present. A single calibration standard will not be suitable for all fermented-hydrolyzed foods. As such, any analytical method that is used to analyze multiple fermentation products must be able to distinguish between the different protein/peptide profiles so the appropriate calibrant can be selected to ensure accurate quantitation.

The novel multiplex-competitive ELISA included gluten specific antibodies from nine different commercial ELISA test kits. Utilizing antibodies that target different gluten epitopes, it was possible to distinguish between the protein/peptide characteristics of several different fermentation processes [**Figure 1**, (56)]. This assay simultaneously detects gliadin, deamidated gliadin and glutenin derived proteins, and peptides. Wheat gluten was used as a calibrant in the assay. Variability in the quantities and proportions of gluten proteins among wheat, rye, and barley cultivars exists and this makes the establishment of a universal standard or reference material problematic (93–96). Although reference materials comprised of both wheat gliadin and barley hordein have been proposed for gluten analysis, currently there is no certified reference material and moreover no suitable reference material is available for the detection of fermented-hydrolyzed gluten (9, 58, 97–100). Wheat gluten was chosen as a calibrant in order to avoid excluding any gluten protein fraction (gliadins or glutenins) from the analysis. Further, the material forms the regulatory basis for the analytical methods employed by the FDA (and several other governments) in assessing gluten content and potential health risk.

Using the multiplex-competitive ELISA, it was possible to distinguish between the wheat beers, barley beers, sourdough breads, and the soy-based sauces using cluster analysis. Of the 26 barley beers analyzed, 25 clustered separately from wheat beers and 24 clustered separately from sourdough breads. Only one barley beer clustered with the majority of soy-based sauces. It was also possible to distinguish samples with similar composition or processing within a particular category of fermented-hydrolyzed food by this method (e.g., some barley

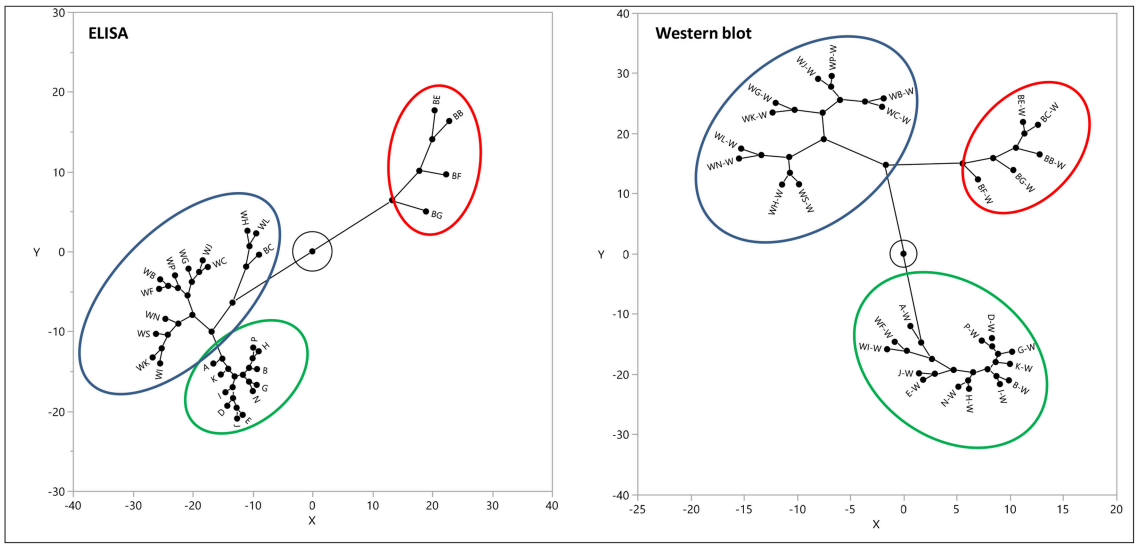


FIGURE 4 | Constellation plot, displaying the clusters, of the apparent gluten concentration values (ppm) obtained by the multiplex-competitive ELISA and the estimated gluten concentration values obtained by western blot for 11 barley beers, 12 wheat beers, and 5 sourdough breads (49, 56). Samples analyzed by western blot have been identified by adding "W" after each sample code.

beers and gluten-reduced beers) (56). The various antibodies used in the multiplex-competitive ELISA may display different cross-reactivity patterns with wheat gluten, barley hordein, and rye secelin. However, these differences don't affect the utility of the multiplex-competitive ELISA. The classification of the peptide profiles is based on empirical observations and as such would appropriately group the fermented-hydrolyzed foods and accordingly enable the proper choice of reference materials that fit the empirical observation.

Western blot analysis utilizing the same gluten specific antibodies used in the multiplex-competitive ELISA confirmed the cluster analysis by the multiplex-competitive ELISA (49, 56). Although soy-based sauces showed non-specific false positive responses with the multiplex-competitive ELISA, it didn't affect the cluster pattern and the assay was still able to differentiate the soy-based sauces from other fermented-hydrolyzed foods. Indeed, the western blot analyses differentiated the false positive responses of soy-based sauces from the presence of antigenic proteinaceous materials. **Figure 4** shows a constellation plot illustrating three different clusters that the barley beers, wheat beers and the sourdough bread generated when analyzed using the multiplex-competitive ELISA and western blot analyses, illustrating the ability of the assays to differentiate the protein/peptide profile characteristic of these three different fermented-hydrolyzed foods, which is essential for selection of appropriate calibration standard specific for each category of fermented-hydrolyzed foods required for accurate gluten quantitation.

It is obvious that further research is needed before accurate quantitation of gluten content in fermented-hydrolyzed foods can be achieved. The multiplex-competitive ELISA provides a first step by making it possible to determine the suitability of different hydrolysates as calibration standards for different fermentation-hydrolysis processes. This method also helps rule out false negative results. For examples, in **Figure 2A**, the apparent gluten concentration values of gluten-reduced beers with the two RIDASCREEN R5 antibodies and the two G12 antibodies were lower compared to the two Neogen Veratox antibodies and the Skerritt antibody. When the G12 or the RIDASCREEN R5 antibodies are used alone, the gluten content of the gluten-reduced beers may seem to be very low; however, the values are higher with both the Neogen Varatox antibodies and the Skerritt antibody, indicating that gluten components reactive to these antibodies are present at higher concentrations in these beers. Another potential utility of the multiplex-competitive ELISA especially with regulatory implications is in classification of an unknown fermented-hydrolyzed food sample into a particular category based on its overall apparent gluten concentration profile, subsequently allowing for the selection of an appropriate calibration standard required for accurate gluten analysis.

CONCLUSION

It is currently impossible to accurately quantitate gluten in fermented/hydrolyzed foods and assess its potential immunopathogenicity using antibody-based methods. This is complicated by the fact that no current commercial antibody-based assay targets all the components of gluten. Further complicating the quantitative analysis of hydrolyzed gluten is the lack of appropriate calibrants that reflect the protein/peptide profiles characteristic of the various forms of fermentation. It is therefore necessary to first distinguish

between the protein/peptide profiles to ensure the use of appropriate calibration standards for accurate quantitation. Mass spectrometry has potential by virtue of its ability to directly detect the peptides and proteins; however, its use as a routine analytical method is still in its infancy. In the meantime, the multiplex-competitive ELISA along with the western blot analysis make it possible to distinguish between the different protein/peptide profiles resulting from different fermentation processes and, ultimately, select appropriate standards for calibration.

AUTHOR CONTRIBUTIONS

EG oversaw the project. RP carried out the experiments, wrote the manuscript. All authors discussed the results and contributed to the final manuscript.

REFERENCES

1. Tonutti E, Bizzaro N. Diagnosis and classification of celiac disease and gluten sensitivity. *Autoimmunity Rev*. (2014). 13:472–6. doi: 10.1016/j.autrev.2014.01.043
2. Lundin KEA, Scott H, Hansen T, Paulsen G, Halstensen TS, Fausa O, et al. Gliadin-specific, HLA-DQ (α1*0501, β1*0201) restricted T cells isolated from the small intenstinal mucosa of celiac disease patients. *J Expr Med*. (1993) 178:187–96. doi: 10.1084/jem.178.1.187
3. van de Wal Y, Kooy Y, van Veelen P, Pena S, Mearin L, Papadopoulos G, et al. Cutting edge: selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity. *J Immunol*. (1998) 161:1585–8.
4. Wieser H. Chemistry of gluten protein. *Food Microbiol*. (2007) 24:115–9. doi: 10.1016/j.fm.2006.07.004
5. Vensel WH, Tanaka CK, Altenbach SB. Protein composition of wheat gluten polymer fractions determined by quantitative two-dimensional gel electrophoresis and tandem mass spectrometry. *Proteome Sci*. (2014) 12:8. doi: 10.1186/1477-5956-12-8
6. Wieser H, Seilmeier W, Belitz H. Comparative investigations of parital aminoacid sequences of prolamins and glutelins from cereals. *Z Lebensm Unter Forsch*. (1988) 187:27–34. doi: 10.1007/BF01454319
7. Konning F. Celiac disease: qunatity metters. *Semin Immunopathol*. (2012) 34:541–9. doi: 10.1007/s00281-012-0321-0
8. Rubio-Tapia A, Ludvigsson JF, Brantner TL, Murraym JA, Everhart JE. The prevalence of celiac disease in the United States. *Am J Gastroenterol*. (2012) 107:1538–44. doi: 10.1038/ajg.2012.219
9. Slot ID, van der Fels-Klerx HJ, Bremer MGEG, Hamer RJ. Immunochemical detection methods for gluten in food products: where do we go from here? *Crit Rev Food Sci Nutr*. (2016) 56:2455–66. doi: 10.1080/10408398.2013.847817
10. FDA Food Labeling. *21 CFR Part 101 [Docket No. FDA-2005-N-0404] RIN 0910-AG84. Federal Register 78*. (150, Monday, August 5, 2013) 47154–47179. (2013). Available online at: http://www.gpo.gov/fdsys/pkg/FR-2013-08-05/pdf/2013-18813.pdf (accessed June 19, 2019).
11. FDA Food Labeling. *Gluten-Free Labeling of Fermented or Hydrolyzed Foods 21 CFR Part 101 [Docket No. FDA−2014–N−1021] RIN 0910–AH00*. Federal Register 80 (222, Wednesday, November 18, 2015) 71990–72006 (2015). Available online at: https://www.gpo.gov/fdsys/pkg/FR-2015-11-18/pdf/2015-29292.pdf (accessed June 19, 2019).
12. Rosell CM, Barro F, Sousa C, Mena MC. Cereals for developing gluten-free profucts and analytical tools for gluten detection. *J Cereal Sci*. (2014) 59:354–64. doi: 10.1016/j.jcs.2013.10.001
13. Scherf KA, Poms RE. Recent developments in analytical methods for tracing gluten. *J Cereal Sci*. (2016) 67:112–22. doi: 10.1016/j.jcs.2015.08.006
14. Miranda-Castro R, de-los-Santos-Alvarez N, Miranda-Ordieres AJ, Lobo-Castanon MJ. Harnessing aptamers to overcome challenges in gluten detection. *Biosensors*. (2016) 6:16. doi: 10.3390/bios6020016
15. Melini F, Melini V. Immunological methods in gluten risk analysis: a snapshot. *Safety*. (2018) 4:56. doi: 10.3390/safety4040056
16. Skerritt JH, Hill AS. Monoclonal antibody sandwich enzyme immunoassay for determination of gluten in foods. *J Agric Food Chem*. (1990) 38:1771–8. doi: 10.1021/jf00098a029
17. Panda R, Fiedler KL, Cho CY, Cheng R, Stutts WL, Jackson LS, et al. Effects of a proline endopeptidase on the detection and quantitation of gluten by antibody-based methods during the fermentation of a model sorghum beer. *J Agric Food Chem*. (2015) 63:10525–35. doi: 10.1021/acs.jafc.5b04205
18. Allred L, Ritter BW. Recognition of gliadin and glutenin fractions in four commercial gluten assays. *J AOAC Int*. (2010) 93:190–6.
19. Mendez E, Vela C, Immer U, Janssen FW. Reports of a collaborative trial to investigate the performance of the R5 enzyme linked immunoassay to determine gliadin in gluten-free food. *Eur J Gastroentero Hepatol*. (2005) 17:1053–63. doi: 10.1097/00042737-200510000-00008
20. Kahlenberg F, Sanchez D, Lachmann I, Tuckova L, Tlaskalova H, Mendez E, et al. Monoclonal antibody R5 for detection of putatively coeliac-toxic gliadin peptides. *Eur Food Res Technol*. (2006) 222:78–82. doi: 10.1007/s00217-005-0100-4
21. Moron B, Cebolla A, Manyani H, lvarez-Maqueda MA, Megías M, Thomas M, et al. Sensitive detection of cereal fractions that are toxic to celiac disease patients by using monoclonal antibodies to a main immunogenic wheat peptide. *Am J Clin Nutr*. (2008) 87:405–14. doi: 10.1093/ajcn/87.2.405
22. Moron B, Bethune MT, Comino I, Manyani H, Ferragud M, Lopez MC, et al. Toward the assessment of food toxicity for celiac patients: characterization of monoclonal antibodies to a main immunogenic gluten peptide. *PLoS ONE*. (2008) 3:e2294. doi: 10.1371/journal.pone.0002294
23. Masiri J, Benoit L, Katepalli M, Meshgi M, Cox D, Nadala C, et al. Novel monoclonal antibody-based immunodiagnostic assay for rapid detection of deamidated gluten residues. *J Agric Food Chem*. (2016) 64:3678–87. doi: 10.1021/acs.jafc.5b06085
24. Benoit L, Cox D, Katepalli M, Masiri J, Nadla C, Gendel S, et al. *Development and Characterization of a Novel Monoclonal Antibody Directed Against Gluten*. (2016). Available online at: https://iafp.confex.com/iafp/2016/webprogram/Paper11379.html (accessed June 19, 2019).
25. Akiyama HK, Isuzugawa N, Harikai H, Watanabe K, Iijima H, Yamakawa Y, et al. Inter-laboratory evaluation studies for development of notified ELISA methods for allergic substances (wheat). *Shokuhin Eiseigaku Zasshi*. (2004) 45:128–34. doi: 10.3358/shokueishi.45.128
26. Matsuda RY, Yoshioka K, Akiyama K, Aburatani Y, Watanabe T, Matsumoto N. Interlaboratory evaluation of two enzyme-linked immunosorbent assay kits for the detection of egg, milk, wheat, buckwheat and peanut in foods. *J AOAC Int*. (2006) 89:1600–8.
27. Immer U, Haas-Lauterbach S. Gliadin as a measure of gluten in foods containing wheat, rye, and barley—enzyme immunoassay method based on a specific monoclonal antibody to the potentially celiac toxic amino acid prolamin sequences: collaborative study. *J AOAC Int*. (2012) 95:1118–24. doi: 10.5740/jaoacint.CS2012_01
28. Halbmayr-Jech E, Hammer E, Fiedler R, Coutts J, Rogers A, Cornish M. Characterization of G12 sandwich ELISA, a next-generation immunoassay for gluten toxicity. *J AOAC Int*. (2012) 95:372–6. doi: 10.5740/jaoacint.SGE_Halbmayr-Jech
29. Panda R, Zoerb HF, Cho CY, Jackson LS, Garber EAE. Detection and quantification of gluten during the brewing and fermentation of beer using antibody-based technologies. *J Food Prot*. (2015) 78:1167–77. doi: 10.4315/0362-028X.JFP-14-546
30. Haraszi R, Chassaigne H, Maquet A, Ulberth F. Analytical methods for detection of gluten in food-method developments in support of food labeling legislation. *J AOAC Int*. (2011) 94:10006–25.
31. Diaz-Amigo C, Popping B. Accuracy of ELISA detection methods for gluten and reference materials: a realistic assessment. *J Aric Food Chem*. (2013) 61:5681–8. doi: 10.1021/jf3046736

32. Wieser H, Koehler P. Is the calculation of the gluten content by multiplying the prolamin content by a factor of 2 valid? *Eur Food Res Technol*. (2009) 229:9–13. doi: 10.1007/s00217-009-1020-5
33. Bamforth C. *Beer. Tap Into the Art and Science of Brewing*. Oxford: Oxford University Press (2003).
34. Comino I, Real A, Moreno Mde L, Montes R, Cebollab A, Sousa C. Immunological determination of gliadin 33-mer equivalent peptides in beers as a specific and practical analytical method to assess safety for celiac patients. *J Sci Food Agric*. (2013) 93:933–43. doi: 10.1002/jsfa.5830
35. Picariello G, Mamone G, Cutignano A, Fontana A, Zurlo L, Addeo F, et al. Proteomics, peptidomics, and immunogenic potential of wheat beer (weissbier). *J Agric Food Chem*. (2015) 63:3579–86. doi: 10.1021/acs.jafc.5b00631
36. Colgrave ML, Goswami H, Blindell M, Howitt CA. Using mass spectrometry to detect hydrolysed gluten in beer that is responsible for false negatives by ELISA. *J Chromatogr A*. (2014) 1370:105–14. doi: 10.1016/j.chroma.2014.10.033
37. Allred LK, Sealey-Voyksner JA, Voyksner RD. Evaluation of Qualitative and quantitative immunoassays to detect barley contamination in gluten-free beer with conformation using LC/MS/MS. *J AOAC Int*. (2014) 97:1615–25. doi: 10.5740/jaoacint.14-058
38. Dostalek P, Hochel I, Mendez E, Hernando A, Gabrovska D. Immunochemical determination of gluten in malts and beers. *Food Addit Contam*. (2006) 23:1074–8. doi: 10.1080/02652030600740637
39. Fiedler KL, Panda R, Croley TR. Analysis of gluten in a wheat-gluten-incurred sorghum beer brewed in the presence of proline endopeptidase by LC/MS/MS. *Anal Chem*. (2018) 90:2111–8. doi: 10.1021/acs.analchem.7b04371
40. Ferranti P, Mamone G, Picariello G, Addeo F. Mass spectrometry analysis of gliadins in celiac disease. *J Mass Spectrum*. (2007) 4:1531–48. doi: 10.1002/jms.1361
41. Kanerva P, Sontag-Strohm T, Lehtonen P. Determination of prolamins in beers by ELISA and SDS-PAGE. *J Inst Brew*. (2005) 111:61–4. doi: 10.1002/j.2050-0416.2005.tb00649.x
42. Picariello G, Mamone G, Nitride C, Addeo F, Camarca A, Vocca I, et al. Shotgun proteome analysis of beer and the immunogenic potential of beer polypeptides. *J Proteomics*. (2012) 75:5872–82. doi: 10.1016/j.jprot.2012.07.038
43. Sealey-Voyksner JA, Khosla C, Vouksner RD, Jorgenson JW. Novel aspects of qunatiation of immunogenic wheat gluten peptides by liquid chormatography-mass spctrometry/mass spectrometry. *J Chormatograp A*. (2010) 1217:4167–83. doi: 10.1016/j.chroma.2010.01.067
44. Hutkins RW. (ed.). Fermentation of foods in the Orient. In: *Microbiology and Technology of Fermented Foods*. Oxford: Blackwell Publishing Ltd. (2006). p. 419–55.
45. Kobayashi M, Hashimoto Y, Taniuchi S, Tanabe S. Degradation of wheat allergen in Japanese soy sauce. *J Mol Med*. (2004) 13:821–7. doi: 10.3892/ijmm.13.6.821
46. Janssen FW, van Klinken F, Immer U, Gowein C. *Analytical Research Reports*. (2012). Available online at: https://glutenfree.files.wordpress.com/2012/04/safety-of-asian-soy-sauce-in-gf-diet-janssen-et-al.pdf (accessed June 19, 2019).
47. Cao W, Watson D, Bakke M, Bedford B, Panda R, Jackson LS, et al. Detection of gluten during the fermentation process to produce soy sauce. *J Food Prot*. (2017) 80:799–808. doi: 10.4315/0362-028X.JFP-16-483
48. Hefle SL, Lambrecht DM, Nordlee JA. Soy sauce retains allergenicity through the fermentation/production process. *J Allerg Clin Immunol*. (2005) 115:S32. doi: 10.1016/j.jaci.2004.12.143
49. Panda R, Garber EAE. Western blot analysis of fermented-hydrolyzed foods utilizing gluten specific antibodies employed in a novel multiplex competitive ELISA. *Anal Bioanal Chem*. (2019). doi: 10.1007/s00216-019-01893-0. [Epub ahead of print].
50. Li H, Byrne K, Galiamov R, Mendoza-porras O, Bose U, Howitt CA, Colgrave ML. Using LC-MS to examine the fermented food products vinegar and soy sauce for the presence of gluten. *Food Chem*. (2018) 254:302–8. doi: 10.1016/j.foodchem.2018.02.023
51. De Angelis M, Rizzello CG, Fasano A, Clemente MG, Simone CD, Silano M, et al. VSL#3 probiotic preparation has the capacity to hydrolyze gliadin polypeptides responsible for celiac sprue. *Biochemica Biophysica Acta*. (2006) 1762:80–93. doi: 10.1016/j.bbadis.2005.09.008
52. Gerez CL, Dallagnol A, Rollan G, de Valdez GF. A combination of two lactic acid bacteria improves the hydrolysis of gliadin during wheat dough fermentation. *Food Microbiol*. (2012) 32:427–30. doi: 10.1016/j.fm.2012.06.007
53. Di Cagno R, De Angelis M, Lavermicocca P, De Vincenzi M, Giovannini C, Faccia M, et al. Proteolysis by sourdough lactic acid bacteria: effects on wheat flour protein fractions and gliadin peptides involved in human cereal intolerance. *Appl Environ Microbiol*. (2002) 68:623–33. doi: 10.1128/AEM.68.2.623-633.2002
54. Di Cagno RD, De Angelis M, Auricchio S, Greco L, Clarke C, Vincenzi MD, et al. Sourdough bread made from wheat and nontoxic flour and started with selected *Lactobacilli* is tolerated in celiac sprue patients. *Appl Environ Microbiol*. (2004) 70:1088–96. doi: 10.1128/AEM.70.2.1088-1096.2004
55. Engstrom N, Sandberg A, Scheers N. Sourdough fermentation of wheat flour doesn't prevent the interaction of Transglutaminase 2 with α2-gliadin or gluten. *Nutrients*. (2015) 7:2134–44. doi: 10.3390/nu7042134
56. Panda R, Boyer M, Garber EAE. A multiplex competitive ELISA for the detection and characterization of gluten in fermented-hydrolyzed foods. *Anal Bioanl Chem*. (2017) 409:6959–73. doi: 10.1007/s00216-017-0677-z
57. Knorr V, Wieser H, Koehler P. Production of gluten-free beer by peptidase treatment. *EurFood Res Technol*. (2016) 242:1129–40. doi: 10.1007/s00217-015-2617-5
58. Tanner GJ, Colgrave ML, Howitt CA. Gluten, celiac disease, and gluten intolerance and the impact of gluten minimization treatments with prolylendopeptidase on the measurement of gluten in beer. *J Am Soc Brew Chem*. (2014) 72:36–50. doi: 10.1094/ASBCJ-2014-0129-01
59. Lacorn M, Weiss T. Partially hydrolyzed gluten in fermented cereal-based products by R5 competitive ELISA: collaborative study, first action 2015.05. *J AOAC Int*. (2015) 98:1346–54. doi: 10.5740/jaoacint.CS2015.15
60. Guerdrum LJ, Bamforth CW. Levels of gliadin in commercial beers. *Food Chem*. (2011) 129:1783–4. doi: 10.1016/j.foodchem.2011.06.021
61. Tanner GJ, Colgrave ML, Blundell MJ, Goswami HP, Howitt CA. Measuring hordein (gluten) in beer-A comparison of ELISA and mass spectrometry. *PLoS ONE*. (2013) 8:e56452. doi: 10.1371/journal.pone.0056452
62. Wieser H, Koehler P. Detoxification of gluten by means of enzymatic treatment. *J AOAC Int*. (2012) 95:356–63. doi: 10.5740/jaoacint.SGE_Wieser
63. Gass J, Vora H, Bethune MT, Gary GM, Khosla C. Effect of barley endoprotease EP-B2 on gluten digestion in the intact rat. *J Pharmacool Exp Ther*. (2006) 318:1178–86. doi: 10.1124/jpet.106.104315
64. Stepniak D, Spaenij-Dekking L, Mitea C, Moesterm de Ru A, Baak-Pablo R, van Veelen P, et al. Highly efficient gluten degradation with a newly identified prolyl endoprotease: implications for celiac disease. *Am J Physio Gastrointest Liver Physiol*. (2006) 291:G621–9. doi: 10.1152/ajpgi.00034.2006
65. Gass J, Bethune MT, Siegel M, Spencer A, Khosla C. Combination enzyme therapy for gastric digestion of dietary gluten in patients with celiac sprue. *Gastroentrol*. (2007) 133:472–80. doi: 10.1053/j.gastro.2007.05.028
66. Mitea C, Havenaar R, Drijfhout JW, Edens L, Dekking L, Koning F. Efficient degradation of gluten by a prolyl endoprotease in a gastrointestinal model: implications for coeliac disease. *Gut*. (2008) 57:25–32. doi: 10.1136/gut.2006.111609
67. Tye-Din JA, Anderson RP, Ffrench RA, Brown GJ, Hodsman P, Siegel M, et al. The effects of ALV003 pre-digestion of gluten on immune response and symptoms in celiac disease *in vivo*. *Clin Immunol*. (2010) 134:289–95. doi: 10.1016/j.clim.2009.11.001
68. Toft-Hansen H, Rasmussen KS, Staal A, Roggen EL, Sollid LM, Lillevang TB, et al. Treatment of both native and deamidated gluten peptides with an endo-peptidase from *Aspergillus niger* prevents stimulation of gut-derived gluten-reactive T cells from either children or adults with celiac disease. *Clin Immunol*. (2014) 153:323–31. doi: 10.1016/j.clim.2014.05.009
69. Montserrat V, Bruins MJ, Edens L, Koning F. Influence of dietary components on *Aspergillus niger* Prolyl endoprotease mediated gluten degradation. *Food Chem*. (2015) 174:440–5. doi: 10.1016/j.foodchem.2014.11.053
70. Wei G, Tian N, Valery AC, Zhong Y, Schuppan D, Helmerhorst EJ. Identification of pseudolysin (lasB) as an aciduric gluten-degrading enzyme

with high therapeutic potential for celiac disease. *Am J Gastroentrol*. (2015) 110:899–908. doi: 10.1038/ajg.2015.97
71. Tack GJ, van de water JMW, Bruins MJ, Kooy-Winkelaar EMC, van Bergen J, Bonnet P, et al. Consumption of gluten with gluten-degrading enzyme by celiac patients: a pilot study. *World J Gastroenterol*. (2013) 19:5837–47. doi: 10.3748/wjg.v19.i35.5837
72. Lahdeaho ML, Kaukinen K, Laurila K, Vuotikka P, Koivurova OP, Karja-Lahdensuu T. Glutenase ALV003 attenuates gluten-induced mucosal injury in patients with celiac disease. *Gastroenterology*. (2014) 146:1649–58. doi: 10.1053/j.gastro.2014.02.031
73. Walter T, Wieser H, Koehler P. Production of gluten-free wheat starch by peptidase treatment. *J Cereal Sci*. (2014) 60:202–9. doi: 10.1016/j.jcs.2014.02.012
74. Walter T, Wieser H, Koehler P. Degradation of gluten in wheat bran and bread drink by means of a proline-specific peptidase. *J Nutr Food Sci*. (2014) 4:1–6. doi: 10.4172/2155-9600.1000293
75. Lopez M, Edens L. Effective prevention of chill-Haze in beer using an acid proline-specific endoprotease from *Aspergillus niger*. *J Agric Food Chem*. (2005) 53:7944–9. doi: 10.1021/jf0506535
76. Guerdrum LJ, Bamforth CW. Prolamin levels through brewing and the impact of prolyl endoproteinase. *J Am Soc Brew Chem*. (2012) 70:35–8. doi: 10.1094/ASBCJ-2012-0130-01
77. Akeroyd M, van Zandycke S, den Hartog J, Mutsaers J, Edens L, van den Berg M, et al. AN-PEP, proline-specific endopeptidase, degrades all known immunostimulatory gluten peptides in beer made from barley malt. *J Am Soc Brew Chem*. (2016) 74:91–9. doi: 10.1094/ASBCJ-2016-2300-01
78. Colgrave ML, Byrne K, Howitt CA. Liquid chromatography-mass spectrometry analysis reveals hydrolyzed gluten in beers crafted to remove gluten. *J Agric Food Chem*. (2017) 65:9715–25. doi: 10.1021/acs.jafc.7b03742
79. Allred LK, Lesko K, McKiernan D, Kupper C, Guandalini S. The celiac patient antibody response to conventional and gluten-removed beer. *J AOAC Int*. (2017) 100:485–91. doi: 10.5740/jaoacint.16-0184
80. Diaz-Amigo C, Popping B. Gluten and gluten-free: issues and considerations of labeling regulations, detection methods, and assay validation. *J AOAC Int*. (2012) 95:337–48. doi: 10.5740/jaoacint.SGE_Diaz-Amigo
81. Rzychon M, Brohee M, Cordeiro F, Harazi R, Ulberth F, O'Connor G. The feasibility of harmonizing gluten ELISA measurement. *Food Chem*. (2017) 234:144–54. doi: 10.1016/j.foodchem.2017.04.092
82. Tye-Din JA, Stewart JA, Dromey JA, Beissbarth T, van Heel DA, Tatham A, et al. Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Sci Transl Med*. (2010) 2:41ra51. doi: 10.1126/scitranslmed.3001012
83. Vader W, Kooy Y, van Veelen P, de Ru A, Harris D, Benckhuijsen W, et al. The gluten response in children with celiac disease is directed towards multiple gliadin and glutenin peptides. *Gatroentrology*. (2002) 122:1729–37. doi: 10.1053/gast.2002.33606
84. Dewar DH, Amato M, Julie Ellis H, Pollock EM, Gonzalez-Cinca N, Wieser H, et al. The toxicity of high molecular weight glutenin subunits of wheat to patients with coeliac disease. *Gastroenterol Hepatol*. (2006) 18:483–91. doi: 10.1097/00042737-200605000-00005
85. Koehler P, Schwalb T, Immer U, Lacorn M, Wehling P, Don C. AACCI approved methods technical committee report: collaborative study on immunochemical determination of partially hydrolyzed gluten using an R5 competitive ELISA. *Cereal Foods World*. (2013) 58:1–6. doi: 10.1094/CFW-58-3-0402
86. Hefle SL, Koppelman S. (eds.). Conclusions. In: *Detecting Allergen in Food*. Boca Raton, FL: Wood head Publishing Limited; CRC Press LLC (2006). p. 405–12.
87. Mattarozzi M, Careri M. The role of incurred materials in method development and validation to account for food processing effects in food allergen analysis. *Anal Bioanl Chem*. (2019) doi: 10.1007/s00216-019-01642-3. [Epub ahead of print].
88. Sollid LM, Qiao S. Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenetics*. (2012) 64:455–60. doi: 10.1007/s00251-012-0599-z
89. De Re V, Magris R, Cannizzaro R. New insight into the pathogenesis of Celiac disease. *Front Med*. (2017) 4:137. doi: 10.3389/fmed.2017.00137
90. Real A, Comino I, de Lourdes Moreno M, Lopez-Casado MA, Lorite P, Torres MI, et al. Identification and *in vitro* reactivity of celiac immunoreactive peptides in an apparent gluten-free beer. *PLoS ONE*. (2014) 9:e100917. doi: 10.1371/journal.pone.0100917
91. Moreno Mde L, Munoz-Suano A, Lopez-Casado MA, Torres MI, Sousa C, Cebolla A. Selective capture of most celiac immunogenic peptides from hydrolyzed gluten proteins. *Food Chem*. (2016) 205:36–42. doi: 10.1016/j.foodchem.2016.02.066
92. Colgrave ML, Goswami H, Howitt CA, Tanner GJ. Wheat is in a beer? Proteomic characterization and relative quantification of hordein (gluten) in beer. *J Proteome Res*. (2012) 11:386–96. doi: 10.1021/pr2008434
93. Prandi B, Mantovani P, Galaverna G, Sforza S. Genetic and environmental factors affecting pathogenicity of wheat as related to celiac disease. *J Cereal Sci*. (2014) 59:62–9. doi: 10.1016/j.jcs.2013.10.006
94. Rallabhandi P, Sharma GM, Pereira M, Williams KM. Immunological characterization of the gluten fractions and their hydrolysates from wheat, rye and barley. *J Agric Food Chem*. (2015) 63:1852–32. doi: 10.1021/jf505716p
95. Pahlavan A, Sharma GM, Pereira M, Williams KM. Effects of grain species and cultivar, thermal processing and enzymatic hydrolysis on gluten quantitation. *Food Chem*. (2016) 208:264–71. doi: 10.1016/j.foodchem.2016.03.092
96. Schalk K, Lexhaller B, Koehler P, Scherf KA. Isolation and characterization of gluten protein types from wheat, rye, barley and oats for use as reference materials. *PLoS ONE*. (2017) 12:e0172819. doi: 10.1371/journal.pone.0172819 eCollection 2017.
97. Huang X, Kanerva P, Salovaara H, Stoddard FL, Sontag-Strohm T. Proposal for C-Hordein as reference material in gluten quantification. *J Aric Food Chem*. (2017) 65:2155–61. doi: 10.1021/acs.jafc.6b05061
98. Hajas L, Scherf KA, Torok K, Bugyi X, Schall E, Poms RE, et al. Variation in protein composition among wheat (*Triticum aestivum* L.) cultivars to identify cultivars suitable as reference material for wheat gluten analysis. *Food Chem*. (2018) 267:387–94. doi: 10.1016/j.foodchem.2017.05.005
99. van Eckert R, Berghofer E, Ciclitira PJ, Chirdo F, Denery-Papini S, Ellis HJ, et al. Towards a new gliadin reference material-isolation and characterization. *J Cereal Sci*. (2006) 43:331–41. doi: 10.1016/j.jcs.2005.12.009
100. Tanner GJ, Blundell MJ, Colgrave ML, Howitt CA. Quantification of hordeins by ELISA: the correct standard makes a magnitude of difference. *PLoS ONE*. (2013) 8:e56456. doi: 10.1371/journal.pone.0056456

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Nutrition

REVIEW
published: 05 July 2019
doi: 10.3389/fnut.2019.00101

What Is Gluten—Why Is It Special?

*Peter Shewry**

Rothamsted Research, Harpenden, Hertfordshire, United Kingdom

Wheat gluten has an immense impact on human nutrition as it largely determines the processing properties of wheat flour, and in particular the ability to make leavened breads, other baked products, pasta and noodles. However, there has been increasing interest in wheat gluten over the past two decades because of its well-established role in triggering coeliac disease, and its perceived role in other adverse reactions to wheat. The literature on wheat gluten is vast and extends back over two centuries, with most studies focusing on the structures of gluten proteins and their role in determining the functional properties of wheat flour and dough. This article provides a concise account of wheat gluten, focusing on properties, and features which are relevant to its role in triggering coeliac disease and, to a lesser extent, other gluten-related disorders. It includes descriptions of the biological role of the gluten proteins, the structures and relationships of gluten protein families, and the presence of related types of protein which may also contribute to functional properties and impacts on health. It therefore provides an understanding of the gluten protein system at the level required by those focusing on its impact on human health.

Keywords: wheat, gluten, coeliac disease, protein, prolamin, gliadin, gluten, ATI

OPEN ACCESS

Edited by:
Katharina Anne Scherf, Technical University of Munich, Germany

Reviewed by:
Sándor Tömösközi, Budapest University of Technology and Economics, Hungary
Mona Elena Popa, University of Agronomic Sciences and Veterinary Medicine, Romania
Luud Gilissen, Wageningen University and Research, Netherlands

***Correspondence:**
Peter Shewry
peter.shewry@rothamsted.ac.uk

Specialty section:
This article was submitted to Nutrition and Food Science Technology, a section of the journal Frontiers in Nutrition

Received: *01 May 2019*
Accepted: *19 June 2019*
Published: *05 July 2019*

Citation:
Shewry P (2019) What Is Gluten—Why Is It Special? Front. Nutr. 6:101. doi: 10.3389/fnut.2019.00101

INTRODUCTION

Wheat gluten was one of the earliest proteins to be studied scientifically, by Jacopo Beccari (Professor of Chemistry at the University of Bologna) in his article "De Frumento" (Concerning Grain) in 1745 (1, 2). It has since been studied in great detail by cereal chemists, because of its role in underpinning the ability to make leavened bread, other baked goods, pasta, and noodles. These properties are only shared to a very limited extent by related cereals (barley and rye). Hence, gluten underpins the production of staple foods for a substantial proportion of the global population, particularly in temperate zones.

Although gluten was identified as the trigger for coeliac disease almost 70 years ago (3), interest in gluten outside the scientific community was limited to those unfortunate enough to suffer from coeliac disease until early in the present century, which has seen an explosion of interest, particularly in the popular press and social media. As an example, a "Google" search carried out in December 2018 gave almost 400 million hits in less than a minute. This interest relates, of course, to the proposed role of gluten in triggering a range of adverse reactions, with substantial proportions of the population in many countries choosing to adopt a gluten-free, or low-gluten, diet. However, despite this massive interest few people have a clear understanding of gluten itself: what is it, what is the origin, why is it special?

This article, which forms part of the Special Research Topic "Gluten, from Plant to Plate: Implications for People with Celiac Disease," therefore, provides a broad account of wheat gluten including its synthesis and deposition in the developing grain, the structures, and evolutionary relationships of its component proteins, and its unique properties which are exploited in grain

processing, focusing on features which are relevant to its role in triggering coeliac disease. It does not cover other impacts of wheat proteins on human health, notably allergy, and non-coeliac gluten sensitivity (NCGS) which are discussed in other recent review articles (4, 5).

WHAT IS GLUTEN?

Gluten Is Defined Based on Its Origin and Solubility

Gluten is classically defined as the largely proteinaceous mass which remains when a dough made from wheat flour and water is gently washed in an excess of water or dilute salt solution to remove most of the starch and soluble material (6). The remaining material, which has been described as "rubbery," comprises about 75–80% protein on a dry matter basis, depending on how well the material is washed. Hence "gluten proteins" are defined as those present in this mass and, because similar material cannot be isolated from doughs made with flours from other cereals, gluten proteins are restricted to the grain of wheat (species of the genus *Triticum*). However, related proteins are present in other cereals (as discussed below) and these are frequently referred to as gluten in the non-specialist literature and the wider popular media.

More correctly, gluten and related proteins from other cereals are classified as "prolamins." This name was coined by T.B. Osborne, the father of plant protein chemistry who worked at the Connecticut Agricultural experiment station from 1886 till 1928. During this period he published some 250 papers, including studies of seed proteins from 32 species. This allowed him to develop a broad classification of proteins based on their extraction in a series of solvents (7). This extraction is often performed sequentially (and called "Osborne fractionation") with the four Osborne fractions being called albumins (soluble in water), globulins (soluble in dilute saline), prolamins (soluble in 60–70% alcohol), and glutelins (insoluble in the other solvents but may be extracted in alkali). The first two fractions are readily distinguished and the names are still in use, while prolamins were recognized as a defined group present only in cereal grains with the name being based on their high contents of proline and amide nitrogen (now known to be derived from glutamine). This fraction is given specific names in different cereal species: gliadin in wheat, hordein in barley, secalin in rye, zein in maize etc.

However, the final fraction (glutelin) is more difficult to define, as it effectively comprises all proteins which are insoluble in the three previous solvents but can be solubilized under conditions of extreme pH. In fact, glutelins are now known to comprise a mixture of unrelated proteins, including insoluble structural and metabolic proteins such as those bound to membranes and cell walls. However, these proteins are only present in small amounts and in wheat (and most other cereals) the major glutelin components are in fact prolamin subunits which are not extractable with alcohol/water mixtures due their presence as high molecular mass polymers stabilized by inter-chain disulphide bonds. In wheat these proteins are called glutenin and are present in about equal amounts to the alcohol-soluble gliadins, the two groups comprising gluten.

Gluten Proteins Are the Major Storage Protein Fraction

Gluten proteins are the major group of proteins which are stored in the grain to support germination and seedling development. They are restricted in distribution to the starchy endosperm cells of the grain, and have not been detected in any other tissues of the grain or plant. Their pathway and mechanisms of synthesis and deposition have been studied in detail [see Tosi (8)] but two points are particularly relevant here. Firstly, they are initially deposited in discrete protein bodies, which fuse during the later stages of grain development to form a continuous matrix surrounding the starch granules (**Figure 1A**). This matrix forms a continuous protein network within the cell, which can be revealed when the starch is removed from a flour particle by enzyme digestion (**Figure 1B**). It is easy to envisage how the protein networks present in the individual cells can be brought together during dough mixing to form the continuous gluten network in dough.

The second important point is that gluten proteins are not uniformly distributed in the starchy endosperm cells, but enriched in the outer 2 to 3 layers of cells (which are called the sub-aleurone cells). This is illustrated in **Figure 1C**, which shows a section of the starchy endosperm cells and outer layers from the lobe of the grain at a late stage of development stained with toluidine blue to show protein. In fact, Kent (11) calculated that the protein content of the cells of the starchy endosperm varies by over 4-fold, from 45% in the sub-aleurone cells to 8% in the central region. Furthermore, the gluten protein composition also varies, with the percentage of high molecular weight glutenin subunits (HMW subunits) increasing and the proportion of low molecular weight (LMW) subunits and gliadins (except for ω-gliadins) decreasing (these protein types are discussed below) (12). These gradients in composition are reflected to some extent in the contents and compositions of gluten proteins in the flour streams produced by commercial roller milling, meaning that these fractions may also vary in their impact on health (13).

Implications for Coeliac Disease

Fractionation by conventional milling combined with pearling (abrasion) or peeling (friction) could lead to flour streams that are enriched or depleted in coeliac-active proteins. The use of vital gluten (which is produced commercially for fortification of food products) also has implications. This will contain all of the gluten proteins present in the flour of origin, but may also contain other biologically active proteins as "co-passengers."

GLUTEN PROTEINS

Gluten Comprises Several Related Families of Proteins Encoded by Multigene Families

The gluten protein fraction comprises a complex mixture of components which can be separated into groups by electrophoresis. Electrophoresis of the gliadins at low pH separates four groups of bands, called (in terms of decreasing mobility) α-gliadins, β-gliadins, γ-gliadins, and ω-gliadins. However, comparisons of amino acid sequences show that the α- and β-gliadins form a single group, sometimes called α-type gliadins.

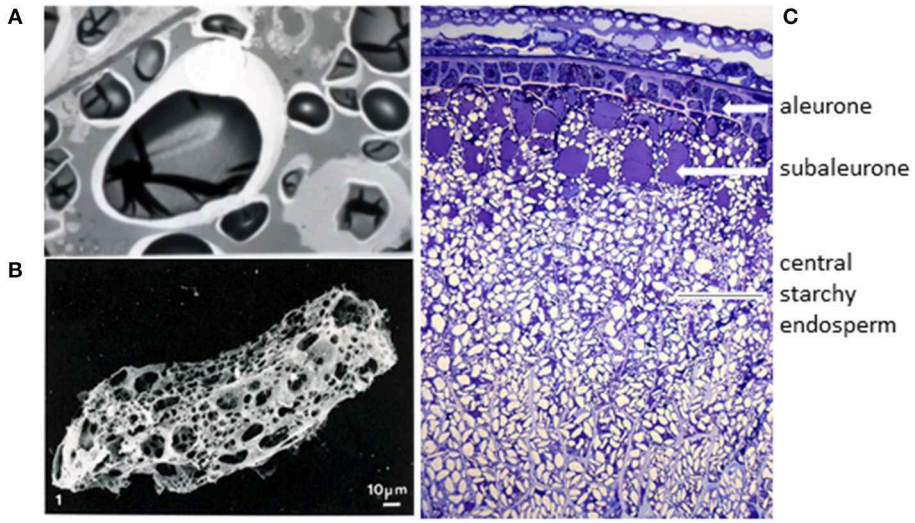


FIGURE 1 | The origin of wheat gluten. **(A)** Transmission electron microscopy of starchy endosperm cells at a late stage of grain development (46 days after anthesis) shows that the individual protein bodies have fused to form a continuous proteinaceous matrix. Taken from Shewry et al. (9) with permission, provided by Dr. M. Parker (IFR, Norwich, UK). **(B)** Digestion of a flour particle to remove starch reveals a continuous proteinaceous network. Taken from Amend and Beauvais (10) with permission. **(C)** Transverse section of the lobe region of a developing wheat grain stained with Toluidine Blue to show the tissue structure and deposited protein (in blue). Figure kindly provided by Cristina Sanchis Gritsch and Paola Tosi (Rothamsted Research).

The glutenin polymers are too big to be separated by conventional electrophoresis, but reduction of the inter-chain disulphide bonds that stabilize the polymers allows the subunits to be separated by sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS-PAGE) into two groups of bands, called the HMW and LMW subunits. The latter group can be further sub-divided into a major group of components (B-type LMW subunits) and two minor groups (C-type and D-type).

Comparisons of amino acid sequences of these groups of gluten protein components clarifies their relationships, showing that the HMW subunits and ω-gliadins form discrete groups, with the α-gliadins, γ-gliadins, and B-type LMW subunits forming a third group. The minor groups of C-type and D-type LMW subunits appear to be modified forms of gliadins in which mutations to form cysteine residues allow their incorporation into glutenin polymers, with the C-type LMW subunits being modified α-gliadins or γ-gliadins and the D-type modified ω-gliadins. This classification is summarized in **Table 1**, which also shows their relative amounts and summarizes their characteristics (molecular masses and partial amino acid compositions).

Table 1 also groups the types of gluten proteins discussed above into three "families" (the HMW, sulfur(S)-rich, and S-poor prolamins), which were defined about 30 years ago based on emerging sequence data (15). This classification remains valid despite the vast increase in our knowledge of gluten protein sequences over the past few decades. For example, in May 2015 Bromilow et al. (16) retrieved over 24,000 sequences related to gluten proteins from the UniProt database. Removal of redundant, partial and mis-assigned sequences allowed the assembly of a curated database of 630 sequences.

The retrieval of over 600 sequences of gluten proteins does not, of course, mean that individual wheat genotypes contain this number of gluten proteins. Although the precise number of gluten proteins present in mature seed has not been determined, examination of two-dimensional (2D) electrophoretic separations indicates that the number of gluten proteins present in detectable amounts is probably between 50 and 100. This is consistent with the recent study of Bromilow et al. (17), who identified 63 gluten proteins in a single cultivar, using mass spectrometry and a curated sequence database (16). However, this study identified eight individual HMW subunit proteins, which is twice the number known to be present in the cultivar studied. This highlights the problems inherent in identifying gluten proteins based on short peptide sequences.

Although the prolamin groups discussed above undoubtedly account for the vast majority of the gluten proteins, recent work has shown that small amounts of a further type of gluten protein are present. These have been defined as δ-gliadins, although sequence comparisons indicate that they form part of the wider family of γ-prolamins (being closest in sequence to the γ3-hordeins of barley) (18, 19). Proteomic analysis indicates that they account for 1.2% of the total normalized spot volume in grain of Chinese Spring wheat (20).

Molecular Basis for Gluten Protein Polymorphism

The large numbers of individual gluten proteins present in single genotypes, and the 10-fold greater number of sequences in databases, arises from three factors: the presence of multigene families, the high level of polymorphism between genotypes and, to a more limited extent, post-translational modification. It is therefore, necessary to consider these factors in turn.

TABLE 1 | Summary of the types and characteristics of wheat gluten proteins [based on Shewry and Halford (14)].

Gluten protein type	Molecular mass	% total gluten fraction	Polymers or monomers?	Partial amino acid composition (mol %)
HMW prolamins				
HMW subunits	65–90,000	6–10	Polymers	30–35% glutamine, 10–16% proline, 15–20% glycine, 0.5–1.5% cysteine, 0.7–1.4% lysine
S-rich prolamins				
α-gliadins	30–45,000	70–80	Monomers	30–40% glutamine, 15–20% proline, 2–3% cysteine, <1% lysine
γ-gliadins				
B-type and C-type LMW subunits			Polymers	
S-poor prolamins				
ω-gliadins	30–75,000	10–20	Monomers	40–50% glutamine, 20–30% proline, 0–0.5% phenyl alanine, 0–0.5% lysine, 0 cysteine, 1 cysteine residue in D-type LMW subunits
D-type LMW subunits			Polymers	

Common wheat (*Triticum aestivum*), which includes modern bread wheat and spelt, is a hexaploid species, with three genomes (called A, B, and D) derived from related wild grasses. Only two of these genomes (A and B) are present in the tetraploid durum (pasta) wheat and emmer (forms of *Triticum turgidum*) while einkorn (*Triticum monococcum*) is diploid with only the A genome. Gluten proteins are encoded by loci on the group 1 and group 6 chromosomes of all three genomes, meaning that the gluten fraction can be expected to comprise more individual protein components in common wheat than in the other species. A detailed discussion of the genetics of gluten proteins is outside the scope of this article, but the reader can refer to Shewry et al. (21) for a detailed account.

Furthermore, all of the gluten protein loci comprise multiple genes. The simplest loci are the *Glu-1* loci which are located on the long arms of the group 1 chromosomes. Each of these loci comprises two genes which encode two types of HMW subunit of glutenin (called x-type and y-type). However, because not all of the *Glu-1* genes are expressed in all genotypes, the number of HMW subunit proteins in cultivars of bread wheat varies from 3 to 5 (22). Because of the simple genetic system, and the fact that the HMW subunits have been studied in more detail than most groups of gluten proteins, it is possible to define alleles at all three loci. Thus, the widely occurring pairs of subunits called 1Dx2 + 1Dy12 and 1Dx5 + 1Dy10 are alleles, while the pairs of subunits called 1Dx2 + 1Dy12 and 1Bx7 + 1By9 are homeoalleles (alleles on different genomes). The greater complexity of other gluten protein loci makes it much more difficult to recognize allelic forms of genes and proteins, although detailed analyses of allelic variation in LMW subunits have been reported [reviewed by Juhász et al. (23)].

However, whereas the individual HMW subunits can be assigned to sequenced genes, this is very difficult, if not impossible, for many other gluten proteins because of the complexity of the loci. For example, Huo et al. (19) assembled sequences of the α-gliadin loci on the three genomes of bread wheat, showing a total of 47 genes of which 26 encoded intact full-length protein products. Similarly, Qi et al. (24) reported the sequences of 29 putatively functional γ-gliadin genes (encoded by genes at the *Gli-1 loci* on the short arms of the group 1 chromosomes) in a single cultivar. Further information on the structures of the gluten protein multigenic loci are being provided by genome analysis [see, for example, (5, 25, 26)].

It is also likely that the numbers of expressed genes vary between genotypes. Thus, the high polymorphism in gluten protein composition observed between genotypes may arise both from variation in the numbers of expressed genes, and variation in the sequences of the encoded proteins.

A third factor which may contribute to protein polymorphism is post-translational modification. Gluten proteins contain between about 20 and 50 mol % of glutamine residues so post-translational deamidation has long been recognized as a possibility. It may, for example, account for the fact that HMW subunits often form "trains" of spots in 2D electrophoresis, while Dupont et al. (27) reported the presence of HMW subunit sequences in 43 spots separated on 2D gels. However, the extent of deamidation has never been quantified. Other proposed modifications, such as glycosylation (28) and phosphorylation (29) have not been substantiated by further studies. Other types of post-translational modification may include cyclisation of N-terminal glutamine to give pyroglutamate (which is likely to be responsible for many gluten proteins having "blocked" N-termini), differential processing of the signal peptide (30) and proteolysis by legumain-like asparaginyl endoproteinase (31).

Finally, the proportions of gluten proteins may also be affected by the environment, including temperature during grain development and availability of nutrients (nitrogen and sulfur) [reviewed by DuPont and Altenbach (32) and Altenbach (33)]. In particular, increases in the proportions of gliadins occur under high nitrogen availability and of ω-gliadins when nitrogen availability is high but sulfur is limiting.

Implications for Coeliac Disease

Protein polymorphism is clearly a challenge for attempts to eliminate "toxic" proteins and to develop coeliac-safe wheats, whether by exploiting natural variation or by genetic engineering/genome editing.

Effects of environment on gluten protein composition will also have impacts on the abundances of specific coeliac disease epitopes.

Gluten Proteins Contain Unique Repetitive Domains

The most important characteristic of wheat gluten proteins in relation to their role in coeliac disease is the presence of protein domains comprising repetitive sequences. The domains vary in extent, but generally account for between about 30 and 50% of the protein sequence in S-rich gliadins and LMW subunits, between 75 and 85% in HMW subunits, and almost the whole protein in ω-gliadins [reviewed by Shewry et al. (34)]. They comprise tandem repeats of short peptides comprising between three and nine amino acid residues, and may be based on tandem repeats of one motif or tandem and interspersed repeats of two or more motifs.

The most widely studied repetitive sequences are those present in the HMW subunits of glutenin. These comprise repeats based on three motifs: the hexapeptide PGQGQQ, the nonapeptide GYYPTSPQQ or GYYPTSLQQ, and in x-type subunits only, a tripeptide GQQ (P, proline; G, glycine; Q, glutamine, Y, tyrosine; P, proline; T, threonine, S, serine; L, leucine) (34). The motifs present in the other groups of gluten proteins are generally less well-conserved and the identification of consensus motifs is more subjective than in the HMW subunits, but all are rich in proline and glutamine, for example, PQQPFPQQ (F, phenyl alanine) in γ-gliadins. It should be noted that these sequences are responsible for the characteristic amino acid compositions of the whole proteins, notably the high contents of glutamine (35–55 mol%) and proline (10–25 mol%) in all groups of prolamins, high glycine in HMW subunits (11–12 mol%), and high phenyl alanine (about 11 mol%) in ω-gliadins [reviewed by Shewry et al. (34)].

The repeated sequences may also be responsible for the unusual solubility properties of gluten proteins. Although glutamine is a hydrophilic amino acid, the regularly repeated glutamine residues in gluten proteins are considered to form protein:protein hydrogen bonds resulting in insolubility in water (as discussed by Belton (35) for HMW subunits). However, in most gluten proteins, all of the cysteine residues, which may form interchain or intrachain disulphide bonds, are located in the non-repetitive domains.

The repetitive sequences also play a crucial role in triggering coeliac disease. In fact, all of the 31 "coeliac disease relevant T-cell epitopes" listed by Sollid et al. (36) are present in the repetitive domains of wheat or related cereals (barley, oats, rye) and all groups of gluten proteins (gliadins and glutenins) contain epitopes. Nevertheless, some individual proteins within these groups may lack recognized coeliac epitopes (although the current list of epitopes is considered to be incomplete). This is illustrated by **Figure 2** (37) and discussed in detail by Shewry and Tatham (37), Gilissen et al. (38), and Juhasz et al. (5).

Implications for Coeliac Disease

As discussed above, all of the coeliac-toxic epitopes in wheat gluten proteins are present in the repeated sequences, with multiple epitopes present in some repetitive domains. This clearly poses a significant challenge for attempts to "remove" epitopes by transgenesis or gene editing.

THE PROLAMIN SUPERFAMILY

The prolamins, including wheat gluten proteins, were historically defined as a unique class of proteins restricted to the grain of cereals and related grass species, based on their unusual amino acid compositions and solubility properties (7) and this dogma was not questioned until the increasing availability of protein sequence data allowed wider comparisons to be made. The first report that prolamins were related to a wider range of proteins was in 1985, when Kreis et al. (39) showed the sequences present in the cysteine-rich non-repetitive regions of prolamins were related to sequences in two other groups of seed proteins: cereal inhibitors of α-amylase and trypsin (now called ATIs) and 2S albumin storage proteins of dicotyledonous seeds. Although these groups of proteins have little sequence identity with each other or with prolamins, the homology was based on very high conservation in the numbers and spacing of cysteine residues. Further comparisons exploiting the vast increase in sequence data have since identified several other groups of related proteins, which are together referred to as the "prolamin superfamily."

The prolamin superfamily includes proteins which are not restricted to cereals and grasses, and present in tissues other than seeds (40). However, several types are present in wheat grain, and may contribute to the functional properties and role in diet and health (34). They are therefore, briefly discussed here and summarized in **Table 2**.

Farinins and Purinins

It has been known for many years that wheat flour contains proteins with molecular masses below 30 kDa which are related to gluten proteins, including types described as globulins, LMW gliadins, and avenin-like proteins. Kasarda et al. (41) have recently discussed the relationships of these proteins and suggested that they should be classified into two types, which they termed farinins and purinins. Both are more closely related to gliadins than the other protein types discussed below, but lack the repeated sequences which are typical of gliadins. Hence they have been classed as globulins based on solubility. The farinins correspond to the avenin-like proteins (defined based on homology with the avenin proteins of oats) with two types called a (which correspond to LMW gliadins) and b (42). These groups differ in that the b-type proteins contain a duplicated sequence of about 120 residues, resulting in a higher molecular weight (about 30 kDa compared with 17 kDa). The b-type proteins are associated with the surface of the starch granule and are post-translationally cleaved to give two subunits (11 and 19 kDa) linked by a single disulphide bond (41). Ma et al. (43) showed that over-expression of a transgene encoding a b-type protein resulted in improved flour mixing properties and an increased proportion of large glutenin polymers, presumably due to their ability to form inter-chain disulphide bonds.

The LMW gliadins/purinins have masses of about 17–19 kDa (44) and are more closely related to the γ-gliadins in sequence (41, 45). They may, perhaps, be considered to be similar to the "ancestral" prolamin proteins, before they diverged due to the development and amplification of the repetitive sequence domains. Mixing of heterologously expressed proteins

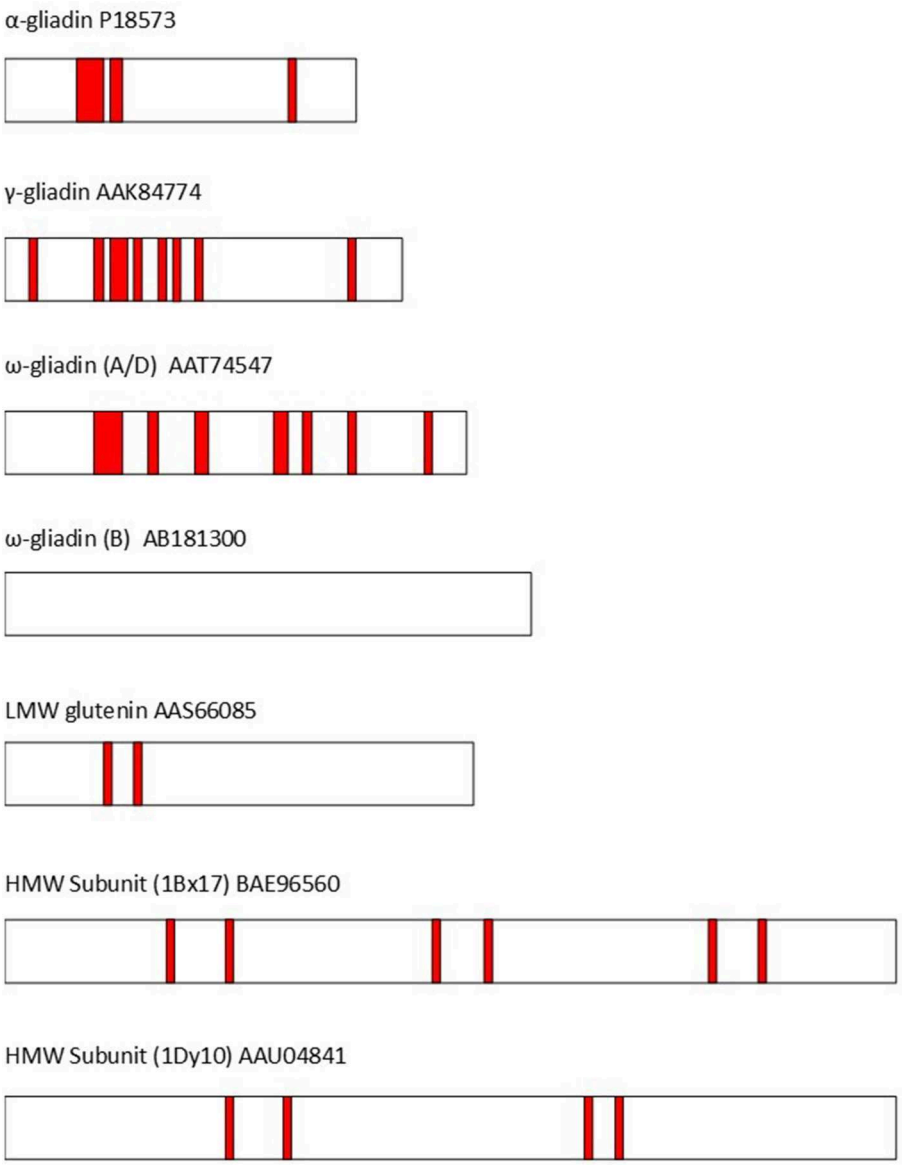


FIGURE 2 | The distribution of T-cell epitopes (shown as red bars) in representative wheat gluten proteins (identified by GenBank accession codes). The epitopes are based on Sollid et al. (36). **α**-gliadin P18573: DQ2.5-glia-α1a, DQ2.5-glia-α1b, DQ2.5-glia-α2, & DQ8-glia-α1. γ-gliadin AAK84774: DQ2.5-glia-ω1/hor-1/sec-1, DQ8-glia-γ1a, DQ8-glia-γ2, DQ8-glia-γ4c, & DQ8-glia-γ5. ω-gliadin (A/D) AAT74547: DQ2.5-glia-γ5, DQ8-glia-γ1a, DQ2.5-glia-ω1/hor-1/sec-1, DQ8-glia- γ1b, & DQ2.5-glia- γ3. ω-gliadin (B) AB181300 no coeliac toxic epitopes present. LMW subunit AAS66085:DQ2.5-glut-L1. HMW Subunit (1Bx17) BAE96560: DQ8.5-glut-H1. HMW Subunit (1Dy10) AAU04841: DQ8.5-glut-H1. Modified from Shewry and Tatham (37).

into dough showed similar effects to the incorporation of gliadins (45).

Puroindolines (Pins) and Grain Softness Protein (GSP)

Hardness is one of the major characteristics used to divide wheat into end use classes. It is determined by the *Hardness* (*Ha*) locus on the short arm of chromosome 5D of bread wheat, although the name is misleading because the encoded genes actually determine softness. This locus is not present in durum wheat which is therefore ultrahard. The *Ha* locus comprises three genes (46), encoding proteins called puroindoline a (Pin a), puroindoline b (Pin b) and grain softness protein (GSP). The mature Pin a and Pin b proteins comprise about 120 amino acid residues including 10 cysteine residues which form inter-chain disulphide bonds. They also contain five (in Pin a) or three (in Pin b) tryptophan residues which are grouped together in the sequences. Comparison of wholemeal flours of 40 wheat cultivars (19 soft and 21 hard) grown on four French sites showed 0.029–0.060 % dry wt of Pin a and 0.004–0.031% dry wt of Pin b (47). Differences in the expression of these proteins, and/or their amino acid sequences, account

TABLE 2 | Wheat grain proteins of the prolamin superfamily (based on literature discussed in the text).

Protein group	Molecular mass	Characteristics	Abundance	Functional properties/impact on health
Farinins	17,000–30,000	Correspond to avenin-like proteins and LMW gliadins	Not determined	Transgenic expression results in improved mixing properties
Purinins (low molecular weight gliadins)	17,000–19,000	Possibly correspond to "ancestral" type of prolamin	Not determined	Behave like gliadins in dough
Puroindolines a and b	13,000	Tryptophan-rich loop region which may be involved in binding to starch granule surface	0.029–0.060 % dry wt of Pin a and 0.004–0.031 % dry wt. of Pin b in wholemeal flour	Determine about 75% of the variation in softness in European wheats
Grain softness protein (GSP)	∼15,000	Associated with the starch granule surface	Not determined	Small effect on grain softness
+				
Arabinogalactan peptide (AGP)	23,000	15 residue peptide *o*-glycosylated with arabinogalactan chains at 3 hydroxyproline residues	0.39% dry wt. white flour	Prebiotic properties *in vitro*
Non-specific lipid-transfer proteins (LTP)	9,000 (LTP1) + 7,000 (LTP2)	Bind and transport lipids *in vitro* Concentrated in aleurone layer and embryo	Not determined	LTP1 is a food and respiratory allergen
α-amylase/trypsin inhibitors (ATIs)	12,000 to 16,000	Monomeric, dimeric, and tetrameric forms, some subunits inhibit trypsin or α-amylase	0.34–0.41% dry wt. of wholemeal flour	Include respiratory and food allergens, putative links to coeliac disease, NCWS, and other adverse reactions to wheat Contribute to pasta-making quality

for about 75% of the variation in grain hardness in bread wheat (48).

The third gene at the *Ha* locus encodes a protein which is cleaved post-translationally, probably in the vacuole by a similar legumain-type asparaginyl endoproteinase to the enzyme(s) responsible for proteolysis of gluten proteins (as discussed above). This releases a 15 residue peptide from the N-terminus (49). This peptide contains three proline residues which are hydroxylated to give hydroxyprolines and then *o*-glycosylated with arabinogalactan chains to give a mass of about 23 kDa (50). The resulting "arabinogalactan peptide" (AGP) accounts for about 0.39% of the dry weight of white flour (50) and is readily fermented by the colonic microflora (51). The remaining part of the protein, termed "grain softness protein" (GSP), may contribute to hardness to a limited extent [by about 10 units measured by the Perten Single Kernal Characterization System (SKCS)] (52), but the biological roles of AGP and GSP are not known.

Non-specific Lipid Transfer Proteins (LTPs)

Unlike the other proteins discussed here, LTPs are not restricted to seed tissues, or to cereals and other grass species. Although they were initially defined on their ability to transfer phospholipids between liposomes and membranes *in vitro*, their true physiological role is unknown with one possible function being to contribute to defense to biotic stresses. They occur in two classes, with masses of about 9 kDa (LTP1) and 7 kDa (LTP2) and are concentrated in the aleurone layer and embryo of the wheat grain [reviewed by Marion et al. (53)]. Many LTPs have been identified as allergens, in seeds, fruit, and pollen (53), with LTP1 of wheat contributing to both food allergy and Bakers' asthma (respiratory allergy to wheat flour) (54, 55).

α-Amylase/Trypsin Inhibitors

Wheat inhibitors of α-amylase and trypsin have been studied for over 40 years, resulting in an extensive and somewhat confusing literature. This results partly from the complexity of the fraction but also from use of different nomenclatures, based on relative electrophoretic mobilities (the major components being called 0.19, 0.28, and 0.53), solubility in chloroform:methanol (called CM1 to CM17) and subunit structure (monomeric, dimeric, and tetrameric forms occurring) (56). Dupont et al. (27) used mass spectrometry of proteins separated by 2D electrophoresis to identify two spots corresponding to forms of the putative monomeric trypsin inhibitor(s) CM1/3, two related to the monomeric amylase inhibitor WMAI, two related to the homodimeric amylase inhibitor WDAI1, and nine related to subunits of the heterotetrameric amylase inhibitor WTAI (1 × CM1, 2 × CM2, 2 × CM3, 2 × CM16, and 2 × CM17). More recently, Geisslitz et al. (57) have used targeted LC-MS to quantify the amounts of the major ATIs (WDAI/0.19 + 0.53; WMAI1/0.28, CM2, CM3, CM16, and CM17), showing that they together accounted for 3.4–4.1 mg/g in wholemeal flour of bread wheat.

Wheat ATIs are well-characterized as wheat allegens, particularly in Bakers' asthma but also on ingestion of food [reviewed by Salcedo et al. (58)]. In addition, they have been studied widely over the past few years because of putative roles in

other adverse reactions to wheat consumption, including coeliac disease, and non-coeliac wheat/gluten sensitivity (as discussed in other contributions to this special section).

ATIs have also been reported to contribute to the cooking quality of pasta, where they were initially reported to be glutenin components (called durum sulfur-rich glutenin, DSG) (59–61).

Implications for Coeliac Disease

Wheat grain contains many other proteins including other families of protease and amylase inhibitors, thionins, ribosome-inactivating proteins, and putative defense-related proteins with unknown functions [reviewed by Shewry et al. (34)]. All of these may be present in food products, present either in flours or as "contaminants" in vital gluten. However, the proteins discussed above share some properties which may be particularly relevant. Firstly, most are small globular proteins which are tightly folded and stabilized by multiple interchain disulphide bonds. Hence, they are particularly stable to heating during food processing and to degradation in the gastro-intestinal tract: although proteolysis may occur, the proteins will not disintegrate because the fragments are held together by the disulphide bonds. Secondly, they may interact strongly with gluten proteins and hence be present in vital gluten. These interactions may be stabilized by non-covalent forces, such as the LMW gliadins/purinins, or by disulphide bonds formed either during grain development and maturation or re-arrangements during processing. Irrespective of the mechanism, the fact that they may be present in "gluten protein" fractions shows that they must be considered when interpreting studies carried out on human responses to wheat proteins.

GLUTEN PROTEINS HAVE UNIQUE BIOPHYSICAL PROPERTIES WHICH UNDERPIN GRAIN PROCESSING

Several factors have contributed to the global success of wheat, one being its wide adaptability. However, the main reason why it is grown in preference to other cereal crops in many countries is the functional properties of wheat flour. As discussed above, wheat is the only cereal which can be baked to give leavened bread and other baked products, as well as pasta and noodles. The quality for these end uses is determined largely by the gluten proteins, which form a continuous network in dough. This network provides the cohesiveness required for making products such as pasta as well as the visco-elasticity required for breadmaking.

Despite a massive literature the molecular basis for the biophysical properties of gluten is still not completely understood, and it is not possible to provide a detailed discussion here. However, two points are particularly relevant. Firstly, the properties depend on the contributions of both the gliadins and glutenins, with the glutenin subunits forming large three dimensional networks stabilized by inter-chain disulphide bonds which interact with gliadins, and with other glutenin networks, by non-covalent forces, particularly hydrogen bonds. Secondly, the polymers are stabilized by a combination of forces. The importance of disulphide bonds is readily demonstrated as these can be disrupted using reducing agents, with catastrophic effects on functionality. The importance of hydrogen bonds is less easy to demonstrate, but Belton (35) has proposed that hydrogen bonds are particularly important in developing optimal protein interactions during dough mixing.

Implications for Coeliac Disease

The clearest implication for coeliac disease is that any drastic modification to the composition of the gluten protein fraction and/or to the sequences of the individual subunits are likely to have effects on functionality. Although these effects are not easy to predict, that fact that bread making wheats have been selected for functional properties for almost a century suggests that most modifications will be detrimental. Thus, although it may be possible to produce "acceptable" loaves from modified lines of wheat in the laboratory and in small scale systems [see, for example, (62, 63)], this is a much greater challenge for large scale commercial production where profit margins are narrow and small differences in parameters such as loaf height, crumb texture, color and shelf life will affect the quality of the product and hence acceptability by consumers.

CONCLUSION

Wheat gluten fulfills an essential biological role as the major grain storage protein fraction, and is the major determinant of the functional (processing) properties of the grain. It is a highly complex mixture of proteins, encoded by multigene families at multiple loci on the three genomes of bread wheat, with a high degree of polymorphism between genotypes. The individual proteins also have unusual structures, including extensive domains of repetitive sequences. In addition, a range of related proteins are present in the grain and may be present in isolated gluten fractions. All of these factors must be considered when studying the role of gluten in coeliac disease and other adverse responses to wheat consumption, and in designing strategies to develop safe types of wheat and wheat products.

AUTHOR CONTRIBUTIONS

PS wrote the whole paper. Part of **Figure 1** was provided by colleagues.

FUNDING

Rothamsted Research receives grant-aided support from the Biotechnology and Biological Sciences Research Council (BBSRC) of the UK and the work forms part of the Designing Future Wheat strategic programme (BB/P016855/1).

ACKNOWLEDGMENTS

I am grateful to Dr. D. D. Kasarda (USDA-ARS, WRRC, Albany CA, USA) for discussions and comments on the text and Dr. S. Altenbach (USDA-ARS, WRRC, Albany CA, USA) for providing a copy of her accepted publication.

REFERENCES

1. Beccari JB. *De Frumento*. Bononia: De Bononiensi Scientifiarurm et Atrium atque Academia Commentarii Tomi Secundi (1745).
2. Bailey CH. A translation of Beccari's "Concerning grain" 1728. *Cereal Chem.* (1941) 18:555–61.
3. Dicke W. K. (1950). *Coeliakie*. (MD Thesis). Utrecht: University of Utrecht.
4. Catassi C, Alaedini A, Bojarski C, Bonaz B, Bouma G, Carroccio A, et al. The overlapping area of non-celiac gluten sensitivity (NCGS) and wheat-sensitive irritable bowel syndrome (IBS): an update. *Nutrients.* (2017) 9:1268. doi: 10.3390/nu9111268
5. Juhász A, Belova T, Florides C, Maulis C, Fischer I, Gell G, et al. Genome mapping of seed-borne allergens and immunoresponsive proteins in wheat. *Sci Adv*. (2018) 4:eaar8602. doi: 10.1126/sciadv.aar8602
6. Wrigley CW, Bietz JA. Proteins and amino acids. In: Pomeranz Y, editor. *Wheat Chemistry and Technology*. 3rd ed. St. Paul Mn: AACC (1988). p.159–275.
7. Osborne TB. *The Vegetable Proteins*. 2nd ed. London UK: Longmans, Green and Co (1924).
8. Tosi P. Trafficking and deposition of prolamins in wheat. *J Cereal Sci*. (2012) 56:81–90 doi: 10.1016/j.jcs.2012.02.004
9. Shewry PR, Tatham AS, Barro F, Barcelo P, Lazzeri P. Biotechnology of breadmaking: unraveling and manipulating the multi-protein gluten complex. *Biotechnology*. (1995) 13:1185–90. doi: 10.1038/nbt1195-1185
10. Amend T, Beauvais F. Der mechanismus der Teigbildung: vorstoß in den molekularen Strukturbereich. *Getreide Mehl Und Brot.* (1995) 49:359–62.
11. Kent NL. Subaleurone endosperm cells of high protein content. *Cereal Chem.* (1966) 43:585–601.
12. He J, Penson S, Powers S, Hawes C, Shewry PR, Tosi P. Spatial patterns of gluten protein and polymer distribution in wheat grain. *J Agric Food Chem.* (2013) 61:6207–15. doi: 10.1021/jf401623d
13. Tosi P, He J, Lovegrove A, Gonzalez-Thuillier I, Penson S, Shewry PR. Gradients in compositions in the starchy endosperm of wheat have implications for milling and processing. *TIFST*. (2018) 82:1–7. doi: 10.1016/j.tifs.2018.09.027
14. Shewry PR, Halford MG. Cereal seed storage proteins: structures, properties and role in grain utilisation. *J Exp Bot.* (2002) 53:947–58. doi: 10.1093/jexbot/53.370.947
15. Shewry PR, Tatham AS, Forde J, Kreis M, Miflin BJ. The classification and nomenclature of wheat gluten proteins: a reassessment. *J Cereal Sci*. (1986) 4:97–106. doi: 10.1016/S0733-5210(86)80012-1
16. Bromilow S, Gethings LA, Buckley M, Bromley M, Shewry PR, Langridge JI, et al. A curated gluten protein sequence database to support development of proteomics methods for determination of gluten in gluten-free foods. *J Proteomics*. (2017) 163:67–75 doi: 10.1016/j.jprot.2017.03.026
17. Bromilow SNL, Gethings LA, Langridge JI, Shewry PR, Buckley M, Bromley ML, et al. Comprehensive proteomic profiling of wheat gluten using a combination of data-independent and data-dependent acquisition. *Front Plant Sci.* (2017) 7:2020. doi: 10.3389/fpls.2016.02020
18. Anderson D, Dong L, Huo N, Gu YQ. A new class of wheat gliadin genes and proteins. *PLoS ONE.* (2012) 7:12e52139. doi: 10.1371/journal.pone.0052139
19. Huo N, Zhang S, Zhu T, Dong L, Wang Y, Mohr T, et al. Gene duplication and evolution dynamics in the homeologous regions harbouring multiple prolamin and disease resistance gene families in hexaploidy wheat. *Front Plant Sci.* (2018) 9:673. doi: 10.3389/fpls.2018.00673
20. Altenbach SB, Chang H-C, Simon-Buss A, Mohr T, Hou N, Gu YQ. Exploiting the reference genome sequence of hexaploid wheat; a proteomic study of flour proteins from the cultivar Chinese Spring. *Func Integr Genom.* (2019). doi: 10.1007/s10142-019-00694-z. [Epub ahead of print].
21. Shewry PR, Halford NG, Lafiandra, D. The genetics of wheat gluten proteins. In: Hall JC, Dunlap JC, Friedman T, editors. *Advances in Genetics*, Vol. 49. London: Academic Press (2003). p. 111–84. doi: 10.1016/S0065-2660(03)01003-4
22. Payne PI. Genetics of wheat storage proteins and the effect of allelic variation on breadmaking quality. *Ann Rev Pl Physiol.* (1987) 38:141–53. doi: 10.1146/annurev.pp.38.060187.001041
23. Juhász A, Gianibelli MC. Low-Molecular-Weight Glutenin Subunits: Insights into This Abundant Subunit Group Present in Glutenin Polymers. In Wrigley CWF, Békés F, Bushuk W, editors. *Gliadin and Glutenin. The Unique Balance of Wheat Quality*. St. Paul Mn: AACC (2006). p.71–212. doi: 10.1094/9781891127519.009
24. Qi P-F, Wei Y-M, Ouellet T, Chen Q, Tan X, Zheng Y-L. The γ-gliadin multigene family in common wheat (*Triticum aestivum*) and its closely related species. *BMC Genom.* (2009) 10:168. doi: 10.1186/1471-2164-10-168
25. Huo N, Zhu T, Altenbach S, Dong L, Wang Y, Mohr T, et al. Dynamic evolution of α-gliadin prolamin gene families in homeologous genomes of hexaploidy wheat. *Sci Rep.* (2018) 8:5181. doi: 10.1038/s41598-018-23570-5
26. Clavijo BJ, Venturini L, Schudoma C, Accinelli GG, Kaithakottil G, Wright J, et al. An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provided gemonic evidence for chromosomal teanslocations. *Genome Res.* (2019) 27:885–96. doi: 10.1101/gr.217117.116
27. Dupont FM, Vensel WH, Tanaka CK, Hurkman WJ, Altenback SB. Deciphering the complexities of the wheat flour proteome using quantitative two-dimensional electrophoresis, three proteases and tandem mass spectrometry. *Proteome Sci.* (2011) 9:10. doi: 10.1186/1477-5956-9-10
28. Tilley KA, Lookhart GL, Hoseney RC, Mawhinney TP. Evidence for glycosylation of the high molecular weight glutenin subunits 2, 7, 8 and 12 from Chinese Spring and TAM 105 wheats. *Cereal Chem*. (1993) 70:602–6.
29. Tilley KA, Schofield JD. Detection of phosphorylation in the high Mr subunits of wheat glutenin. *J Cereal Sci.* (1995) 22:17–9 doi: 10.1016/S0733-5210(05)80003-7
30. Masci S, D'Ovidio R, Lafiandra D, Kasarda D. 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. doi: 10.1104/pp.118.4.1147
31. DuPont FM, Vensel WH, Chan R, Kasarda DD. Similarities of omega gliadins from *Triticum urartu* to those encoded on chromosome 1A of hexaploid wheat and evidence for their post-translational processing. *Theor Appl Genet*. (2004) 108:1299–308. doi: 10.1007/s00122-003-1565-9
32. DuPont FM, Altenbach SB. Molecular and biochemical impacts of environmental factors op wheat grain development and protein synthesis. *J Cereal Sci*. (2002) 38:133–46. doi: 10.1016/S0733-5210(03)00030-4
33. Altenbach SB. New insights into the effects of high temperature, drought and post-anthesis fertiliser on wheat grain development. *J Cereal Sci.* (2012) 56:39–50. doi: 10.1016/j.jcs.2011.12.012
34. Shewry PR, D'Ovidio R, Lafiandra D, Jenkins JA, Mills ENC, Bekes F. Wheat grain proteins. In: Khan K, Shewry PR, editors. *Wheat: Chemistry and Technology*. 4th ed. St. Paul Mn: AACC (2009), p. 223–98. doi: 10.1094/9781891127557.008
35. Belton PS. On the elasticity of wheat gluten. *J Cereal Sci.* (1999) 29:103–7. doi: 10.1006/jcrs.1998.0227
36. Sollid LM, Qiao S-W, Anderson RP, Gianfrani C, Konig F. Nomenclature and listings of celiac disease relevant gluten T-cell epotopies restricted by HLA-DQ molecules. *Immunogenetics*. (2012) 64:455–60. doi: 10.1007/s00251-012-0599-z
37. Shewry PR, Tatham AS. Improving wheat to remove coeliac epitopes but retain functionality. *J Cereal Sci.* (2016) 67:12–21. doi: 10.1016/j.jcs.2015.06.005
38. Gilissen LJWJ, van der Meer IM, Smulders MJ. Reducing the incidence of allergy and intolerance to cereals. *J Cereal Sci.* (2014) 59:337–53. doi: 10.1016/j.jcs.2014.01.005
39. Kreis M, Forde BG, Rahman S, Miflin BJ, Shewry PR. Molecular evolution of the seed storage proteins of barley, rye and wheat. *J Mol Biol.* (1985) 183:499-502 doi: 10.1016/0022-2836(85)90017-8
40. Shewry PR, Jenkins J, Beaudoin F, Mills ENC. The classification, functions and evolutionary relationships of plant proteins in relation to food allergens. In: Mills ENC, Shewry PR editors. *Plant Food Allergens.* Oxford: Blackwell Science (2004). 24–41. doi: 10.1002/9780470995174.ch2
41. Kasarda DD, Adalsteins E, Lew EJ-L, Lazo GR, Altenbach B. Farinin: characterisation of a novel wheat endosperm protein belonging to the prolamin superfamily. *J Agric Food Chem*. (2013) 61:2407–17. doi: 10.1021/jf3053466

42. Kan Y, Wan Y, Mitchell RAC, Beaudoin F, Leader DJ, Edwards KJ, et al. Transcriptome analysis reveals novel storage protein transcripts in seeds of *Aegilops* and wheat. *J Cereal Sci.* (2006) 44:75–85. doi: 10.1016/j.jcs.2006.04.004
43. Ma F, Li M, Yu L, Li Y, Liu Y, Li T, et al. Transformation of common wheat (*Triticum aestivum* L.) with *avenin-like b* gene improves flour mixing properties. *Mol Breed.* (2013) 32:853–65. doi: 10.1007/s11032-013-9913-1
44. Salcedo G, Prada J, Aragoncillo C. Low MW gliadin-like proteins from wheat endosperm. *Phytochem*. (1979) 18:725–7. doi: 10.1016/0031-9422(79)80003-5
45. Clarke BC, Phongkham T, Gianibelli MC, Beasley H, Békés F. The characterisation and mapping of a family of LMW-gliadin genes: effects on dough properties and bread volume. *Theor Appl Genet*. (2003) 106:629–35. doi: 10.1007/s00122-002-1091-1
46. Chantret N, Salse J, Sabot F, Rahman S, Bellec A, Laubin B, et al. Molecular basis of evolutionary events that shaped the *Hardness* locus in diploid and polyploidy wheat species (*Triticum* and *Aegilops*). *The Plant Cell.* (2005) 17:1033–45. doi: 10.1105/tpc.104.029181
47. Igrejas G, Gaborit T, Oury F-X, Chiron H, Marion D, Branlard G. Genetic and environmental effects on puroindoline-a and puroindoline-b content and their relationship to technological properties in French bread wheats. *J Cereal Sci*. (2001) 34:37–47. doi: 10.1006/jcrs.2000.0381
48. Turner AS, Bradburne RP, Fish L, Snape JW. New quantitative trait loci influencing grain texture and protein content in bread wheat. *J Cereal Sci*. (2004) 40:51–60. doi: 10.1016/j.jcs.2004.03.001
49. Van den Bulck K, Loosveld A-M A, Courtin CM, Proost P, Van Damme J, Robben J, et al. Amino acid sequence of wheat flour arabinogalactan-peptide, identical to part of grain softness protein GSP-1, leads to improved structural model. *Cereal Chem*. (2002) 79:329–31. doi: 10.1094/CCHEM.2002.79.3.329
50. Van den Bulck K, Swennen K, Loosveld A-MA, Courtin CM, Brijs K, Proost P, et al. Isolation of cereal arabinogalactan-peptides and structural comparison of their carbohydrate and peptide moieties *J Cereal Sci.* (2005) 41:59–67. doi: 10.1016/j.jcs.2004.10.001
51. Harris S, Powers S, Monteagudo-Mera A, Kosik O, Lovegrove A, Shewry PR, et al. Determination of the prebiotic activity of wheat arabinogalactan peptide (AGP) using batch culture fermentation. *Euro J Nutr Press.* (2019) 1–11. doi: 10.1007/s00394-019-01908-7. [Epub ahead of print].
52. Wilkinson MD, Tosi P, Lovegrove A, Corol DI, Ward JL, Palmer R, et al. The *Gsp-1* genes encode the wheat arabinogalactan peptide. *J Cereal Sci.* (2017) 74:155–64. doi: 10.1016/j.jcs.2017.02.006
53. Marion D, Douliez J-P, Gautier M-F, Elmorjani K. Plant lipid transfer proteins: relationships between allergenicity, and structural, biological and technological properties. In: Mills ENC, Shewry PR, editors. *Plant Food Allergens.* Oxford: Blackwell Science (2004). p. 57–86. doi: 10.1002/9780470995174.ch4
54. Pastorello EA, Farioli L, Conti A, Pravettoni V, Bonomi S, Iametti S, et al. Wheat IgE-mediated food allergy in European patients: a-amylase inhibitors, lipid transfer proteins and low molecular weight glutenins. Allergenic molecules revealed by double-blind, placebo-controlled food challenges. *Int Arch Allergy Immunol*. (2007) 144:10–22. doi: 10.1159/000102609
55. Palacin A, Quirce S, Armentaria A, Fernandez-Nieto M, Pacios LF, Asencio T, et al. Wheat lipid transfer protein is a major allergen associated with bakers' asthma. *J Allergy Clin Immunol.* (2007) 120:1132–8. doi: 10.1016/j.jaci.2007.07.008
56. Carbonero P, García-Olmedo F. A multigene family of trypsin/α-amylase inhibitors from cereals. In: Shewry PR and Casey R, editors. *Seed Proteins.* Dordrecht: Kluwer Academic Publishers (1999). p 617–33. doi: 10.1007/978-94-011-4431-5_26
57. Geisslitz S, Ludwig C, Scherf KA, Koehler P. Targeted LC–MS/MS reveals similar contents of α-amylase/trypsin-inhibitors as putative triggers of nonceliac gluten sensitivity in all wheat species except Einkorn. *J Agric Food Chem*. (2018) 66:12395–403. doi: 10.1021/acs.jafc.8b04411
58. Salcedo G, Sanchez-Monge R, Garcia-Casado G, Armentaria A, Gomez L, Barber D. The cereal α-amylase/trypsin inhibitor family associated with Bakers' asthma and food allergy. In: Mills ENC, Shewry PR editors. *Plant Food Allergens. Oxford UK: Blackwell Science.* (2004). p.70–86. doi: 10.1002/9780470995174.ch5
59. Kobrehel K, Alary R. Isolation and partial characterisation of two low-molecular weight durum-wheat (*Triticum durum*) glutenins. *J Sci Food Agric.* (1989) 48:441–52. doi: 10.1002/jsfa.2740480406
60. Kobrehel K, Alary R. The role of a low-molecular weight glutenin fraction in the cooking quality of durum-wheat pasta. *J Sci Food Agric.* (1989) 47:487–500. doi: 10.1002/jsfa.2740470409
61. Gautier M-F, Alary R, Kobrehel K, Joudrier P. Chloroform methanol-soluble proteins are the main components of *Triticum durum* sulfur-rich glutenin fractions. *Cereal Chem.* (1989) 66:535.
62. Gil-Humanes J, Piston F, Barro F, Rosell C. The shutdown of celiac disease-related gliadin epitopes in bread wheat provides flours with increased stability and better tolerance to over-mixing). *PLoS ONE.* (2014) 9:e91931. doi: 10.1371/journal.pone.0091931
63. Gil-Humanes J, Piston F, Altamirano-Fortoul R, Real A, Comino S, Sousa C, et al. Reduced-gliadin wheat bread: an alternative to the gluten-free diet for consumers suffering gluten-related pathologies. *PLoS ONE.* (2014) 9:e90898. doi: 10.1371/journal.pone.0090898

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Nutrition

REVIEW
published: 01 October 2019
doi: 10.3389/fnut.2019.00153

Toward New Paradigms in the Follow Up of Adult Patients With Celiac Disease on a Gluten-Free Diet

Maria I. Pinto-Sanchez[1,2]* *and Julio C. Bai*[3,4]

[1] Department of Medicine, Gastroenterology Division, McMaster University Medical Center, Hamilton, ON, Canada, [2] Farncombe Family Digestive Health Research Institute, Hamilton, ON, Canada, [3] Research Institute, Universidad del Salvador, Buenos Aires, Argentina, [4] Hospital de Gastroenterologia Dr. C. Bonorino Udaondo, Buenos Aires, Argentina

Gluten free diet is the only available treatment for celiac disease (CeD). Patients with CeD who do not adhere to a strict gluten-free diet (GFD) have been found to have complications involving nutritional deficiencies, increased risk of bone fractures, increased risk of mortality, and certain types of cancers. Complete removal of gluten from the diet in a patient with CeD often results in symptomatic, serologic, and histologic remission. However, strict compliance with the diet is challenging. Long-term follow-up care is needed to assure treatment compliance and positive health outcomes. Monitoring celiac specific serology, nutrient deficiencies, bone mineral density, and assessment of GFD compliance have been recommended in clinical practice. However, there is no consensus on which specific tests and how often they should be performed during the follow up. Here, we have performed a review of the literature on current strategies to follow up patients with CeD. There are new tools for monitoring adherence to the GFD which could change some paradigms in following up treated patients.

Keywords: celiac disease, follow-up, antibodies, gluten-free diet, biopsy

OPEN ACCESS

Edited by:
Katharina Anne Scherf, Karlsruhe Institute of Technology (KIT), Germany

Reviewed by:
Harald Engelhardt, Children's Hospital St. Mary's, Germany
Govind K. Makharia, All India Institute of Medical Sciences, India

***Correspondence:**
Maria I. Pinto-Sanchez
pintosm@mcmaster.ca

Specialty section:
This article was submitted to Nutrition and Food Science Technology, a section of the journal Frontiers in Nutrition

Received: *01 July 2019*
Accepted: *11 September 2019*
Published: *01 October 2019*

Citation:
Pinto-Sanchez MI and Bai JC (2019) Toward New Paradigms in the Follow Up of Adult Patients With Celiac Disease on a Gluten-Free Diet. Front. Nutr. 6:153. doi: 10.3389/fnut.2019.00153

BACKGROUND

Celiac disease (CeD) is a chronic systemic, immune-mediated condition precipitated by exposure to dietary gluten in genetically pre-disposed individuals (1). It is a relatively common disorder which affects around 1% of the population worldwide, and the prevalence has been increasing in the last years (2–4). The hallmark of CeD is enteropathy immune mediated, with characteristic villous atrophy in the proximal small intestine. CeD often presents with malabsorptive symptoms, including diarrhea and weight loss; with non-specific symptoms, such as abdominal pain, anemia, or osteopenia; or may be completely asymptomatic (3). Independently on the type of presentation, untreated, or partially treated celiac disease is associated with persistent symptoms and complications including nutritional deficiencies, osteoporosis, infertility, increased malignancies, and increased mortality (5).

The only available therapy for CeD is a strict, lifelong, gluten-free diet (GFD), which requires the complete removal of all wheat (gluten), rye (secalin), and barley (hordein) products. It is known that 50 mg of gluten (6–8), which could be found in a few crumbs of bread or a small piece of pasta, can perpetuate the enteropathy in patients with CeD. Due to accidental or intentional gluten exposure (contamination with gluten), it is not possible for some people to remain totally gluten-free. Therefore, most of patients with CeD are restricted gluten diet rather than gluten-free. Clinical studies using methods for indirect assessment of GFD compliance, such as food interviews, dietary

self-report, or follow-up serology showed that 17–80% of patients with CeD are not compliant with the GFD (9). Not surprisingly, their symptoms persist, and their small bowel does not heal (10). The negative psycho-social aspects of diet that is highly restricted, the need of permanent vigilance to avoid gluten, and the high frequency of inadvertent gluten exposure lead to low patient satisfaction and significant disease burden (11, 12). Patients with CeD often report decreased health-related quality of life (13) and a high treatment burden compared to those with other chronic diseases, including inflammatory bowel disease and type 1 diabetes; which are often perceived as more severe than celiac disease. For this reason, ideally immediately after diagnosis, patients with CeD should receive dietary counseling by an expert dietician in celiac disease and GFD compliance monitored in the follow up.

Assessment of Disease Activity After the Diagnosis

Celiac disease (CeD) is a systemic inflammatory condition, and may lead to serious complications if not adequately controlled. Even though it has been recommended that patients with CeD visit regularly the clinic, and specific markers of celiac disease are monitored after the diagnosis (14–19); patients with CeD are not followed up consistently (20). Improving understanding of the role of symptoms and tests in the follow-up of patients with CeD could positively impact on disease management.

Role of Symptoms/Signs in the Follow Up of CeD Patients

A substantial proportion of patients with CeD (∼30%) have recurrent or persistent symptoms despite being on a GFD (21), and the most common cause are continued or intermittent, purposeful or inadvertent gluten ingestion (20). Other causes of non-responsive celiac disease could be related to exocrine pancreatic insufficiency, bacterial overgrowth, microscopic colitis, carbohydrates (fructose/lactose) intolerance, or functional disorders (16, 21) (**Figure 1**). However, symptoms are not always present to alert for gluten ingestion, and some patients with CeD may persist with enteropathy for years without been aware (22). Independently of the presence or absence of symptoms, celiac patients with CeD with persistent enteropathy are at increased risk of complications, including lymphoproliferative malignancy, compared to those with mucosal healing (HR 2.26; 95% CI, 1.18–4.34) (22, 23). To prevent complications, current guidelines recommend regular follow up and monitoring of GFD compliance in both symptomatic and asymptomatic patients with CeD. There is general agreement among guidelines (16) that patients with CeD should be monitored at least two times in the first year after diagnosis, to assess disease activity, nutrition, dietary adherence, and bone health status (**Table 1**).

CeD Specific Serology in the Follow Up

IgA antibodies to TG2 and to deamidated gliadin peptides (DGPs) are commonly used to monitor celiac disease activity in the follow up (24). Although it takes several months for CeD specific serology to become under the normal cutoff level, a significant decrease in serology levels over the first year is suggestive of GFD adherence, and patients with CeD whose serologic features do not improve should be re-assessed for gluten exposure (25). However, negativity of CeD specific serology does not reflect strict compliance with GFD. In adult patients with CeD on a GFD, CeD serology is poor predictor of dietary transgressions (26). Although the CeD antibody tests show a high accuracy for the diagnosis of CeD, these tests are not as reliable in the follow up as they don't correlate well with histological findings or symptoms either (23). However, it is important to highlight that a negative CeD specific serology in a treated patient, does not necessarily guarantee intestinal mucosal healing (23, 26).

Even though CeD specific serology is imperfect test, guidelines recommend to assess CeD serology (anti- tissue transglutaminase; tTG IgA, or DGP IgA) every 3–6 months in the first year after the diagnosis or until stabilization, and then annually in the long term to monitor CeD activity (16). In cases of IgA deficiency, DGP IgG, and tTG IgG are recommended (14–16, 19).

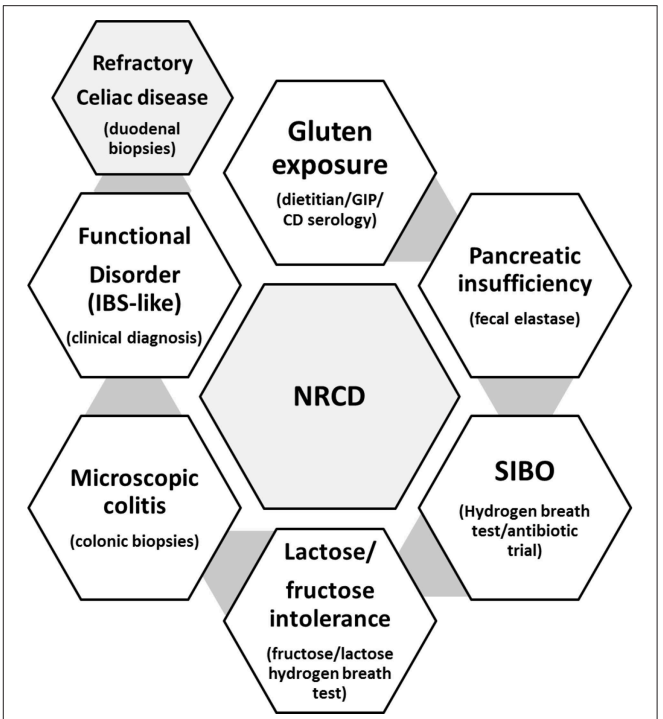


FIGURE 1 | Common causes for persistent symptoms in the follow-up of patients with CeD (non-responsive celiac disease; NRCD), and common tests used in clinical practice for the diagnosis of each concomitant condition.

Role of Endoscopy in the Follow Up

Repeated endoscopy with duodenal biopsies in the follow up has been controverted. There is currently no evidence indicating that performing routine follow-up biopsy is needed for all patients with CeD (17). Endoscopy is expensive, relatively invasive, and impractical procedure for regular disease activity monitoring. It has been suggested that biopsies should be repeated in the

TABLE 1 | Comparison of guidelines recommendations for follow up of adult patients with CeD.

Assessment*	ACG (16)	BSG (18)	WGO (19)	Kelly et al. (14)	AGA (15)	ESsCD (17)
Clinical Short term Long term	Every 6 months Annually	No specific recommendations	Annually after the 1st year in adults	Annually or if recurrent symptoms	No specific recommendations	3–4 months Annually
CeD serology Short term Long term	Every 6 months Annually	1 year	tTG IgA or DGP IgA Every 3–6 months until normal, then/1–2 years	Serology every 3–6 months until normal, then/1–2 years	Every 6 month Annually	tTG IgA 3–4 months Annually
Duodenal biopsy Short term Long term	If persistent Reasonable 1–2 year**	Not mandatory if asymptomatic	In symptomatic seronegatives at follow-up. Unclear in asymptomatic	Not mandatory. Consider 1–2 year. after diagnosis	If symptomatic If persistent enteropathy	If symptomatic Reasonable 1–2 year**
Screening for autoimmune Short term Long term	No specific recommendations	No specific recommendations	No specific recommendations	At diagnosis, then ev 1–2 year	No specific recommendations	At diagnosis E. 1–2 year
BMD Short term Long term	unclear unclear	In high risk for osteoporosis. Repeat if abnormal.	Baseline. Repeat if abnormal or at meno-andropause if normal	Baseline. Repeat if abnormal	No specific recommendations	Baseline If abnormal
GFD compliance Short term Long term	Every 3–6 month Annually	No specific recommendations	Every 3–6 months until normal, then/1–2 years. Potential use of GIP	No specific recommendations	No specific recommendations	3–4 month E. 1–2 year
Nutritional Short term Long term	Every 3 months until normal Annually	No specific recommendations	Every 3–6 months until normal, then/1–2 years	Every 3–6 months until normal, then/1–2 years	No specific recommendations	Baseline Annually
Vaccine	No specific recommendations	Pneumococci, in Hyposplenism *H. influenzae* Unclear	Pneumococci, H. influenzae, and meningococci should be performed	No specific recommendations	No specific recommendations	Pneumococci, in Hyposplenism *H. influenzae* Unclear

**Short term, <2 year after diagnosis; long-term, >2 year after diagnosis; DGP, Deaminated gliadin peptides.*

***It may be reasonable to do a follow-up biopsy in adults after 1–2 years of starting a GFD to assess for mucosal healing, especially in patients older than 40 years or in those having initially severe presentations.*

follow up of patients with CeD 2 years after the diagnosis to confirm mucosal healing (5, 15–17, 23). However, others have discouraged this practice based on previous demonstration of persistent damage in adults for years despite strict compliance with GFD (27, 28). There is general consensus that patients with persistent or newly developed symptoms without clear explanation, should undergo endoscopic biopsies to assess mucosal healing even if TG2-IgA levels are within normal range (23). Even though mucosal healing is likely in asymptomatic patients with negative serology on a GFD, studies suggested increased risk of lymphoma and mortality in this population with persistent inflammation (27). Therefore, current guidelines find reasonable a follow-up biopsy after 1–2 years of GFD, with the idea to assess mucosal healing, especially in patients over the age of 40 years or in those with severe presentations (16, 28). However, these recommendations are based on expert advice, and evidence on benefit of this strategy on long term outcomes is still lacking.

Nutritional Deficiencies in the Follow Up

Nutritional deficiencies in CeD may be directly related to celiac enteropathy, or could develop as a consequence of nutrients restriction associated to the GFD; or a combination of both factors (16, 29). The most common micronutrient deficiency is iron; however, iron stores typically improve on a GFD. Iron supplementation may be needed in a subset of patients with CeD. Folate, vitamin B12, vitamin D, and zinc are commonly deficient in patients with CeD in the follow up and often require supplementation (16). **Table 2** denotes the most frequent micronutrient deficiencies in celiac disease, and suggested supplementation.

It is strongly recommended that patients with CeD is assessed by an expert dietitian, to provide education on GFD and develop dietary strategies to help with symptoms management (16, 29).

Bones Disease in the Follow Up

Bone health can be negatively affected in CeD owing to the inflammatory process and malabsorption of calcium and vitamin D (30, 31). Osteopenia and osteoporosis and bone fractures are the most common complications associated with celiac disease (32). The risk of bone fractures is increased in celiac disease (33) regardless of the presence of symptoms; and the excess risk is reduced with adherence to GFD (34).

Testing of BMD should be performed at diagnosis of celiac disease before deciding on further management (35). In those with osteoporosis or osteopenia at diagnosis or those who

TABLE 2 | Common nutrient deficiencies in the follow up of adult patients with CeD and recommended oral supplementation.

Nutrient	Supplementation dose	Comments
Iron	*Oral supplements* Ferrous gluconate: 300 mg (35 mg) 1–3 tab. bid to tid Ferrous fumarate: 300 mg (100 mg) 1 tab. bid Ferrous sulfate : 300 mg (60 mg) 1 tab tid Heme Iron : 398 mg (11 mg Heme) 1 tab tid Polysaccharide: 150 mg (150 mg) 1 caps. Daily IV iron Iron sucrose: 200–300 mg 3–5 doses Iron dextran: 510 mg weekly × 2 doses	*Iron and ferritin at diagnosis *Vitamin C (500 units) may increase iron absorption *Zinc decrease absorption *IV iron should be considered in severe cases or intolerance to oral supplementation
Vitamin D	1,000–2,000 IU/day	*Taken with calcium to increase absorption
Folate	400–800 mcg/day	*Increased needs in pregnancy
B12	1,000–1,200 mcg/day	*Sublingual formulation available
Zinc	25–50 mg/day	*High zinc supplementation may lead to copper deficiency
Copper	2–4 mg/day	*Zinc and iron decrease copper absorption
Calcium	1,000–1,500 mg/day	*taken with vitamin D to increase absorption
Fiber	25–30 g/day	*Psyllium and Inulin most common encourage fluids
Chromium	200 mcg/day	*Interaction with PPIs, NSAIDS, and levothyroxine

**Testing for nutrients is recommended at diagnosis and if abnormal, repeat every 3–6 months until normal. Then once every 1–2 years.*

do not adhere to a GFD, a follow up BMD after at least 1 year of supplementation with calcium and vitamin D is recommended (31).

In addition to ensure strict GFD, it is prudent to ensure adequate calcium and vitamin D intake for all patients with CeD. If after 1–2 years of adhering to a GFD and including appropriate calcium and vitamin D supplementation the patient continues to show signs of osteoporosis, the addition of specific osteoactive treatments should be considered (31); despite no clear evidence on the magnitude of the benefit compared to the strict GFD alone. A recent study (30) has shown that a strict GFD improves the microstructural parameters of the bones, which is often difficult to reach, even with osteoactive treatment.

Monitoring Thyroid Function in the Follow Up

Celiac disease (CeD) has been associated to other autoimmune conditions, being the most frequent type 1 diabetes and autoimmune thyroiditis (36).

Autoimmune thyroid disease, especially Hashimoto's hypothyroidism is more frequent in patients with CeD (37). However, we need to consider that low-titer false-positive anti-tTG may occur in patients with thyroid disease (19).

There has been discussion on whether a gluten-free diet in CeD protects against thyroid disease or modifies the natural history of the disease. At least two studies (38, 39) suggest that gluten-free diet compliance does not influence on the development of thyroid disease. Regardless of the degree of compliance with the diet, experts recommend to monitor for thyroid disease in the follow up of patients with CeD (40). How frequent the thyroid tests should be ordered in the follow up of patients with CeD is not clearly stated.

Challenges of Monitoring of GFD Compliance

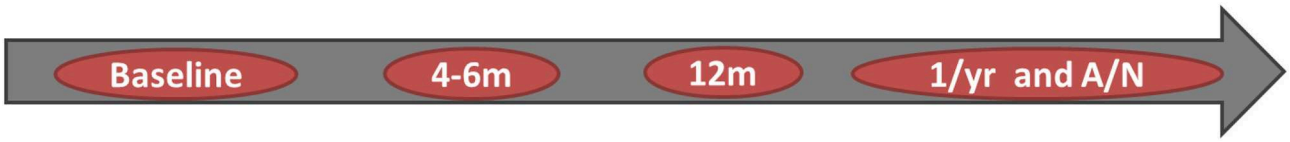
The management and follow-up of patients with CeD is preferentially performed with a team-based approach in which the dietician has an important role (15, 16) in the practical advice on lifestyle and choice of foods. It is well-known that 50 mg of gluten, which is equivalent to a few crumbs of bread or pasta, can produce symptoms and/or increase intestinal inflammation in patients with asymptomatic CeD; therefore, maintaining a lifelong GFD is necessary for all patients (25). The compliance with the diet could be impaired either with inadvertent or purposely gluten intake. Inadvertent gluten intake could be due to lack of proper knowledge, or lack of control on contamination; for example, when eating outside home.

A dietary assessment by an expert dietitian, generally based on an interview or food diary/food frequency questionnaire, is considered an objective, non-invasive, and low-cost way to measure adherence to a GFD (15, 16). However, a detailed dietary review for assessment of compliance with the diet is time consuming (between 45 min and 1 h), expensive to the healthcare system and limited by the lack of expert dietitians. Therefore, due to limited resources, it is not commonly performed in the community; with consequent limitations in the management of patients with CeD. In addition, individuals are not very accurate when reporting their adherence level, and whether intentionally or not, dietary review may not identify involuntary infringements. Identifying immunogenic peptides (9) either in stool, urine, or in food is a promising new tool to assess inadvertent gluten ingestion when patients are not under control of preparing their meals.

Gluten Immunogenic Peptides

There is an increasing interest on the role of certain gluten immunogenic peptides (GIP), such as 33-mer, that are resistant to digestion and are recognized triggers of immune reaction in celiac disease. In their study, Comino et al. (9) described a relatively new method to monitor GFD adherence by detection of GIP in stool samples 6–48 h after any intake of gluten by using the G12 monoclonal antibody. GIPs are excreted in feces after gluten is ingested; therefore, detection in stools of patients with CeD on a GFD reflects gluten exposure. GIPs could be detected in stool after ingestion of as little as 50 mg of gluten (equivalent to a penne noodle). This amount is clinically relevant as estimated ingestion of that amount of gluten per day has been proven to induce mucosal damage in patients

TABLE 3 | Recommended follow up for patients with CeD.



Baseline	4-6m	12m	1/yr and A/N
✓ Wt,Ht	✓ Wt,Ht,	✓ Wt,Ht,	✓ Wt,Ht,
✓ PE	✓ PE	✓ PE	✓ PE
✓ Ed. GFD	✓ Ed. GFD	✓ Ed. GFD	✓ Ed.GFD
✓ RD	✓ RD (by request)	✓ RD (by request)	✓ RD (by request)
✓ CCA	✓ Serology	✓ Serology	✓ Serology
✓ Nutrients	✓ Lab (if abn)	✓ Lab (if abn)	✓ Lab (if abn)
✓ Serology			✓ BMD every 2 year (if abn)
✓ Liver			
✓ TSH			
✓ BMD			
		*Offer GIP test	

**RD, registered dietitian; wt ht, weight and height; PE, physical examination; Ed.GFD, education on gluten free diet; TSH, thyroid stimulant hormone levels; BMD, bone mineral density; Lab laboratory; if abn, if abnormal; 1/yr, once per year; A/N, as needed.*

with CeD. The sensitivity and specificity of GIP testing in stool demonstrated in recent studies were 98.5 and 100%, which highlight the potential clinical usefulness of this new method as a marker of adherence to GFD in adults and children with CeD (41). Fecal GIP analysis has been proposed as a non-invasive and accurate method for a direct and quantitative assessment of gluten exposure. More recently, new tools for detection of GIP in stool and urine has been developed based on lateral flow immunoassays and the point-of-care technology. Based on these new tools, Costa et al. (41) have explored their utility for detecting GFD indiscretions in comparison with three-day dietary reports. The new tools for exploring GIP in stool is more sensitive than dietary reports in detecting short-term gluten exposure in patients with CeD on GFD, regardless of symptoms. Therefore, fecal GIP testing may help to guide patients with CeD during the treatment, as they often are exposed to gluten in the follow up, probably due to decreased awareness for cross contamination as the treatment progresses. These methods can complement the dietitian assessment of GFD compliance and clinical management of CeD.

When Gluten Free Diet Is Not Sufficient: Non-responsive and Refractory Celiac Disease

A great proportion of patients during the follow up present symptoms despite adhering to the gluten free diet, and this is known as Non-Responsive Celiac Disease (NRCD) (21). The most common reason for NRCD is the persistent stimulation by gluten1 (5, 21). Dietitian assessment plays a key role in identifying sources of unaware contamination with gluten. In the case of strict compliance with the diet, other concomitant conditions including small intestinal bacterial overgrowth, pancreatic insufficiency, parasite infections, or functional disorders such as IBS-like symptoms should be investigated (13, 15). The presence of persistent enteropathy in duodenal biopsies after 1 year of strict gluten free diet may suggest a rare complication known as refractory celiac disease (RCD) (15, 42). Further investigations of immunohistochemistry, PCR and flow cytometry will help to differentiate between refractory type 1 and 2. This differentiation is important, as RCD Type 2 is associated with worse prognosis and increased rates of mortality (42).

A strict GFD should be encouraged and monitored in patients with NRCD and RCD. Additional therapies will be required to treat the concomitant condition leading to persistent symptoms in NRCD, such as courses of antibiotics for SIBO, pancreatic enzymes for pancreatic insufficiency or motility agents for IBS-like symptoms. For RCD, treatment with budesonide or other immunosuppressants will be needed to control inflammation, as well as repeated biopsies in the follow up and images to monitor the disease and rule out further complications (42, 43). Patients with RCD will benefit from the referral to a specialized center for further management of their condition (43, 44).

What Are the Benefits of Following Up of Patients With CeD and Monitoring Their GFD Compliance?

Celiac disease is a chronic inflammatory condition, and persistence of inflammatory state may lead to complications including nutritional deficiencies, osteoporosis and increased risk of certain types of cancer (45). The risk for complications is increased in persistent active disease, regardless of the presence or absence of symptoms. It is well-known that a compliance with the GFD will lead to disease control in a great majority of patients with CeD, and consequently, decreased risk of complications and mortality (23). A strict gluten free diet is difficult to follow, and patients often are exposed to gluten in the follow up. Therefore, guidelines recommend adequate follow up to monitor for GFD compliance to prevent serious complications associated to the condition. **Table 3** summarizes recommendations for follow up of patients with CeD.

Who Should Follow Celiac Patients?

There is no consensus on whom and how should patients with CeD be followed-up, and there were several studies attempting to clarify this issue. Whether a great proportion of patients preferred to be followed-up by both a dietitian and a doctor (46) a study from Finland, demonstrated that follow-up by primary care providers is also effective (47). If experienced, primary care physicians should be responsible of following up patients with celiac disease.

CONCLUSION

Patients with CeD should be monitored in the short and long term to ensure an adequate control on disease activity; regardless symptoms are present or not. Even though there is consensus on the need of clinical, serological, nutritional, and bone health status assessment in the follow up, there are still areas of uncertainty. The development of new tools will lead to changes in strategies to explore adherence to treatment in patients with CeD. Studies involving long term follow up are encouraged to clarify the role of endoscopy and of new tools to monitor GFD compliance on disease outcomes.

AUTHOR CONTRIBUTIONS

MP-S and JB designed the review, wrote the manuscript, and approved the final version.

FUNDING

MP-S holds a Medicine Internal Career Research Award, and Farncombe Family Digestive Health Research Institute award.

REFERENCES

1. Ludvigsson JF, Leffler DA, Bai JC, Biagi F, Fasano A, Green PHR, et al. The Oslo definitions for coeliac disease and related terms. *Gut*. (2013) 62:43–52. doi: 10.1136/gutjnl-2011-301346
2. Lebwohl B, Ludvigsson JF, Green PHR. Celiac disease and non-celiac gluten sensitivity. *BMJ*. (2015) 351:h4347. doi: 10.1136/bmj.h4347
3. Leonard MM, Sapone A, Catassi C, Fasano A. Celiac disease and nonceliac gluten sensitivity: a review. *JAMA*. (2017) 318:647–56. doi: 10.1001/jama.2017.9730
4. Ludvigsson JF, Rubio-Tapia A, van Dyke CT, Melton LJ, Zinsmeister AR, Lahr BD, et al. Increasing incidence of celiac disease in a North American population. *Am J Gastroenterol*. (2013) 108:818–24. doi: 10.1038/ajg.2013.60
5. Rubio-Tapia A, Rahim MW, See JA, Lahr BD, Wu TT, Murray JA. Mucosal recovery and mortality in adults with celiac disease after treatment with a gluten-free diet. *Am J Gastroenterol*. (2010) 105:1412–20. doi: 10.1038/ajg.2010.10
6. El Khoury D, Balfour-Ducharme S, Joye IJ. A review on the gluten-free diet: technological and nutritional challenges. *Nutrients*. (2018) 10:E1410. doi: 10.3390/nu10101410
7. Catassi C, Fabiani E, Iacono G, D'Agate C, Francavilla R, Biagi F, et al. A prospective, double-blind, placebo-controlled trial to establish a safe gluten threshold for patients with celiac disease. *Am J Clin Nutr*. (2007) 85:160–6. doi: 10.1093/ajcn/85.1.160
8. Akobeng AK, Thomas AG. Systematic review: tolerable amount of gluten for people with coeliac disease. *Aliment Pharmacol Ther*. (2008) 27:1044–52. doi: 10.1111/j.1365-2036.2008.03669.x
9. Comino I, Fernández-Bañares F, Esteve M, Ortigosa L, Castillejo G, Fambuena B, et al. Fecal gluten peptides reveal limitations of serological tests and food questionnaires for monitoring gluten-free diet in celiac disease patients. *Am J Gastroenterol*. (2016) 111:1456–65. doi: 10.1038/ajg.2016.439
10. Moreno ML, Cebolla Á, Muñoz-Suano A, Carrillo-Carrion C, Comino I, Pizarro Á, et al. Detection of gluten immunogenic peptides in the urine of patients with coeliac disease reveals transgressions in the gluten-free diet and incomplete mucosal healing. *Gut*. (2015) 2015:1–8. doi: 10.1136/gutjnl-2015-310148
11. Shah S, Akbari M, Vanga R, Kelly CP, Hansen J, Theethira T, et al. Patient perception of treatment burden is high in celiac disease compared with other common conditions. *Am J Gastroenterol*. (2014) 109:1304–11. doi: 10.1038/ajg.2014.29
12. Hall NJ, Rubin G, Charnock A. Systematic review: adherence to a gluten-free diet in adult patients with coeliac disease. *Aliment Pharmacol Ther*. (2009) 30:315–30. doi: 10.1111/j.1365-2036.2009.04053.x
13. Nachman F, del Campo MP, González A, Corzo L, Vázquez H, Sfoggia C, et al. Long-term deterioration of quality of life in adult patients with celiac disease is associated with treatment noncompliance. *Dig Liver Dis*. (2010) 42:685–91. doi: 10.1016/j.dld.2010.03.004
14. Kelly CP, Bai JC, Liu E, Leffler DA. Advances in diagnosis and management of celiac disease. *Gastroenterology*. (2015) 148:1175–86. doi: 10.1053/j.gastro.2015.01.044
15. Husby S, Murray J, Katzka DA. AGA clinical practice update on diagnosis and monitoring of celiac disease—changing utility of serology and histologic measures: expert review. *Gastroenterology*. (2019) 156:885–9. doi: 10.1053/j.gastro.2018.12.010
16. Rubio-Tapia A, Hill ID, Kelly CP, Calderwood AH, Murray JA. ACG clinical guidelines: diagnosis and management of celiac disease. *Am J Gastroenterol*. (2013) 108:656–76. doi: 10.1038/ajg.2013.79
17. Al-Toma A, Volta U, Auricchio R, Castillejo G, Sanders DS, Cellier C, et al. European Society for the Study of Coeliac Disease (ESsCD) guideline for coeliac disease and other gluten-related disorders. *United Euro. Gastroenterol J.* (2019) 7:583–613. doi: 10.1177/2050640619844125
18. Ludvigsson JF, Bai JC, Biagi F, Card TR, Ciacci C, Ciclitira PJ, et al. Diagnosis and management of adult coeliac disease: guidelines from the British Society of Gastroenterology. *Gut*. (2014) 63:1210–28. doi: 10.1136/gutjnl-2013-306578
19. Bai JC, Ciacci C. World gastroenterology organization global guidelines: celiac disease February 2017. *J Clin Gastroenterol*. (2017) 51:755–68. doi: 10.1097/MCG.0000000000000919
20. Herman ML, Rubio-Tapia A, Larson JJ, Larson JJ, Van Dyke CT, Murray JA. Patients with celiac disease are not followed up adequately. *Clin Gastroenterol Hepatol*. (2012) 10:893–9. doi: 10.1016/j.cgh.2012.05.007
21. Leffler DA, Dennis M, Hyett B, Kelly E, Schuppan D, Kelly CP. Etiologies and predictors of diagnosis in nonresponsive celiac disease. *Clin Gastroenterol Hepatol*. (2007) 5:445–50. doi: 10.1016/j.cgh.2006.12.006
22. Mahadev S, Murray JA, Wu TT, Chandan VS, Torbenson MS, Kelly CP, et al. Factors associated with villus atrophy in symptomatic coeliac disease patients on a gluten-free diet. *Aliment Pharmacol Ther*. (2017) 45:1084–93. doi: 10.1111/apt.13988
23. Ludvigsson JF. Mortality and malignancy in celiac disease. *Gastrointest Endosc Clin N Am*. (2012) 22:705–22. doi: 10.1016/j.giec.2012.07.005
24. Ludvigsson JF, Ciacci C, Green PH, Kaukinen K, Korponay-Szabo IR, Kurppa K, et al. Outcome measures in coeliac disease trials: the Tampere recommendations. *Gut*. (2018) 67:1410–24. doi: 10.1136/gutjnl-2017-314853
25. Husby S, Bai JC. Follow up of celiac disease. *Gastroenterol Clin N Am*. (2018) 48:127–36. doi: 10.1016/j.gtc.2018.09.009
26. Vahedi K, Mascart F, Mary JY, Laberenne JE, Bouhnik Y, Morin MC, et al. Reliability of antitransglutaminase antibodies as predictors of gluten-free diet compliance in adult celiac disease. *AmJ G*. (2003) 98:1079–87. doi: 10.1111/j.1572-0241.2003.07284.x

27. Sharkey LM, Corbett G, Currie E, Lee J, Sweeney N, Woodward JM. Optimising delivery of care in coeliac disease–comparison of the benefits of repeat biopsy and serological follow-up. *Aliment Pharmacol Ther*. (2013) 38:1278–91. doi: 10.1111/apt.12510
28. Lebwohl B, Granath F, Ekbom A, Montgomery SM, Murray JA, Rubio-Tapia A, et al. Mucosal healing and mortality in coeliac disease. *Aliment Pharmacol Ther*. (2013) 37:332–9. doi: 10.1111/apt.12164
29. Theethira TG, Dennis M Leffler D. Nutritional consequences of celiac disease and the gluten-free diet. *Expert Rev Gastroenterol. Hepatol.* (2014) 8:123–9. doi: 10.1586/17474124.2014.876360
30. Zanchetta MB, Costa AF, Longobardi V, Mazure R, Silveira F, Temprano MP, et al. Improved bone microarchitecture in patients with celiac disease after 3 years on a gluten-free diet. *Clin Gastroenterol Hepatol*. (2018) 16:774–5. doi: 10.1016/j.cgh.2017.09.054
31. Duerksen D, Pinto-Sanchez MI, Anca A, Schnetzler J, Case S, Zelin J, et al. Management of bone health in patients with celiac disease. Practical guide for clinicians. *Can Fam Phys.* (2018) 64:433–8.
32. Zanchetta MB, Longobardi V, Bai JC. Bone and celiac disease. *Curr Osteoporos Rep.* (2016) 14:43–8. doi: 10.1007/s11914-016-0304-5
33. Moreno ML, Vazquez H, Mazure R, Smecuol E, Niveloni S, Pedreira S, et al. Stratification of bone fracture risk in patients with celiac disease. *Clin Gastroenterol Hepatol.* (2004) 2:127–34. doi: 10.1016/S1542-3565(03)00320-3
34. Sánchez MI, Mohaidle A, Baistrocchi A, Matoso D, Vázquez H, Gonzalez A, et al. Risk of fracture in celiac disease: gender, dietary compliance, or both? *World J Gastroenterol.* (2011) 17:3035–42. doi: 10.3748/wjg.v17.i25.3035
35. Fouda MA, Khan AA, Sultan MS, Rios LP, McAssey K, Armstrong D. Evaluation and management of skeletal health in celiac disease: position statement. *Can J Gastroenterol*. (2012) 26:819–29. Erratum in: *Can J Gastroenterol Hepatol.* (2017) 2017:1323607. doi: 10.1155/2012/823648
36. Omar NN, Azeez G, Roger D. Smalligan and Steven urban review and practice guidelines for celiac disease in 2014. *Postgrad. Med.* (2015) 27:259–65. doi: 10.1080/00325481.2015.1015926
37. Sun X, Lu L, Yang R, Li Y, Shan L, Wang Y. Increased incidence of thyroid disease in patients with celiac disease: a systematic review and meta-analysis. *PLoS ONE.* (2016) 11:e0168708. doi: 10.1371/journal.pone.0168708
38. Cooper BT, Holmes GK, Cooke WT. Coeliac disease and immunological disorders. *Br Med J*. (1978) 1:537–39. doi: 10.1136/bmj.1.6112.537
39. Mainardi E, Montanelli A, Dotti M, Nano R, Moscato G. Thyroid-related autoantibodies and celiac disease: a role for a gluten-free diet? *J Clin Gastroenterol*. (2002) 35:245–8. doi: 10.1097/00004836-200209000-00009
40. Fasano A, Catassi C. Celiac disease. *N Engl J Med*. (2012) 367:2419–26. doi: 10.1056/NEJMcp1113994
41. Costa AF, Sugai E, Temprano MP, Niveloni SI, Vázquez H, Moreno ML, et al. Gluten immunogenic peptide excretion detects dietary transgressions in treated celiac disease patients. *World J Gastroenterol*. (2019) 25:1409–20. doi: 10.3748/wjg.v25.i11.1409
42. Krauss N, Schuppan D. Monitoring nonresponsive patients who have celiac disease. *Gastrointest Endosc Clin N Am*. (2006) 16:317–27. doi: 10.1016/j.giec.2006.07.001
43. Mauriño E, Niveloni S, Cherñavsky AC, Sugai E, Vázquez H, Pedreira S, et al. Clinical characteristics and long-term outcome of patients with refractory sprue diagnosed at a single institution. *Acta Gastroenterol Latinoam*. (2006) 36:10–22.
44. van Wanrooij RL, Bouma G, Bontkes HJ, Neefjes-Borst A, van Grieken NC, von Blomberg BM, et al. Outcome of referrals for non-responsive celiac disease in a tertiary center: low incidence of refractory celiac disease in the Netherlands. *Clin Transl Gastroenterol*. (2017) 8:e218. doi: 10.1038/ctg.2016.70
45. Corrao G, Corazza GR, Bagnardi V, Brusco G, Ciacci C, Cottone M, et al. Mortality in patients with coeliac disease and their relatives: a cohort study. *Lancet*. (2001) 358:356–61. doi: 10.1016/S0140-6736(01)05554-4
46. Hughey JJ, Ray BK, Lee AR, Voorhees KN, Kelly CP, Schuppan D. Self-reported dietary adherence, disease-specific symptoms, and quality of life are associated with healthcare provider follow-up in celiac disease. *BMC Gastroenterol*. (2017) 17:156. doi: 10.1186/s12876-017-0713-7
47. Kurppa K, Lauronen O, Collin P, Ukkola A, Laurila K, Huhtala H, et al. Factors associated with dietary adherence in celiac disease: a nationwide study. *Digestion*. (2012) 86:309–14. doi: 10.1159/000341416

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 15 October 2019
doi: 10.3389/fnut.2019.00162

Preparation and Characterization of Avenin-Enriched Oat Protein by Chill Precipitation for Feeding Trials in Celiac Disease

Greg Tanner[1†], Angéla Juhász[2†], Christakis George Florides[3], Mitchell Nye-Wood[2], Frank Békés[4], Michelle L. Colgrave[2], Amy K. Russell[5,6], Melinda Y. Hardy[5,6] and Jason A. Tye-Din[5,6,7,8]*

[1] School of Biosciences, University of Melbourne, Melbourne, VIC, Australia, [2] School of Science, Edith Cowan University, Joondalup, WA, Australia, [3] School of Veterinary and Life Sciences, Murdoch University, Murdoch, WA, Australia, [4] FBFD PTY LTD, Sydney, NSW, Australia, [5] Immunology Division, The Walter and Eliza Hall Institute, Melbourne, VIC, Australia, [6] Department of Medical Biology, The University of Melbourne, Melbourne, VIC, Australia, [7] Department of Gastroenterology, The Royal Melbourne Hospital, Parkville, VIC, Australia, [8] Murdoch Children's Research Institute, Royal Children's Hospital, Melbourne, VIC, Australia

OPEN ACCESS

Edited by:
Satyanarayan R. S. Dev, Florida A&M University, United States

Reviewed by:
Sachin Rustgi, Clemson University, United States
Venkatesan Arul, Pondicherry University, India

***Correspondence:**
Greg Tanner
gregory.tanner@unimelb.edu.au

†Joint first authors

Specialty section:
This article was submitted to Nutrition and Food Science Technology, a section of the journal Frontiers in Nutrition

Received: *31 July 2019*
Accepted: *26 September 2019*
Published: *15 October 2019*

Citation:
Tanner G, Juhász A, Florides CG, Nye-Wood M, Békés F, Colgrave ML, Russell AK, Hardy MY and Tye-Din JA (2019) Preparation and Characterization of Avenin-Enriched Oat Protein by Chill Precipitation for Feeding Trials in Celiac Disease. Front. Nutr. 6:162. doi: 10.3389/fnut.2019.00162

The safety of oats for people with celiac disease remains unresolved. While oats have attractive nutritional properties that can improve the quality and palatability of the restrictive, low fiber gluten-free diet, rigorous feeding studies to address their safety in celiac disease are needed. Assessing the oat prolamin proteins (avenins) in isolation and controlling for gluten contamination and other oat components such as fiber that can cause non-specific effects and symptoms is crucial. Further, the avenin should contain all reported immunogenic T cell epitopes, and be deliverable at a dose that enables biological responses to be correlated with clinical effects. To date, isolation of a purified food-grade avenin in sufficient quantities for feeding studies has not been feasible. Here, we report a new gluten isolation technique that enabled 2 kg of avenin to be extracted from 400 kg of wheat-free oats under rigorous gluten-free and food grade conditions. The extract consisted of 85% protein of which 96% of the protein was avenin. The concentration of starch (1.8% dry weight), β-glucan (0.2% dry weight), and free sugars (1.8% dry weight) were all low in the final avenin preparation. Other sugars including oligosaccharides, small fructans, and other complex sugars were also low at 2.8% dry weight. Liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis of the proteins in these preparations showed they consisted only of oat proteins and were uncontaminated by gluten containing cereals including wheat, barley or rye. Proteomic analysis of the avenin enriched samples detected more avenin subtypes and fewer other proteins compared to samples obtained using other extraction procedures. The identified proteins represented five main groups, four containing known immune-stimulatory avenin peptides. All five groups were identified in the 50% (v/v) ethanol extract however the group harboring the epitope DQ2.5-ave-1b was less represented. The avenin-enriched protein fractions were quantitatively collected by reversed phase HPLC and analyzed by

MALDI-TOF mass spectrometry. Three reverse phase HPLC peaks, representing ∼40% of the protein content, were enriched in proteins containing DQ2.5-ave-1a epitope. The resultant high quality avenin will facilitate controlled and definitive feeding studies to establish the safety of oat consumption by people with celiac disease.

Keywords: oats, avenin, gluten-free diet, LC-MS, MALDI-TOF, celiac disease

INTRODUCTION

Establishing a safe gluten-free diet (GFD) for people with celiac disease (CD) is important for the health of the 1.4% of affected sufferers globally (1). CD is a chronic immune illness with features closely related to autoimmunity, and its pathogenesis is strongly linked to $CD4^+$ T cells that are activated by dietary gluten peptides (2). Since the 1950s, after gluten was identified as the causative antigenic trigger, its sole treatment has been strict and lifelong gluten exclusion by removing wheat, barley, and rye. In this clinical context, the term gluten encompasses all of the pathogenic prolamin proteins from wheat (gliadin and glutenin), barley (hordein) and rye (secalin) that are immunotoxic in CD.

Oats contain a gluten-like prolamin protein called avenin and early food challenge studies indicated they can cause clinical relapse in some people with CD (3). More recent oat feeding studies that control for the confounding effect of wheat and barley contamination have indicated that pure oats are safe in CD although methodological limitations of these studies have been identified (4). A few clinical studies have shown mucosal inflammation (5), raised intra-epithelial lymphocytes (6, 7), or villous atrophy (8) induced by oats, suggestive of an adverse CD effect. A variety of *in vitro* studies have shown oat fractions can induce pro-inflammatory immune effects (9–11). Importantly, key avenin peptides that stimulate the pathogenic gluten-specific T cells in CD patients *in vivo* have been defined (12, 13). These peptides contain the immunodominant T cell epitopes DQ2.5-ave-1a (PYPEQEEPF), DQ2.5-ave-1b (PYPEQEQPF), DQ2.5-ave-1c (PYPEQEQPI), and DQ2.5-ave-2 (PYPEQQPF) with close sequence homology to barley T cell epitopes immunotoxic in CD such as DQ2.5-hor-3 (PIPEQPQPY) (14). The collective uncertainty of these findings regarding the true clinical safety of oats in CD has translated into different feeding recommendations; while Australia and New Zealand mandate the exclusion of oats from the GFD, most countries do not.

Oats are the sixth most significant cereal crop in the world, with production exceeding 24 million tons annually, and *Avena sativa* is the most important crop (15). Oats, wheat, barley and rye belong to the same Poaceae family but oats are sub-classified into the Aveneae tribe, while the other cereals belong to the Triticeae tribe. This phylogenetic relationship is exemplified by the homology between oat avenin sequences with those in α- and γ-gliadins of wheat, the B-hordeins of barley and the γ-secalins of rye. Notably, there is no homology with the 33 mer peptide from wheat α-gliadin that encompasses several highly immunostimulatory T cell epitopes in CD. An important distinction is that within the Triticeae tribe, gluten protein makes up 75–80% of the protein in wheat, 45–50% in rye, and 50–55% in barley, but in the Aveneae tribe the equivalent prolamin protein (avenin) makes up only 10–15% of the protein i.e., 1% of the flour (16). This has particular relevance for feeding studies in CD. To illustrate, the consumption of 3–7 g wheat gluten daily induces typical clinical effects in most CD patients after 2 weeks (17), however the equivalent avenin prolamin dose would require the consumption of ∼300–700 g oats per day. As a standard serving size is ∼30–40 g, this is impractical and makes oat feeding studies aiming to reliably induce and assess biological effects near impossible.

A strict GFD is essential to ensure mucosal healing in CD and a failure to heal is correlated with higher morbidity and mortality (18, 19). This is an onerous, costly and restrictive treatment and its ability to induce CD remission is compromised by poor dietary adherence (20). The GFD is generally lower in dietary fiber and frequently higher in simple carbohydrates and fat (21). The introduction of oats to the GFD increases the range of foods that can be consumed, provides an excellent source of fiber and increases GFD palatability and adherence. Soluble fiber and β-glucan found in oats have been associated with a range of health benefits including reduced serum cholesterol (22, 23). Oats may support the 8–10% of CD patients who also suffer Type 1 diabetes mellitus by lowering post-prandial glycaemia and improving glycaemic control. Long-term oat consumption in CD may also improve quality of life (24).

In the face of these extremely positive nutritional attributes for oats in people with CD yet with the uncertain issue of their safety, there is a strong medical need to resolve this issue. It has become clear from trials in a separate clinical entity, non-celiac wheat sensitivity (NCWS), that properly controlled feeding studies to assess the clinical effects of gluten need to feed purified gluten, and not whole wheat (25). Wheat contains a broad mix of proteins [including gliadins, glutenins, amylase trypsin inhibitors (ATIs), and globulins] and carbohydrates. Recent controlled feeding studies indicate the fermentable carbohydrate component, fructan, is the driver of adverse gastrointestinal symptomatology in many people with NCWS (26). While oats, unlike wheat, is considered low in fermentable carbohydrates (27), it is high in fiber which could

Abbreviations: AACC, American Association of Cereal Chemists; ATIs, α-amylase/trypsin inhibitors; CD, coeliac disease; cRAP, common Repository of Adventitious Proteins; DTT, dithiothreitol; FDA, US Food and Drug Administration; GFD, gluten-free diet; HRP, horseradish peroxidase; IPA, propan-2-ol; IPA/DTT, 50% (v/v) IPA, 1% (w/v) DTT; LC-MS/MS, liquid chromatography tandem mass spectrometry; LOD, limit of detection; MALDI-TOF-MS, matrix-assisted laser desorption ionization time-of-flight mass spectrometry; NCWS, non-celiac wheat sensitivity; RP-HPLC, reverse phase high pressure liquid chromatography; Urea/DTT, 8 M urea, 1% (w/v) DTT, 20 mM triethylamine-HCl (pH 6).

trigger adverse gastrointestinal symptoms such as bloating or abdominal discomfort independent of its avenin content. With remarkable foresight, researchers in 1958 reported the challenge of getting children with CD to consume enough oats to establish clinical safety, and identified the need to extract oat protein i.e., avenin to address this challenge (28). At the time, the author lamented, "The possibility of using oat protein itself was explored but although several methods of preparation were attempted, considerable difficulties were encountered." Thus, there exists a need for purified and uncontaminated food-grade avenin in sufficient quantities for CD feeding studies. We present here a simple method capable of isolating purified, avenin-enriched protein. We demonstrate the avenin is of food-grade standard, uncontaminated by agricultural chemicals and heavy metals, and suitable for human feeding trials.

METHODS

Protein Content Determination

Protein content was determined by the method of Bradford (29).

Proximate Composition Analysis

All proximate analyses were completed by a National Association of Testing Authorities accredited food testing facility Agrifood Technology (Werribee, Australia). In brief, protein content was determined using the Kjeldahl method [American Association of Cereal Chemists (AACC) Methods 70-20A and 70-70]; starch and total free sugars (fructose, glucose, lactose, maltose, and sucrose) by in-house developed liquid chromatography—mass spectrometry (LC-MS); β-glucan using enzymatic method after hydrolysis by β-glucosidase (AOAC International Official Method 995.16 and AACC Method 32-23); and water soluble carbohydrates by water extraction and the anthrone method (AFIA 1.11R).

Urea-SDS-PAGE and Western Blot Analysis

One dimensional urea-sodium dodecyl sulfate polyacrylamide protein gels and Western blots were run as previously described (30). The Sigma-anti-gliadin-horseradish peroxidase (HRP) antibody used here, has been shown to be a general anti-gluten antibody, detecting all gluten protein types present in wheat, barley, rye, and oats (31).

Oat Cultivation and Purity

Two crops of oats (cv. Wandering, 200 kg) were grown in Williams, in the south-east of Western Australia, using dedicated wheat-free machinery and cropland, and harvested in September 2016 and December 2017. The oats were transported to Melbourne in sealed "bulka" bags and processed in two batches of 200 kg. Prior to grinding of each batch, sequential lots of 100 g of oats was spread thinly on a tray and examined for other grains. No wheat, barley or ryegrass were detected in any samples. No other material was detected, confirming the purity of the oats. The oat grain was ground to pass a 40 hole/in screen, in a dedicated gluten-free hammer mill kindly supplied by Wards Mackenzie (Altona, Australia). The flour was captured in two batches of eight 25 kg bags. Before each bag was sealed, a 100 g flour sample was taken from each bag and screened for accidental contamination with herbicides, pesticides and common aflatoxins (**Supplementary Table S1**) by Agrifood Technology (Werribee, Australia). Purified avenin was also screened for heavy metal contamination. The flour was extracted using food grade procedures and ethanol in a lab decontaminated to remove traces of wheat flour (Manildra Group, Nowra). All containers used for solvent storage and extraction were Food and Drug Approved and bisphenol A-free (Bunnings, Australia).

Effect of Solvent Polarity on Small Scale Avenin Precipitation

Oat flour (500 g) was extracted twice in 50% (v/v) ethanol (750 mL) and the extracts pooled. Duplicate 10 mL aliquots of the pooled 50% (v/v) ethanol extract containing 5.2 mg protein/mL was subject to varied total ethanol concentration by diluting the 50% (v/v) ethanol extract with either water, to achieve a final ethanol concentrations of 10–41% (v/v), or with ethanol to achieve final ethanol concentrations of 66–90% (v/v) (**Figure 1A**). In addition the 50% (v/v) ethanol extract was also chilled at 4°C and centrifuged as below (**Figure 1A**). Ethanol extracts were centrifuged at 3000 × g for 10 min at room temperature and pellets were dissolved in 8 M urea, 1% (w/v) dithiotreitol (DTT), 20 mM triethylamine-HCl (pH 6) (Urea/DTT) during overnight incubation at room temperature. The protein content of the resultant solutions was measured by Bradford, and 20 and 2 μg protein aliquots were subjected to Urea-SDS-PAGE and western blot respectively (**Figures 1B,C**).

Chill Precipitation of a Range of Gluten Proteins

Gluten proteins were isolated from wheat, barley and oat flour by extracting 5 g flour in 15 mL of 50% (v/v) ethanol, vortexing regularly over 1 h, and centrifuging at 3,000x g for 1 h. The 50% ethanol (v/v) supernatants were chilled at 4°C overnight, centrifuged as above and the pellets redissolved in Urea/DTT (wheat 10 mL, barley 10 mL, and oats 1 mL). The prolamin content of the chill precipitates were compared to duplicate extracts of 50 mg (wheat, barley) or 100 mg oats in 1 mL of 50% (v/v) propan-2-ol (IPA), 1% (w/v) DTT (IPA/DTT) which was extracted by violent reciprocal shaking in a Savant bead beater at 30 movements s^{-1} for 1 min and centrifuged at 13,000x g for 5 min. The protein content in the IPA/DTT supernatants and the chill-precipitated pellets were determined and either 20 or 2 μg protein were loaded on each lane of a protein gel and western blot, respectively (**Figures 2A,B**).

Avenin Isolation From 500 g Oats

Oat flour (500 g) was shaken regularly over 2 min, 90 min, or for one or 2 days in 750 mL of 50% (v/v) ethanol and then centrifuged at 500 g for 5 min and the supernatant reserved. Pellets were resuspended in 750 mL 50% (v/v) ethanol, re-centrifuged and the process repeated. The yield of protein was determined in the pooled supernatants (**Figures S1A,B**).

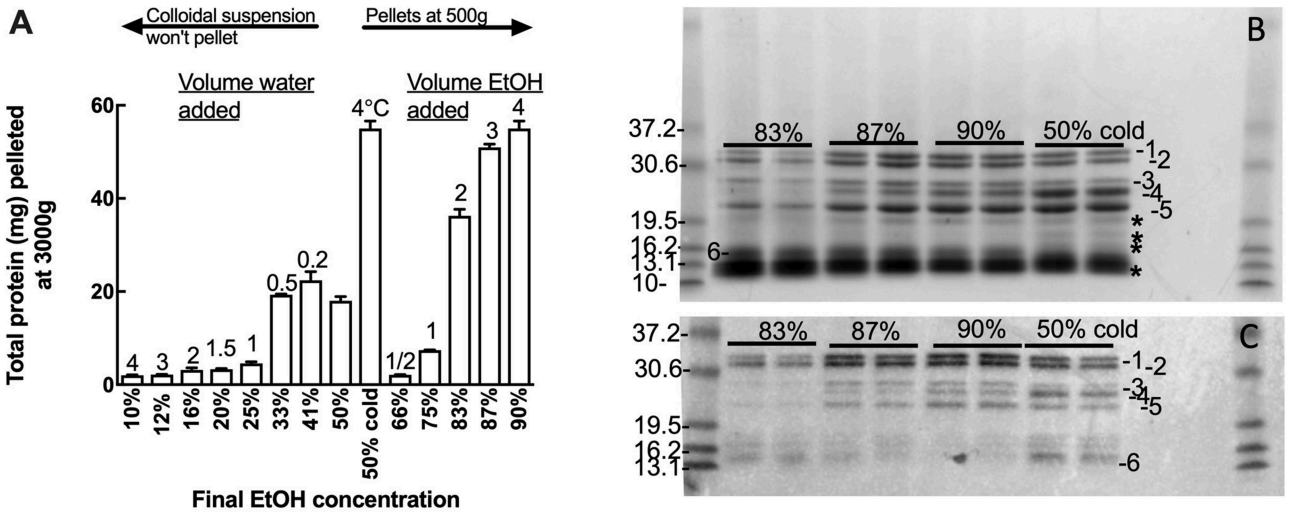


FIGURE 1 | Effect of solvent concentration on avenin precipitation. Adding water or ethanol (EtOH) precipitated avenin, however water-induced precipitate could not be conveniently pelleted. Chill precipitation isolated as much protein as could be isolated by solvent precipitation with a final concentration of 90% ethanol. **(B,C)** Effect of ethanol concentration (83, 87, 90, or 50% chill precipitation) on avenin precipitation by **(B)** protein gel or **(C)** western blot with Sigma anti-gliadin, this antibody has previously been shown to identify all major gluten proteins including avenin (30). Protein bands not corresponding to western bands are indicated on **(B)**, *. Prestained 10 kDa ladder (Benchmark, Invitrogen) is shown to the left of **(B,C)**. The pre-stained standards were themselves calibrated against Benchmark unstained standards which have accurately designated molecular weights.

Large Scale (Sequential) Oat Extraction

Over the course of 9 days in January 2018 and May 2018, two lots of 200 kg of oat flour were ground in a blender to a fine flour, and extracted with 50% (v/v) ethanol as follows. First, 8 kg lots of oat flour were soaked in 12 L of 50% (v/v) ethanol overnight with occasional mixing at room temperature. Tap water was used for all solutions except where noted. In the morning, the oat flour suspension was stirred and decanted into successive 6 × 500 mL buckets and centrifuged at 800 × g for 5 min at 20°C in a Sigma 6-16S centrifuge to give a firm pellet. The clear supernatants were pooled in a 30 L bottle and chilled at 4°C for 1–2 days to selectively precipitate the avenins. The bulk of the avenin settled after 2 days storage at 4°C and was removed from the bottom of the storage container by decanting the supernatant. The avenin precipitate which remained in the supernatant was collected by centrifugation at 5,000x g for 10 min at 4°C and formed a clear honey-like pellet which, with the bulk of the avenin above, was resuspended in a minimum volume of 10% (v/v) ethanol, made with reagent grade 18 MΩ water, and stored at 4°C. Clumps of precipitated avenin were dispersed with an overhead blender, frozen, and freeze-dried in a dedicated gluten-free facility to yield a white powder which was stored dry at −20°C until required. A final yield of 0.9 and 1.2 kg of freeze dried avenin was recovered from each 200 kg oat flour (**Figures S2A–C**).

Reversed Phase High Pressure Liquid Chromatography (RP-HPLC) Analysis

Purified, freeze dried avenin (10 mg) was re-solubilized using 70% (v/v) ethanol and vortexed for 30 min (MO BIO Laboratories, Inc. Vortex-Genie® 2). Samples were prepared in triplicate and were centrifuged for 20 min at 15870x g in an Eppendorf Centrifuge 5424. The supernatant was filtered using a 0.45 μm filter. The protein extracts were separated using an Agilent 1200 LC system (Agilent Technologies) using a modified method (32). An aliquot (10 μl) of extract was injected into a C18 reversed-phase Zorbax 300SB-C18 column (4.6 × 150 mm, 5 μm, 300 Å, Agilent Technologies) maintained at 60°C. The eluents used were ultrapure water (solvent A) and acetonitrile (solvent B), each containing 0.1% trifluoroacetic acid (TFA) (HPLC grade, Sigma Aldrich). The separation was carried out using a linear gradient from 33 to 80% solvent B over 65 min at a flow rate of 1 mL/min.

Protein Profiling Using Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF-MS) Analysis

Mass spectra representing the protein composition of the prolamin-enriched fraction were obtained from both flour and the purified avenin, and the RP-HPLC fractions collected from the purified avenin sample using MALDI-TOF-MS. Briefly, 60 mg flour sample or 10 mg purified avenin was extracted in triplicate using 300 μL of 70% (v/v) ethanol with vortex mixing for 30 min at room temperature. The protein extract (200 μL) as well as the equivalent amount of eluent from each collected RP-HPLC peak were lyophilized and resuspended in matrix consisting of 40 mg of sinapinic acid (SA), 600 μL of acetonitrile, 360 μL of methanol and 80 μL of water. An aliquot (1 μL) of this matrix was spotted on a 100 spot MALDI-TOF plate and an additional 1 μL sample layer was applied. An Applied Biosystems Voyager DE Pro MALDI-TOF mass spectrometer was operated in linear high mass positive mode using 2050 V laser intensity, an

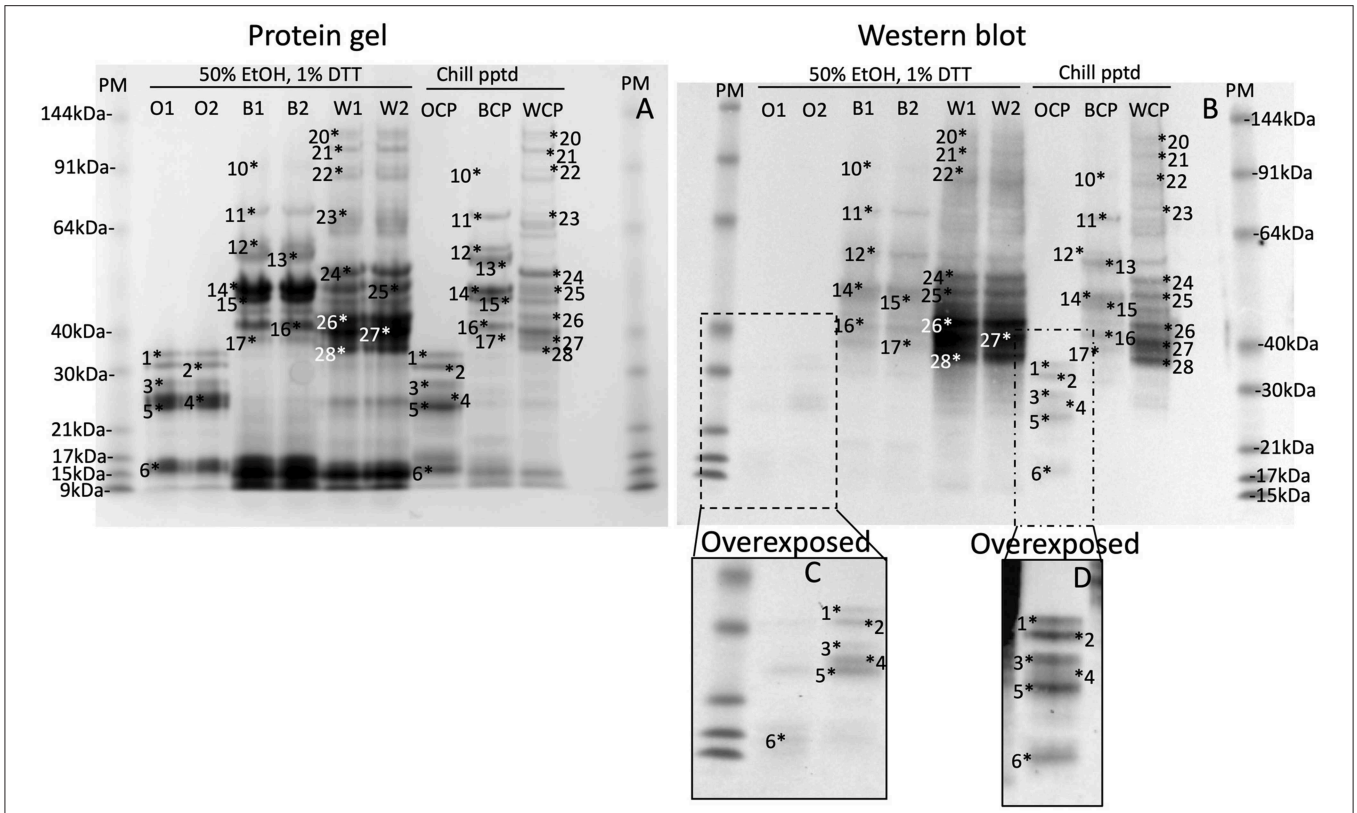


FIGURE 2 | Chill induced precipitation is a general method for all gluten proteins. Gluten proteins were isolated from wheat, barley, and oats by chill precipitation (WCP, BCP, OCP respectively, *Chill pptd*) and compared to gluten proteins freshly isolated by extraction of wheat, barley and oats in 50% EtOH, 1% DTT (W1, W2, B1, B2, O1, and O2, respectively, 50% EtOH, 1% DTT) by SDS-PAGE (**A**; 20 μg per lane), or western blot (**B**; 2 μg protein per lane). Corresponding protein bands, calibrated against pre-stained standards are numbered 1–28 in both images. The inset **(C,D)** shows an overdeveloped section of the western blot to highlight faint bands.

acceleration voltage of 25 kV, grid at 93% and guide wire at 0.2 settings, with 700 ns delay time. A total of 1,000 laser shots were averaged per spectra with three technical replicates analyzed. The detection mass range was set between *m/z* 10,000–60,000.

LC-MS/MS Analysis of the Collected RP HPLC Peaks

Protein samples were digested by chymotrypsin and peptides were extracted according to standard techniques (33). The peptide samples were analyzed by LC-MS on an Agilent 1260 Infinity HPLC system coupled to an Agilent 1260 Chipcube Nanospray interface on an Agilent 6540 mass spectrometer. Peptides were loaded onto a ProtID-Chip-150 C18 column (Agilent) and separated with a linear gradient of solvent A (2% acetonitrile/97.9% water/0.1% formic acid v/v/v) and solvent B (98% acetonitrile/1.9% water/0.1% formic acid v/v/v) from 2 to 98% solvent B over 18 min. Spectra were analyzed to identify proteins of interest using *in silico* proteolytic digests of Poales subset of UniProt-KB database (accessed 23/05/2019) supplemented with an oat seed transcriptome dataset (34) and appended with the common repository of adventitious proteins (cRAP). The number of protein sequences in the database was 861,955.

Protein Identification From the Enriched Avenin and Oat Flour

Proteins were extracted in four replicates using the finely ground oat cv Wandering flour and solvents suitable to obtain gluten protein enriched extracts: A −50% (v/v) ethanol; B −55% (v/v) IPA + 2% (w/v) DTT in water following the protocols of Colgrave et al. (31). The purified avenin sample was resolubilized in 50% (v/v) ethanol. From these, 100 μL aliquots of extract were applied to a 10 kDa molecular weight cut-off filter (Millipore, Sydney, Australia), alkylated using iodoacetamide and digested overnight using sequencing-grade trypsin (Promega, Madison, USA) or chymotrypsin (Promega, Madison, USA) as described in Fallahbaghery et al. (35). After digestion, samples were centrifuged at 20,000x g for 15 min and the digested filtrates were dried in a Speedvac (Thermo Fisher Scientific) and the obtained peptides were stored at −20°C until analysis.

Prior to LC-MS analysis the peptides were resuspended in 100 μL of 1% (v/v) formic acid. An aliquot (4 μL) of the peptide digest solutions were separated on an Eksigent NanoLC 415 system (SCIEX, Redwood City, USA) using a trap-elute configuration and 10 μL/min flow rate for trapping and 5 μL/min flow rate for separation. A Protecol C18 120Å trapping column (3 μm particle size, 10 mm × 300 μm, Trajan Scientific, Australia) and an

Eksigent ChromXP C18 120Å analytical column (3 μm particle size, 150 mm × 300 μm) was used with a linear gradient from 3 to 25% solvent B over 68 min. Mobile phases consisted of solvent A [0.1% (v/v) formic acid/5% (v/v) dimethyl sulfoxide/94.9% (v/v) water] and solvent B [0.1% (v/v) formic acid/5% (v/v) dimethyl sulfoxide/84.9% (v/v) acetonitrile/10% (v/v) water], and 0.1% (v/v) formic acid/99.9% (v/v) water was used to load the trap column. The eluate was directed into a TripleTOF 6600 MS (SCIEX), operating in information dependent acquisition mode over mass range *m/z* 100–2000. The Paragon algorithm of ProteinPilot 5.0.2 Software (SCIEX) was used for protein identification (36). The tandem mass spectrometry data were searched against the *in silico* proteolytic digests of Poales subset of UniProt database (accessed 23/05/2019) supplemented with the oat seed transcriptome data (34) and appended with the common Repository of Adventitious Proteins (cRAP) database. The database search results were examined and identifications were confirmed if they passed a 1% global false discovery rate (FDR) threshold as determined by the built-in FDR tool within ProteinPilot software (37). The resulting protein dataset was analyzed to identify avenins using the representative prolamin characteristics as described (38). Pfam domains, cysteine residues and major known T cell epitopes (DQ2.5-ave-1a, 1b, 1c, and 2) were mapped to the identified protein sequences using CLC Genomics Workbench v12 (Qiagen, Aarhus Denmark). Protein sequences were aligned using ClustalW algorithm and phylogenetic analysis was performed as reported elsewhere (39).

RESULTS

Effect of Solvent Concentration on Avenin Precipitation

Many gluten proteins can be dissolved in 50% (v/v) ethanol or 2-propanol and precipitated by dilution with either water or alcohol. The polarity of the 50% (v/v) ethanol avenin extract was varied by diluting the 50% ethanol extract with water to achieve final concentrations of 10–41% ethanol (v/v), or with ethanol to achieve final concentrations of 66–90% ethanol (v/v) (**Figure 1A**). All additions (water or ethanol) to the 50% ethanol extract produced a milky white precipitate; however only those precipitates produced by increasing the ethanol addition could be conveniently precipitated. The cloudy precipitate produced by adding water was extremely difficult to spin down and resisted precipitation at 5,000 g, beyond the capacity of the centrifuge used for the large scale avenin preparation. Fortuitously, chilling the 50% ethanol extract at 4°C for 10 min was noted to selectively precipitate avenin, producing a milky white precipitate that could be precipitated at either 500 g or 3000 g over 10 min (**Figure S1B**) or that settled at 1 g overnight. The chill-induced precipitation of avenin commenced below 15°C and could be reversed by warming, resulting in a clear solution that could be reproduced at least 10 times. The purity of the precipitates produced by varying the ethanol concentration, or chilling, was investigated by comparing proteins (**Figure 1B**) corresponding with avenin bands on a western blot (**Figure 1C**). All of the major bands present in the protein gel (**Figure 1B**, 1–6) correspond to avenins identified in the western blot (**Figure 1C**, 1-6). Some non-avenin proteins were present in the protein gel (**Figure 1B**, *). In addition the intensity of both protein and western bands was a maximum in the chill precipitation from 50% ethanol (**Figures 1B,C**). This indicated that chill precipitation produced an avenin precipitate of high purity showing higher specificity for avenins.

Chilling the water-induced precipitates at 4°C did not help the suspensions to precipitate, possibly due to protein-lipid binding causing the avenins to float. Defatting the oat flour with common defatting solvents (i.e., butanol, ether, or hexane) was not possible due to food safety concerns. For instance, n-hexane biodegrades to form a neurotoxin that must be avoided to maintain food grade standard. Successful and complete defatting of oat flour with 100% ethanol has been reported (40). However, in this study ethanol defatting did not allow protein precipitation at lower centrifugal force, suggesting that the difficulty in pelleting the water-induced precipitates was not due to lipid interaction.

Chill Precipitation Is a General Method Applicable to a Range of Gluten Proteins

Chill-induced precipitation was assessed for its utility as a general method for isolating gluten proteins. Gluten proteins were isolated from oats, barley and wheat in 50% (v/v) ethanol extracts, followed by chill precipitation (**Figures 2A,B**, *Chill ppted,* Oats Chill Precipitated (*OCP)*, Barley Chill Precipitated (*BCP)*, Wheat Chill Precipitated (*WCP),* respectively), and compared to gluten proteins isolated by duplicate extraction of oats, barley and wheat in 50% (v/v) ethanol, 1% (w/v) DTT (**Figures 2A,B**, *50% ethanol/DTT*, *O, B, W,* respectively). Previous studies have demonstrated that IPA/DTT is able to efficiently extract the majority of gluten proteins (35). In each case gluten proteins, numbered 1–28 in both images, and by definition extracted by 50% (v/v) ethanol, 1% (w/v) DTT were present in the corresponding 50% (v/v) ethanol chill-precipitated fractions. This was observed for Coomassie stained protein bands (**Figure 2A**) and gluten-specific proteins by western blot (**Figure 2B**, highlighted in **Figures 2C,D**) confirming that chill precipitation isolated the same range of gluten proteins that were isolated from wheat, barley, and oats as extraction in 50% (v/v) ethanol/1% (w/v) DTT, thus confirming the general nature of the chill precipitation method.

Avenin Yield and Purity

The maximum protein yield was produced by extracting 500 g oat flour in 750 mL of 50% (v/v) ethanol over 1–2 days (**Figure S1A**). The protein purity of the large scale avenin preparation was firstly examined by resolving protein bands with detection using a general protein stain (Coomassie Blue G250; **Figure 3A**). The protein bands were then compared to avenin bands identified by a general anti-gluten antibody on a Western blot (**Figure 3B**) for large scale avenin preparations 1 and 2 (*Prep1* and *Prep2)*. Using calibrated pre-stained standards on the blot and protein gel, avenin bands on both the protein gel and western blot were shown to be of the same molecular weight (**Table 1**). It is clear that the protein bands (**Figure 3A**) were due to the dominant avenin bands on the western blot (**Figure 3B**), except for one

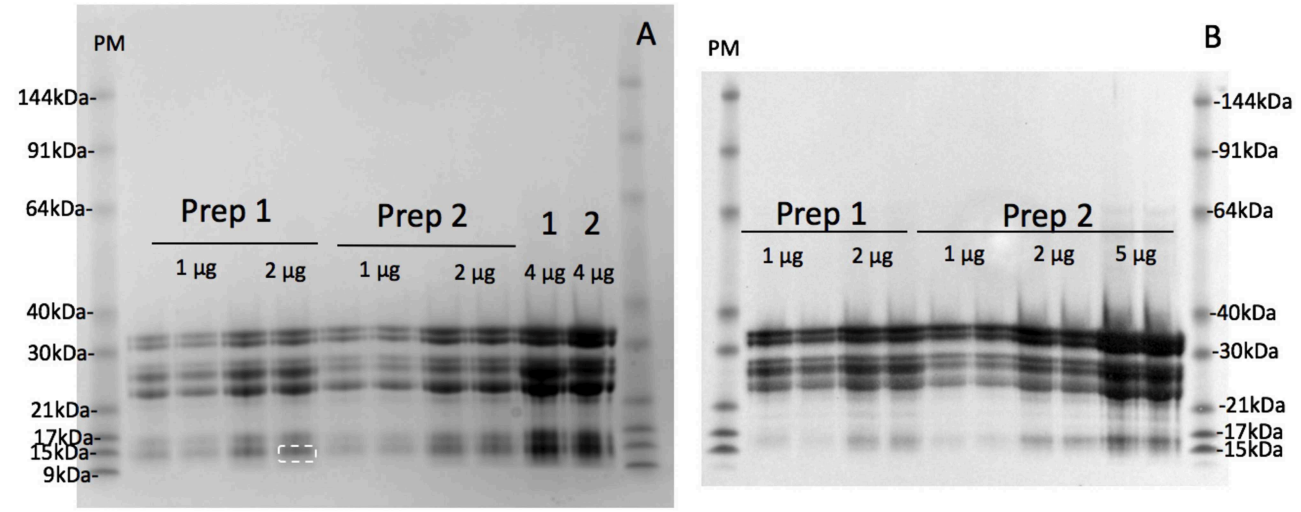


FIGURE 3 | Purity of avenin isolation from 200 kg avenin preparations by SDS-PAGE protein gel **(A)** or western blot **(B)**. Two hundred kilograms of oat flour was successively purified to yield two lots of purified avenin (*Prep 1*) and (*Prep 2*) are shown calibrated against pre-stained 10 kDa ladder (PM). Avenin was dissolved in 8 M urea, 1% DTT, 20 mM triethylamine (pH6), protein measured and 1, 2, 4, or 5 ug loaded per lane as shown. A single non-avenin protein band was present at 14.1 kDa (**Figure 2A**, *white broken rectangle*) corresponding to 4.2% of total band intensity. This does not correspond to an avenin band in the western blot 2B, and was probably due to α-amylase/trypsin inhibitors. The protein composition of the initial 50% (v/v) ethanol extract of preparation 1 and 2 are shown (**A**, *1* & *2*).

TABLE 1 | Comparison of avenin molecular weights calculated from protein gel and western blots.

Avenin band	Molecular weight (blot) kDa ± SE	Molecular weight (gel) kDa ± SE
1	32.6 ± 0.02	33.0 ± 0.02
2	31.5 ± 0.02	31.5 ± 0.03
3	28.1 ± 0.07	28.6 ± 0.02
4	27.3 ± 0.08	27.3 ± 0.02
5	25.3 ± 0.09	25.3 ± 0.04

*Molecular weights were compared from the protein gel and western blot proteins bands in **Figures 3A,B** and calibrated against prestained standards.*

protein band at ∼14 kDa, which did not correspond to an avenin band in the western blot (**Figure 3A**, marked by a broken white rectangle). Note that the 9 kDa prestained marker did not bind to the membrane and did not appear in the Western blot (**Figure 3B**, PM). The protein purity was calculated from the percentage of the protein load attributed to avenin bands in the protein gel. The average purity of both preparations shown for the four 2 µg lanes on the protein gel was 95.8 ± 0.01%.

Proximate Composition Analysis

The protein content was assessed by the Kjeldahl method revealing 85.4% protein. The remainder of the preparation consisted of low concentrations of starch, β-glucan, free sugars and water soluble carbohydrates. The starch content was 1.8% on a dry weight (DW) basis. The β-glucan content was 0.2% DW, total free sugars were 1.8% DW and the total water-soluble carbohydrate (WSC) content was 4.6% DW. As starch is insoluble, it does not contribute to the WSC content. The total free sugars (at 1.8% DW) include fructose, glucose, lactose, maltose and sucrose, and are soluble and are thus included in the WSC measurement, leaving 2.8% DW attributed to other sugars, such as oligosaccharides, small fructans, and other complex sugars.

Food grade purity was confirmed by nil detection of herbicides and/or pesticides (**Supplementary Table S1**). No inadvertent chemical contamination was detected in either of the oat crops. Both oat crops were tested for the presence of aflatoxins (**Supplementary Table S1**), and all were below the limit of detection (LOD listed in **Supplementary Table S1**). The purified avenin was also tested for heavy metal contamination. Mercury, chromium and lead were all below the limit of detection. Copper was reported to be 5.1 and aluminum 58 mg/kg. These are within acceptable safety limits. The US Food and Drug Administration (FDA) reports that 10–100 mg aluminum per day is acceptable (41). The proposed feeding trial will involve a daily intake of 0.35 mg aluminum in a 6 g serve of avenin. The FDA does not define a limit on copper safety but these levels and much higher are common in other foods. The Food Standard Australia New Zealand code does not specify levels for copper or aluminum.

Comparative Proteomic Analysis of the Enriched Avenin and Oat Flour

Previous studies have shown that gluten proteins can be efficiently extracted from gluten using alcohol in the presence of reducing agents. Such solvents are not applicable to production of food grade extracts, but are suitable for biochemical analyses that aim to identify the protein complement in grains, flours, and food products.

Extraction using IPA/DTT showed the highest number of proteins in the oat flour sample after LC-MS/MS analysis of the trypsin digested sample, with 276 proteins identified at the

TABLE 2 | Number of detected protein types at 95 and 99% confidence levels.

Protein type	Protein confidence	Flour				Purified sample	
Solvent		IPA/DTT[a]		50% (v/v) ethanol			
Enzyme		TR[b]	CTR[c]	TR	CTR	TR	CTR
Avenin	99%	18	20	17	19	18	18
	95%	20	25	18	20	21	20
ATI	99%	18	17	15	16	19	16
	95%	24	17	19	16	29	17
Vromindoline and GSP	99%	14	8	10	5	11	6
	95%	15	9	11	5	13	7
LTP and nsLTP	99%	8	4	7	4	6	1
	95%	11	5	9	4	8	3
Globulin-like proteins	99%	32	24	12	8	4	1
	95%	38	26	13	12	6	2
Enzymes and metabolic proteins	99%	186	33	97	25	59	19
	95%	288	54	152	48	103	28
Total number of proteins		397	136	223	105	180	77

[a]*55% IPA + 2% DTT extraction buffer;* [b]*trypsin;* [c]*chymotrypsin.*

99% protein confidence level and additional 121 proteins were identified when the confidence level was lowered to 95%. It should be noted that LC-MS/MS is a very sensitive technique capable of detecting proteins that vary in abundance over four orders of magnitude. Detection of a number of non-avenin proteins does not imply reduced purity, but that the technique employed was capable of identifying a range of co-extracted proteins that were not obvious on the gel images (**Figure 3**). This protein set included 20 avenin sequences, 18 of which were identified with 99% confidence (**Table 2**). Additionally, there were 25 ATIs, 11 lipid transfer proteins (LTPs) and 15 vromindolines and grain softness proteins identified. A further 38 proteins were identified with the cupin-1 domain that is characteristic of globulins and germins expressed in seeds. The remaining 288 proteins represented a range of enzymes and metabolic proteins. The chymotrypsin digested IPA/DTT extract revealed 136 proteins, of which 106 were detected with 99% protein confidence level. Within this set, 25 avenins (20 with 99% confidence level), 17 ATIs, 5 LTPs, 9 vromindolines and 26 globulins and germins were identified (**Table 2**). Comparing the trypsin and chymotrypsin digests yielded 10 avenins, 14 ATIs, 7 vromindolines and 4 LTPs that were commonly identified when IPA/DTT extraction was used during the sample preparation.

To obtain comprehensive protein information of the avenin-enriched protein fractions, 50% ethanol extracts of the oat flour and purified avenin were digested using trypsin and chymotrypsin separately, and analyzed. Within the tryptic digests (**Table 2**), 223 proteins were identified in the flour sample and 180 proteins were identified in the resolubilized avenin sample. Of these, 96 proteins were commonly detected between the flour and the purified avenin extracts. Of the detected proteins in the flour sample 8.1% were avenins, while other prolamin superfamily member proteins were also detected, including ATIs (9%), kernel structure-related vromindoline and grain-softness proteins (4.9%) and LTPs (4.1%). The remaining 164 proteins in the flour extract represented seed storage globulins and proteins with various enzyme and metabolic functions. In the large-scale purified avenin sample, 12% of detected proteins represented avenins, showing a qualitative enrichment in avenin proteins. Of the remaining proteins, 16.6% were ATIs, 7.4% were vromindolines and grain softness proteins and 4.5% were LTPs. The remaining 104 proteins showed a similar composition to that of the flour extract. The chymotryptic peptides yielded 105 protein identifications in the flour sample and 77 proteins in the purified avenin sample with 42 proteins commonly detected between the flour and the purified avenin sample. Identified avenin sequences represented 19% of the detected proteins in the flour sample and 27% in the purified protein sample. Similar to the trypsin digested extracts, ATIs, vromindolines and LTPs were also detected, though at lower numbers than in trypsin digested protein samples (**Table 2**).

Utilizing chymotrypsin for proteolytic digestion and subsequent LC-MS analysis resulted in an increased number of unique avenin protein identifications independent of the extraction solution used in the analysis. In total, 18% of the identified proteins in the IPA/DTT flour extract were avenins and a similar number (19%) were identified from the 50% ethanol extract of flour. Using trypsin the abundance of avenins within the identified protein pool was significantly lower: 5% for the IPA-DTT and 8% for the 50% ethanol extract (**Table 2**). Of the trypsin digested avenin proteins 50% were also identified from the chymotrypsin digests. Comparing the number of theoretically detectable tryptic and chymotryptic peptides, a higher number of detectable chymotryptic peptides can be seen in most of the sequences (**Supplementary Table S2**, **Figure S3**). This resulted in the detection of an additional 15 avenin proteins in the IPA/DTT and a further 11 avenin sequences in the 50% ethanol extracts when chymotrypsin was used.

Both the oat flour and the purified avenin samples were screened for the presence of wheat and barley contamination using published methods (31, 42) and no evidence was found.

Protein Molecular Weight and Hydrophobicity Distribution Within the Avenin-Enriched Extract

Altogether 25 protein peaks were identified in the undigested purified extract using MALDI-TOF MS with centroid mass values ranging between 19.3 and 32.4 kDa (**Figure 4**). The proteins were grouped into three distinct groups: three proteins represented proteins below 22.4 kDa, eight proteins with centroid mass values between 22.4 and 24.4 kDa and 13 proteins with a centroid mass range of 25.8 and 32.4 kDa.

The purified protein sample was also separated based on hydrophobicity using RP-HPLC (**Figure 5**). The obtained chromatogram peaks were grouped into two retention time ranges. The first group represented peaks between retention time values of 27.0 and 31.3 min and accounted for ∼57% of the total protein. In this subgroup peaks between 29.6

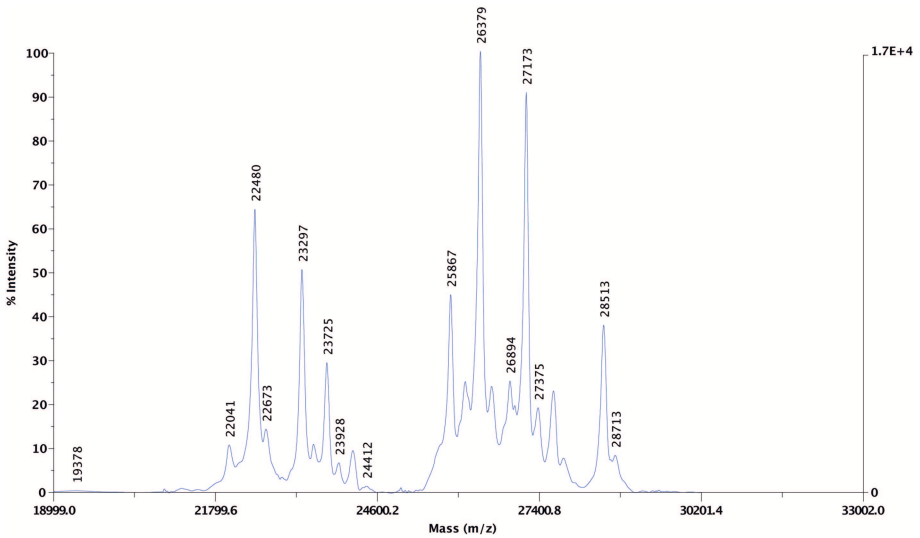


FIGURE 4 | Representative MALDI-TOF MS spectra of the undigested purified protein dissolved in 70% ethanol. The left y-axis represents signal intensity in arbitrary units, the right y-axis shows the maximum intensity value. The x-axis represents the mass-to-charge ratios (*m/z*). The *m/z* values detected from the 70% ethanol resolubilized purified protein extract are highlighted above the peaks.

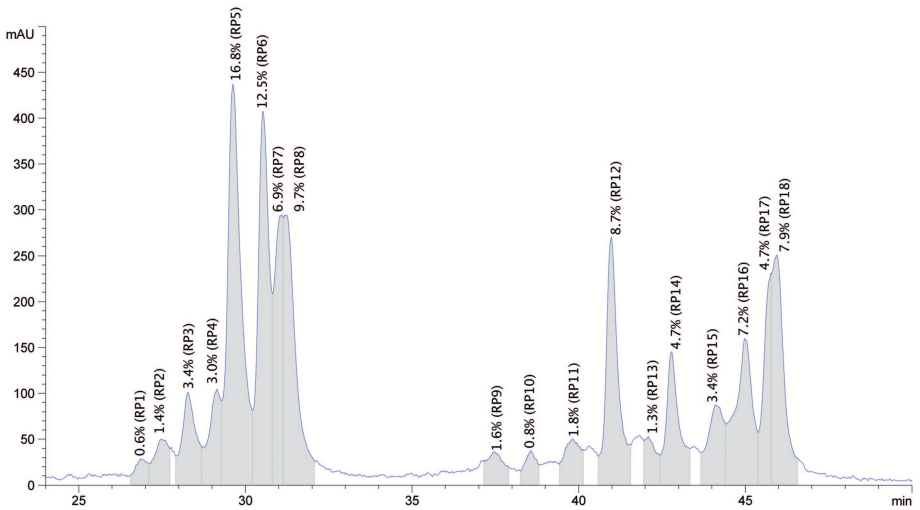


FIGURE 5 | Representative reversed phase HPLC chromatogram of the purified oat protein sample resolubilized in 70% ethanol. The x-axis represents the retention time in min, y-axis shows the relative amount of protein fractions in mAU units. Collected RP peaks and the % peak areas are labeled from RP1 to RP18 and highlighted in gray.

and 31.3 min represented the most abundant peaks. The second subgroup that accounted for 43% of the total protein content was composed of 13 peaks and retention time values ranged between 37.6 and 46.0 min. Peaks with retention time values of 41.1, 46.0, and 45.1 min were the most abundant. Based on the obtained chromatogram there were 18 peaks collected and the undigested samples were subjected to protein profiling using MALDI-TOF MS. In parallel to this analysis proteins present in the chymotrypsin-digested RP-HPLC peaks were identified using LC-MS/MS after chymotrypsin digestion.

The RP-HPLC peaks were also quantified by calculating the mean Area (mAU * s) values of the three replicates × three injections. The absolute amount of protein within each peak was estimated based on the obtained protein content using the Kjeldahl method as described in section Proximate Composition Analysis. The quantitative results, including retention time ranges, calculated mean values as absolute amounts of protein/100 g purified protein are provided in **Table 3**.

The RP-HPLC peaks RP3-RP7 comprised proteins with the larger molecular mass range (above 26 kDa). Peaks RP1 and

TABLE 3 | Identified proteins and their calculated amounts of the collected 18 RP-HPLC peaks.

RP peak #	Retention time peak (min)	Collection time range (min)	Mean g/100 g protein	Range g/100 g protein
1	26.85	26.6 – 26.9	0.57	0.37–0.72
2	27.38	27.1 – 27.8	1.90	1.65–2.05
3	28.18	28.0 – 28.6	3.29	3.18–3.42
4	29.04	28.7 – 29.1	2.72	2.39–3.04
5	29.54	29.2 – 30.0	17.05	16.64–17.55
6	30.47	30.1 – 30.6	13.13	12.52–13.55
7	30.93	30.7 – 31.0	6.57	6.20–6.89
8	31.15	31.0 – 31.7	10.69	10.10–11.20
9	37.42	37.2 – 37.7	0.99	0.73–1.59
10	38.48	38.2 – 38.6	0.63	0.33–0.85
11	39.72	39.3 – 40.4	1.58	1.40–1.83
12	40.95	40.4 – 41.6	9.29	8.65–9.74
13	41.69	41.9 – 42.3	2.54	1.10–1.28
14	42.77	42.3 – 43.2	5.16	4.50–4.86
15	44.16	43.8 – 44.4	2.99	2.66–3.38
16	44.50	44.4 – 45.3	7.05	6.65–7.53
17	45.94	45.4 – 45.9	4.74	4.56–4.96
18	46.20	46.0 – 46.4	9.11	8.67–9.56

RP8 were comprised of proteins spanning the entire mass range while peaks RP9-RP12 mainly included proteins with smaller molecular mass values (below 22.5 kDa). Peaks RP15-RP18 comprised proteins between 22.5 and 24 kDa and above 26 kDa. The number of representative avenin protein types per peak is shown in **Table 4**. Proteins present in the 18 peaks consisted of avenins and gliadin-like avenins, however peptides characteristic of ATIs were also confidently detected in peak RP2 (**Table 4**). RP-HPLC analysis indicated that the proteins that eluted between 26.9 and 31.2 min exclusively contained proteins with the DQ2.5-ave-1a epitope and this fraction represented approximately 55.9 g in 100 g purified protein. While DQ2.5 ave-1b epitope containing proteins were detected in the retention time range of 38.5–45.9 min, proteins with DQ2.5-ave-2 epitope were only detected in the RP peak 11.

Characterization of the Identified Avenin Types

As the purpose of the avenin feeding trials is to provoke a T cell response to assess the clinical significance of oats in CD, it was important for us to confirm the presence of the previously reported and potentially pathogenic avenin peptides that encompass T cell epitopes (13). Avenin protein sequences that were detected in any of the analyzed samples were used for phylogenetic analysis and epitope mapping. The identity of the known T cell epitopes and the protein identification analyses of oat flour and purified avenin samples along with RP peak analysis were used to annotate the phylogenetic tree and compare the characteristics of the sequence groups. The sequences were grouped into major clusters according to the presence of conserved avenin-specific T cell epitopes (**Figure 6**, **Supplementary Table S2**). Group 1 included proteins with the epitope DQ2.5-ave-1b, the majority of the group 2 sequences contained DQ2.5-ave-1c, group 3, the largest group, was typified by the presence of DQ2.5-ave-1a and group 4 was typified by the presence of DQ2.5-ave-2. The sub-branch without T cell epitopes (group 5) represents protein sequences mostly similar to high and low molecular weight glutenin-like sequences. The results clearly indicate that the 50% ethanol extraction method is suitable to extract the major epitope containing avenin types. The resultant purified protein contains all of the reported immunogenic avenin peptides, and therefore should be suitable as a way to assess their clinical toxicity in feeding studies.

DISCUSSION

Determining if oats should be excluded or included in the GFD is an important issue with medical and societal implications. Dedicated, controlled feeding studies are needed to definitively resolve this issue. Here, we demonstrate an approach that enables the production of a highly purified oat avenin preparation, containing all the known immunogenic avenin T cell epitopes, which contains minimal levels of other carbohydrates or proteins which could otherwise confound the interpretation of the feeding trial. This avenin preparation will facilitate immune and feeding studies to test the suitability of oats in CD. Studies can now be undertaken to assess the immunogenicity of this protein using patient-derived gluten-specific T cells, and employ feeding studies in patients with CD to determine its biological effects.

Despite the many nutritional benefits of oats, contradictory clinical feeding studies, and lack of a clear scientific rationale for dietary guidelines have led to differing views on the safety of oats consumption in CD. A systematic review on studies of oat safety in CD (4) identified several limitations including (i) small sample sizes: uncommon patients with oats sensitivity may not have been included; (ii) assessment of GFD adherence: often not reported; (iii) adverse symptoms: cannot determine if they are related to oats avenin or the fiber load of oats itself; and (vi) oat cultivars: usually not reported. Furthermore, study withdrawals which were often due to adverse gastrointestinal symptoms and/or the inability to maintain the oats diet, may have underestimated the adverse impact of oats intake, and recruitment bias, may lead to oats-sensitive patients avoiding participating in oats feeding studies. The authors concluded: "Our confidence is limited by the low quality and limited geographic distribution of the data" and "Rigorous double-blind, placebo-controlled, randomized controlled trials, using commonly available oats sourced from different regions, are needed." (4).

An important consideration is that testing for cereal contamination over the last two decades has highlighted a high frequency of commercial oats brands contaminated by wheat, barley and/or rye prolamin in amounts toxic to CD patients (43–45). Contamination can occur since wheat, barley, and oats are grown in the same areas and often harvested and transported with the same machinery. The presence of a single grain of wheat in 200 g of oats can result in the wheat gluten level of >100 ppm; well above the 20 ppm level set in most legislations as the upper

TABLE 4 | Protein characterization of the identified RP-HPLC peaks.

Fraction #	No. MS peaks in HPLC fraction*	Observed mass range (m/z)	Most intense peaks in MALDI profile	Accession (Uniprot/ transcript ID)*	Confidence level	Monoisotopic mass (signal peptide removed)	Epitope
1	15	20543–28907	28604	L0L6J0	99%	30789	DQ2.5-ave-1a
			27783	Asat-prolamin54	99%	23496	DQ2.5-ave-1a
2	11	21321–28711	27760	L0L6J0	99%	30789	DQ2.5-ave-1a
			21321	L0L6K1	95%	22031	DQ2.5-ave-1a
			28481				
3	9	21327–28652	28467	L0L6J0	99%	30789	DQ2.5-ave-1a
4	6	22038–28675	28498	Asat-prolamin2	99%	24703	DQ2.5-ave-1a
				Q09072	95%	23539	DQ2.5-ave-1a
5	5	22049–28704	28486	Asat-Prolamin10	99%	25189	DQ2.5-ave-1a
				L0L6K1	95%	22031	
6	4	27609–28659	28478	Asat-Prolamin54	99%	23496	DQ2.5-ave-1a
				L0L6K1	99%	22031	DQ2.5-ave-1a
				Asat-Prolamin10	99%	25189	DQ2.5-ave-1a
7	4	22009–28481	27595	L0L6J0	99%	30789	DQ2.5-ave-1a
				L0L6K1	99%	22031	DQ2.5-ave-1a
8	10	21240–28516	24184	Asat-Prolamin10	99%	25189	DQ2.5-ave-1a
			22040	Q09072	99%	23539	DQ2.5-ave-1a
			27652				
9	2	22513–22711	22513	L0L4J7	99%	22481	DQ2.5-ave-1c
10	4	21893–25833	22465	L0L4J7	99%	22481	DQ2.5-ave-1c
				L0L6J0	99%	30789	DQ2.5-ave-1a
11	2	21792–22495	22495	L0L4J7	99%	22481	DQ2.5-ave-1c
				Q09072	99%	23539	DQ2.5-ave-1a
				I4EP88	99%	25862	DQ2.5-ave-1b
				I4EP86	95%	21915	DQ2.5-ave-2
				L0L6J0	95%	30789	DQ2.5-ave-1a
12	3	22495–22703	22495	L0L4J7	99%	22481	DQ2.5-ave-1c
			22598	I4EP58	95%	27966	DQ2.5-ave-1b
13	5	22481–25855	22481	L0L4J7	99%	22481	DQ2.5-ave-1c
				Asat-Prolamin71	99%	22752	DQ2.5-ave-1b
				L0L6J0	95%	30789	DQ2.5-ave-1a
14	5	22474–26051	25850	I4EP58	99%	27966	DQ2.5-ave-1b
15	11	22642–27149					
			25853	Asat-Prolamin71	99%	22752	DQ2.5-ave-1b
			23314	G8ZCU7	99%	21628	DQ2.5-ave-1c
				Q09097	99%	Fragment	
16	7	23317–27187	23735				
				I4EP54	99%	26895	DQ2.5-ave-1b
				L0L6J0	99%	30789	DQ2.5-ave-1a
				Asat-Prolamin15	95%	20556	DQ2.5-ave-1c
17	10	23317–27328	23729				
				I4EP58	99%	27966	DQ2.5-ave-1b
				Asat-Prolamin15	99%	20556	DQ2.5-ave-1c
18	8	23294–27363		I4EP57	99%	27948	DQ2.5-ave-1b
			23294				
			26370				

**Peaks above 10% intensity in MALDI-TOF MS were considered.*

limit for gluten-free food status. Failure to provide harvesting, transport and milling facilities dedicated to oats may easily result in significant inadvertent contamination by wheat grains.

Another challenge with oat feeding studies may arise from supplying insufficient avenin to provoke a measurable response in people with CD. With a short-term oral oat challenge of 100 g daily, 8% of patients with CD had pro-inflammatory T cell responses detectable in the bloodstream (13) in contrast to the 75–80% that would be seen after just four slices of wheat bread daily (46). To fully assess the clinical safety of oats we argue it will be important to deliver oats avenin at a "dose" sufficient to trigger both immune and biologic responses. While immune studies suggest oats avenin contain immunogenic sequences capable of stimulating gluten-specific T cells in CD, such a study may help

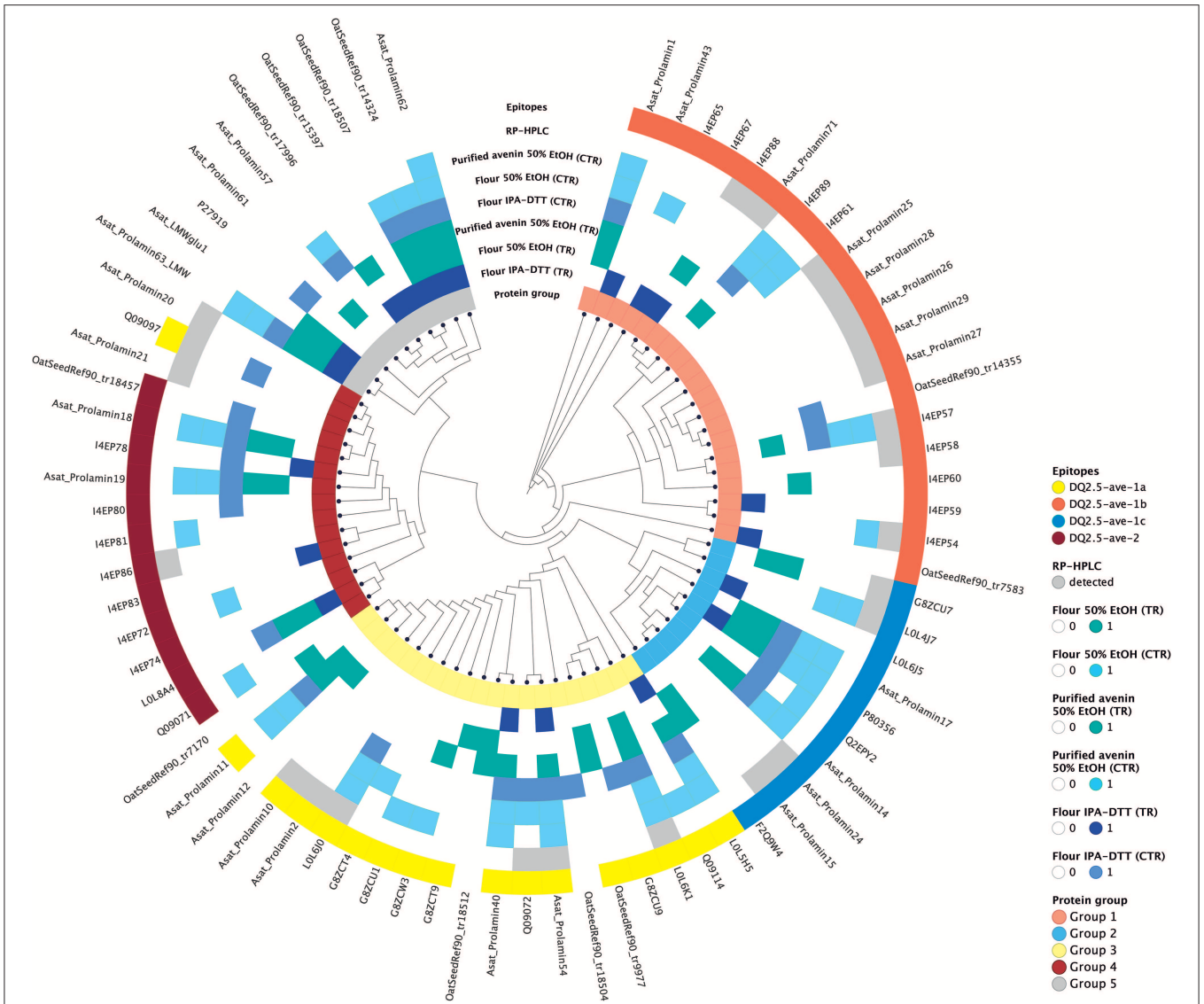


FIGURE 6 | Epitope mapping and phylogenetic analysis of avenin sequences identified from the different protein extracts. Avenin sequences identified from any of the analyses are shown in the tree and were grouped into five groups (labeled in the inner circle). The sequences that contain epitopes DQ2.5-ave-1a, DQ2.5-ave-1b, DQ2.5-ave-1c, and DQ2.5-ave-2 are highlighted yellow, red, blue, and dark red, respectively, in the outer ring. Sequences within the same phylogenetic groups without these epitopes are shown as empty blocks. Those proteins identified using the different protein extraction methods and enzymes are labeled with blue and teal colored blocks in intermediate rings. Proteins identified from the individual RP HPLC peaks are labeled in gray in the outer ring.

establish a safe threshold dose of oats for consumption in CD by correlating these findings with clinically important endpoints.

Here, we demonstrate that it is possible to produce 85% pure oat grain storage protein extracts from oat flour using repeated 8 kg scale avenin-enriched extractions. The IPA/DTT method combined with trypsin digestion yields a large number of gluten proteins from cereal grains such as wheat (31). This analysis in oats demonstrated that using 50% (v/v) ethanol was sufficient to identify a majority of the avenin protein types. The least represented group was that of proteins with the DQ2.5-ave-1b epitope. This protein type however, was extracted in higher number using the IPA/DTT solvent. Similarly, when reducing agent was added during the purified avenin resolubilization the protein group was better represented (data not shown). This indicates that these proteins with 9 cysteine residues in their sequence might be part of the protein polymer and therefore less soluble in the absence of reducing agent (**Supplementary Table S2**).

The final avenin preparation contains the avenin bands present in the initial crude 50% (v/v) ethanol extract. Thus, the majority of avenins are recovered by this process. Ninety six percent of the protein in the final protein preparation was due to authentic avenins which were identified by western blotting. The content of starch, β-glucan, sugars, and fatty acids were low

in the final avenin preparation. Using repeated extractions it is possible to process 200 kg of oat flour in 9 days to produce ∼1 kg of food-grade, avenin-enriched freeze-dried powder.

The RP-HPLC peak analysis indicated the presence of ATIs in the purified sample and this was verified with the comparative analysis using various extraction methods. The number of ATIs detected in the 50% (v/v) ethanol purified protein was in the same range as the avenins. Western blots show the presence of ATI is <4.2% however the precise amount requires further analysis. Other prolamin proteins were also detected. The number of LTPs and kernel structure-related vromindolines was in the same range as of the avenins. Furthermore, globulins, enzymes, and metabolic proteins were also present in large number. Comparison of the IPA/DTT extracted flour sample and the protein preparate clearly shows the depletion of other protein types in the protein preparate and the increase of detected avenin sequences. Using RP-HPLC analysis combined with the MALDI-TOF MS and LC-MS confirmed the presence of significant amount of avenins in the sample. RP-HPLC peaks 5, 6, 8 altogether represent about 40% of the measured protein amount. These peaks were primarily enriched in DQ2.5-ave-1a containing avenins.

Although CD is not a food allergy, controlling for potential confounding effects from ATIs or LTPs in the subsequent analysis of the avenin preparation will be important as these proteins have been associated with human disease including food allergies and Baker's asthma. ATIs have also been implicated in *in vitro* studies to contribute to intestinal inflammation via activation of innate immune pathways (47, 48). Utilizing readouts specific to CD and gluten such as gluten-specific T cells will ensure assessment is focussed on biological effects driven by the avenin and not other components.

The public databases are relatively poor in oat specific storage protein sequences, therefore wild relatives of *Avenae* were also included in the analysis. In total, 185 avenin or avenin-like gliadin sequences were included in the data background. Due to this poor avenin protein sequence representation a precise protein identification is rather challenging but our results clearly indicated that the purified protein is enriched in avenin and gliadin-like proteins. Importantly, using an expanded prolamin database that also includes prolamin sequences of other Poales species, including wheat and barley, we have excluded the possibility of wheat and barley contamination in the oat purified protein.

Comparison of the MALDI-TOF and MicroLC-TripleTOF MS analysis results also revealed that, similar to wheat gliadins, the avenins show a significant level of genetic variability, indicating the presence of multiple gene copies with largely similar protein size and hydrophobicity within the used cultivar. It also indicates the possible presence of avenin alleles with slightly different protein mass values in the different *Avena sativa* cultivars and *Avena* species. Some of the protein peaks were present in multiple adjacent RP-HPLC peaks which might be due to the resolution limitations of RP-HPLC. Mass differences <100 Da observed in adjacent RP peaks can be explained by the presence of post-translational modifications and highly similar sequences with amino acid substitutions or insertions/deletions. Although the overlapping protein set between the different extraction protocols and the RP HPLC peak analysis was small, the avenin sequence analysis confirmed that proteins from the same avenin sub-types were detected. The background database used included <200 avenin sequences representing a rather narrow genetic variability of only a few *Avena* species and *Avena sativa* cultivars. The avenin sequences within the analysis set share 70–99% sequence identity within the groups and 24–45% sequence identity between the groups. Only a few avenin proteins were identified with nearly complete peptide coverage. This also demonstrates that the avenin sequences present in the cultivar Wandering are different from those in the public databases.

CONCLUSION

Resolving the issue of oat safety in people with CD will require feeding studies to assess the clinical and immune effects of pure oat avenin. The purity of oat avenin can be confirmed by rigorous proteomic characterization to control for confounding factors such as contamination or potential cross reactivity. A requirement for purified oat avenin for feeding studies was identified as far back as the 1950's, but until recently, production of this at a scale and purity suitable for human feeding studies was not possible. With the method reported here, it is possible for the first time to generate highly pure avenin suitable for controlled feedings studies in CD. This will allow the biological effects of oats avenin to be assessed devoid of the confounding effects from other oat proteins, sugars or fiber, or non-oat sources of gluten. We believe this is a crucial advance that will allow this issue to be definitively assessed and resolved.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the manuscript/**Supplementary Files**.

AUTHOR CONTRIBUTIONS

JT-D conceived the study. The method was developed by GT and FB. The large-scale purification was completed by GT and FB. Protein analysis was completed by GT. RP-HPLC and MALDI-TOF analysis was carried out by CF and quantitatively evaluated by FB. Manuscript analysis and proteomic data analysis was carried out by AJ, MN-W, and MC. Bioinformatic analysis was completed by AJ. Proximate analysis was supervised by AR. MH contributed to study design and epitope analysis. All authors contributed to the writing of the manuscript.

FUNDING

JT-D was supported by the Mathison Centenary Fellowship (University of Melbourne). The Tye-Din Coeliac Research Lab are grateful for research support provided by Coeliac Australia and Coeliac UK.

ACKNOWLEDGMENTS

This work would not have been possible without the generous support of key individuals and businesses. Avenin was purified from wheat-free oats (cv Wandering), harvested in Sep 2016 and Dec 2017 by Kim and Danielle Rintoul in Williams, Western Australia. The oats were ground in a gluten-free, food-grade plant at Ward Mackenzie at Altona (Victoria), with the assistance of Sam Crispin. The avenin purification was carried out in food grade laboratories of the Manildra Group (Nowra, NSW) with the kind assistance of Amy Barrie and Mark Basinski. The final avenin preparation was freeze-dried in food-grade facilities at BioTech Freeze Drying, Knoxfield with the assistance of Patrick Dunne. James Tanner assisted with the second large scale avenin purification.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2019.00162/full#supplementary-material

REFERENCES

1. Singh P, Arora A, Strand TA, Leffler DA, Catassi C, Green PH, et al. Global prevalence of celiac disease: systematic review and meta-analysis. *Clin Gastroenterol Hepatol.* (2018) 16:823–6. doi: 10.1007/978-3-319-91316-2
2. Hardy MY, Tye-Din JA. Coeliac disease: a unique model for investigating broken tolerance in autoimmunity. *Clin Transl Immunol.* (2016) 5:e112. doi: 10.1038/cti.2016.58
3. Dicke WK, Weijers HA, Van De Kamer JH. Coeliac disease. II The presence in wheat of a factor having a deleterious effect in cases of coeliac disease. *Acta Paediatr.* (1953) 42:34–42. doi: 10.1111/j.1651-2227.1953.tb05563.x
4. Pinto-Sanchez MI, Causada-Calo N, Bercik P, Ford AC, Murray JA, et al. Safety of adding oats to a gluten-free diet for patients with celiac disease: systematic review and meta-analysis of clinical and observational studies. *Gastroenterology.* (2017) 153:395–409 e393. doi: 10.1053/j.gastro.2017.04.009
5. Sjoberg V, Hollen E, Pietz G, Magnusson KE, Falth-Magnusson K, Sundstrom M, et al. Noncontaminated dietary oats may hamper normalization of the intestinal immune status in childhood celiac disease. *Clin Transl Gastroenterol.* (2014) 5:e58. doi: 10.1038/ctg.2014.9
6. Peraaho M, Kaukinen K, Mustalahti K, Vuolteenaho N, Maki M, Laippala P, et al. Effect of an oats-containing gluten-free diet on symptoms and quality of life in coeliac disease. A Randomized Study Scandinavian. *J Gastroenterol.* (2004) 39:27–31. doi: 10.1080/00365520310007783
7. Tuire I, Marja-Leena L, Teea S, Katri H, Jukka P, Paivi S, et al. Persistent duodenal intraepithelial lymphocytosis despite a long-term strict gluten-free diet in celiac disease. *Am J Gastroenterol.* (2012) 107:1563–9. doi: 10.1038/ajg.2012.220
8. Lundin KE, Nilsen EM, Scott HG, Loberg EM, Gjoen A, Bratlie J, et al. Oats induced villous atrophy in coeliac disease. *Gut.* (2003) 52:1649–52. doi: 10.1136/gut.52.11.1649
9. Silano M, Di Benedetto R, Maialetti F, De Vincenzi A, Calcaterra R, Cornell HJ, et al. Avenins from different cultivars of oats elicit response by coeliac peripheral lymphocytes. *Scand J Gastroenterol.* (2007) 42:1302–5. doi: 10.1080/00365520701420750
10. Comino I, Real A, De Lorenzo L, Cornell H, Lopez-Casado MA, Barro F, et al. Diversity in oat potential immunogenicity: basis for the selection of oat varieties with no toxicity in coeliac disease. *Gut.* (2011) 60:915–22. doi: 10.1136/gut.2010.225268
11. Comino I, Bernardo D, Bancel E, De Lourdes Moreno M, Sanchez B, Barro F, et al. Identification and molecular characterization of oat peptides implicated on coeliac immune response. *Food Nutr Res.* (2016) 60:30324. doi: 10.3402/fnr.v60.30324
12. Arentz-Hansen H, Fleckenstein B, Molberg O, Scott H, Koning F, Jung G, et al. The molecular basis for oat intolerance in patients with celiac disease. *PLoS Med.* (2004) 1:e1. doi: 10.1371/journal.pmed.0010001
13. Hardy MY, Tye-Din JA, Stewart JA, Schmitz F, Dudek NL, Hanchapola I, et al. Ingestion of oats and barley in patients with celiac disease mobilizes cross-reactive T cells activated by avenin peptides and immuno-dominant hordein peptides. *J Autoimmun.* (2015) 56:56–65. doi: 10.1016/j.jaut.2014.10.003
14. Tye-Din JA, Stewart JA, Dromey JA, Beissbarth T, Van Heel DA, Tatham A, et al. Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Sci Transl Med.* (2010) 2:1–14. doi: 10.1126/scitranslmed.3001012
15. Food and Agriculture Organization. *Oats Production in 2016, Crops/World Regions/Production Quantity from Pick Lists.* Rome: FAOSTAT (2017).
16. Schalk K, Lexhaller B, Koehler P, Scherf KA. Isolation and characterization of gluten protein types from wheat, rye, barley and oats for use as reference materials. *PLoS ONE.* (2017) 12:e0172819. doi: 10.1371/journal.pone.0172819
17. Leffler D, Schuppan D, Pallav K, Najarian R, Goldsmith JD, Hansen J, et al. Kinetics of the histological, serological and symptomatic responses to gluten challenge in adults with coeliac disease. *Gut.* (2013) 62:996–1004. doi: 10.1136/gutjnl-2012-302196
18. Metzger M-H, Heier M, Mäki M, Bravi E, Schneider A, Löwel H, et al. Mortality excess in individuals with elevated iga anti-transglutaminase antibodies: the KORA/MONICA Augsburg Cohort Study 1989–1998. *Eur J Epidemiol.* (2006) 21:359–65. doi: 10.1007/s10654-006-9002-4
19. Rubio-Tapia A, Rahim MW, See JA, Lahr BD, Wu TT, Murray JA. Mucosal recovery and mortality in adults with celiac disease after treatment with a gluten-free diet. *Am J Gastroenterol.* (2010) 105:1412–20. doi: 10.1038/ajg.2010.10
20. Mearns ES, Taylor A, Boulanger T, Craig KJ, Gerber M, Leffler DA, et al. Systematic literature review of the economic burden of celiac disease. *Pharmacoeconomics.* (2019) 37:45–61. doi: 10.1007/s40273-018-0707-5
21. Dennis M, Lee AR, Mccarthy T. Nutritional considerations of the gluten-free diet. *Gastroenterol Clin North Am.* (2019) 48:53–72. doi: 10.1016/j.gtc.2018.09.002
22. Whitehead A, Beck EJ, Tosh S, Wolever TM. Cholesterol-lowering effects of oat beta-glucan: a meta-analysis of randomized controlled trials. *Am J Clin Nutr.* (2014) 100:1413–21. doi: 10.3945/ajcn.114.086108
23. Rasane P, Jha A, Sabikhi L, Kumar A, Unnikrishnan VS. Nutritional advantages of oats and opportunities for its processing as value added foods - a review. *J Food Sci Technol.* (2015) 52:662–75. doi: 10.1007/s13197-013-1072-1
24. Aaltonen K, Laurikka P, Huhtala H, Maki M, Kaukinen K, Kurppa K. The long-term consumption of oats in celiac disease patients is safe: a large cross-sectional study. *Nutrients.* (2017) 9:611. doi: 10.3390/nu9060611
25. Khan A, Suarez MG, Murray JA. Non-celiac wheat/gluten sensitivity. *Clin Gastroenterol Hepatol.* (in press). doi: 10.1016/j.cgh.2019.04.009
26. Skodje GI, Sarna VK, Minelle IH, Rolfsen KL, Muir JG, Gibson PR, et al. Fructan, rather than gluten, induces symptoms in patients with self-reported non-celiac gluten sensitivity. *Gastroenterology.* (2018) 154:529–39 e522. doi: 10.1053/j.gastro.2017.10.040
27. Varney J, Barrett J, Scarlata K, Catsos P, Gibson PR, Muir JG. FODMAPs: food composition, defining cutoff values and international application. *J Gastroenterol Hepatol.* (2017) 32:53–61. doi: 10.1111/jgh.13698
28. Moulton ALC. The place of oats in the coeliac diet. *Arch Dis Childhood.* (1959) 34:51–5. doi: 10.1136/adc.34.173.51
29. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principle of dye binding. *Anal Chem.* (1976) 72:248–54. doi: 10.1006/abio.1976.9999
30. Tanner GJ, Blundell MJ, Colgrave ML, Howitt CA. Quantification of hordeins by ELISA: the correct standard makes a magnitude of difference. *PLoS ONE.* (2013) 8:e56456. doi: 10.1371/journal.pone.0056456
31. Colgrave ML, Goswami H, Byrne K, Blundell M, Howitt CA, Tanner GJ. Proteomic profiling of 16 cereal grains and the application of targeted

proteomics to detect wheat contamination. *J Prot Res.* (2015) 14:2659–68. doi: 10.1021/acs.jproteome.5b00187
32. Larroque OR, Bekes F, Wrigley CW, Rathmell WG. Analysis of gluten proteins in grain and flour blends by RP-HPLC. In: Shewry PR, Tatham AR, editors. *Seventh International Workshop on Gluten Proteins*. Bristol: Institute for Arable Crops Research (2000). p. 136–139. doi: 10.1039/9781847552372-00136
33. Bringans S, Eriksen S, Kendrick T, Gopalakrishnakone P, Livk A, Lock R, Lipscombe R. Proteomic analysis of the venom of Heterometrus longimanus (Asian black scorpion). *Proteomics.* (2008) 8:1081–96. doi: 10.1002/pmic.200700948
34. Gutierrez-Gonzalez JJ, Tu ZJ, Garvin DF. Analysis and annotation of the hexaploid oat seed transcriptome. *BMC Genomics.* (2013) 14:471. doi: 10.1186/1471-2164-14-471
35. Fallahbaghery A, Zou W, Byrne K, Howitt CA, Colgrave ML. Comparison of gluten extraction protocols assessed by LC-MS/MS analysis. *J Agric Food Chem.* (2017) 65:2857–66. doi: 10.1021/acs.jafc.7b00063
36. Shilov IV, Seymour SL, Patel AA, Loboda A, Tang WH, Keating SP, et al. The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol Cell Proteomics.* (2007) 6:1638–55. doi: 10.1074/mcp.T600050-MCP200
37. Tang WH, Shilov IV, Seymour SL. Nonlinear fitting method for determining local false discovery rates from decoy database searches. *J Proteome Res.* (2008) 7:3661–7. doi: 10.1021/pr070492f
38. Juhász A, Belova T, Florides C, Maulis C, Fischer I, Gell G, et al. Genome mapping of seed borne allergens and immune-responsive proteins in wheat. *Sci Adv.* (2018) 4:eaar8602. doi: 10.1126/sciadv.aar8602
39. Iwgsc. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science.* (2018) 361:eaar7191. doi: 10.1126/science.aar7191
40. Zhou M, Robards K, Glennie-Holmes M, Helliwell S. Analysis of volatile compounds and their contribution to flavor in cereals. *J Agric Food Chem.* (1999) 47:3941–53. doi: 10.1021/jf990428l
41. *Metals.* (2019). Available online at: https://www.fda.gov/food/chemicals-metals-pesticides-food/metals (accessed July 16, 2019).
42. Colgrave ML, Byrne K, Blundell M, Heidelberger S, Lane CS, Tanner GJ, et al. Comparing Multiple reaction monitoring and sequential window acquisition of all theoretical mass spectra for the relative quantification of barley gluten in selectively bred barley lines. *Anal Chem.* (2016) 88:9127–35. doi: 10.1021/acs.analchem.6b02108
43. Thompson T. Gluten contamination of commercial oat products in the United States. *N Engl J Med.* (2004) 351:2021–2. doi: 10.1056/NEJM200411043511924
44. Hernando A, Mujico JR, Mena MC, Lombardia M, Mendez E. Measurement of wheat gluten and barley hordeins in contaminated oats from Europe, the United States and Canada by Sandwich R5 ELISA. *Eur J Gastroenterol Hepatol.* (2008) 20:545–54. doi: 10.1097/MEG.0b013e3282f46597
45. Koerner TB, Cleroux C, Poirier C, Cantin I, Alimkulov A, Elamparo H. Gluten contamination in the Canadian commercial oat supply. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess.* (2011) 28:705–10. doi: 10.1080/19440049.2011.579626
46. Anderson RP, Degano P, Godkin AJ, Jewell DP, Hill AVS. *In vivo* antigen challenge in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope. *Nat Med.* (2000) 6:337–42. doi: 10.1038/73200
47. Junker Y, Zeissig S, Kim S-J, Barisani D, Wieser H, Leffler DA, et al. Wheat amylase trypsin inhibitors drive intestinal inflammation via activation of toll-like receptor 4. *J Exp Med.* (2012) 209:2395–408. doi: 10.1084/jem.20102660
48. Zevallos VF, Raker V, Tenzer S, Jimenez-Calvente C, Ashfaq-Khan M, Russel N, et al. Nutritional wheat amylase-trypsin inhibitors promote intestinal inflammation via activation of myeloid cells. *Gastroenterology.* (2017) 152:1100. doi: 10.1053/j.gastro.2016.12.006

Conflict of Interest: GT and JT-D are co-authors on a patent to use chill precipitation for gluten isolation (Provisional Patent Application no. 16/291,326) and Australia (Patent Application No. 2019201465). JT-D and MH are co-inventors of patents pertaining to the use of gluten peptides in therapeutics, diagnostics, and non-toxic gluten in celiac disease. JT-D holds shares in Nexpep Pty. Ltd. and is on the Scientific Advisory Board of ImmusanT, Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

REVIEW
published: 14 November 2019
doi: 10.3389/fpls.2019.01470

Modern Approaches in the Identification and Quantification of Immunogenic Peptides in Cereals by LC-MS/MS

Thais O. Alves[1,2], *Carolina T. S. D'Almeida*[1,2], *Katharina A. Scherf*[3] *and Mariana S. L. Ferreira*[1,2]*

[1] Food and Nutrition Graduate Program (PPGAN), Laboratory of Bioactives, Federal University of the State of Rio de Janeiro (UNIRIO), Rio de Janeiro, Brazil, [2] Laboratory of Protein Biochemistry—Center of Innovation in Mass Spectrometry (LBP-IMasS), UNIRIO, Rio de Janeiro, Brazil, [3] Department of Bioactive and Functional Food Chemistry, Institute of Applied Biosciences, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

OPEN ACCESS

Edited by:
Rainer Cramer, University of Reading, United Kingdom

Reviewed by:
Stefano Sforza, University of Parma, Italy
Stefanie Wienkoop, University of Vienna, Austria

***Correspondence:**
Mariana S. L. Ferreira
mariana.ferreira@unirio.br

Specialty section:
This article was submitted to Plant Proteomics, a section of the journal Frontiers in Plant Science

Received: *17 August 2019*
Accepted: *22 October 2019*
Published: *14 November 2019*

Citation:
Alves TO, D'Almeida CTS, Scherf KA and Ferreira MSL (2019) Modern Approaches in the Identification and Quantification of Immunogenic Peptides in Cereals by LC-MS/MS. Front. Plant Sci. 10:1470. doi: 10.3389/fpls.2019.01470

Celiac disease (CD) is an immunogenic disorder that affects the small intestine. It is caused by the ingestion of gluten, a protein network formed by prolamins and glutelins from cereals such as wheat, barley, rye and, possibly, oats. For predisposed people, gluten presents epitopes able to stimulate T-cells causing symptoms like nausea, vomiting, diarrhea, among others unrelated to the gastrointestinal system. The only treatment for CD is to maintain a gluten-free diet, not exceeding 20 mg/kg of gluten, what is generally considered the safe amount for celiacs. Due to this context, it is very important to identify and quantify the gluten content of food products. ELISA is the most commonly used method to detect gluten traces in food. However, by detecting only prolamins, the results of ELISA tests may be underestimated. For this reason, more reliable and sensitive assays are needed to improve gluten quantification. Because of high sensitivity and the ability to detect even trace amounts of peptides in complex matrices, the most promising approaches to verify the presence of gluten peptides in food are non-immunological techniques, like liquid chromatography coupled to mass spectrometry. Different methodologies using this approach have been developed and described in the last years, ranging from non-targeted and exploratory analysis to targeted and specific methods depending on the purpose of interest. Non-targeted analyses aim to define the proteomic profile of the sample, while targeted analyses allow the search for specific peptides, making it possible to quantify them. This review aims to gather and summarize the main proteomic techniques used in the identification and quantitation of gluten peptides related to CD-activity and gluten-related allergies.

Keywords: allergenic peptides, cereals, gluten, LC coupled to mass spectrometry, multiple reaction monitoring, prolamins, proteomics

INTRODUCTION

Cereals are one of the main food sources in the world. The nutrients provided by this group represent about 50% of the recommended daily intake (RDI) of carbohydrates and one third of the RDI for proteins. Cereal grains are also considered a good source of minerals and vitamins, especially complex B vitamins (Belitz et al., 2009). According to updated FAO data (2018), the cereal production, including non-food uses specially for maize, in the last year exceeded 2,600 million tons, with a slight decrease in production expected for 2019.

Wheat is one of the most important cereals in the world for human consumption, and is considered the most suitable raw material for bread and pasta making. Its production has remained constant over the years, currently only behind maize and followed by rice (FAOSTAT, 2018). In recent data reported by USDA (2018), world wheat production reached 733 million tons, whereas the estimated consumption is about 745 million tons. Barley, rye, and oats also have large production and consumption, but not so expressive as wheat, their production corresponds to about 25% of that of wheat. Rye is mostly applied for baking, while barley is applied in beer production and oats essentially commercialized as flour, bran, and other products for immediate consumption (Owusu-Apenten, 2002).

The search for practical ways in the preparation and consumption of meals combined with the promotion of healthier eating habits, sparked an increase in research for new processes for products (Pfeifer et al., 2014). Grain processing involves techniques that can alter protein structure, causing changes in solubility, viscoelastic properties, spatial conformation of proteins, and other changes (Hayta and Alpaslan, 2001). Among the main treatments used in cereal processing, extrusion and cooking can be highlighted, as well as baking and pasta production. However, there is a lack of studies to elucidate how processing may alter not only technological characteristics, but also nutritional and health implications, since cereal proteins, especially wheat, have a high allergenic potential in susceptible individuals.

The allergenic potential of cereals has been mainly related to gluten, a complex mixture of storage proteins found in cereals that is composed mainly of prolamins (responsible for the cohesiveness and extensibility of the gluten) and glutelins (maintenance of the elasticity and strength of the gluten). Gluten proteins have common structural characteristics. Their primary structure is subdivided into distinct domains that may exhibit repetitive sequences rich in the amino acids proline (P) and glutamine (Q) (Shewry and Halford, 2002), but low in amino acids with charged side groups. Different compositions in amino acids can be responsible for different reactivity associated with celiac disease (CD) (Belitz et al., 2009; Colgrave et al., 2015). Grains belonging to the *Triticeae* subtribe (wheat, barley, and rye) contain significantly higher levels of Q and P, being the main cereal grains responsible for triggering the immune response in celiacs (Colgrave et al., 2015). Cysteines represent only 2% of the amino acids of gluten proteins, but are extremely important for their structure and functionality, since they allow the formation of disulfide bonds, responsible for gluten polymerization (Wieser, 2007).

The disorders associated to gluten consumption are known as GRD (gluten-related diseases) and are classified into three types according to the response triggered in the body: autoimmune, allergic, and neither autoimmune nor allergic (Sapone et al., 2012). Examples of autoimmune diseases are dermatitis herpetiformis, gluten-induced ataxia, and CD. Among IgE antibody-mediated allergies, WDEIA (wheat-dependent exercise-induced anaphylaxis), contact urticaria, food allergy, and respiratory allergies are prominent. The respiratory allergies are related to the proteins of the albumin and globulin fractions, and are known as "baker's asthma" (Weiss et al., 1997). There are also disorders of non-allergic and non-autoimmune origin known as non-celiac gluten sensitivity or intolerance (Sapone et al., 2012).

In all cases of GRD, diagnosed patients cannot consume foods containing gluten or its traces, since even minimal amounts can trigger the reaction, causing variable symptoms, ranging from abdominal pain, bloating, and diarrhea, to osteoporosis and long-term infertility. The severity of the reaction is due to the degree of intolerance of each individual (Pietzak and Fasano, 2005; Banerjee, 2010). Therefore, it is extremely important to correctly identify the presence of immunogenic proteins in cereal products, in order to guarantee the safety of their consumption by the patients. One major problem for patients are the "hidden sources of gluten" that may be present in foods due to inadequate labeling or cross contamination during manufacturing or transportation. There is also concern about the presence of gluten due to the tendency of its incorporation into foods that traditionally do not contain wheat in its composition (*e.g.* sausages, nuggets, meatballs) (Day et al., 2006).

Some authors indicated the natural genetic variability as a strategy to be further exploited for the development of wheat varieties with lower levels of immunogenic epitopes (Spaenij–Dekking et al., 2005). By using the R5-based quantitation of immunodominant toxic epitopes as the trait of interest, Ribeiro et al. (2016) demonstrated that tetraploid varieties had a lower amount of toxic epitopes than hexaploid varieties, especially when compared to *Triticum aestivum* landraces, which were not subjected to breeding practices. Despite the advances in the study of genetic variability of wheat toxicity, at present there is no common hexaploid wheat that might be safe for CD patients. Furthermore, considering the wide range of *in vivo* immunoresponse between celiac patients and the limitation of

Abbreviations: APCI, atmospheric pressure chemical ionization; APPI, atmospheric pressure photoionization; BLAST, Basic Local Alignment Search Tool; CCS, collision cross-section; CD, celiac disease; DIA, data independent analysis; EBI, European Bioinformatics Institute; ELISA, enzyme-linked immunosorbent assay; ESI, electrospray ionization; FT-ICR, Fourier transform ioncyclotron resonance; GMO, genetic modified organisms; GP-HPLC-FLD, high performance liquid chromatography with gel permeation and fluorescence detection; GRD, gluten related disorders; IMS, ion mobility spectrometry; IT, ion trap; MRM, multiple reaction monitoring; MS^E, multiplex data-independent acquisition; LC-MS/MS, liquid chromatography coupled to mass spectrometry; MALDI, matrix-assisted laser desorption ionization; PIR, Protein Information Resource; PRM, parallel reaction monitoring; QMC, quartz crystal microbalance; Q, quadrupole; QqQ, triple quadrupole; SIB, Swiss Institute of Bioinformatics; SPR, surface plasmon resonance; SRM, single or selected reaction monitoring; ToF, time of flight; $UDMS^E$, ultra definition mass spectrometry; WDEIA, wheat-dependent exercise-induced anaphylaxis.

the immunological techniques for quantifying gluten proteins, the quantification and identification of cereal reactive proteins and peptides has been a complex task requiring constant analytical improvements.

Currently, the gold standard method to detect and quantify gluten in foods is the R5 ELISA and it is recommended by the Codex Alimentarius Commission (2008). More recently, the G12 ELISA was accepted by AOAC International as an official method of analysis, first action (Halbmayr-Jech et al., 2015). ELISAs are based on the immune reaction between specific antibodies that have been raised to detect the antigen to be determined, such as gluten. Due to their sensitivity, adequate recovery, repeatability, and reproducibility as demonstrated by collaborative studies, ELISAs are most commonly used to check for the presence of gluten in gluten-free raw materials and products. However, in some cases, ELISAs may give false negative results, because the monoclonal antibodies have been raised against prolamins (R5: raised against a rye extract and G12: raised against the α-gliadin 33-mer peptide) and are not suitable for all gluten protein types. As a consequence, the quantification can be compromised since the result is converted to gluten amount by multiplying the prolamin content by two, assuming the prolamin/glutelin ratio to be constant (Thompson and Méndez, 2008; Wieser and Koehler, 2009). ELISA methods currently cannot distinguish between the different gluten-containing cereals and are affected by the cross-reactivity of antibodies (Wieser and Koehler, 2009; Diaz-Amigo and Popping, 2013; Martínez-Esteso et al., 2017).

In this context, proteomic approaches appear to be more sensitive and reliable techniques than the currently used assays to identify gluten proteins, which present high amino acid sequence similarity and are difficult to distinguish. Especially when applying modern *in tandem* tools, proteomics can undoubtedly provide additional information to ELISA results, such as the confirmation of specific proteins by unraveling the peptide sequences (Martínez-Esteso et al., 2017).

A general workflow for cereal proteomics, as shown in **Figure 1**, should first consider the appropriate extraction taking into account the solubility of gluten proteins (Osborne, 1907) that usually requires the use of reducing (*e.g.* DTT—dithiothreitol, DTE—dithioerythritol, and TCEP-Tris2-Carboxyethyl phosphine hydrochloride) and denaturing agents (*e.g.* SDS or urea) (Schalk et al., 2018a; Schalk et al., 2018b). The enzymatic digestion is the crucial step in bottom-up proteomics. This high-throughput analysis is based on the detection of peptides to assign the proteins. The digestion is important, because the sensitivity of methods depends on the optimal size of peptides, considering the ability to be ionized and fragmented. Trypsin is the most commonly used enzyme due to its specific cleavage on the C-terminal side of lysine and arginine residues. However, due to the small number of these proteolytic cleavage sites in gluten proteins, a multiple enzymatic digestion or less specific enzymes have been used for cereal proteomics (Vensel et al., 2011; Fiedler et al., 2014). After that, the peptides can be separated by electrophoresis or liquid chromatography (LC).

LC coupled to mass spectrometry (LC-MS) is the most important tool for the identification and quantification of immunoreactive cereal proteins (Alves et al., 2017). One of the major contributions of proteomics in the study of CD has been the identification of the immunogenic epitope sequences of gluten peptides. The application of LC-MS methods makes it possible to identify the cereal species, the protein subunit, and to quantify thousands of peptides and proteins in the same experiment. Having a well-curated database that includes all possible proteins present in that organism is a great advantage for the identification of the sequences. However, peptide sequences may also be identified by *de novo* sequencing (Ferreira et al., 2014).

Other aspects, such as ionization source and type of MS analyzer, also influence the analysis and consequently the identification and quantification of the proteins. All of these topics will be briefly covered in this review. With the use of this information, significant advances in the understanding of GRD mechanisms, such as aspects related to resistance to proteolysis of these proteins and influence of cereal processing can be clarified, contributing to various aspects from the development of peptide detection and quantification methods to the selection of less reactive genotypes for better tolerability of these cereals.

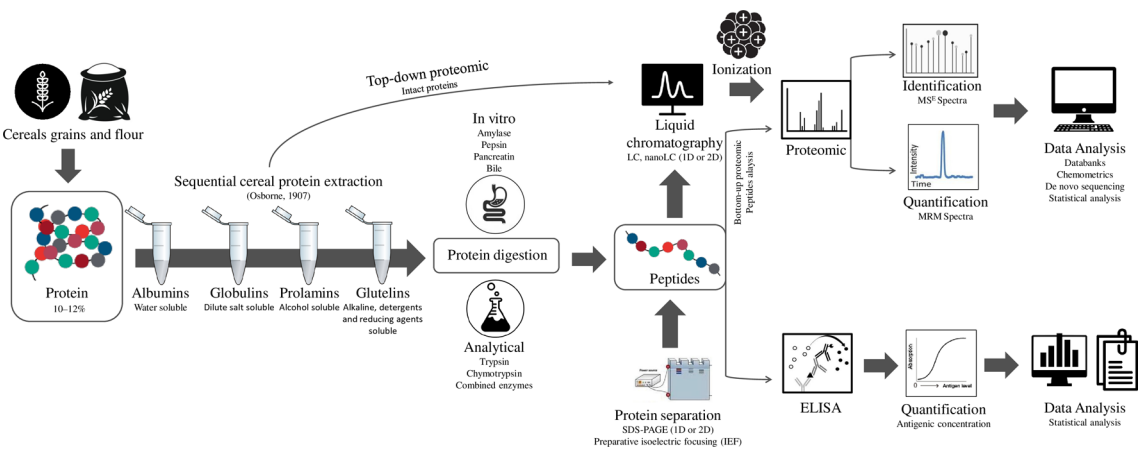


FIGURE 1 | General workflow for cereal proteomic analysis.

AVAILABLE GLUTEN PROTEIN AND CUSTOMIZED DATABASES

For LC-MS/MS analysis is important to define and use a well-curated gluten protein sequence database to improve the identification of immunogenic peptides. For this, it may be necessary to build a custom database based on an existing general database.

To provide the scientific community with a high quality protein knowledge base, the Swiss Institute of Bioinformatics (SIB), the European Bioinformatics Institute (EBI), and the Protein Information Resource (PIR) group have joined forces and created the UniProt consortium in 2002 (https://www.uniprot.org/). The UniProt Knowledgebase (UniProtKB), the main product of this consortium, combines UniProtKB/Swiss-Prot (contains over 560,823 sequences that have been created by experimental information extracted from the literature, organized and summarized, 379 belonging to *Triticum aestivum*—accessed Oct. 2019) and UniProtKB/TrEMBL (171,501,488 sequences that have been largely derived from high throughput DNA sequencing, 142,558 belong to wheat) (The UniProt Consortium, 2017). Besides this, the UniProt consortium also produces and maintains UniRef (which consists of clusters of sequences sharing 100%, 90%, or 50% of identity), UniParc (a highly redundant archive that contains original protein sequences retrieved from several different sources), or UniMES (a collection of metagenomic and environmental sequences) (Schneider et al., 2009). All known sequences can be BLAST searched against the entire database or a part of it and the resulting sequence of high homology can be downloaded from UniProt in FASTA format.

To customize a database, other softwares should be applied. Clustal Omega (Goujon et al., 2010) and Jalview (Waterhouse et al., 2009) are used in multiple sequence alignments. Clustal Omega is an online software tool that allows protein sequences to be entered in a text file format, with optional output formats (msf output format). Jalview is a desktop program or online software for editing, visualizing, and analyzing multiple sequence alignments using Clustal Omega. Lastly, it is necessary to count the number of sequences within the file and remove redundant sequences with DBtoolkit software (Martens et al., 2005). A custom database (GluPro V1.0) of wheat gluten proteins containing 630 unique protein sequences was created to be used in LC-MS/MS data analysis to identify the presence of immunoreactive gluten peptides in foods (Bromilow et al., 2017a). All software tools mentioned above were used to create this database and it provides more reliable protein IDs compared to the general database (*Viridiplantae*).

Juhász et al. (2015) also collected datasets from various public databases (UniprotKB, IEDB, NCBI GenBank) to create a specific database addressed to cereal prolamin protein families. The ProPepper database contains 2,484 unique and complete prolamin sequences, but also their peptides obtained with single- and multi-enzyme *in silico* digestions and specific epitopes that are responsible for wheat-related food disorders. Accordingly, is provided 667,402 unique digestion events, but also including redundant protein–peptide connections due to the simultaneous presence of some protein sequences in many genotypes and the frequency of the same peptide within a protein. Besides to be highly specific in the identification of protein sequences, this database provides specific information, such as the possible disease associated with the sequence.

Developed in 2005, Allergen Online database provides a updated peer reviewed allergen list and sequence searchable dataset to offer a risk assessment tool for evaluating the potential allergenicity of new food proteins produced by genetically modified organisms (GMO) and novel protein ingredients in processed foods (Goodman et al., 2016). The main goal is identify proteins that may present a potential risk of allergenic cross-reactivity. This database currently presents a list of 72 proteins known to induce CD together with a downloadable list containing more than 1,000 CD-active peptide sequences. However, this function cannot be used to search mass spectrometry (MS) data directly due to the restrictive size and not adapted format of the database (*e.g.* not available in FASTA format).

PROTEOMICS AS A TOOL FOR THE SCREENING FOR IMMUNOGENIC PEPTIDES

The "omic" suffix means collectively considering all constituents. Proteomics consists of the analysis of the set of proteins encoded by the genome and its component molecules responsible for the control of almost all biological processes (Graves and Haystead, 2002). The use of proteomics in food analysis has become a key technological tool for the characterization and quantification of proteins and peptides, especially when it comes to the evaluation of biological markers (Carr and Anderson, 2008; Herrero et al., 2012). The coupling of the chromatographic separation and mass spectrometer detection techniques (LC-MS) increases the speed of the analyses, allowing a large number of samples to be analyzed in a short period of time (Alves et al., 2017). In these studies, the amount of data generated is enormous and requires an important computational analytical effort to process data in a systemic and comparative way in order to deliver a practical conclusion and application (Victorio et al., 2018).

MS analyses can be divided into two types: untargeted and targeted approaches. While untargeted approaches aim to establish a comprehensive profile of the proteome of the sample, the targeted analysis allows the selection of specific molecules to be screened and studied in the sample (Saghatelian and Cravatt, 2005). Both types follow a standard workflow, where the sample is ionized *via* an ion source; the ions are separated according to their mass-to-charge ratios (m/z) and monitored by a mass analyzer prior to detection. *In tandem* MS (MS/MS) these precursor ions are then introduced into a collision cell where they undergo specific fragmentation through collision-induced dissociation (CID) by an inert gas, usually nitrogen or argon, resulting in the formation of product ions (Lovric, 2011). MS/MS is usually applied for complex samples, where identified peptides are selected and subjected to fragmentation to decipher the amino acid sequence, allowing the identification of sequences

that differ from each other by a single amino acid (Graves and Haystead, 2002).

The ionization source significantly impacts MS analysis as there are many ionization techniques and each has its advantages and ideal applications. The selection of the ideal ionization technique should be made based on the structure of the analyte of interest as well as the desired application (Buse et al., 2014). Various ionization techniques have been used with MS, including Electrospray Ionization (ESI), Atmospheric Pressure Chemical Ionization (APCI), Atmospheric Pressure Photoionization (APPI), and Matrix-Assisted Laser Desorption Ionization (MALDI). For the ion source, it is important to be efficient, but at the same time sensitive and "soft", to avoid the destruction of the analyte by unwanted fragmentation in-source (Lovric, 2011). Of these, the techniques most commonly used for having this feature are ESI and MALDI (El-Aneed et al., 2009). MALDI ionization essentially generates monocharged ions and thus does not require any deconvolution step. This technique emerged as an alternative to characterize wheat storage proteins due to its robustness and ability to ionize intact proteins and tolerate the presence of contaminants, such as detergents (SDS) commonly used for gluten extraction (Ferreira et al., 2014). However, this technique cannot be hyphenated directly to LC.

Conversely, ESI is powerful technique for the analysis of complex protein and peptide mixtures that benefit from the additional separation. Jira and Münch (2019) used LC-ESI-MS/MS for the simultaneous MS detection of the six most important grain species (barley, maize, oats, rice, rye, and wheat) in meat products based on marker peptides. ESI was also suitable to detect traces of immunogenic gluten marker peptides in a variety of foods (Sealey-Voyksner et al., 2010) and gluten marker peptides (e.g., Manfredi et al., 2015; Colgrave et al., 2016; Schalk et al., 2018a; Schalk et al., 2018b).

A miniaturized version of ESI, termed nanospray, has become the preferred method of introducing large peptides into the mass spectrometer in case peptide contents are suspected to be low to very low (Nadler et al., 2017; Hopper et al., 2019). nanoLC-ESI-MS/MS was efficient to identify 29 immunogenic peptides from wheat flour carrying a high number of epitopes (Alves et al., 2018). Droplets produced from nanoESI are smaller than in conventional ESI (of the order of a few hundred nanometers), greatly improving the sensitivity and explaining the predominance of this technique in quantitative large-scale proteomics. The use of nanoLC to analyze complex peptide mixtures, especially when combined to orthogonal separation such as 2D RP/RP separation prior to MS/MS analysis, improves the resolution facilitating the identification and quantification of peptides containing CD immunogenic epitopes even at low femtomolar levels of detection (van den Broeck et al., 2015). When sample amounts are limited, nanoLC remains the best option due to the increased analytical sensitivity, otherwise UPLC or even HPLC separation is also useful for gluten detection.

Quadrupole is one of the most common type of mass analyzer, which four parallel metal rods are opposite connected electrically and voltage is applied to the diagonally placed pair of rods, resulting in an electrical field that causes the ions to travel forward. Nonetheless, a set of mass analyzers can be used for this purpose, such as ToF (Time of Flight), IT (ion trap), Orbitrap®, or FT-ICR (Fourier transform ion cyclotron resonance), they can also be combined to improve the sensitivity of the method (Herrero et al., 2012).

MS-Based Identification of Immunogenic Peptides

The variability of cereal protein composition caused by the different species and varieties (genetic variability) and by growing conditions (environmental variability) leads to methodological difficulties for the analysis of immunoreactive peptides and also for the selection of genotypes (Juhász et al., 2015). In addition, the high amount of repetitive units and the similarity of the amino acid sequences of the different prolamins, with limitations in the available methodologies, make it difficult to accurately identify peptides that cause diseases related to cereal consumption, as well as their genotype frequency, variability, and stability (Juhász et al., 2015).

As mentioned, MS is considered to be the golden standard for the analysis of biomolecules in complex samples, such as food matrices, because it presents high levels of sensitivity and specificity, and has been increasingly used in food analysis (Colgrave, 2017). In cereal proteins, multiple acquisition methods or DIA (data independent acquisition), such as MS^E allow minimizing data loss (*e.g.*, non-fragmented precursors) (Victorio et al., 2018). In MS^E methods, all ions generated in the source are transmitted to the collision chamber, which alternates between low and high energy, sending precursors and fragments *quasi*-simultaneously to the TOF (Time of Flight) analyzer (Egertson et al., 2015). In DIA methods there is no previous selection of precursors or a threshold of ion intensity to undergo fragmentation, while for DDA typically the three most intensive single or multiple charged ions eluting from the column are selected for fragmentation (van den Broeck et al., 2015).

The use of label-free acquisition methods, such as the multiplex MS^E method, takes advantage of a data collection approach that focuses on maximizing peptide fragmentation and then improving the identification and proteome coverage (Victorio et al., 2018). MS^E methods have been applied to gluten protein identification and quantitation (Uvackova et al., 2013; van den Broeck et al., 2015; Bromilow et al., 2017b). Label-free absolute quantification is based on the relationship between MS signal response and protein peptide concentration: the average MS signal response for the three most intense tryptic peptides per mole of protein (top 3) is constant (CV < 10%) and this relationship is used to calculate a universal signal response factor given an internal standard (Silva et al., 2006). However, due to data complexity many steps of data processing are required in DIA such as peak alignment, ion detection, clustering, and normalization prior to peptide matching by search algorithms from a database of protein sequences.

In general, there are two possible approaches when applying LC-MS/MS for gluten detection, both of which are valid, but depend on the question to be answered. The first option is to specifically detect known CD-immunogenic peptides in order to estimate the immunogenicity of gluten. This has been reported

for a selection of α- and γ-gliadin peptides (Sealey-Voyksner et al., 2010), α-gliadin peptides (van den Broeck et al., 2015), the 33-mer peptide (Schalk et al., 2017), and various gluten-derived peptides (Alves et al., 2018; Malalgoda et al., 2018). In contrast, the second option is to look for the presence of gluten, but not necessarily for CD-immunogenic peptides. Due to their length of at least nine amino acids, the poor enzymatic digestibility of the corresponding repetitive sequences, and their high contents of glutamine and proline, CD-immunogenic peptides often have properties unfavorable for MS detection, whereas other gluten peptides might be more abundant. With the overall aim to detect gluten, this approach was also used to identify marker peptides in wheat, rye, barley, and oats (Manfredi et al., 2015; Schalk et al., 2018a; Schalk et al., 2018b).

Recent examples demonstrating the successful application of proteomics in the evaluation of the presence of gluten marker peptides, include the detection of the presence of gluten in beers (Tanner et al., 2013; Allred et al., 2014). Tanner et al. (2013) also made a comparison between two different gluten detection methods, reinforcing the superiority of LC-MS/MS to detect gluten peptides in relation to the ELISA due to its higher sensibility and the ability to detect both, glutelin and prolamins, and not only prolamins as ELISA. This fact can be corroborated by Colgrave et al. (2014), where MS was used to detect and confirm the presence of hydrolyzed gluten proteins in beers which had been previously estimated as gluten-free by ELISA. A set of barley-specific peptide markers was also proposed to evaluate the contamination of processed food, ensuring the food safety for CD patients (Colgrave et al., 2016).

In fact, MS has been effectively applied to define a set of specific analytical targets, such as signature peptides specific to prolamins or cereal-containing gluten proteins. The main interest of these works is to apply new methodologies that can overcome food adulteration and mislabeling or to check authenticity of cereal based-products (**Table 1**). Bönick et al. (2017) reported an analytical strategy, based on *in silico* steps and LC-MS/MS, to check the authenticity of wheat, spelt, and rye addition in bread products. MS has been reported as a promising alternative to ELISA, in particular for the detection but also quantification of proteins in contaminated food, as it can target multiple and very specific analytes (Martínez-Esteso et al., 2016).

Fiedler et al. (2014) identified a list of specific grain peptides of wheat, barley, rye, and oats for the detection of gluten contamination in several types of commercial flours. Specifically, targeted MS/MS method enabled the detection of two wheat peptide markers at a level of 10 ppm of wheat flour spiked into gluten-free oat flour. Martínez-Esteso et al. (2016) identified a set of unique wheat gluten peptides and proposed their use as markers of the presence of gluten related to the manifestation of CD symptoms. The authors reinforce the idea that this strategy can be applied to other allergens and that this is the first step toward the standardization of a new methodology, using LC-MS techniques, to evaluate the immunogenicity of different food matrices but also to produce reference materials, since the establishment of a set of markers is the first step to infer the presence of gluten and that enable the quantity of gluten present to be determined.

In the last decade, ion mobility spectrometry (IMS) has appeared as an analytical separation technique, especially important to the analysis of primary structures with a high degree of homology, such as gluten proteins. The IMS consists of an orthogonal separation technique, where for each value of *m/z* a spectrum of drift time is added. The drift time corresponds to the time the ion takes to cross the ionic mobility cell where an inert gas is inserted, allowing the determination of shock sections, or collision cross-section (Michaelevski et al., 2010). Thus, the ions can be further differentiated by size, shape, and charge, which allow separating by the three-dimensional conformation even peptides that present the same *m/z* or reverse peptides. In this way, the IMS can be applied to improve LC-MS and GC-MS workflows, since it increases method sensitivity by isolating the compounds of interest from background noise, improving confidence of identification, either in targeted or non-targeted approaches (Hernández-Mesa et al., 2019).

Wheat allergens from the non-gluten soluble protein fraction (albumins and globulins) have also been reported and identified by MS (Larre et al., 2011). Samples of diploid and hexaploid wheat were used to incite immunological reaction with human sera and then were subsequently analyzed and identified by MS. The analysis of 2D spots revealed by immunoblotting leads to the MS-based identification of 39 IgE-binding proteins, some of them unknown thus far as wheat allergens. A recent study evaluated albumins and globulins from different genotypes of Brazilian wheat flour through the application of MS^E and IMS, called $UDMS^E$ (Ultra Definition Mass Spectrometry). Collectively, about 5,900 proteins and 45,000 peptides (Victorio et al., 2018) were identified in the dataset and relatively quantified with 8 peptides/protein. Alves et al. (2018) reported that some of these proteins found have been previously described and associated with the development of respiratory allergies such as baker's asthma. Serpins, purinins, α-amylase/protease inhibitors, globulins, and farinins have also been associated with the humoral response to celiac disease (Huebener et al., 2014).

Following the same approach, Alves et al. (2018) evaluated the allergenic potential of nine wheat flours of different technological qualities by assessment of their immunogenic profiles. Peptides responsible for the manifestation of CD and other wheat-related allergies were identified in both gluten and soluble protein fractions. This work points to a relation between the variability in the expression of allergens and the technological quality of wheat flour, showing a distinct proteomic profile in flours of inferior technological quality, concluding that they can be more immunoreactive than the other qualities, especially due to the highest expression of two isoforms of serpins.

It is important to highlight that, to reach the identification of the peptide sequences by proteomic tools, the peptides must be present in the databases, so that the results obtained in the analyses can be cross-checked with those already consolidated (Altenbach et al., 2010). One of the major limitations to conducting proteomic studies in wheat was the lack of complete sequencing of the wheat genome (Bromilow et al., 2017a). It is important to note that a high percentage of non annotated proteins makes difficult the functional classification based on the basis of gene ontology. From the 414 soluble proteins found differentially

TABLE 1 | Overview of studies using LC-MS to detect gluten in foods.

Title	Food matrix	Techniques/methods	Reference
Novel aspects of quantitation of immunogenic wheat gluten peptides by liquid chromatography–mass spectrometry/mass spectrometry	Quinoa flour; whole grain corn flour; whole grain soy flour; vital wheat gluten flour; whole wheat flour; rye flour; barley flour; rice flour; oat flour; powdered ice tea mix; pasta; orzo; cheerios; hot sauce; bread; goldfish crackers; white vinegar; toothpaste; body lotion; body wash; beer; gin; vodka; rum; red wine; white wine and GF product	HPLC-ESI-TQS-MS/MS	Sealey-Voyksner et al. (2010)
Assessment of allergenicity of diploid and hexaploid wheat genotypes: identification of allergens in the albumin/globulin fraction	Wheat; human sera	ELISA; SDS-PAGE; immunoblotting; LC-MS/MS	Larre et al. (2011)
Measuring hordein (gluten) in beer—a comparison of ELISA and mass spectrometry	Beer	Western blot; ELISA sandwich; MRM-MS	Tanner et al. (2013)
MS^E based multiplex protein analysis quantified important allergenic proteins and detected relevant peptides carrying known epitopes in wheat grain extracts	Wheat	NanoUPLC-QTOF-MS/MS	Uvackova et al. (2013)
The MS^E-proteomic analysis of gliadins and glutenins in wheat grain identifies and quantifies proteins associated with celiac disease and baker's asthma	Wheat	NanoUPLC-QTOF-MS/MS	Uvackova et al., 2013
Evaluation of qualitative and quantitative immunoassays to detect barley contamination in gluten-free beer with confirmation using LC-MS/MS	Barley; GF beer	EZ Gluten assay; AllerTek Gluten ELISA; LC-Qtof-MS/MS	Allred et al. (2014)
Characterization of grain-specific peptide markers for the detection of gluten by mass spectrometry	Gluten; wheat flour; barley flour; rye flour; oat flour	NanoHPLC-ESI-pSMR; MS/MS	Fiedler et al. (2014)
Assessment of the allergenicity of soluble fractions from GM and commercial genotypes of wheats	Wheat; GM wheat (*T. aestivum* and *T. durum*); human sera	SDS-PAGE; western blot; immunoblotting; nanoLC-Qtof-MS/MS	Lupi et al. (2014)
Specific nongluten proteins of wheat are novel target antigens in celiac disease humoral response	Wheat; Human sera	ELISA; SDS-PAGE; immunoblotting; MS/MS	Huebener et al. (2014)
Using mass spectrometry to detect hydrolysed gluten in beer that is responsible for false negatives by ELISA	Beer	ELISA; nanoHPLC-ESI-MRM-MS	Colgrave et al. (2014)
Qualitative and quantitative determination of peptides related to celiac disease in mixtures derived from different methods of simulated gastrointestinal digestion of wheat products	Durum wheat (ground kernels; semolina; dough; extruded pasta; dried pasta and cooked pasta)	LC-ESI-MS	Prandi et al. (2014)
Label free targeted detection and quantification of celiac disease immunogenic epitopes by mass spectrometry	Wheat	On-line 2D nanoLC–MS/MS; UPLC-MRM-MS/MS	van den Broeck et al. (2015)
Allergen relative abundance in several wheat varieties as revealed *via* a targeted quantitative approach using MS	Wheat (*T. aestivum, T. durum, T. monococcum*)	LC-MS/MS	Rogniaux et al. (2015)
Proteomic profiling of 16 cereal grains and the application of targeted proteomics to detect wheat contamination	Barley; wheat; rye; oats; green wheat; amaranth; chia; quinoa; sorghum; tef; buckwheat; soy; millet; maize	SDS-PAGE; western blot; nanoUPLC-ESI-MRM-MS	Colgrave et al. (2015)
Multiplex liquid chromatography–tandem mass spectrometry for the detection of wheat, oat, barley and rye prolamins towards the assessment of gluten-free product safety	Flour; seeds; pasta; biscuits; cookies; crackers; beverages; breads; breakfast cereals; snacks	HPLC-IonTrap-MS/MS	Manfredi et al. (2015)
Defining the wheat gluten peptide fingerprint *via* a discovery and targeted proteomics approach	Wheat gluten; GluVital®	ELISA; nanoUPLC-ESI-MS/MS	Martínez-Esteso et al. (2016)
Identification of barley-specific peptide markers that persist in processed foods and are capable of detecting barley contamination by LC-MS/MS	Barley; wheat; rye; oats; green wheat; amaranth; chia; quinoa; sorghum; tef; buckwheat; soy; millet; maize; breakfast cereals	nanoUPLC-ESI-MRM-MS	Colgrave et al. (2016)
Quantitation of the immunodominant 33-mer peptide from α-gliadin in wheat flours by liquid chromatography tandem mass spectrometry	Wheat flour	RP-HPLC; 1H qNMR; untargeted MS/MC; ESI-MRM-MS/MS	Schalk et al. (2017)
Determination of wheat, rye and spelt authenticity in bread by targeted peptide biomarkers	Wheat; spelt; emmel wheat; einkorn wheat; barley; maize; oat; rye	UPLC-ESI-MRM-MS/MS	Bönick et al. (2017)
Peptides from gluten digestion: a comparison between old and modern wheat varieties	Wheat (*T. aestivum*, *T. durum*, *T. monococcum*, *T. dicoccum*, *T. spelta*)	UPLC-ESI-MS; HPLC-ESI-MS/MS	Prandi et al. (2017)

(Continued)

TABLE 1 | Continued

Title	Food matrix	Techniques/methods	Reference
Development and validation of the detection method for wheat and barley glutens using mass spectrometry in processed foods	Seeds; flour; beers; cookies; beverages; GF products (GF flour; corn flour; apple wine; rice wine)	ELISA; LC-ESI-MRM-MS	Liao et al. (2017)
Using LC-MS to examine the fermented food products vinegar and soy sauce for the presence of gluten	Vinegar; malt vinegar; soy sauce	ELISA; UHPLC-MRM-MS/MS	Li et al. (2018)
Differential expression of albumins and globulins of wheat flours of different technological qualities revealed by nanoUPLC-UDMS^E	Wheat flour	nanoUPLC-HDMSE; nanoUPLC-UDMSE	Victorio et al. (2018)
Immunogenic and allergenic profile of wheat flours from different technological qualities revealed by ion mobility mass spectrometry	Wheat flour	nanoUPLC-MSE; nanoUPLC-UDMSE	Alves et al. (2018)
Detection and quantitation of immunogenic epitopes related to celiac disease in historical and modern hard red spring wheat cultivars	Wheat	RP-HPLC; SDS-PAGE; SRM-MS	Malalgoda et al. (2018)
Targeted liquid chromatography tandem mass spectrometry to quantitate wheat gluten using well-defined reference proteins	Wheat	RP-HPLC; untargeted MS/MS; MRM-MS	Schalk et al. (2018b)
Quantitation of specific barley, rye, and oat marker peptides by targeted liquid chromatography–mass spectrometry to determine gluten concentrations	Barley; rye; oat	RP-HPLC; untargeted MS/MS; MRM-MS; competitive R5-ELISA; SDS-PAGE	Schalk et al. (2018a)
A complete mass spectrometry (MS)-based peptidomic description of gluten peptides generated during *in vitro* gastrointestinal digestion of durum wheat: implication for celiac disease	Durum wheat	SDS-PAGE; UHPLC-ESI-MS/MS; UPLC-ESI-MS	Boukid et al. (2019)

expressed in common wheat flours, 85% proteins were not yet described, according to their biological function (Victorio et al., 2018). An alternative to reduce the misidentification of sequences is the use of *de novo* sequencing to assemble wheat gluten gene sequences (Zhang et al., 2014). However, recently, the complete wheat genome was released, making it possible to improve the identifications of the proteins present in this cereal, since more peptides will be annotated in the proteomic databases (Ramírez-González et al., 2018).

MS-Based Quantification of Immunogenic Peptides

MS can also be applied for the selection and quantification of specific peptides by methods called MRM (multiple reaction monitoring) (Anderson and Hunter, 2006) or also called SRM (selected reaction monitoring) or PRM (parallel reaction monitoring) (Peterson et al., 2012), depending on the instrument manufacturer, which allow a targeted analysis of these peptides and their quantification even at minimum or trace concentrations. A set of strategies has been developed to measure the allergenic potential of various cereal species and LC-MRM/MS technology has been useful for the identification and quantification of peptides containing immunogenic epitopes at low levels of detection, such as femtomolar (van den Broeck et al., 2015). Different approaches can be used to quantify these peptides, like label-free quantification combined with external calibration.

This methodology was used by van den Broeck et al. (2015) to quantify CD immunogenic epitopes in three varieties of wheat (two hexaploid and one tetraploid). A list of nine peptides was proposed to create the calibration curves that quantified the amount of glia-α2 and glia-α20 in gluten extracts from the samples (**Table 2**). The reliability of the results depends on optimal digestion conditions and limit of detection and/or ionization properties of the peptides. Malalgoda et al. (2018) used the same approach to quantify immunogenic peptides from old and modern hard red spring wheat cultivars. Even though, it was not possible to associate the year of harvesting with the amounts of immunogenic epitopes and α-gliadin since it was randomly detected in all samples analyzed.

Schalk et al. (2017) developed a targeted LC-MS/MS method to quantify the immunodominant gluten peptide called 33-mer (LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF), which contains three different overlapping T-cell epitopes (PFPQPQLPY; PYPQPQLPY; PQPQLPYPQ) that initiate a strong immunological response (Shan et al., 2002). In this study, the quantitative data on contents of 33-mer peptide in different

TABLE 2 | List of gluten peptides selected for the creation of calibration curves (van den Broeck et al., 2015).

Peptide sequence
LQLQ**PFPQPQLPY**
LQLQ**PFPQPQLPYPQ**PQPF
LQLQ**PFPQPQLPYPQ***P/H*LPYPQPQPF
LQLQ**PFPQPQLPYPQPQLPYPQ**PQPF
LQLQ**PFPQPQLPYPQPQLPYPQPQLPYPQ**PQPF
RPQQPYPQPQPQY
RPQQPYPQSQPQY
QQQLIPCRDVVL
QQILQQQLIPCRDVVL

CD-epitope sequences within the peptides are shown in bold

wheat cultivars was carried out by combining a stable isotope dilution assay with LC-MS/MS, as first reported for peptides by Stöcklin et al. (1997). The authors detected the presence of this peptide in 23 common wheat flours and in two spelt flours (*T. spelta*), but it was absent in tetraploid and diploid wheat flours. No obvious cluster formation between modern and old wheat cultivars and no correlations between contents of 33-mer and those of α-gliadins, gliadins, gluten, or crude protein were observed. Indeed, the harvest year had a higher influence on 33-mer contents than the cultivar. It is important to highlight that this was the first study that accurately quantitated the 33-mer peptide in wheat flours.

Recent studies use the combination of untargeted and targeted methods as a strategy to quantify gluten marker peptides in cereals and determine gluten concentrations in different types of samples (Schalk et al., 2018a; Schalk et al., 2018b). Schalk et al. (2018b) developed a methodology that allowed the simultaneous determination of 33 marker peptides, 16 for wheat, seven for rye, seven for barley, and three for oats using LC-MS/MS in MRM mode, using a labeled peptide as internal standard. Furthermore, they compared the LC-MS/MS results with those of R5 ELISA RP-HPLC and GP-HPLC-FLD (gel-permeation high-performance liquid chromatography with fluorescence detection) and found a strong correlation between LC-MS/MS and the other methods. When analyzing wheat starch samples, the LC-MS/MS and ELISA results agreed well in four out of seven cases, but there were two samples where LC-MS/MS found substantially higher and one with lower gluten contents than ELISA. The lower values obtained by LC-MS/MS may be explained by the presence of other gluten peptides that were not monitored with the targeted method, whereas the higher values may be due to variable gliadin/glutenin ratios in wheat starches that may lead to an underestimation of gluten contents by ELISA (Schalk et al., 2018a).

One of the most important considerations when using targeted LC-MS/MS is the careful selection of gluten marker peptides, because only these pre-defined peptides will be monitored. Even a single amino acid substitution, deletion, insertion, or post-translational modification will result in that marker peptide not being detected anymore, even if the sample may still contain other gluten-derived and possibly immunogenic peptides. While it is possible to use stable isotope labeled peptides or concatamers as internal standards to precisely quantify the selected peptides, the conversion of gluten peptide contents to gluten contents is far from being trivial. Legislation requires the result to be expressed as mg gluten/kg of the food, so that the correspondence between the amount of gluten and the resulting peptides needs to be established by careful calibration, also considering the whole sample preparation procedure. One of the most important points to verify is the extent of enzymatic hydrolysis. Matrix-matched calibration has been applied in many cases (Fiedler et al., 2014; Manfredi et al., 2015), but the use of well-defined gluten reference materials revealed the complexity of converting marker peptide contents to gluten contents (Schalk et al., 2018a; Schalk et al., 2018b). Further pro's and con's of using ELISA and LC-MS/MS for gluten detection are given in **Figure 2**.

A quantitative approach was also used to compare the relative abundance of 12 allergens in the albumins/globulins fraction in seven wheat varieties (Rogniaux et al., 2015). Allergens were monitored by targeted investigation of one to two proteotypic peptides (single protein peptides), and the abundance of some allergens was found to be quite stable among genotypes, while others, such as α-amylase inhibitors, showed clear differences depending on the wheat species, revealing themselves as possible markers of allergenicity in wheat. The content of allergenic polypeptides from these fractions was also investigated in common and genetically modified wheat (Lupi et al., 2014) revealing a large variation in the amounts of these allergens. The lack of information on the peptide sequences and epitopes responsible for the allergies triggered by albumins/globulins render targeted studies in this protein fraction even more complicated.

OTHER STRATEGIES TO UNRAVEL AND TO DETECT GLUTEN PEPTIDES

Even with all the benefits of LC-MS/MS, such as the identification and quantification of specific proteins and peptides, new techniques have also been highlighted, such as the use of biosensors. Soler et al. (2016) used Surface Plasmon Resonance (SPR), a biosensor able to detect and quantify chemical and biological analytes quickly, sensitively, and specifically in complex field samples. SPR was able to detect gluten toxic peptides in the urine of CD patients and directly quantify the small digestive peptides without the need for prior extraction or purification procedures, so that the assay can be performed in 20 min. White et al. (2018) developed a floating gate transistor biosensor with longer analysis time (1.5 h), but it was still able to quantify wheat proteins faster than ELISA.

In addition to the shortest analysis time, biosensors also have high sensitivity at low detection limits and low cost. Chu et al. (2012) used a quartz crystal microbalance (QCM) immunosensor to detect gliadin in foods and had high sensitivity, being able to detect 8 ppb of this protein. In addition, the cost of materials for biosensor analyses is estimated to be approximately threefold less than the cost of a single ELISA kit (Soler et al., 2016). In the future, immunosensors may be promising alternatives for existing immunochemical tests, such as ELISAs, because of their specificity and sensitivity (Scherf et al., 2016). However, this method does not allow the characterization of proteins and their respective peptides, as in LC-MS/MS. In addition, the type of sensor that is the best candidate to replace the ELISA still needs to be evaluated.

LC-MRM/MS analysis can also be linked to genomics to improve our understanding of the genes responsible for expressing allergenic proteins, culminating in the development of wheat varieties with a lower allergenic potential (Salentijn et al., 2013), increasing the variety of food options that can be consumed by GRD patients by ensuring food safety. Moreover, the studies about authenticity requires also an approach towards a well-defined "proteogenomic annotation" looking carefully

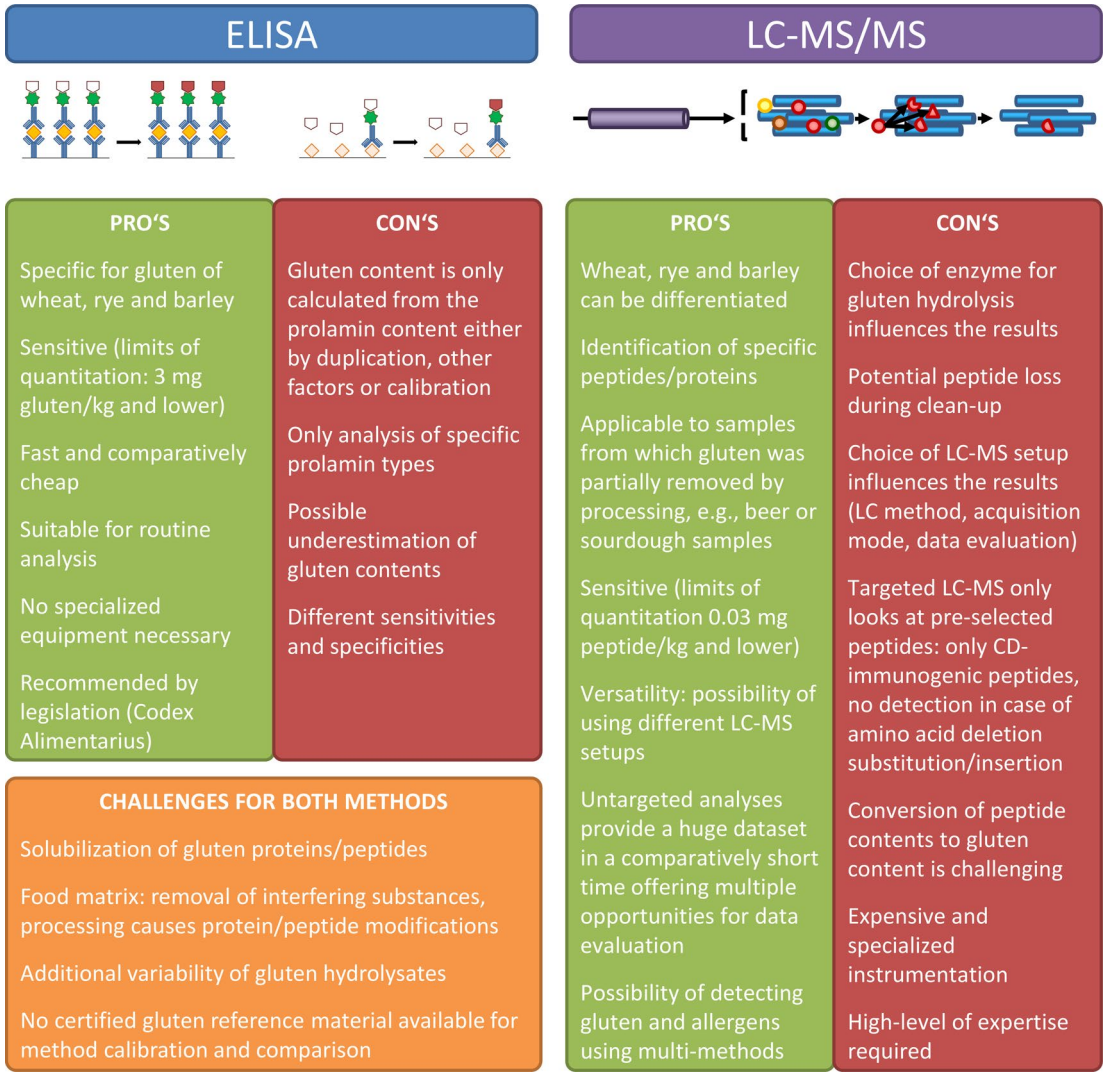


FIGURE 2 | Overview of advantages and disadvantages of ELISA and LC-MS/MS for gluten detection.

at the specific peptide candidates from an enzymatic digest (Bönick et al., 2017).

CONCLUDING REMARKS AND PERSPECTIVES

The use of LC-MS/MS strategies is the most useful and promising path to improve the identification and quantification of immunogenic peptides. Despite the methodological difficulties, it proves to be a fast, sensitive, and reproducible method. In addition, it can be extended to several other allergenic food matrices, like dairy, nuts, and seafood. Thus, knowing the profile of allergenic proteins of cereals is necessary as a basis, not only for future applications of MS in the quantification of gluten in food, but also to ensure the safety of consumers regarding food labeled cereal- or gluten-free.

Although the declaration of gluten-containing cereals on products labeled gluten-free is mandatory worldwide, there is no certified reference material available for gluten. The available reference material contains only gliadins that underestimate the gluten content, besides the problem of reproducing a new batch with similar properties and composition. The majority of MS-based studies have been conducted with the final objective to establish a reference material for gluten analysis starting from the study of specific grain peptide markers. Therefore, targeted high-resolution MS/MS methods allowed the quantification of low levels of specific marker peptides from different species and protein types.

When comparing LC-MS/MS methods to ELISA for gluten detection, ELISA still remains the method of choice in most cases, because it is fast, comparatively cheap, suitable for routine analyses, and does not require highly specialized equipment. However, several studies have shown that ELISA may

underestimate gluten contents especially in processed foods that have been extensively heat-treated or hydrolyzed. Untargeted LC-MS/MS is recommended to screen for the presence of gluten-derived peptides in products such as beer, malt vinegar, and fermented sauces. However, there are some points that will equally all analytical methods, because gluten extractability has been shown to be reduced substantially in heat-treated foods and processing-induced post-translational protein modifications will lead to reduced gluten detectability irrespective of the analytical method used.

The use of modern MS-based techniques, combining orthogonal separations with high sensitivity and reliable certified references materials will hopefully help to better comprehend the effect of food processing or plant breeding on gluten immunogenicity. Continued efforts in this area will also help to solve the questions about the selection of relevant target epitopes and even antibodies, taking account the high protein polymorphism and the fact that patients react individually to different proteins and present variable sensitivities.

AUTHOR CONTRIBUTIONS

TA organized and wrote the manuscript and the summarized table, CD'A complemented the writing and designed the figure, KS reviewed the manuscript, and MF supervised and reviewed the manuscript.

FUNDING

This work was funded by Federal University of State of Rio de Janeiro (UNIRIO), FAPERJ (E-26/202.709/2018), CNPq (427116/2018-0), and in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001.

REFERENCES

Allred, L. K., Sealey Voyksner, J. A., and Voyksner, R. D. (2014). Evaluation of qualitative and quantitative immunoassays to detect barley contamination in gluten-free beer with confirmation using LC-MS/MS. *J. AOAC Int.* 97 (6), 1615–1625. doi: 10.5740/jaoacint.14-058

Altenbach, S. B., Vensel, W. H., and DuPont, F. M. (2010). Integration of transcriptomic and proteomic data from a single wheat cultivar provides new tools for understanding the roles of individual alpha gliadin proteins in flour quality and celiac disease. *J. Cereal Sci.* 52 (2), 143–151. doi: 10.1016/j.jcs.2010.04.006

Alves, T. O., D'Almeida, C. T. S., and Ferreira, M. S. L. (2017). "*Celiac Disease and Non-Celiac Gluten Sensitivity*," in *Determination of Gluten Peptides Associated with Celiac Disease by Mass Spectrometry* (London, UK: InTech). doi: 10.5772/67547

Alves, T. O., D'Almeida, C. T. S., Victorio, V. C. M., Souza, G. H. M. F., Cameron, L. C., and Ferreira, M. S. L. (2018). Immunogenic and allergenic profile of wheat flours from different technological qualities revealed by ion mobility mass spectrometry. *J. Food Composition Anal.* 73, 67–75. doi: 10.1016/j.jfca.2018.07.012

Anderson, L., and Hunter, C. L. (2006). Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol. Cell. Proteomics* 5 (4), 573–588. doi: 10.1074/mcp.M500331-MCP200

Awad, H., Khamis, M. M., and El-Aneed, A. (2015). Mass Spectrometry, Review of the Basics: Ionization. *Appl. Spectrosc. Rev.* 50 (2), 158–175. doi: 10.1080/05704928.2014.954046

Banerjee, B. (2010). *Nutritional management of digestive disorders*. Boca Raton, Florida, US: CRC Press. doi: 10.1201/9781420086553

Belitz, H. D., Grosch, W., and Schieberle, P. (2009). *Food chemistry, 4th revised and extended edn*.

Bönick, J., Huschek, G., and Rawel, H. M. (2017). Determination of wheat, rye and spelt authenticity in bread by targeted peptide biomarkers. *J. Food Composition Anal.* 58, 82–91. doi: 10.1016/j.jfca.2017.01.019

Boukid, F., Prandi, B., Faccini, A., and Sforza, S. (2019). A Complete Mass Spectrometry (MS)-Based Peptidomic Description of Gluten Peptides Generated During In Vitro Gastrointestinal Digestion of Durum Wheat: Implication for Celiac Disease. *J. Am. Soc. Mass Spectrom.*, 30, 1–10. doi: 10.1007/s13361-019-02212-8

Bromilow, S., Gethings, L. A., Buckley, M., Bromley, M., Shewry, P. R., Langridge, J. I., et al. (2017a). A curated gluten protein sequence database to support development of proteomics methods for determination of gluten in gluten-free foods. *J. Proteomics*. 163, 67–75. doi: 10.1016/j.jprot.2017.03.026

Bromilow, S. N., Gethings, L. A., Langridge, J. I., Shewry, P. R., Buckley, M., Bromley, M. J., et al. (2017b). Comprehensive proteomic profiling of wheat gluten using a combination of data-independent and data-dependent acquisition. *Front. Plant Sci.* 7, 2020.

Buse, J., Purves, R. W., Verrall, R. E., Badea, I., Zhang, H., Mulligan, C. C., et al. (2014). The development and assessment of high-throughput mass spectrometry-based methods for the quantification of a nanoparticle drug delivery agent in cellular lysate. *J. Mass Spectrom.* 49 (11), 1171–1180. doi: 10.1002/jms.3444

Carr, S. A., and Anderson, L. (2008). Protein Quantitation Through Targeted Mass Spectrometry: the Way Out of Biomarker Purgatory? *Clin. Chem.* 54 (11), 1749–1752. doi: 10.1373/clinchem.2008.114686

Chu, P. T., Lin, C. S., Chen, W. J., Chen, C. F., and Wen, H. W. (2012). Detection of gliadin in foods using a quartz crystal microbalance biosensor that incorporates gold nanoparticles. *J. Agric. Food Chem.* 60 (26), 6483–6492. doi: 10.1021/jf2047866

Codex Alimentarius Committee (2008). Standard for foods for special dietary use for persons intolerant to gluten. *Codex Stan*, 118-1979.

Colgrave, M. L. (2017). *Proteomics in Food Science - From Farm to Fork*.

Colgrave, M. L., Byrne, K., Blundell, M., and Howitt, C. A. (2016). Identification of barley-specific peptide markers that persist in processed foods and are capable of detecting barley contamination by LC-MS/MS. *J. Proteomics* 147, 169–176. doi: 10.1016/j.jprot.2016.03.045

Colgrave, M. L., Goswami, H., Blundell, M., Howitt, C. A., and Tanner, G. J. (2014). Using mass spectrometry to detect hydrolysed gluten in beer that is responsible for false negatives by ELISA. *J. Chromatography A* 1370, 105–114. doi: 10.1016/j.chroma.2014.10.033

Colgrave, M. L., Goswami, H., Byrne, K., Blundell, M., Howitt, C. A., and Tanner, G. J. (2015). Proteomic profiling of 16 cereal grains and the application of targeted proteomics to detect wheat contamination. *J. Proteome Res.* 14 (6), 2659–2668. doi: 10.1021/acs.jproteome.5b00187

Day, L., Augustin, M. A., Batey, I. L., and Wrigley, C. W. (2006). Wheat-gluten uses and industry needs. *Trends Food Sci. Technol.* 17 (2), 82–90. doi: 10.1016/j.tifs.2005.10.003

Diaz-Amigo, C., and Popping, B. (2013). Accuracy of ELISA detection methods for gluten and reference materials: a realistic assessment. *J. Agric. Food Chem.* 61 (24), 5681–5688. doi: 10.1021/jf3046736

Egertson, J. D., MacLean, B., Johnson, R., Xuan, Y., and MacCoss, M. J. (2015). Multiplexed peptide analysis using data-independent acquisition and Skyline. *Nat. Protoc.* 10 (6), 887–903. doi: 10.1038/nprot.2015.055

El-Aneed, A., Cohen, A., and Banoub, J. (2009). *Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers.* doi: 10.1080/05704920902717872

FAO. (2018). *World Food Situation - FAO Cereal Supply and Demand Brief* [Online]. Available: http://www.fao.org/worldfoodsituation/csdb/en/[Accessed January 28 2019].

FAOSTAT. (2018). "Food and agricultural organization of the United Nations", (ed.) F.c.s.a.d. brief.).

Ferreira, M. S. L., Mangavel, C., Rogniaux, H., Bonicel, J., Samson, M.-F., and Morel, M.-H. (2014). A MALDI-TOF based study of the in-vivo assembly of glutenin polymers of durum wheat. *Food Res. Int.* 63, 89–99. doi: 10.1016/j.foodres.2013.12.025

Fiedler, K. L., McGrath, S. C., Callahan, J. H., and Ross, M. M. (2014). Characterization of grain-specific peptide markers for the detection of gluten by mass spectrometry. *J. Agric. Food Chem.* 62 (25), 5835–5844. doi: 10.1021/jf500997j

Goodman, R. E., Ebisawa, M., Ferreira, F., Sampson, H. A., van Ree, R., Vieths, S., et al. (2016). AllergenOnline: A peer-reviewed, curated allergen database to assess novel food proteins for potential cross-reactivity. *Mol. Nutr. Food Res.* 60 (5), 1183–1198. doi: 10.1002/mnfr.201500769

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., et al. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* 38, W695–W699. doi: 10.1093/nar/gkq313

Graves, P. R., and Haystead, T. A. J. (2002). Molecular biologist's guide to proteomics. *Microbiol. Mol. Biol. Rev.* 66 (1), 39–63. doi: 10.1128/MMBR.66.1.39-63.2002

Halbmayr-Jech, E., Rogers, A., Don, C., and Prinster, M. (2015). Gluten in rice flour and baked rice products by G12 sandwich ELISA: First action 2014.03. *J. AOAC Int.* 98 (1), 103–111. doi: 10.5740/jaoacint.14-197

Hayta, M., and Alpaslan, M. (2001). Effects of processing on biochemical and rheological properties of wheat gluten proteins. *Nahrung/Food* 45 (5), 304. doi: 10.1002/1521-3803(20011001)45:5<304::AID-FOOD304>3.0.CO;2-K

Hernández-Mesa, M., Ropartz, D., García-Campaña, A. M., Rogniaux, H., Dervilly-Pinel, G., and Le Bizec, B. (2019). Ion Mobility Spectrometry in Food Analysis: Principles, Current Applications and Future Trends. *Molecules* 24 (15), 2706. doi: 10.3390/molecules24152706

Herrero, M., Simó, C., García-Cañas, V., Ibáñez, E., and Cifuentes, A. (2012). Foodomics: MS-based strategies in modern food science and nutrition. *Mass Spectrom. Rev.* 31 (1), 49–69. doi: 10.1002/mas.20335

Hopper, J. T., and Robinson, C. V. (2019). "*Proteomics for Biological Discovery*," in *Mass Spectrometry of Intact Protein Complexes*. Wiley Online Library, John Wiley & Sons. doi: 10.1002/9781119081661.ch6

Huebener, S., Tanaka, C. K., Uhde, M., Zone, J. J., Vensel, W. H., Kasarda, D. D., et al. (2014). Specific nongluten proteins of wheat are novel target antigens in celiac disease humoral response. *J. Proteome Res.* 14 (1), 503–511. doi: 10.1021/pr500809b

Jira, W., and Münch, S. (2019). A sensitive HPLC-MS/MS screening method for the simultaneous detection of barley, maize, oats, rice, rye and wheat proteins in meat products. *Food Chem.* 275, 214–223. doi: 10.1016/j.foodchem.2018.09.041

Juhász, A., Haraszi, R., and Maulis, C. (2015). ProPepper: a curated database for identification and analysis of peptide and immune-responsive epitope composition of cereal grain protein families. *Database* 2015, bav100. doi: 10.1093/database/bav100

Larre, C., Lupi, R., Gombaud, G., Brossard, C., Branlard, G., Moneret-Vautrin, D. A., et al. (2011). Assessment of allergenicity of diploid and hexaploid wheat genotypes: identification of allergens in the albumin/globulin fraction. *J. Proteomics* 74 (8), 1279–1289. doi: 10.1016/j.jprot.2011.03.014

Li, H., Byrne, K., Galiamov, R., Mendoza-Porras, O., Bose, U., Howitt, C. A., et al. (2018). Using LC-MS to examine the fermented food products vinegar and soy sauce for the presence of gluten. *Food Chem.* 254, 302–308. doi: 10.1016/j.foodchem.2018.02.023

Liao, Y.-S., Kuo, J.-H., Chen, B.-L., Tsuei, H.-W., Lin, C.-Y., Lin, H.-Y., et al. (2017). Development and validation of the detection method for wheat and barley glutens using mass spectrometry in processed foods. *Food Analyt. Methods* 10 (8), 2839–2847. doi: 10.1007/s12161-017-0827-0

Lovric, J. (2011). *Introducing proteomics: from concepts to sample separation, mass spectrometry and data analysis*. John Wiley & Sons. 296. doi: 10.1002/pmic.201190118

Lupi, R., Masci, S., Rogniaux, H., Tranquet, O., Brossard, C., Lafiandra, D., et al. (2014). Assessment of the allergenicity of soluble fractions from GM and commercial genotypes of wheats. *J. Cereal Sci.* 60 (1), 179–186. doi: 10.1016/j.jcs.2014.02.009

Malalgoda, M., Meinhardt, S. W., and Simsek, S. (2018). Detection and quantitation of immunogenic epitopes related to celiac disease in historical and modern hard red spring wheat cultivars. *Food Chem.* 264, 101–107. doi: 10.1016/j.foodchem.2018.04.131

Manfredi, A., Mattarozzi, M., Giannetto, M., and Careri, M. (2015). Multiplex liquid chromatography-tandem mass spectrometry for the detection of wheat, oat, barley and rye prolamins towards the assessment of gluten-free product safety. *Analyt. Chem. Acta* 895, 62–70. doi: 10.1016/j.aca.2015.09.008

Martens, L., Vandekerckhove, J., and Gevaert, K. (2005). DBToolkit: processing protein databases for peptide-centric proteomics. *Bioinformatics* 21 (17), 3584–3585. doi: 10.1093/bioinformatics/bti588

Martínez-Esteso, M. J., Nørgaard, J., Brohée, M., Haraszi, R., Maquet, A., and O'Connor, G. (2016). Defining the wheat gluten peptide fingerprint via a discovery and targeted proteomics approach. *J. Proteomics*. 147, 156–168. doi: 10.1016/j.jprot.2016.03.015

Martínez-Esteso, M. J., Brohée, M., Nørgaard, J., and O'Connor, G. (2017). Label-Free Proteomic Analysis of Wheat Gluten Proteins and Their Immunoreactivity to ELISA Antibodies. *Cereal Chem.* 94 (5), 820–826. doi: 10.1094/CCHEM-11-16-0266-R

Michaelevski, I., Kirshenbaum, N., and Sharon, M. (2010). T-wave ion mobility-mass spectrometry: basic experimental procedures for protein complex analysis. *J. Visualized Experiments* 41 (41), e1985–e1985. doi: 10.3791/1985

Nadler, W. M., Waidelich, D., Kerner, A., Hanke, S., Berg, R., Trumpp, A., et al. (2017). MALDI versus ESI: The Impact of the Ion Source on Peptide Identification. *J. Proteome Res.* 16 (3), 1207–1215. doi: 10.1021/acs.jproteome.6b00805

Osborne, T. B. (1907). *The proteins of the wheat kernel*. Wasington, DC, US: Carnegie Institution of Washington. doi: 10.5962/bhl.title.26152

Owusu-Apenten, R. (2002). *Food protein analysis: quantitative effects on processing*. Boca Raton, Florida, US: CRC Press. doi: 10.1201/9780203910580

Peterson, A. C., Russell, J. D., Bailey, D. J., Westphall, M. S., and Coon, J. J. (2012). Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol. Cell. Proteomics* 11 (11), 1475–1488. doi: 10.1074/mcp.O112.020131

Pfeifer, M., Kugler, K. G., Sandve, S. R., Zhan, B., Rudi, H., Hvidsten, T. R., et al. (2014). Genome interplay in the grain transcriptome of hexaploid bread wheat. *Science* 345 (6194), 1250091. doi: 10.1126/science.1250091

Pietzak, M. M., and Fasano, A., (2005). "*Reviews in Food and Nutrition Toxicity*," in *Celiac disease: a new paradigm of an immune-mediated disorder due to dietary gluten* (Boca Raton, Florida, US: Citeseer), 243.

Prandi, B., Faccini, A., Tedeschi, T., Cammerata, A., Sgrulletta, D., D'Egidio, M. G., et al. (2014). Qualitative and quantitative determination of peptides related to celiac disease in mixtures derived from different methods of simulated gastrointestinal digestion of wheat products. *Analyt. Bioanalyt. Chem.* 406 (19), 4765–4775. doi: 10.1007/s00216-014-7858-9

Prandi, B., Tedeschi, T., Folloni, S., Galaverna, G., and Sforza, S. (2017). Peptides from gluten digestion: A comparison between old and modern wheat varieties. *Food Res. Int.* 91, 92–102. doi: 10.1016/j.foodres.2016.11.034

Ramírez-González, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., et al. (2018). The transcriptional landscape of polyploid wheat. *Science* 361 (6403), eaar6089. doi: 10.1126/science.aar6089

Ribeiro, M., Rodriguez-Quijano, M., Nunes, F. M., Carrillo, J. M., Branlard, G., and Igrejas, G. (2016). New insights into wheat toxicity: Breeding did not seem to contribute to a prevalence of potential celiac disease's immunostimulatory epitopes. *Food Chem.* 213, 8–18. doi: 10.1016/j.foodchem.2016.06.043

Rogniaux, H., Pavlovic, M., Lupi, R., Lollier, V., Joint, M., Mameri, H., et al. (2015). Allergen relative abundance in several wheat varieties as revealed via a targeted quantitative approach using MS. *Proteomics* 15 (10), 1736–1745. doi: 10.1002/pmic.201400416

Saghatelian, A., and Cravatt, B. F. (2005). Global strategies to integrate the proteome and metabolome. *Curr. Opin. Chem. Biol.* 9 (1), 62–68. doi: 10.1016/j.cbpa.2004.12.004

Salentijn, E. M. J., Esselink, D. G., Goryunova, S. V., van der Meer, I. M., Gilissen, L. J. W. J., and Smulders, M. J. M. (2013). Quantitative and qualitative differences in celiac disease epitopes among durum wheat varieties identified through deep RNA-amplicon sequencing. *BMC Genomics* 14 (1), 905. doi: 10.1186/1471-2164-14-905

Sapone, A., Bai, J. C., Ciacci, C., Dolinsek, J., Green, P. H. R., Hadjivassiliou, M., et al. (2012). Spectrum of gluten-related disorders: consensus on new nomenclature and classification. *BMC Med.* 10 (1), 13. doi: 10.1186/1741-7015-10-13

Schalk, K., Koehler, P., and Scherf, K. A. (2018a). Quantitation of Specific Barley, Rye, and Oat Marker Peptides by Targeted Liquid Chromatography–Mass Spectrometry To Determine Gluten Concentrations. *J. Agric. Food Chem.* 66 (13), 3581–3592. doi: 10.1021/acs.jafc.7b05286

Schalk, K., Koehler, P., and Scherf, K. A. (2018b). Targeted liquid chromatography tandem mass spectrometry to quantitate wheat gluten using well-defined reference proteins. *PloS One* 13 (2), e0192804. doi: 10.1371/journal.pone.0192804

Schalk, K., Lang, C., Wieser, H., Koehler, P., and Scherf, K. A. (2017). Quantitation of the immunodominant 33-mer peptide from α-gliadin in wheat flours by liquid chromatography tandem mass spectrometry. *Sci. Rep.* 7, 45092. doi: 10.1038/srep45092

Scherf, K. A., Koehler, P., and Wieser, H. (2016). Gluten and wheat sensitivities–an overview. *J. Cereal Sci.* 67, 2-11.

Schneider, M., Lane, L., Boutet, E., Lieberherr, D., Tognolli, M., Bougueleret, L., et al. (2009). The UniProtKB/Swiss-Prot knowledgebase and its Plant Proteome Annotation Program. *J. Proteomics* 72 (3), 567–573. doi: 10.1016/j.jprot.2008.11.010

Sealey-Voyksner, J. A., Khosla, C., Voyksner, R. D., and Jorgenson, J. W. (2010). Novel aspects of quantitation of immunogenic wheat gluten peptides by liquid chromatography–mass spectrometry/mass spectrometry. *J. Chromatography A* 1217 (25), 4167–4183. doi: 10.1016/j.chroma.2010.01.067

Shan, L., Molberg, Ø., Parrot, I., Hausch, F., Filiz, F., Gray, G. M., et al. (2002). Structural basis for gluten intolerance in celiac sprue. *Science* 297 (5590), 2275–2279. doi: 10.1126/science.1074129

Shewry, P. R., and Halford, N. G. (2002). Cereal seed storage proteins: structures, properties and role in grain utilization. *J. Exp. Bot.* 53 (370), 947–958. doi: 10.1093/jexbot/53.370.947

Silva, J. C., Gorenstein, M. V., Li, G.-Z., Vissers, J. P. C., and Geromanos, S. J. (2006). Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Mol. Cell. Proteomics* 5 (1), 144–156. doi: 10.1074/mcp.M500230-MCP200

Soler, M., Estevez, M. C., Moreno Mde, L., Cebolla, A., and Lechuga, L. M. (2016). Label-free SPR detection of gluten peptides in urine for non-invasive celiac disease follow-up. *Biosens Bioelectron.* 79, 158–164. doi: 10.1016/j.bios.2015.11.097

Spaenij–Dekking, L., Kooy–Winkelaar, Y., van Veelen, P., Drijfhout, J. W., Jonker, H., van Soest, L., et al. (2005). Natural variation in toxicity of wheat: potential for selection of nontoxic varieties for celiac disease patients. *Gastroenterology* 129 (3), 797–806. doi: 10.1053/j.gastro.2005.06.017

Stöcklin, R., Vu, L., Vadas, L., Cerini, F., Kippen, A. D., Offord, R. E., et al. (1997). A stable isotope dilution assay for the in vivo determination of insulin levels in humans by mass spectrometry. *Diabetes* 46 (1), 44–50. doi: 10.2337/diab.46.1.44

Tanner, G. J., Colgrave, M. L., Blundell, M. J., Goswami, H. P., and Howitt, C. A. (2013). Measuring hordein (gluten) in beer – a comparison of ELISA and mass spectrometry. *PloS One* 8 (2), e56452. doi: 10.1371/journal.pone.0056452

The UniProt Consortium. (2017). UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 45 (D1), D158–D169. doi: 10.1093/nar/gkw1099

Thompson, T., and Méndez, E. (2008). Commercial assays to assess gluten content of gluten-free foods: why they are not created equal. *J. Am. Dietetic Assoc.* 108 (10), 1682–1687. doi: 10.1016/j.jada.2008.07.012

USDA. (2018). *Wheat Data* [Online]. Available: https://www.ers.usda.gov/data-products/wheat-data/ [Accessed January 28 2019].

Uvackova, L., Skultety, L., Bekesova, S., McClain, S., and Hajduch, M. (2013). MSE based multiplex protein analysis quantified important allergenic proteins and detected relevant peptides carrying known epitopes in wheat grain extracts. *J. Proteome Res.* 12 (11), 4862–4869. doi: 10.1021/pr400336f

van den Broeck, H. C., Cordewener, J. H. G., Nessen, M. A., America, A. H. P., and van der Meer, I. M. (2015). Label free targeted detection and quantification of celiac disease immunogenic epitopes by mass spectrometry. *J. Chromatography A* 1391, 60–71. doi: 10.1016/j.chroma.2015.02.070

Vensel, W. H., Dupont, F. M., Sloane, S., and Altenbach, S. B. (2011). Effect of cleavage enzyme, search algorithm and decoy database on mass spectrometric identification of wheat gluten proteins. *Phytochemistry* 72 (10), 1154–1161.

Victorio, V. C. M., Souza, G. H. M. F., Santos, M. C. B., Vega, A. R., Cameron, L. C., and Ferreira, M. S. L. (2018). Differential expression of albumins and globulins of wheat flours of different technological qualities revealed by nanoUPLC-UDMSE. *Food Chem.* 239, 1027–1036. doi: 10.1016/j.foodchem.2017.07.049

Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009). Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25 (9), 1189–1191. doi: 10.1093/bioinformatics/btp033

Weiss, W., Huber, G., Engel, K. H., Pethran, A., Dunn, M. J., Gooley, A. A., et al. (1997). Identification and characterization of wheat grain albumin/globulin allergens. *Electrophoresis* 18 (5), 826–833. doi: 10.1002/elps.1150180529

White, S. P., Frisbie, C. D., and Dorfman, K. D. (2018). Detection and Sourcing of Gluten in Grain with Multiple Floating-Gate Transistor Biosensors. *ACS Sensors* 3 (2), 395–402. doi: 10.1021/acssensors.7b00810

Wieser, H. (2007). Chemistry of gluten proteins. *Food Microbiol.* 24 (2), 115–119. doi: 10.1016/j.fm.2006.07.004

Wieser, H., and Koehler, P. (2009). Is the calculation of the gluten content by multiplying the prolamin content by a factor of 2 valid? *Eur. Food Res. Technol.* 229 (1), 9–13. doi: 10.1007/s00217-009-1020-5

Zhang, W., Ciclitira, P., and Messing, J. (2014). PacBio sequencing of gene families—A case study with wheat gluten genes. *Gene* 533 (2), 541–546. doi: 10.1016/j.gene.2013.10.009

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 13 December 2019
doi: 10.3389/fpls.2019.01530

Characterization and Relative Quantitation of Wheat, Rye, and Barley Gluten Protein Types by Liquid Chromatography–Tandem Mass Spectrometry

*Barbara Lexhaller[1], Michelle L. Colgrave[2,3] and Katharina A. Scherf[1,4]**

[1] Leibniz-Institute for Food Systems Biology at the Technical University of Munich, Freising, Germany, [2] CSIRO Agriculture and Food, St Lucia, QLD, Australia, [3] School of Science, Edith Cowan University, Joondalup, WA, Australia, [4] Department of Bioactive and Functional Food Chemistry, Institute of Applied Biosciences, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

OPEN ACCESS

Edited by:
Nicolas L. Taylor, University of Western Australia, Australia

Reviewed by:
Robert Winkler, Center for Research and Advanced Studies (CINVESTAV), Mexico
Barbara Prandi, University of Parma, Italy

***Correspondence:**
Katharina A. Scherf
k.scherf.leibniz-lsb@tum.de

Specialty section:
This article was submitted to Plant Proteomics, a section of the journal Frontiers in Plant Science

Received: *02 August 2019*
Accepted: *01 November 2019*
Published: *13 December 2019*

Citation:
Lexhaller B, Colgrave ML and Scherf KA (2019) Characterization and Relative Quantitation of Wheat, Rye, and Barley Gluten Protein Types by Liquid Chromatography–Tandem Mass Spectrometry. Front. Plant Sci. 10:1530. doi: 10.3389/fpls.2019.01530

The consumption of wheat, rye, and barley may cause adverse reactions to wheat such as celiac disease, non-celiac gluten/wheat sensitivity, or wheat allergy. The storage proteins (gluten) are known as major triggers, but also other functional protein groups such as α-amylase/trypsin-inhibitors or enzymes are possibly harmful for people suffering of adverse reactions to wheat. Gluten is widely used as a collective term for the complex protein mixture of wheat, rye or barley and can be subdivided into the following gluten protein types (GPTs): α-gliadins, γ-gliadins, ω5-gliadins, ω1,2-gliadins, high- and low-molecular-weight glutenin subunits of wheat, ω-secalins, high-molecular-weight secalins, γ-75k-secalins and γ-40k-secalins of rye, and C-hordeins, γ-hordeins, B-hordeins, and D-hordeins of barley. GPTs isolated from the flours are useful as reference materials for clinical studies, diagnostics or in food analyses and to elucidate disease mechanisms. A combined strategy of protein separation according to solubility followed by preparative reversed-phase high-performance liquid chromatography was employed to purify the GPTs according to hydrophobicity. Due to the heterogeneity of gluten proteins and their partly polymeric nature, it is a challenge to obtain highly purified GPTs with only one protein group. Therefore, it is essential to characterize and identify the proteins and their proportions in each GPT. In this study, the complexity of gluten from wheat, rye, and barley was demonstrated by identification of the individual proteins employing an undirected proteomics strategy involving liquid chromatography–tandem mass spectrometry of tryptic and chymotryptic hydrolysates of the GPTs. Different protein groups were obtained and the relative composition of the GPTs was revealed. Multiple reaction monitoring liquid chromatography–tandem mass spectrometry was used for the relative quantitation of the most abundant gluten proteins. These analyses also allowed the identification of known wheat allergens and celiac disease-active peptides. Combined with functional assays,

these findings may shed light on the mechanisms of gluten/wheat-related disorders and may be useful to characterize reference materials for analytical or diagnostic assays more precisely.

Keywords: allergy, amylase/trypsin-inhibitor, celiac disease, gliadin, gluten, mass spectrometry, non-celiac gluten sensitivity, proteomics

INTRODUCTION

Cereals including wheat, rice, and maize are the most important staple foods for mankind worldwide. However, the consumption of wheat and the closely related cereals rye and barley may cause adverse reactions to wheat such as celiac disease (CD), non-celiac gluten sensitivity (NCGS), or wheat allergy (Sapone et al., 2012; Ludvigsson et al., 2013; Catassi et al., 2017, for review). The triggers are mainly the storage proteins (gluten), but non-gluten proteins like α-amylase/trypsin-inhibitors (ATIs), lipid transfer proteins, puroindolines, or β-amylases are also immunoreactive (Tatham and Shewry, 2008; Scherf, 2019, for review). Gluten is widely used as a collective term for the complex protein mixture of wheat, rye, or barley, which is not soluble in water or salt solution (Codex Alimentarius Commission, 2015). Traditionally, cereal proteins are classified into the so-called Osborne fractions that can be obtained with salt solution (albumins/globulins), 60% aqueous ethanol (prolamins), and a reducing solution of 50% propanol and Tris-hydrochloride buffer (Tris-HCl) (glutelins).

Albumins/globulins are mainly protective or metabolic proteins whereas prolamins and glutelins constitute the storage proteins called gluten. Gluten is composed of gliadins (prolamins) and glutenins (glutelins) in wheat, secalins in rye and hordeins in barley (Scherf et al., 2016). Each gluten fraction can be further subdivided into the respective gluten protein types (GPTs) by preparative reversed-phase high-performance liquid chromatography (RP-HPLC) according to their characteristic retention times. The GPTs of wheat prolamins are α-gliadins, γ-gliadins, ω1,2-gliadins, and ω5-gliadins, and wheat glutelins are divided into high- (HMW-GS) and low-molecular-weight glutenin subunits (LMW-GS). The GPTs of rye are called ω-secalins, HMW-secalins, γ-75k-secalins, and γ-40k-secalins and the barley GPTs are B-hordeins, C-hordeins, D-hordeins, and γ-hordeins (Scherf et al., 2016). These GPTs can be classified into three different groups according to their homologous amino acid sequences and similar molecular weights: LMW group, medium-molecular-weight group and HMW group (**Table 1**). Each GPT contains numerous different proteins, which differ partly only by exchange, deletion or insertion of single amino acids in their sequences. Proteins of the HMW group occur in the glutelin fraction as polymers linked by interchain disulfide bonds. Previous studies revealed similar molecular weights (70–90 kDa) and homologous amino acid sequences of D-hordeins, HMW-secalins and HMW-GS (Field et al., 1982; Shewry et al., 1988; Gellrich et al., 2003). The amino acid sequences contain repetitive units such as QQPGQG, YYPTSP, or QQP and QPG. Differences between the proteins result from modifications of single amino acids or the arrangement and number of the repetitive units. The medium-molecular-weight group proteins mainly occur as monomers in the prolamin fraction and have molecular weights around 40–50 kDa, with the exception of ω5-gliadins (60–68 kDa) that are unique for wheat. The typical repetitive unit for ω5-gliadins is QQQPF, and QPQQPFP is characteristic for ω1,2-gliadins, ω-secalins and C-hordeins. The LMW group consists of monomeric (α-gliadins, γ-gliadins, γ-40k-secalins, and γ-hordeins) and polymeric proteins (LMW-GS, γ-75k-secalins, and B-hordeins). Their molecular weights range from 28 to 35 kDa, except for γ-75k-secalins with a molecular weight around 50 kDa. The proteins of the LMW group comprise unique repetitive units such as QPQPFPPQQPY (α-gliadins), QQPQQPFP (γ-gliadins, γ-75k-secalins, and B-hordeins), and QQPPFS (LMW-GS).

These characteristic features of the GPTs are known to contribute to the CD-immunoreactivity of wheat, rye, and barley, because most CD-active peptides are derived from these repetitive units. For example, the T-cell epitopes QG**YYPTSP**Q (DQ8.5-glut-H1), **QQPQQPFP**Q (DQ2.5-glia-γ4c), or **QQPQQPFP**Q (DQ8-glia-γ1a) contain typical repetitive units highlighted in bold (Sollid et al., 2012). Beside CD, a wide range of wheat, rye, and barley proteins are potential allergens or triggers of innate immunity in NCGS. The recently published reference sequence RefSeq v1.0 of the hexaploid common wheat genome (International Wheat Genome Sequencing Consortium (IWGSC), 2018) provides further insights as the first reference to which known immunoreactive gluten and non-gluten proteins can be annotated (Juhasz et al., 2018).

Numerous studies have demonstrated the complexity of gluten as a mixture of closely related, but distinct proteins (Arentz-Hansen et al., 2000; Dupont et al., 2011; Colgrave et al., 2013; Schalk et al., 2017). Their similarity poses major difficulties in clearly separating gluten into well-defined gluten protein fractions, GPTs and especially individual gluten

TABLE 1 | Gluten protein types and their classification according to molecular weight (Scherf et al., 2016).

Group	Wheat	Rye	Barley
HMW	HMW-GS	HMW-secalins	D-hordeins
MMW	ω1,2-gliadins	ω-secalins	C-hordeins
	ω5-gliadins	-	-
LMW	LMW-GS	γ-75k-secalins	B-hordeins
	γ-gliadins	γ-40k-secalins	γ-hordeins
	α-gliadins	-	-

GS, glutenin subunits; HMW, high-molecular-weight; MMW, medium-molecular-weight; LMW, low-molecular-weight.

proteins (Mamone et al., 2009; Ellis et al., 2011; Lagrain et al., 2013). One strategy is to combine separation according to solubility (Osborne fractionation) with subsequent fractionation according to polarity by preparative RP-HPLC. However, the ultraviolet signal at a specific retention time during preparative RP-HPLC does not provide any further information on the identity of the proteins being collected. Considering the highly variable immunoreactivities of wheat, rye and barley proteins it is essential to know the exact composition of the GPT isolates, especially when trying to gain further insights into pathogenic cascades of CD, NCGS, and wheat allergies (Vader et al., 2002; Matsuo et al., 2005; Scherf et al., 2019). For example, wheat ATIs were only identified as triggers of innate immunity *via* the toll-like receptor 4 in NCGS, because they were co-purified within the ω-gliadin fraction (Junker et al., 2012). Therefore, it is crucial to identify the individual proteins within each GPT isolate and undertake relative quantitation of the highly abundant proteins by liquid chromatography–mass spectrometry (LC-MS/MS).

In the current fundamental study, LC-MS/MS analysis was applied to all isolated GPTs of wheat, rye, and barley to precisely determine the identities of the proteins in each isolate as well as their relative abundances to provide a detailed assessment of the molecular composition. A special focus was placed on the identification of known CD-immunoreactive and allergenic peptides and proteins.

MATERIAL AND METHODS

Material

All chemicals and solvents were at least HPLC or LC-MS grade. Formic acid (FA), ammonium bicarbonate (Ambic), dithiothreitol (DTT), and iodoacetamide (IAM), were purchased from Sigma-Aldrich (Sydney, NSW, Australia). Trypsin (sequencing grade, V511A; specific activity: 15,282 units/mg) and chymotrypsin (sequencing grade, V106A; specific activity: at least 70 units/mg by N-benzoyl-L-tyrosine ethyl ester assay) were purchased from Promega (Sydney, NSW, Australia).

Grain Samples

Grains of wheat [cultivar (cv.) Akteur, harvest year 2011, I.G. Pflanzenzucht, Munich, Germany], rye (cv. Visello, harvest year 2013, KWS Lochow, Bergen, Germany), and barley (cv. Marthe, harvest year 2009, Nordsaat Saatzucht, Langenstein, Germany) grown in Germany were milled into white flour using a Quadrumat Junior mill (Brabender, Duisburg, Germany). Subsequently, the flours were sieved to a particle size of 200 µm and allowed to rest for 2 weeks. The choice of these cultivars was based on production shares in Germany for conventional farming to ensure that these cultivars were of economic relevance and, therefore, deemed to be representative for each grain.

Analysis of Moisture and Crude Protein Contents

The determination of moisture and crude protein (CP) contents (conversion factor N × 5.7) was carried out according to International Association for Cereal Science and Technology Standards 110/1 and 167.

Preparation of Gluten Protein Types

The α-gliadins, γ-gliadins, ω1,2-gliadins, ω5-gliadins, HMW-GS and LMW-GS of wheat, ω-secalins, HMW-secalins, γ-75k-secalins, and γ-40k-secalins of rye, and B-hordeins, C-hordeins, D-hordeins, and γ-hordeins were isolated by modified Osborne fractionation and preparative RP-HPLC (Schalk et al., 2017) from the flours after a maximum of 6 weeks storage after milling in the respective year. The flours of wheat, rye, and barley (4 × 50 g) were extracted step-wise three times each with 200 ml salt solution (0.4 mol/l NaCl with 0.067 mol/l Na_2HPO_4/KH_2PO_4, pH 7.6) for 10 min at 22°C, centrifuged and the supernatant containing albumins/globulins was discarded. The sediments were extracted with ethanol/water (60/40, v/v) (3 × 200 ml) for 10 min at 22°C to obtain the prolamin fractions. For the glutelins, the resulting sediments were extracted three times each with 200 ml 2-propanol/water (50/50, v/v)/0.1 mol/l Tris-HCl, pH 7.5, containing 2 mol/l (w/v) urea and 0.06 mol/l (w/v) DTT for 30 min at 60°C under nitrogen. The supernatants of each prolamin and glutelin fraction were combined, concentrated, lyophilized and stored at -20°C until use. This whole extraction procedure was performed on four independent batches to give enough material for further analyses.

For preparative RP-HPLC, the wheat, rye, and barley prolamin fractions (200 mg) were dissolved in 10 ml ethanol/water and the glutelin fractions (1,000 mg) in 10 ml of the glutelin extraction solution. The solutions were filtered (0.45 µm) and separated on a Jasco HPLC (Jasco, Gross-Umstadt, Germany) according to their retention times, collected from several runs, pooled and lyophilized as described previously (Schalk et al., 2017). The isolated GPTs were again stored at -20°C until use. Long-term experience with storage of the Prolamin Working Group-gliadin reference material (Van Eckert et al., 2006) in our laboratory since its isolation in the early 2000s indicates that protein isolates are stable for several years or even decades when kept frozen at -20°C or, ideally, at -80°C.

Enzymatic Cleavage of GPTs

The GPT hydrolysates were prepared as reported in Colgrave et al. (2016a; 2016b). Briefly, each GPT (n = 3) was dissolved in 50 mmol/l Ambic buffer with a concentration of 2 mg/ml and applied to a 10 kDa molecular weight cutoff filter (Millipore, Australia). The GPT solutions were washed with washing solution (2 × 100 µl; 8 mol/l urea; 100 mmol/l Tris-HCl; pH 8.5) and the filters were centrifuged. For reduction, DTT solution (10 mmol/l) was added; the filters were incubated for 40 min at room temperature and then centrifuged. For cysteine alkylation, 100 µl of IAM solution (25 mmol/l; in 8 mol/l urea; 100 mmol/l Tris-HCl) was added and the solution was incubated at room temperature in the dark for 20 min. The filters were centrifuged and washing solution was added (2 × 100 µl). To exchange the buffer, two times 200 µl of Ambic buffer was added and centrifuged. The 10 kDa filters were transferred to fresh centrifuge tubes, the digestion enzyme (trypsin or chymotrypsin:

200 µl; 250 µg/ml in 50 mmol/l Ambic; 1 mmol/l $CaCl_2$; enzyme/substrate ratio of 1/4 (w/w); respectively) was added, and the mixture was incubated overnight at 37°C. The filtrates with the enzymatically cleaved peptides were collected by centrifugation, the filters were washed again with 200 µL of Ambic, and the filtrates and the washing solution were combined separately for each replicate and lyophilized. For LC-MS/MS analysis the peptides were resuspended in 100 µl 1% FA.

Undirected LC-MS/MS Analysis

Aliquots (5 µl) of each GPT replicate were pooled for analysis. The LC-MS/MS analysis was performed on an Ekspert nanoLC415 (Eksigent, Dublin, CA, United States) directly coupled to a TripleTOF 6600 MS (SCIEX, Redwood City, CA, United States) with the following parameters: Trap column: ChromXP C18 (3 µm, 12 nm, 10 × 0.3 mm); flow rate: 10 µl/min solvent A; 5 min; column: ChromXP C18 (3 µm, 12 nm, 150 mm × 0.3 mm); flow rate: 5 µl/min; solvents: (A) 5% DMSO, 0.1% FA, 94.9% water; (B) 5% DMSO, 0.1% FA, 90% acetonitrile, 4.9% water; linear gradient from 3 to 25% solvent B over 68 min, followed by a second linear step from 25–35% solvent B over 5 min, followed by a third linear step from 35–80% B over 2 min; a 3 min hold at 80% B; return to 3% B over 1 min; 8 min of re-equilibration; injection volume: 2 µl. DMSO was added as it enhances ionization and increases the signal-to-noise ratio (Hahne et al., 2013). The eluent from the HPLC was directly coupled to the DuoSpray source of the TripleTOF 6600 MS. The MS settings were as follows: Ion spray voltage: 5,500 V; curtain gas: 138 kPa (20 psi); ion source gas 1 and 2 (GS1 and GS2): 103 and 138 kPa (15 and 20 psi); heated interface temperature: 100°C. The MS was operated in the information-dependent acquisition (IDA) mode. The IDA method consisted of a high-resolution time-of-flight-MS survey scan followed by 30 MS/MS scans, each with an accumulation time of 40 ms. The mass-to-charge (*m/z*) range of the acquisition of the MS1 spectra in positive ion mode was 400–1,250 with a 0.25 s accumulation time. MS2 spectra were acquired on precursor ions that exceeded 150 counts/s with charge states 2+ to 5+ and over the mass range of *m/z* 100–1,500 using the manufacturer's rolling collision energy based on the size and charge of the precursor ion and a collision energy spread of 5 V for optimum peptide fragmentation. Analysis was carried out with dynamic ion exclusion of precursor ions with a 15 s interval after one occurrence and a mass tolerance of 100 ppm, and peaks within 6 Da of the precursor mass were excluded.

Data Analysis for Protein Identification

For protein identification, the SCIEX.wiff raw files were directly used as input in the ProteinPilot 5.0 software (SCIEX) with the Paragon algorithm (Shilov et al., 2007). The raw data were searched against a database comprising UniProtKB-*Poaceae* proteins (https://www.uniprot.org; version 2018/02) appended with cRAP (http://www.thegpm.org/crap/), the common repository of adventitious proteins (1,601,923 sequences). The settings used were: IAM as the alkylating agent; trypsin, chymotrypsin, or no enzyme as the cleavage enzyme. ProteinPilot automatically considers enzyme cleavage specificity rules and all UniMod modifications, including e.g., oxidation of methionine and deamidation of asparagine and glutamine, and uses a probability-based approach that considers sample treatment conditions. A 1% global false discovery rate (FDR) was applied for the protein identifications. The detected proteins were classified according to Dupont et al. (2011) into the following groups: gluten proteins, ATIs, globulins, β-amylase, other enzymes, farinins, serpins, grain softness proteins and puroindolines (GSPs+PINs), avenin-like proteins, other inhibitors, uncharacterized proteins (name of entries in the database UniProtKB) and others. The group "others" contains all identified proteins, which could not be assigned to any of the aforementioned groups. All proteins identified as "uncharacterized" and "predicted" were manually reviewed using the basic local alignment search tool (BLAST) (Altschul et al., 1990) on the UniProtKB webpage with the target database UniProtKB reference proteomes plus SwissProt (parameters: identity >70%, except for hits with names of a group or from the subfamily *Pooideae*). Due to the challenge of having different terms and often uncurated and incomplete protein sequences in the UniProtKB *Poaceae* database, the protein names for gluten proteins were summarized in the group "gluten proteins", which comprise gliadins, glutelins, glutenins and prolamins for wheat, secalins, glutelins, glutenins and prolamins for rye and hordeins, glutelins, glutenins and prolamins for barley. By means of the rank for the specified protein given by the Paragon algorithm in ProteinPilot, the detected proteins are sorted relative to all other ones. The proportion in each different group was calculated as the number of identified proteins per group multiplied by the number of distinct peptides with a >95% confidence level by which these proteins were identified to have a weighting factor for the rank of the specific protein relative to all other proteins

Preparation of the Multiple Reaction Monitoring Methods Using Skyline

Within each GPT, the identified proteins were selected according to the following parameters: belonging to the family *Poaceae*, the subfamily *Pooideae* and to gluten; 1% global FDR; confidence score > 99% and unused score > 2.0. The manually curated FASTA files list and the results of the undirected LC-MS/MS experiments were imported into Skyline (version 4.2.0.19072). Multiple reaction monitoring (MRM) transitions were determined for each peptide predicted with precursor ion (Q1) with *m/z* (50–1,500) and charge (2+; 3+) and fragment ion (Q3) *m/z* values using the data collected in the undirected LC-MS/MS experiments (Colgrave et al., 2012). Up to six transitions were used in the preliminary analyses and the MRM transitions were refined and the top four MRM transitions were selected per peptide for use in the final method. In the subsequent experiments scheduled MRM transitions were used for analysis in triplicate.

Multiple Reaction Monitoring Mass Spectrometry for Relative Protein Quantitation

Scheduled MRM experiments were used for quantitation of the reduced and alkylated tryptic and chymotryptic peptides of each GPT in triplicate, respectively. The LC-MS/MS analysis was

performed on an UHPLC system (Shimadzu Nexera, Sydney, Australia) directly coupled to a QTRAP 6500 mass spectrometer (SCIEX). The cycle time was set to 0.3 s, and the MRM transitions were scheduled to be monitored within 60 s of their expected retention time (± 30 s) (Colgrave et al., 2017a).

Relative Protein Quantitation

The peaks were integrated using Skyline. The relative quantitation of the proteins within each GPT was performed by using the "best flyer methodology" (Ludwig et al., 2012), in which the peak areas of four transitions of one peptide (average of three replicates) were summarized. One peptide is used to represent one protein and the values of the peak area of each peptide were assigned to the respective protein. The datasets from the tryptic and chymotryptic digests were combined by removing the duplicate protein with the lower value. Then, the areas of all proteins from the same category according to their UniProtKB accession were summarized. The calculations were done in Microsoft Excel and the graphical images were done in Origin (version 2018b (9.55), OriginLab Northampton, MA, USA).

RESULTS

General Characterization of Gluten Protein Types

The moisture contents of the flours were 14.59 ± 0.01% for wheat, 11.42 ± 0.01% for rye and 12.09 ± 0.06% for barley. The contents of CP, albumin/globulin, prolamin, and glutenin fractions in the flours are given in **Table S1**. **Table S2** lists the CP contents of the GPTs isolated from wheat, rye and barley flours and the proportions of each GPT within total gluten. The Osborne fraction values are based on flour weight; the proportions of GPTs are based on total gluten content (Lexhaller et al., 2016; Lexhaller et al., 2017). The results corresponded well to those reported previously (Gellrich et al., 2003; Kerpes et al., 2016; Schalk et al., 2017)

Identification of Protein Groups in the Gluten Protein Types

The Osborne fractions (prolamins and glutelins) extracted from the flours were separated into the GPTs by preparative RP-HPLC. These purified GPTs were reduced, alkylated and subjected to tryptic (T) and chymotryptic (C) hydrolysis, respectively. The GPT hydrolysates were analyzed by LC-MS/MS to identify the complete suite of proteins present in each GPT. Proteins with identical sequences were used once. For each GPT, the suite of proteins identified after tryptic digest (**Table S3**) and after chymotryptic digest (**Table S4**) were recorded. All proteins originally identified as "uncharacterized" or "predicted" were manually searched again using the BLAST tool available from the UniProtKB webpage. According to the data of the undirected LC-MS/MS experiments, **Figure 1** shows the qualitative composition and proportion of the proteins in each GPT.

Wheat

A similar composition with mainly gluten proteins (87% and 85%, respectively) and 6–7% ATIs was detected in the α- and γ-gliadin-GPTs. The ω5-gliadin-GPT was composed of 77% gluten proteins and 14% ATIs, whereas the ω1,2-gliadin-GPT contained about 58% gluten proteins, 26% ATIs and 6% GSPs+PINs. HMW- and LMW-GS-GPTs showed a comparable composition with about 78% or 81% gluten proteins, respectively (**Figure 1A**).

Rye

The ω-secalin-GPT consisted of 79% gluten proteins, 10% ATIs, and 6% GSPs+PINs. In the HMW-secalin-GPT, 4% farinins, 3% other enzymes, and 3% globulins were identified besides 76% gluten proteins. The γ-75k-secalin-GPT was composed of 58% gluten proteins, 5% ATIs and more than 10% other enzymes. The composition of the γ-40k-secalin-GPT included only 23% gluten proteins, 23% other enzymes and about 23% others. It should be noted that 21% of the identified proteins were uncharacterized ones (**Figure 1B**).

Barley

The C-hordein-GPT consisted mainly of 62% gluten proteins, 10% ATIs and 7% GSPs+PINs. The γ-hordein-GPT was composed of over 92% gluten proteins and 4% ATIs and the residual groups amounted only to 4% altogether. The compositions of B- and D-hordein-GPTs were similar, but the B-hordein-GPT had a greater diversity of enzymes (15% in total) and contained 11% uncharacterized proteins. In the D-hordein-GPT (**Figure 1C**) high proportions of other proteins (24%) were present.

Identification of Single Proteins in the Gluten Protein Types

Tables S3 and **S4** list all identified proteins with their UniProtKB accession number, name, organism, rank, score, sequence coverage and number of identified peptides. As an overview of the qualitative data, the three proteins with the highest ranks identified in the tryptic (**Table 2**) and in the chymotryptic (**Table 3**) hydrolysates, respectively, of each GPT according to the rank are summarized. The rank of each specified protein is relative to all identified proteins in the fraction and contaminant proteins, such as the proteases used and/or keratins from sample preparation were excluded.

Wheat

The high-scoring proteins detected in the tryptic hydrolysates of the α-gliadin-GPT and the γ-gliadin-GPT represented gluten proteins, except one α-amylase-inhibitor (**Table 2**). The top-ranked proteins often did not match those of the corresponding protein type, whereas the matching proteins appeared at lower ranks, e.g., γ-gliadins (D0ES80; H8Y0P9) at ranks five and seven in the γ-gliadin-GPT with similar scores and peptide numbers. The chymotryptic hydrolysates (**Table 3**) showed similar compositions. The tryptic hydrolysate of the ω5-gliadin-GPT contained mainly HMW-GS proteins, but an ω-gliadin (A0A0B5J8A9) was identified based on eight peptides at rank 12. Surprisingly, no ω-gliadin was identified in the chymotryptic hydrolysate of the ω5-gliadin-GPT. The tryptic hydrolysate of the ω1,2-gliadin-GPT was composed of different types of proteins

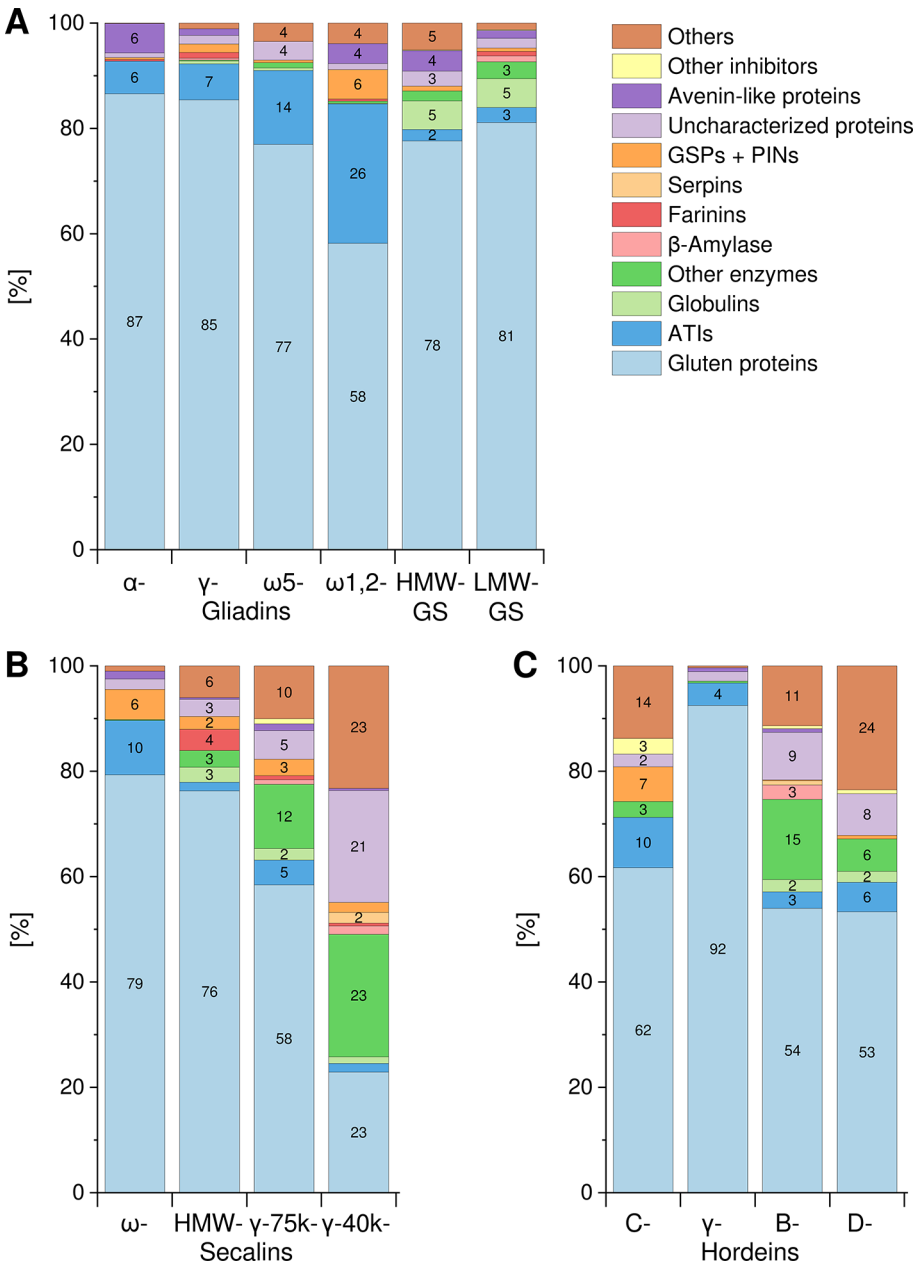


FIGURE 1 | Composition and proportions of proteins in each GPT. Classification of identified proteins into the following groups for wheat **(A)**, rye **(B)**, and barley **(C)** gluten protein types: gluten proteins, α-amylase/trypsin-inhibitors (ATIs), globulins, other enzymes, β-amylase, farinins, serpins, grain softness proteins, and puroindolines (GSPs+PINs), uncharacterized proteins, avenin-like proteins, other inhibitors, and others. When a group is missing in individual GPT, no proteins were identified. Groups without number represent less than 2%. GS, glutenin subunits; HMW, high-molecular-weight; LMW, low-molecular-weight.

representing the two main groups of this GPT (**Figure 1A**). The chymotryptic hydrolysate contained an ω-gliadin protein (A0A060N0S6) at rank 1 with by far the highest score and the most identified peptides (89). In the tryptic and chymotryptic hydrolysates of the HMW-GS-GPT the highest ranked proteins were HMW-GS. The high-scoring proteins in the tryptic LMW-GS-GPT were the 12S seed storage globulin (M7ZK46), which belongs to the cupin super-family with nutrient reservoir activity (Dunwell, 1998) and one LMW-GS, which was identified with the highest number of peptides. These proteins represent the main group, gluten proteins, and the second main group in this GPT, the globulins (**Figure 1A**). Globulins are known to polymerize *via* interchain disulfide bonds and may thus appear in the high-molecular-weight group (Vensel et al., 2014).

Rye

The three proteins with the highest scores in the tryptic ω-secalin-GPT hydrolysate (**Table 2**) were an ω-secalin, a trypsin inhibitor

TABLE 2 | High-scoring proteins (top 3) identified in each gluten protein type (GPT) after tryptic cleavage.

GPT	Rank[a]	UniProtKB accession	UniProtKB name	Score[b]	Peptides
α-gliadins	1	R9XUM8	Alpha-gliadin	30.03	96
	2	B2Y2Q4	Low molecular weight GS	17.12	55
	3	P17314	Alpha amylase inhibitor CM3	13.52	18
γ-gliadins	1	Q41553	HMW-GS Ax2	27.66	31
	3	I3XHQ1	LMW-9	24.08	38
	4	W6AX70	HMW-GS	20.79	25
ω5-gliadins	1	Q41553	HMW-GS Ax2	52.89	56
	4	V9TRL3	HMW-GS 1Dy	21.91	24
	6	P10388	HMW-GS Dx5	16.73	17
ω1,2-gliadins	5	G9I1R7	Alpha-gliadin Gli-M2	24.59	29
	7	C8CAI4	Dimeric alpha-amylase inhibitor	24.10	28
	8	B9VRI3	Alpha-amylase inhibitor CM16	20.01	28
HMW-GS	1	W6AX70	HMW subunit	71.28	113
	2	A0A060MZP1	HMW subunit	54.59	123
	3	P10388	HMW subunit	41.05	97
LMW-GS	1	M7ZK46	Seed storage globulin 1	26.70	32
	2	A0A060MZP1	HMW glutenin subunit	18.51	31
	4	D6RVY4	LMW glutenin subunit	16.56	73
ω-secalins	2	A0A159KI56	Omega-secalin	23.75	79
	5	Q7M220	Trypsin inhibitor	16.31	22
	6	W6AW98	HMW-GS x	14.17	19
HMW-secalins	1	W6AW92	HMW-GS y	39.73	221
	2	Q93WF0	HMW-GS x	30.31	109
	4	W8NKZ9	B-type farinin protein	10.70	27
γ-75k-secalins	1	E5KZQ2	75k gamma secalin	53.31	165
	2	B9A8E2	Protein disulfide-isomerase	11.52	30
	3	Q9ZSR6	Heat shock protein HSP26[c]	2.60	12
γ-40k-secalins	1	A0A1D5U769	Sucrose synthase	28.53	25
	2	M8ASF1	Actin-2[d]	27.76	23
	4	H8Y0K4	Gamma prolamin	24.32	83
C-hordeins	2	Q84LE9	D-hordein	17.38	32
	3	Q5IUH1	Hordoindoline-B 1	10.09	6
	4	Q41518	RNA-binding protein[e]	8.14	7
γ-hordeins	2	I6TMW4	B3-hordein	33.90	60
	3	P06470	B1-hordein	16.30	64
	4	P80198	Gamma-hordein-3	13.20	16
B-hordeins	1	F2D284	Protein disulfide-isomerase	31.05	22
	2	M7ZK46	12S seed storage globulin 1[f]	22.68	17
	3	I6TMW4	B3-hordein	21.90	102
D-hordeins	1	I6TRS8	D-hordein	35.20	209
	3	Q41350	Osmotin-like protein[g]	13.57	8
	4	Q41518	RNA-binding protein[h]	11.63	10

[a]The rank of the specified protein is relative to all other proteins in the list of detected proteins, [b]Unused ProtScore, defined as a measure of the protein confidence for a detected protein, calculated from the peptide confidence for peptides from spectra that are not already completely "used" by higher scoring winning proteins, thus reflecting the amount of total, unique peptide evidence related to a given protein, [c]after BLAST search (identified as uncharacterized protein: R7W8L3), [d]after BLAST search (identified as uncharacterized protein: W5AHI2), [e]after BLAST search (identified as predicted protein: F2CR90), [f]after BLAST search (identified as predicted protein: F2E9N0), [g]after BLAST search (identified as predicted protein: F2DZW3), [h]after BLAST search (identified as predicted protein: F2CR90).

and a HMW-GS, which represent the two main groups of the ω-secalin-GPT in **Figure 1B**. Only two proteins passing the 1% FDR threshold were identified in the chymotryptic hydrolysate of the ω-secalin-GPT (**Table 3**). In the tryptic and chymotryptic hydrolysates of the HMW-secalin-GPT, the highest ranked proteins were a HMW-secalin (Q93WF0; rank 2) and a wheat HMW-GS protein (W6AW92; rank 1), which is, however, very similar to the HMW-secalin protein D3XQB8 (95.8% identity). The tryptic hydrolysate of the γ-75k-secalin-GPT consisted mainly of the 75k gamma secalin protein E5KZQ2. The high scoring proteins represent the three main groups in the γ-75k-secalin-GPT (**Figure 1B**). Another 75k γ-secalin protein (E5KZQ6) was also identified with a high number of peptides, but a lower score. In the chymotryptic hydrolysate, the protein identified with the most peptides (49) was the 75k γ-secalin E5KZQ1 at rank 3. In case of the γ-40k-secalin-GPT, only one γ-prolamin protein was identified in the tryptic hydrolysate at rank 3. A sucrose synthase and an uncharacterized protein (W5AHI2) ranked first and second, respectively. The BLAST search identified an actin-2 protein (M8ASF1) with 100% identity to this uncharacterized protein. Uncharacterized proteins represented one of the largest groups in the γ-40k-secalin-GPT (**Figure 1B**), probably due to missing reference protein sequences. The chymotryptic hydrolysate showed a similar proportion with a formate dehydrogenase and two uncharacterized proteins as the three high-scoring proteins.

TABLE 3 | High-scoring proteins (top 3) identified in each gluten protein type (GPT) after chymotryptic cleavage.

GPT	Rank[a]	UniProtKB accession	UniProtKB name	Score[b]	Peptides
α-gliadins	1	J7I026	Alpha-gliadin	19.29	34
	2	A0A0U2P410	Low molecular weight GS	14.37	19
	3	I3XHQ1	Low molecular weight GS	8.51	7
γ-gliadins	2	Q9XGF0	Low molecular weight GS	4.97	4
	3	B6UKM7	Gamma gliadin	2.97	1
	4	P94021	LMM glutenin 2 (Fragment)	2.92	1
ω5-gliadins	1	D6RVY4	LMW-GS (Fragment)	6.66	10
	2	P10387	HMW-GS Dy10	6.45	9
	4	Q41553	HMW-GS Ax2	2.49	4
ω1,2-gliadins	1	A0A060N0S6	Omega-gliadin	17.11	89
	2	P10388	HMW-GS Dx5	4.84	6
	3	P10385	Low molecular glutenin subunit[c]	4.53	3
HMW-GS	1	C0SUC3	HMW glutenin subunit x5	27.39	39
	2	P10387	Glutenin, HMW subunit Dy10	13.58	44
	3	Q03872	HMW subunit 1Ax1	12.04	30
LMW-GS	1	D6RVY4	Low molecular glutenin subunit	13.32	41
	2	I3XHQ1	LMW glutenin subunit LMW-9	12.30	19
	3	A0A0S2GJT4	LMW glutenin subunit	9.11	24
ω-secalins	1	A0A159KI90	Omega-secalin	10.57	68
	2	W6W98	HMW-GS x	4.45	8
HMW-secalins	1	W6AW92	HMW-GS y	17.33	38
	2	Q93WF0	HMW-GS x	14.61	58
	3	Q43639	Sec1	3.96	12
γ-75k-secalins	1	P52589	Protein disulfide-isomerase	8.38	4
	2	Q94IL2	HMW-GS x	5.47	3
	3	E5KZQ1	75k gamma-secalin	4.68	49
γ-40k-secalins	1	W5IA32	Formate dehydrogenase	7.41	3
	2	K3ZAI0	Uncharacterized protein	7.31	8
	3	W4ZSH7	Uncharacterized protein	5.98	5
C-hordeins	1	P06472	C-hordein[d]	7.44	19
γ-hordeins	2	P06470	B1-hordein[e]	3.95	9
	4	A0A287Q402	Uncharacterized protein	2.00	1
B-hordeins	1	Q84LE9	D-hordein	13.59	21
	2	P06470	B1-hordein	11.49	38
	3	F2D284	Protein disulfide-isomerase	9.79	7
D-hordeins	1	I6SW34	D-hordein	33.73	99
	2	P07597	Non-specific lipid-transfer protein	5.07	4
	3	P02864	C-hordein[f]	4.30	2

[a]The rank of the specified protein is relative to all other proteins in the list of detected proteins, [b]Unused ProtScore, defined as a measure of the protein confidence for a detected protein, calculated from the peptide confidence for peptides from spectra that are not already completely "used" by higher scoring winning proteins, thus reflecting the amount of total, unique peptide evidence related to a given protein, [c]after BLAST search (identified as uncharacterized protein: T1LG74), [d]after BLAST search (identified as uncharacterized protein: A0A287EIM7), [e]after BLAST search (identified as uncharacterized protein: A0A287EFG2), [f]after BLAST search (identified as uncharacterized protein: A0A287EEX5).

Barley

The high-scoring proteins detected in the tryptic hydrolysate of the C-hordein-GPT (**Table 2**) corresponded to the three main groups of this GPT, the gluten proteins, the group of others and the group of GSPs+PINs (**Figure 1C**). A C-hordein (Q40055) was identified at rank 23. An uncharacterized protein of *Hordeum vulgare* subsp. *vulgare* (A0A287EIM7) sharing 99.0% homology with the C-hordein (P06472) was present in the chymotryptic hydrolysate of the C-hordein-GPT (**Table 3**). Two B-hordeins and the previously reported γ3-hordein (P80198) (Colgrave et al., 2012) were detected with a high number of peptides in the tryptic hydrolysate of the γ-hordein-GPT. Only two uncharacterized proteins from *Hordeum vulgare* subsp. *vulgare* were identified in the chymotryptic γ-hordein-GPT hydrolysate. The highest ranked protein was identified as a B1-hordein (P06470) with an identity of 94.6% after the BLAST search. The tryptic and chymotryptic hydrolysates of the B-hordein-GPT contained the B3-hordein I6TMW4 with 102 peptides and the two other B-hordeins with a high peptide number, B1-hordein (P06470) and B hordein (Q40026). D-hordein (I6TRS8, 209 peptides detected) was the highest ranking protein in the tryptic hydrolysate of the D-hordein-GPT. The D-hordein (I6SW34, 99 peptides) and an uncharacterized protein (A0A287EEX5, 2 peptides), which was identified as a C-hordein (P02864) with 50% identity were identified in the chymotryptic hydrolysate. Moreover, D-hordeins were detected in all other hordein GPTs with high sequence coverage.

The best three protein hits of each GPT are summarized in **Tables 2** and **3**, according to their ranking of identification. The total numbers of gluten proteins identified using either trypsin or chymotrypsin are presented in **Table 4**. The numbers of identified

proteins were between 2- to 10-fold higher in all GPT hydrolysates using the so-called gold standard proteolytic enzyme trypsin as compared to chymotrypsin. The numbers of identified gluten proteins were 2- to 8-fold higher in the tryptic hydrolysates, except for HMW-GS and LMW-GS. Chymotrypsin revealed as many gluten proteins as trypsin for HMW-GS and more gluten proteins were identified in the chymotryptic hydrolysate of LMW-GS than with trypsin. The total numbers of identified proteins differed from 24 for the γ-hordeins up to 317 for the γ-40k-secalins in the tryptic hydrolysates and from 4 (ω5-gliadins) to 58 (γ-40k-secalins) in the chymotryptic hydrolysates. The ratio of the numbers of all identified proteins to the numbers of identified gluten proteins ranged from 2 for α-gliadins up to 29 for γ-40k-secalins in the tryptic hydrolysates and from 1 for α-gliadins, ω5-gliadins and ω1,2-gliadins to 19 for γ-40k-secalins in the chymotryptic hydrolysates. It should be noted that 18 gluten proteins, but no GPT-specific proteins were identified (73 proteins in total) in the tryptic digest of the ω1,2-gliadin-GPT. In contrast, only seven gluten proteins were identified in the chymotryptic hydrolysate, but among which three of them were ω-gliadin proteins. The same findings were observed for the LMW-GS, for which 22 LMW-GS proteins of 27 gluten proteins were identified in the chymotryptic hydrolysate, but only 2 LMW-GS-proteins within 20 gluten proteins in the tryptic hydrolysate. For the hordeins, the data shows that the enrichment is more specific and that the trypsin data for these GPTs is misleading, because in the chymotryptic hydrolysates less gluten proteins were identified, but more of them corresponded to their appropriate GPT. When looking at the other GPTs, more GPT-specific proteins were identified in the tryptic than in the chymotryptic hydrolysates.

Identification of Immunoreactive Proteins

Various gluten and non-gluten proteins of wheat, rye and barley have been identified as triggers of adverse reactions. The proteomic characterization of the GPTs also provided an insight into the presence of immunoreactive proteins. All identified proteins of the GPTs were searched for the UniProtKB accession based on the allergen code of the World Health Organization/International Union of Immunological Societies and for the name of the immunoreactive proteins. The identified allergens with their allergen code, molecular weight and identification parameters are shown in **Table 5**. Some of the allergens were identified only in one GPT with a small number of peptides (profilin in the LMW-GS-GPT or serpin in the γ-40k-secalin-GPT), but especially ATIs and gluten proteins were very abundant and present in more than one GPT. However, it should be noted that most of the allergens were enriched in one GPT. The WDEIA allergen tri a 19 "ω5-gliadin" was identified only in the appropriate GPT.

Beside the shown exemplary allergens, many identified proteins contained peptides with known CD-active sequences. Immunoreactive peptides carrying known, non-deamidated peptide-binding motifs of gluten-specific T-cells are shown in **Table 6**. CD-active peptides were identified in all wheat GPTs, except ω5-gliadins. The list of T-cell epitopes according to Sollid et al. (2012) contains 31 entries that are reduced to 21 different motifs after reversal of deamidation and removal of duplicates. One of these motifs is specific to oats that were not studied, leaving 20 possible motifs. Of these, five epitopes were not identified (DQ2.5-glia-α3, DQ2.5-glia-γ4a, DQ2.5-glia-γ4b, DQ2.5-glia-γ4d, DQ8-glia- α1), but 15 motifs were detected, especially in the ω1,2-gliadin-, LMW-GS-, and HMW-GS-GPTs. The findings were comparable for the rye GPTs, where similar numbers of peptides were identified in the ω- and HMW-secalin-GPTs as in the γ-75k-secalin-GPT, with the exception of the γ-40k-secalin-GPT with just two epitopes. In the γ-, B-, and D-hordein-GPTs just one peptide-binding motif was detected, but six different peptides were identified in the C-hordein-GPT. The DQ2.5-glia-γ4c peptide-binding motif QQPQQPFPQ

TABLE 4 | Total numbers of identified proteins, gluten proteins, and gluten protein type (GPT)-specific proteins in each GPT digested with trypsin or chymotrypsin, respectively.

GPT	Tryptic			Chymotryptic			Total gluten proteins[d]
	Proteins[a]	Gluten proteins[b]	GPT-specific proteins[c]	Proteins[a]	Gluten proteins[b]	GPT-specific proteins[c]	
α-gliadins	48	20	7	11	11	2	31
γ-gliadins	61	21	4	19	6	2	27
ω5-gliadins	37	8	1	4	3	0	11
ω1,2-gliadins	73	19	0	8	7	3	26
HMW-GS	117	19	10	37	16	6	35
LMW-GS	78	20	2	52	27	22	47
ω-secalins	56	10	6	10	2	1	12
HMW-secalins	96	10	3	18	6	1	16
γ-75k-secalins	244	13	3	43	3	1	16
γ-40k-secalins	317	11	2	58	3	1	14
C-hordeins	37	7	1	11	1	1	8
γ-hordeins	24	8	1	7	1	0	9
B-hordeins	152	7	3	15	3	2	10
D-hordeins	130	8	1	24	3	1	11

[a]Global FDR = 1%; [b]proteins from all Poaceae included; [c]only proteins from appropriate GPTs included; [d]numbers of gluten proteins identified in tryptic and chymotryptic hydrolysates summed without duplicates.

TABLE 5 | Identified allergens of wheat (Tri), rye (Sec), and barley (Hor), their allergen code according to the World Health Organization/International Union of Immunological Societies allergen nomenclature, their UniProtKB accession number and name, the gluten protein type (GPT), in which they were identified and their identification parameters.

Allergen name	UniProtKB accession	Name	MW[a] [kDa]	GPT	T/ C[b]	Score[c]	Peptides (>95%)
Tri a 12	D0PRB5	Profilin	14	LMW-GS	T	2.00	1
Tri a 15	P01083	α-Amylase-inhibitor 0.28	17	γ-40k-secalins	T	2.00	1
Tri a 19	A0A0B5J8A9	ω5-Gliadin	40	ω5-gliadins	T	8.00	8
Tri a 20	D0ES80	γ-Gliadin	34	γ-gliadins	T	20.59	51
				α-gliadins	T	5.85	9
Tri a 21	I0IT55	α-/β-Gliadin	34	α-gliadins	T	2.27	77
	P04727		36	γ-gliadins	T	2.71	10
	I0IT62		38	ω1,2-gliadins	T	2.00	26
	Q41546		36	HMW-GS	T	2.00	7
	P04721		30	LMW-GS	T	4.38	11
Tri a 26	P10388	High molecular weight glutenin subunit Dx5	88	γ-gliadins	T	3.23	6
				ω5-gliadins	T	16.73	17
				ω1,2-gliadins	T	6.24	24
				HMW-GS	T	41.05	97
	Q45R38	High molecular weight glutenin subunit Bx7	85	HMW-GS	T	39.96	64
				HMW-GS	C	8.16	33
Tri a 28	P01085	α-Amylase-inhibitor 0.19	13	ω5-gliadins	T	6.00	8
				LMW-GS	T	2.00	2
	Q5MD68	α-Amylase-inhibitor 0.19	13	ω1,2-gliadins	T	8.01	20
	P01084	α-Amylase-inhibitor 0.53	13	α-gliadins	T	8.78	6
				γ-gliadins	T	3.94	4
Tri a 30	P17314	Tetrameric alpha-amylase inhibitor CM3	16	α-gliadins	T	13.52	18
				γ-gliadins	T	10.00	5
				ω1,2-gliadins	T	8.31	4
				HMW-GS	T	2.03	1
Tri a 31	P46226[d]	Triosephosphate-isomerase	27	γ-75k-secalins	T	2.00	1
				γ-40k-secalins	T	9.85	5
Tri a 32	Q6W8Q2	1-cys-peroxiredoxin		LMW-GS	T	2.47	1
Tri a 33	Q9ST57	Serpin		γ-40k-secalins	C	2.02	1
Tri a 34	C7C4X1	Glyceraldehyde-3-phosphate-dehydrogenase		γ-40k-secalins	C	2.02	1
Tri a 36	B2Y2Q4	LMW glutenin subunit	42	α-gliadins	T	17.12	55
	I3XHQ1	LMW glutenin subunit		γ-gliadins	T	24.08	38
	Q9XGF0	LMW-9		ω5-gliadins	T	6.47	4
	Q8W3V1	LMW GS group 11 type VI		ω1,2-gliadins	T	17.69	35
	A0A165R8I1	S-type LMW GS		LMW-GS	T	14.25	22
	D6RVY4	LMW GS		HMW-GS	T	16.56	73
	R4JFB5	LMW GS		ω-secalins	T	6.02	5
	Q6PKM2	LMW GS		γ-75k-secalins	T	4.24	5
				γ-hordeins	T	2.00	10
Tri a 37	Q9T0P1	Alpha purothionin	12	HMW-GS	T	2.00	2
Tri a 40	Q41540	Chloroform/methanol-soluble (CM) 17 protein [alpha amylase inhibitor]	16	ω1,2-gliadins	T	8.00	24
				HMW-GS	T	4.39	4
				LMW-GS	T	4.53	3
Tri a 44	A0A0G3F720	Endosperm transfer cell specific PR60 precursor		HMW-secalins	T	4.06	3
Sec c 38	Q9S8H2	Dimeric alpha-amylase/trypsin inhibitor	13.5	γ-75k-secalins	T	2.00	1
				γ-40k-secalins	T	2.00	1
Hor v 15	P16968	Alpha-amylase inhibitor	14.5	γ-75k-secalins	T	2.00	1
	P28041	BMAI-1 precursor Alpha-amylase/trypsin inhibitor CMa		γ-hordeins	T	2.00	1
				B-hordeins	T	4.59	3
				D-hordeins	T	11.35	16
Hor v 17	P16098	Beta-amylase	60	B-hordeins	T	19.68	17
	U5NJ12		60	B-hordeins	C	4.65	3
Hor v 20	I6TEV2	Gamma 3 hordein		C-hordeins	T	7.85	8
	P80198	Gamma-hordein 3	34	γ-hordeins	T	2.00	1

[a]Molecular weight according to UniProtKB accession, [b]T, tryptic digest, C, chymotryptic digest, [c]Unused ProtScore, defined as a measure of the protein confidence for a detected protein, calculated from the peptide confidence for peptides from spectra that are not already completely "used" by higher scoring winning proteins, thus reflecting the amount of total, unique peptide evidence related to a given protein, [d]96% identity to Q9FS79 Triticum aestivum.

TABLE 6 | Celiac disease relevant T-cell epitopes (nomenclature according to Sollid et al., 2012) identified in the gluten protein types, respectively.

Epitope	Peptide-binding motif	Reference	Gluten protein type[a]
DQ2.5-glia-α1a	PFPQPQLPY	Arentz-Hansen et al., 2000	α-gliadins LMW-GS HMW-GS
DQ2.5-glia-α1b	PYPQPQLPY	Arentz-Hansen et al., 2002	LMW-GS
DQ2.5-glia-α2	PQPQLPYPQ	Arentz-Hansen et al., 2000	LMW-GS HMW-GS
DQ2.5-glia- γ1	PQQSFPQQQ	Sjöström et al., 1998	ω-secalins
DQ2.5-glia- γ2	IQPQQPAQL	Qiao et al., 2005; Vader et al., 2002	α-gliadins γ-gliadins LMW-GS
DQ2.5-glia- γ3	QQPQQPYPQ	Arentz-Hansen et al., 2002	ω-secalins γ-75k-secalins
DQ2.5-glia- γ4c	QQPQQPFPQ	Arentz-Hansen et al., 2002	ω1,2-gliadins HMW-GS LMW-GS ω-secalins HMW-secalins γ-75k-secalins γ-40k-secalins C-hordeins
DQ2.5-glia- γ5	QQPFPQQPQ	Arentz-Hansen et al., 2002	HMW-GS ω-secalins HMW-secalins γ-75k-secalins γ-40k-secalins C-hordeins
DQ2.5-glia-ω1 DQ2.5-hor-1 DQ2.5-sec-1	PFPQPQQPF	Tye-Din et al., 2010; Vader et al., 2003	ω1,2-gliadins HMW-GS ω-secalins HMW-secalins γ-75k-secalins C-hordeins γ-hordeins
DQ2.5-glia- ω2	PQPQQPFPW	Tye-Din et al., 2010	ω1,2-gliadins HMW-GS C-hordeins
DQ2.5-glut-L2	FSQQQQSPF	Vader et al., 2002; Stepniak et al., 2005	LMW-GS
DQ2.5-hor-2 DQ2.5-sec-2	PQPQQPFPQ	Vader et al., 2003	γ-75k-secalins
DQ2.5-hor-3	PIPQQPQPY	Tye-Din et al., 2010	γ-hordeins B-hordeins D-hordeins
DQ2.2-glut-L1	PFSQQQQPV	Bodd et al., 2012	α-gliadins HMW-GS
DQ8-glut-H1	QGYYPTSPQ	van de Wal et al., 1999	γ-gliadins ω1,2-gliadins HMW-GS LMW-GS HMW-secalins

[a]*Gluten protein types in which the peptides were identified.*

was detected in the ω1,2-, HMW-, and LMW-GS-GPTs, in all four rye GPTs and in the C-hordein-GPT. The DQ2.5-glia-γ5 motif QQPFPQQPQ was also identified in all rye GPTs and in the HMW-GS-GPT. The most frequently detected peptide-binding motif was PFPQPQQPF (DQ2.5-glia-ω1, DQ2.5-hor-1, DQ2.5-sec-1).

Relative Quantitation of Proteins Within Gluten Protein Types

The tryptic and chymotryptic GPT hydrolysates were then subjected to relative quantitation to monitor the relative abundance of the peptides. Only peptides of gluten-derived proteins were selected for the MRM analysis. According to the "best-flyer method" of Ludwig et al. (2012), the peak areas of the four most intense transitions of the best flying peptide per protein (TopPep1/TopTra4) were summed. The model TopPep1/TopTra4 was selected, because only one peptide was detected for many gluten proteins in the undirected LC-MS/MS experiments and it is indicated that this model is as reasonable and robust as the others. The peak areas cannot be compared between peptides, because the MS response is dependent on the amino acid sequence, but the peak areas of the same peptide may be compared between the GPTs. The peak areas of the peptides were summed according to their categories (**Figure 2**). To estimate the enrichment of each category in every GPT the peak areas of each category were converted to a percentage relative to the summed peak area of the respective category for ease of data comparison.

Wheat

For the wheat GPTs, the single proteins were grouped according to their UniProtKB names into the categories LMW-GS, α-, γ-, and ω-gliadins, HMW-GS and avenin-like proteins. LMW-GS constituted the main proportion in the appropriate LMW-GS-GPT, but they were also enriched in the α- and γ-gliadin-GPTs and were present in the other wheat GPTs (**Figure 2A**). Vice versa, a large share of α-gliadins was detected in the α-gliadin- (≈42% of total α-gliadins) and HMW-GS-GPT (≈40% of total α-gliadins). The percentages always refer to 100% of total protein type summed over all wheat, rye or barley GPTs, respectively, e.g., to 100% of total α-gliadins summed over all wheat GPTs. Smaller proportions of α-gliadins were detected in the ω1,2-, γ-gliadin-, and LMW-GS-GPTs. The γ-gliadins were detected in almost all GPTs, except the ω-gliadin-GPTs, but were noticeably enriched in the γ-gliadin-GPT (≈66% of total γ-gliadins). The ω-gliadins were present almost only in the ω1,2-gliadin-GPT (≈76% of total ω-gliadins). HMW-GS accounted for a small proportion in each wheat GPT, but the HMW-GS-GPT had the highest proportion of HMW-GS (≈77% of total HMW-GS), as expected. The ω5-gliadin-GPT showed low proportions of the analyzed proteins of HMW-GS, LMW-GS and ω-gliadins. The avenin-like proteins were present in small amounts in almost all wheat GPTs, except the ω5-gliadin-GPT. The technical variation was assessed by examining the mean (combining GPTs of wheat) coefficient of variation (CV) for each peptide with an overall average of 13% for the cleavage with trypsin and 12% for the cleavage with chymotrypsin.

Rye

For the rye GPTs, the proteins were categorized according to their UniProtKB names into γ-75k-secalins, γ-prolamins, HMW-secalins, ω-secalins, LMW-GS, and avenin-like proteins (**Figure 2B**). The ω-secalins were almost only detected in the ω-secalin-GPT (≈99% of total ω-secalins). HMW-secalins were detected in all rye GPTs, but with a noticeable enrichment in the appropriate HMW-secalin-GPT (≈96% of total HMW-secalins). The HMW-secalin-GPT contained almost only HMW-secalins. The γ-75k-secalin-GPT contained a very high proportion of γ-75k-secalins (≈95% of total γ-75k-secalins) and lower amounts of HMW-secalins, avenin-like proteins and LMW-GS. In comparison, the γ-40k-secalin-GPT comprised mainly γ-prolamins and γ-75k-secalins with a lower proportion

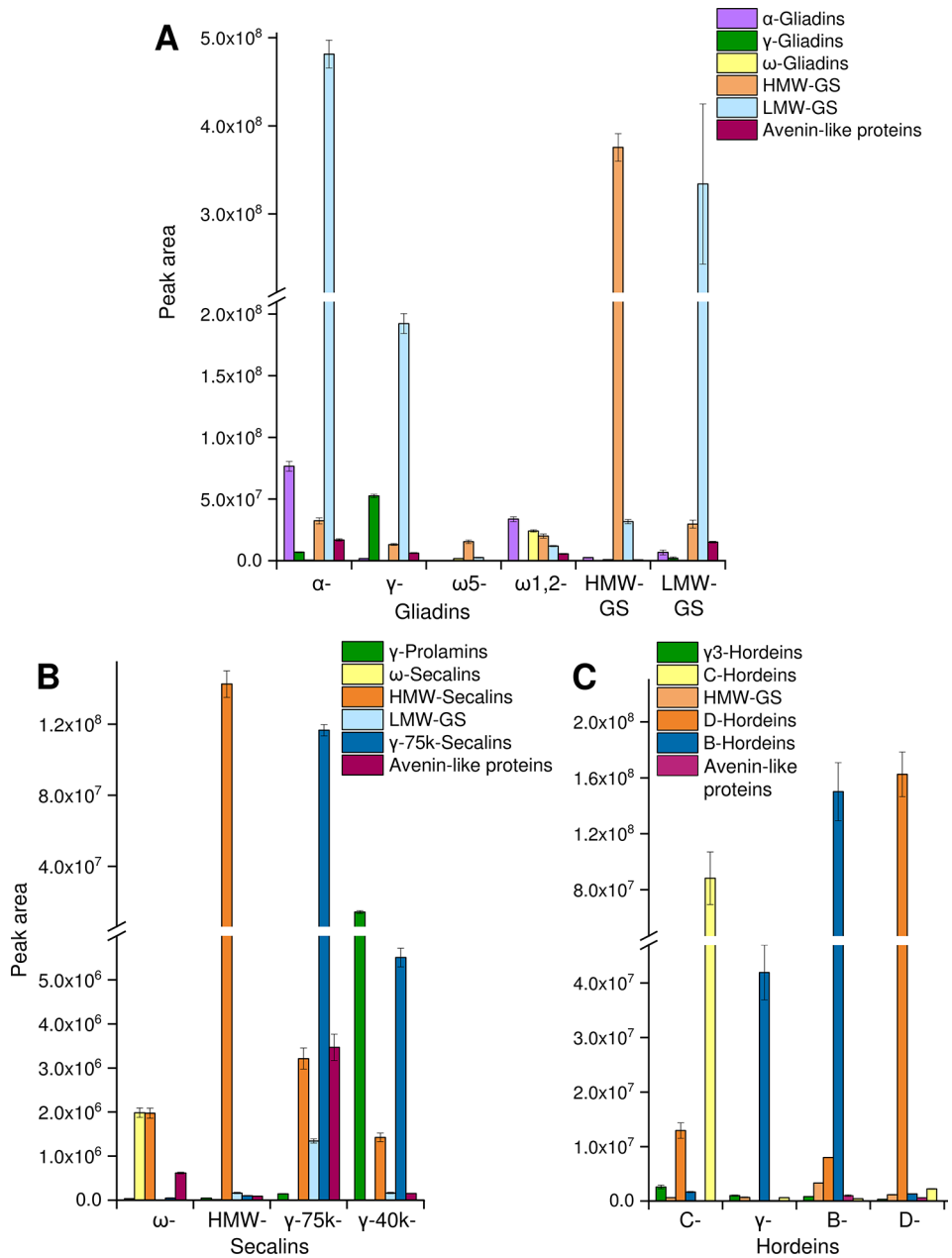


FIGURE 2 | Relative protein quantification in GPTs. The summed peak areas of selected tryptic and chymotryptic peptides of the most abundant proteins representing protein groups in individual GPTs: peak areas of peptides representing α-gliadins, γ-gliadins, ω-gliadins, HMW-GS, LMW-GS, and avenin-like proteins in the GPTs of wheat **(A)**, peak areas of peptides representing γ-prolamins, ω-secalins, HMW-secalins, LMW-GS, γ-75k-secalins, and avenin-like proteins in the GPTs of rye **(B)**, peak areas of peptides representing γ3-hordeins, HMW-GS, D-hordeins, B-hordeins, C-hordeins, and avenin-like proteins in the GPTs of barley **(C)**. Data is plotted as the mean ± standard deviation (n = 3).

of HMW-secalins. The avenin-like proteins were enriched in the γ-75k-secalin-GPT. The average CV for the tryptic cleavage of the GPTs of rye was 10% and for the chymotryptic cleavage 6%.

Barley

The barley GPTs were grouped into the following categories: D-hordeins, B-hordeins, γ3-hordeins, C-hordeins, avenin-like proteins, and HMW-GS from *Triticum aestivum* and a similar tribe (*C*) in the family *Poaceae*. In comparison with the other barley GPTs, the C-hordein-GPT contained the highest amount of C-hordeins (≈96% of total C-hordeins) and a high proportion of D-hordeins. The D-hordeins were also detected in the B-hordein-GPT, but they accounted for the largest share of their appropriate GPT (≈90% of total D-hordeins). B- and γ-hordein-GPTs were mainly composed of B-hordeins, whereas the B-hordein-GPT

showed noticeably higher proportions of the B-hordeins (≈77% of total B-hordeins) and also of proteins of the other groups analyzed (**Figure 2C**). The γ-hordein-GPT showed a clear enrichment of the B-hordeins. For the tryptic cleavage of the barley GPTs the average CV was 9% and for the chymotryptic cleavage 10%.

DISCUSSION

In this study, we provided novel insights into the complexity of gluten from wheat, rye, and barley by identification of the individual proteins and relative quantitation of the most abundant gluten proteins in the GPTs. A preparative strategy (Schalk et al., 2017) was used to isolate the GPTs from wheat, rye and barley flours according to solubility and hydrophobicity. The LC-MS/MS experiments confirmed an enrichment of the expected gluten proteins in their corresponding GPTs in most cases. The application of high-resolution MS allowed a much more detailed and accurate insight into the composition of the isolated GPTs compared to our earlier low-resolution MS analyses (Schalk et al., 2017). The data of the undirected LC-MS/MS experiments showed the qualitative composition of the GPTs, according to the number of peptides identified and revealed a first assumption of the total composition of each GPT. All GPTs contained gluten proteins other than those derived from the known RP-HPLC retention times as well as ATIs, enzymes or uncharacterized proteins. These findings underline the incomplete separation of prolamins and glutelins according to solubility and show that even the separation by preparative RP-HPLC is not clear-cut enough to separate individual GPTs without co-purifying other components, such as ATIs (Junker et al., 2012).

The undirected LC-MS/MS experiments revealed that the group of gluten proteins constituted the highest proportion in the wheat GPTs followed by the second largest group of ATIs, which were present especially in the ω5- and ω1,2-gliadin-GPTs. The MRM data showed that the group of gluten proteins had different compositions of α-, γ-, ω-gliadins, LMW-GS, and HMW-GS, mostly enriched in their appropriate GPTs. However, we found that the LMW-GS were detected in all wheat GPTs. Recently, the presence of LMW-GS in the gliadin fraction has been reported as well (Boukid et al., 2019). Due to their polymeric nature (Shewry, 2019), their similarity to α-gliadins in molecular weight and also to γ-gliadins in RP-HPLC retention times, it may not be possible to achieve a clear-cut separation between those GPTs. Thus, small proportions of LMW-GS were contained in all wheat GPTs.

The ω- and HMW-secalin-GPTs showed high proportions of gluten proteins in the undirected LC-MS/MS analysis. The subsequent MRM analyses revealed that the gluten protein fractions were highly enriched with the expected protein types. As described in previous studies, HMW-secalins were detected with notably high proportions in the other rye GPTs. In case of the ω-secalin-GPT this may be due to the reduction of the disulfide bonds of the HMW-secalins, which then co-eluted in the ω-secalin-GPT (Gellrich et al., 2003). When fractionating rye gluten proteins, we observed that the separation according to solubility is even less complete than in wheat. This led to a higher co-mingling of the individual GPTs even after preparative RP-HPLC. The detection of LMW-GS and avenin-like proteins beside the main group γ-75k-secalins in this GPT may give another hint for the similarity of those GPTs due to the close genetic relationship of rye and wheat (Kasarda et al., 1983). There was no reliable reference sequence available for the γ-40k-secalins (June 2019), but the group named γ-prolamins was only detected in the γ-40k-secalin-GPT. Although the molecular weight (UniProtKB database) of the γ-prolamins detected was somewhat too low compared to the generally known mass range for γ-40k-secalins, the assignment to this GPT would be possible due to amino acid sequence, organism and similarity to other rye proteins. This fact showed the incompleteness of the rye protein entries in the UniProtKB database, because these γ-prolamins were very similar to previously identified ones (Schalk et al., 2017).

The same separation issue as for the rye GPTs appeared for barley GPTs. As stated by Schalk et al. (2017), γ/B-hordeins from the prolamin fraction contained the monomeric γ-hordeins and partly the disulfide-bound B-hordeins. The B/γ-hordeins prepared from glutelin fraction showed the opposite case with the majority of oligomeric or polymeric B-hordeins. Similar results were obtained in this study, except that the γ-hordeins were detected with similar proportions in all barley GPTs. The same applied to the D-hordeins, which were clearly enriched in the D-hordein-GPT, but also identified with noticeably high amounts in the other GPTs. This may also be traced back to the customized separation technique. The identification of hordeins revealed again the challenge with incomplete or unannotated protein entries in the database (Colgrave et al., 2013). Especially the number of entries for barley and rye were low and many proteins were matched as uncharacterized proteins. Reliable protein reference sequences, especially for the *Hordeum* sp. and *Secale* sp. are urgently needed, because the proteomics results are likely to be affected by the drastically different number of protein sequences available.

One limitation of the current study is that the results are based on the analysis of GPTs isolated from one single cultivar of each grain grown in one year. Although the choice of the cultivars was done carefully to select representative samples, genetic and environmental factors and their interaction are known to influence the proteome composition of cereals (Hajas et al., 2018; Juhasz et al., 2018; Malalgoda et al., 2018; Geisslitz et al., 2019). The results obtained here thus only provide one snapshot and are expected to change depending on the flour sample. The overall procedure from milling to collecting sufficient amounts of GPTs after preparative RP-HPLC is rather time-consuming as well as cost- and labor-intensive, so that it is impossible to do this for more than a very limited number of samples. This is why the current study first focused on determining the efficiency of fractionation of the various GPTs, prior to studying the variability arising from different factors.

This study also revealed that trypsin is preferred for the identification experiments for almost all GPTs, except for ω1,2-gliadins and LMW-GS, which were better characterized using the chymotryptic hydrolysate to increase sequence coverage. This may be in part due to the fact that ω1,2-gliadins are more resistant to trypsin and have less K/R (trypsin cleavage sites), so these will be under-represented compared to "other" proteins that have higher K/R and hence more tryptic peptides, such as HMW-GS (Alves et al., 2018). However, for the identification of specific gluten

proteins, chymotrypsin yielded more results, because it is shown that the enrichment is more specific and that the trypsin data for some GPTs might be misleading. In general, gluten contains few lysine and arginine residues, but it seems that trypsin was still mostly superior to chymotrypsin due to its cleavage specificity, efficiency and delivery of peptides with favorable chromatographic and MS properties in terms of ionization and fragmentation, as has been reported before (Colgrave et al., 2017b). Most peptides were tryptic, but some were also generated from aspecific cleavage sites. We also observed that the identified proteins and their ranks change depending on the cleavage enzyme used. Due to a number of confounding factors, it is hard to make an assessment which enzyme is more representative of the truth, which is why the results of both approaches were combined in **Figure 2**. Further experiments would be necessary using additional enzymes with different cleavage specificities to investigate this in more detail. The undirected LC-MS/MS analysis of the chymotryptic hydrolysates seemed to be more suitable for the detection of peptides with CD-active epitopes, because significantly more of these peptides were identified than after tryptic hydrolysis. It is known that peptides containing CD-active epitopes are typically resistant to cleavage by trypsin and may therefore be identified in a low amount (Shan et al., 2005). In total, 15 out of 20 different CD-active epitopes were detected. Of the five that were not detected, two (DQ2.5-glia-γ4a, DQ2.5-glia-γ4d) were not present either in historical and modern spring wheat cultivars (Malalgoda et al., 2018).

To conclude, the combination of discovery proteomics and relative quantitation of gluten proteins provided novel insights into the relative amounts of the individual proteins in purified GPTs. These well-defined materials are suitable for a wide range of applications and have already been used as reference materials to quantitate gluten from wheat, rye and barley using targeted LC-MS/MS (Schalk et al., 2018a; Schalk et al., 2018b), as stimulatory agents for epitope mapping (Röckendorf et al., 2017) and for recognition profiling of monoclonal antibodies (Lexhaller et al., 2017). Further potential uses are a variety of functional assays to study mechanisms of immune activation. Our findings raise awareness of the challenges of obtaining "pure" GPTs for analytical purposes and clinical studies on disease mechanisms. Especially when applying gluten or gluten fractions in studies on pathomechanisms of, e.g., CD, NCGS, or WDEIA, it is essential to know which proteins are present in the fractions of interest to establish relationships between structure, functionality and bioactivity.

DATA AVAILABILITY STATEMENT

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) with the dataset identifier PXD016065 and are publicly available on Panorama Public (https://panoramaweb.org/nOlizr.url).

AUTHOR CONTRIBUTIONS

BL planned and performed the experiments, analyzed the data, designed the figures and wrote the original draft. MC provided access to the LC-MS/MS instruments, contributed to proteomics data analysis and study design. KS was responsible for study conceptualization, contributed to funding acquisition and editing of the manuscript. All authors reviewed and edited the manuscript and approved the final version.

FUNDING

This research project (No. 250645717) was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG, Bonn). BL received additional funding from the Technical University of Munich through the TUM Graduate School Partnership Mobility Grant, the TUM Graduate School Internationalization Grant and a travel grant from the Silesia-Clemens-Hanke Stiftung.

ACKNOWLEDGMENTS

The authors would like to thank Ms. Alexandra Axthelm and Ms. Angelika Grassl (Leibniz-LSB@TUM) for excellent technical assistance, Ms. Keren Byrne (CSIRO) for help with LC-MS experiments, Prof. Dr. Peter Köhler (biotask AG) and Dr. Herbert Wieser for scientific advice and Prof. Dr. Thomas Hofmann (TUM, Leibniz-LSB@TUM) for promoting the Silesia travel grant.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01530/full#supplementary-material

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Alves, T. O., D'Almeida, C. T. S., Victoria, V. C. M., Souza, G. H. M. F., Cameron, L. C., and Ferreira, M. S. L. (2018). Immunogenic and allergenic profile of wheat flours from different technological qualities revealed by ion mobility mass spectrometry. *J. Food Compos Anal.* 73, 67–75. doi: 10.1016/j.jfca.2018.07.012

Arentz-Hansen, H., Körner, R., Molberg, Ø., Quarsten, H., Vader, W., Kooy, Y. M. C., et al. (2000). The intestinal T cell response to α-gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. *J. Exp. Med.* 191, 603–612. doi: 10.1084/jem.191.4.603

Arentz-Hansen, H., McAdam, S. N., Molberg, F. B., Lundin, K. E., Jorgensen, T. J., Jung, G., et al. (2002). Celiac lesion T cells recognize epitopes that cluster in regions of gliadins rich in proline residues. *Gastroenterol* 123, 803–809. doi: 10.1053/gast.2002.35381

Bodd, M., Kim, C. Y., Lundin, K. E., and Sollid, L. M. (2012). T-Cell response to gluten in patients with HLA-DQ2.2 reveals requirement of peptide-MHC stability in celiac disease. *Gastroenterol* 142, 552–561. doi: 10.1053/j.gastro.2011.11.021

Boukid, F., Prandi, B., Faccini, A., and Sforza, S. (2019). A complete mass spectrometry (MS)-based peptidomic description of gluten peptides generated during *in vitro*

gastrointestinal digestion of durum wheat: Implication for celiac disease. *J. Am. Soc. Mass Spectrom.* 30, 1481. doi: doi.org/10.1007/s13361-019-02212-8

Catassi, C., Alaedini, A., Bojarski, C., Bonaz, B., Bouma, G., Carroccio, A., et al. (2017). The overlapping area of non-celiac gluten sensitivity (NCGS) and wheat-sensitive irritable bowel syndrome (IBS): an update. *Nutrients* 9, 1268. doi: 10.3390/nu9111268

Codex Alimentarius Commission (2015). Codex Standard 118-1979: *Codex Standard for Foods for Special Dietary use for Persons Intolerant to Gluten*. Revision 1. Amendment 2.

Colgrave, M. L., Goswami, H., Howitt, C. A., and Tanner, G. J. (2012). What is in a beer? Proteomic characterization and relative quantification of hordein (gluten) in beer. *J. Proteome Res.* 11, 386–396. doi: 10.1021/pr2008434

Colgrave, M. L., Goswami, H., Howitt, C. A., and Tanner, G. J. (2013). Proteomics as a tool to understand the complexity of beer. *Food Res. Int.* 54, 1001–1012. doi: 10.1016/j.foodres.2012.09.043

Colgrave, M. L., Byrne, K., Blundell, M., and Howitt, C. A. (2016a). Identification of barley- specific peptide markers that persist in processed foods and are capable of detecting barley contamination by LC-MS/MS. *J. Prot* 147, 169–176. doi: 10.1016/j.jprot.2016.03.045

Colgrave, M. L., Byrne, K., Blundell, M., Heidelberger, S., Lane, C. S., Tanner, G. J., et al. (2016b). Comparing multiple reaction monitoring and sequential window acquisition of all theoretical mass spectra for the relative quantification of barley gluten in selectively bred barley lines. *Anal. Chem.* 88, 9127–9135. doi: 10.1021/acs.analchem.6b02108

Colgrave, M. L., Byrne, K., and Howitt, C. A. (2017a). Liquid chromatography–Mass spectrometry analysis reveals hydrolyzed gluten in beers crafted to remove gluten. *J. Agric. Food Chem.* 65, 9715–9725. doi: 10.1021/acs.jafc.7b03742

Colgrave, M. L., Byrne, K., and Howitt, C. A. (2017b). Food for thought: selecting the right enzyme for the digestion of gluten. *Food Chem.* 234, 389–397. doi: 10.1016/j.foodchem.2017.05.008

Dunwell, J. M. (1998). Cupins: a new superfamily of functionally diverse proteins that include germins and plant storage proteins. *Biotechnol. Genet. Eng Rev.* 15, 1–32. doi: 10.1080/02648725.1998.10647950

Dupont, F. M., Vensel, W. H., Tanaka, C. T., Hurkman, W. J., and Altenbach, S. B. (2011). Deciphering the complexities of the wheat flour proteome using quantitative two- dimensional electrophoresis, three proteases and tandem mass spectrometry. *Proteome Sci.* 9, 10. doi: 10.1186/1477-5956-9-10

Ellis, H. J., Lozano-Sanchez, P., Bermudo-Redondo, C., Suligoj, T., Biagi, F., Bianchi, P. I., et al. (2011). Antibodies to wheat high-molecular-weight glutenin subunits in patients with celiac disease. *Int. Arch. Allergy Immunol.* 159, 428–434. doi: 10.1159/000338284

Ferranti, P., Mamone, G., Picariello, G., and Addeo, F. (2007). Mass spectrometry analysis of gliadins in celiac disease. *J. Mass Spectrom.* 42, 1531–1548. doi: 10.1002/jms.1361

Field, J. M., Shewry, P. R., and Miflin, B. J. (1982). The purification and characterization of homologous high molecular weight storage proteins from grain of wheat, rye and barley. *Theor. Appl. Genet.* 62, 329–336. doi: 10.1007/BF00275097

Geisslitz, S., Longin, C. F. H., Scherf, K. A., and Koehler, P. (2019). Comparative study on gluten protein composition of ancient (einkorn, emmer and spelt) and modern wheat species (durum and common wheat). *Foods* 8, 409. doi: 10.3390/foods8090409

Gellrich, C., Schieberle, P., and Wieser, H. (2003). Biochemical characterization and quantification of the storage protein (secalin) types in rye flour. *Cereal Chem.* 80, 102– 109. doi: 10.1094/cchem.2003.80.1.102

Hahne, H., Pachl, F., Ruprecht, B., Maier, S. K., Klaeger, S., Helm, D., et al. (2013). DMSO enhances electrospray response, boosting sensitivity of proteomic experiments. *Nat. Methods* 10, 989–992. doi: 10.1038/nmeth.2610

Hajas, L., Scherf, K. A., Török, K., Bugyi, Z., Schall, E., Poms, R. E., et al. (2018). Variation in protein composition among wheat (*Triticum aestivum* L.) cultivars to identify cultivars suitable as reference material for wheat gluten analysis. *Food Chem.* 267, 387–394. doi: 10.1016/j.foodchem.2017.05.005

International Association for Cereal Science and Technology (1976). ICC Standard No. 110/1. Determination of the moisture content of cereals and cereal products (practical method).

International Association for Cereal Science and Technology (2000). ICC Standard No. 167. Determination of crude protein in grain and grain products for food and feed by the Dumas combustion principle.

International Wheat Genome Sequencing Consortium (IWGSC) (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Sci* 361, eaar7191. doi: 10.1126/science.aar7191

Juhasz, A., Belova, T., Florides, C. G., Maulis, C., Fischer, I., Gell, G., et al. (2018). Genome mapping of seed-borne allergens and immunoresponsive proteins in wheat. *Sci. Adv.* 4, eaar8602. doi: 10.1126/sciadv.aar8602

Junker, Y., Zeissig, S., Kim, S.-J., Barisani, D., Wieser, H., Leffler, D. A., et al. (2012). Wheat amylase trypsin inhibitors drive intestinal inflammation *via* activation of toll-like receptor 4. *J. Exp. Med.* 209, 2395–2408. doi: 10.1084/jem.20102660

Kasarda, D. D., Autran, J.-C., Lew, E. J.-C., Nimmo, C. C., and Shewry, P. R. (1983). N-terminal amino acid sequences of ω-gliadins and ω-secalins. Implications for the evolution of prolamin genes. *Biochim. Biophys. Acta* 747, 138–150. doi: 10.1016/0167-4838(83)90132-2

Kerpes, R., Knorr, V., Procopio, S., Koehler, P., and Becker, T. (2016). Gluten-specific peptidase activity of barley as affected by germination and its impact on gluten degradation. *J. Cereal Sci.* 68, 93–99. doi: 10.1016/j.jcs.2016.01.004

Lagrain, B., Brunnbauer, M., Rombouts, I., and Koehler, P. (2013). Identification of intact high molecular weight glutenin subunits from the wheat proteome using combined liquid chromatography-electrospray ionization mass spectrometry. *PloS One* 8, e58682. doi: 10.1371/journal.pone.0058682

Lexhaller, B., Tompos, C., and Scherf, K. A. (2016). Comparative analysis of prolamin and glutelin fractions from wheat, rye, and barley with five sandwich ELISA test kits. *Anal. Bioanal Chem.* 408, 6093–6104. doi: 10.1007/s00216-016-9721-7

Lexhaller, B., Tompos, C., and Scherf, K. A. (2017). Fundamental study on reactivities of gluten protein types from wheat, rye and barley with five sandwich ELISA test kits. *Food Chem.* 237, 320–330. doi: 10.1016/j.foodchem.2017.05.121

Ludvigsson, J. F., Leffler, D. A., Bai, J., Biagi, F., Fasano, A., Green, P. H., et al. (2013). The Oslo definitions for coeliac disease and related terms. *Gut.* 62, 43–52. doi: 10.1136/gutjnl- 2011-301346

Ludwig, C., Claassen, M., Schmidt, A., and Aebersold, R. (2012). Estimation of absolute protein quantities of unlabeled samples by selected reaction monitoring mass spectrometry. *Mol. Cell Proteomics.* 11, 1–16. doi: 10.1074/mcp.M111.013987

Malalgoda, M., Meinhardt, S. W., and Simsek, S. (2018). Detection and quantitation of immunogenic epitopes related to celiac disease in historical and modern hard red spring wheat cultivars. *Food Chem.* 264, 101–107. doi: 10.1016/j.foodchem.2018.04.131

Mamone, G., De Caro, S., Di Luccia, A., Addeo, F., and Ferranti, P. (2009). Proteomic-based analytical approach for the characterization of glutenin subunits in durum wheat. *J. Mass Spectrom.* 44, 1709–1723. doi: 10.1002/jms.1680

Matsuo, H., Kohno, K., and Morita, E. (2005). Molecular cloning, recombinant expression and IgE-binding epitope of ω5-gliadin, a major allergen in wheat-dependent exercise-induced anaphylaxis. *FEBS J.* 272, 4431–4438. doi: 10.1111/j.1742-4658.2005.04858.x

Qiao, S. W., Bergseng, E., Molberg, O., Jung, G., Fleckenstein, B., and Sollid, L. M. (2005). Refining the rules of gliadin T cell epitope binding to the disease-associated DQ2 molecule in celiac disease: importance of proline spacing and glutamine deamidation. *J. Immunol.* 175, 254– 261. doi: 10.4049/jimmunol.175.1.254

Röckendorf, N., Meckelein, B., Scherf, K. A., Schalk, K., Koehler, P., and Frey, A. (2017). Identification of novel antibody-reactive detection sites for comprehensive gluten monitoring. *PloS One* 12, e0181566. doi: 10.1371/journal.pone.0181566

Sapone, A., Bai, J. C., Ciacci, C., Dolinsek, J., Green, P. H., Hadjivassiliou M., et al. (2012). Spectrum of gluten-related disorders: consensus on new nomenclature and classification. *BMC Med.* 10, 13. doi: 10.1186/1741-7015-10-13

Schalk, K., Lexhaller, B., Koehler, P., and Scherf, K. A. (2017). Isolation and characterization of gluten protein types from wheat, rye, barley and oats for use as reference materials. *PloS One* 12, 2. doi: 10.1371/journal.pone.0172819

Schalk, K., Koehler, P., and Scherf, K. A. (2018a). Quantitation of specific barley, rye and oat marker peptides by targeted liquid chromatography - mass spectrometry to determine gluten concentrations. *J. Agric. Food Chem.* 66, 3581–3592. doi: 10.1021/acs.jafc.7b05286

Schalk, K., Koehler, P., and Scherf, K. A. (2018b). Targeted liquid chromatography tandem mass spectrometry to quantitate wheat gluten using well-defined reference proteins. *PloS One* 13, e0192804. doi: 10.1371/journal.pone.0192804

Scherf, K. A., Koehler, P., and Wieser, H. (2016). Gluten and wheat sensitivities – an overview. *J. Cereal Sci.* 67, 2–11. doi: 10.1016/j.jcs.2015.07.008

Scherf, K. A., Lindenau, A.-C., Valentini, L., Collado, M. C., García-Mantrana, I., Christensen, M. J., et al. (2019). Cofactors of wheat dependent exercise-induced anaphylaxis do not increase highly individual gliadin absorption in healthy volunteers. *Clin. Transl. Allergy* 9, 19. doi: 10.1186/s13601-019-0260-0

Scherf, K. A. (2019). Immunoreactive cereal proteins in wheat allergy, non-celiac gluten/wheat sensitivity (NCGS) and celiac disease. *Curr. Opin. Food Sci.* 25, 35–41. doi: 10.1016/j.cofs.2019.02.003

Shan, L., Qiao, S.-W., Arentz-Hansen, H., Molberg, Ø., Gray, G. M., Sollid, L. M., et al. (2005). Identification and Analysis of Multivalent Proteolytically Resistant Peptides from Gluten: Implications for Celiac Sprue. *J. Proteome Res.* 4, 1732–1741. doi: 10.1021/pr050173t

Shewry, P. R., Tatham, A. S., Pappin, D. J., and Keen, J. N. (1988). Terminal amino acid sequences show that D hordein of barley and high molecular weight (HMW) secalins of rye are homologous with HMW glutenin subunits of wheat. *Cereal Chem.* 65, 610–611.

Shewry, P. R. (2019). What is gluten – why is it special? *Front. Nutr.* 6, 101. doi: 10.3389/fnut.2019.00101

Shilov, I. V., Seymour, S. L., Patel, A. A., Loboda, A., Tang, W. H., Keating, S. P., et al. (2007). The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol. Cell. Proteomics* 6, 1638–1655. doi: 10.1074/mcp.T600050-MCP200

Sjöström, H., Lundin, K. E. A., Molberg, K. R., McAdam, S. N., Anthonsen, D., and Quarsten, H., et al. (1998). Identification of a gliadin T-cell epitope in coeliac disease: general importance of gliadin deamidation for intestinal T-cell recognition. *Scand. J. Immunol.* 48, 111–115. doi: 10.1046/j.1365-3083.1998.00397.x

Sollid, L. M., Qiao, S.-W., Anderson, R. P., Gianfrani, C., and Koning, F. (2012). Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenet* 64, 455–460. doi: 10.1007/s00251-012-0599-z

Stepniak, D., Vader, L. W., Kooy, Y., van Veelen, P. A., Moustakas, A., and Papandreou, N. A. (2005). T-cell recognition of HLA-DQ2-bound gluten peptides can be influenced by an N-terminal proline at p-1. *Immunogenet* 57, 8–15. doi: 10.1007/s00251-005-0780-8

Tatham, A. S., and Shewry, P. R. (2008). Allergens in wheat and related cereals. *Clin. Exp. Allergy* 38, 1712–1726. doi: 10.1111/j.1365-2222.2008.03101.x

Tye-Din, J. A., Stewart, J. A., Dromey, J. A., Beissbarth, T., van Heel, D. A., and Tatham, A., et al. (2010). Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Sci. Transl. Med.* 2, 41ra51. doi: 10.1126/scitranslmed.3001012

Vader, L. W., Kooy, Y., van Veelen, P., de Ru, A., Harris, D., Benckhuijsen, W., et al. (2002). The gluten response in children with celiac disease is directed toward multiple gliadin and glutenin peptides. *Gastroenterol* 122, 1729–1737. doi: 10.1053/gast.2002.33606

Vader, L. W., Stepniak, D. T., Bunnik, E. M., Kooy, Y. M. C., de Hann, W., Drijfhout, J. W., et al. (2003). Characterization of cereal toxicity for celiac disease patients based on protein homology in grains. *Gastroenterol* 125, 1105–1113. doi: 10.1016/S0016-5085(03)01204-6

van de Wal, Y., Kooy, Y. M., van Veelen, P., Vader, W., August, S. A., Drijfhout, J. W., et al. (1999). Glutenin is involved in the gluten-driven mucosal T cell response. *Eur. J. Immunol.* 29, 3133-3139. doi: 10.1002/(SICI)1521-4141(199910)29:10<3133::AID-IMMU3133>3.0.CO;2-G

Van Eckert, R., Berghofer, E., Ciclitira, P. J., Chirdo, F., Denery-Papini, S., Ellis, H. J., et al. (2006). Towards a new gliadin reference material – isolation and characterisation. *J. Cereal Sci.* 43, 331–341. doi: 10.1016/j.jcs.2005.12.009

Vensel, W. H., Tanaka, C. K., and Altenbach, S. B. (2014). Protein composition of wheat gluten polymer fractions determined by quantitative two-dimensional gel electrophoresis and tandem mass spectrometry. *Proteome Sci.* 12, 1–13. doi: 10.1186/1477-5956-12-8

Wieser, H., Antes, S., and Seilmeier, W. (1998). Quantitative determination of gluten protein types in wheat flour by reversed-phase high-performance liquid chromatography. *Cereal Chem.* 75, 644–650. doi: 10.1094/cchem.1998.75.5.644

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 18 February 2020
doi: 10.3389/fnut.2020.00011

Directed-Mutagenesis of *Flavobacterium meningosepticum* Prolyl-Oligopeptidase and a Glutamine-Specific Endopeptidase From Barley

Claudia E. Osorio[1,2], *Nuan Wen*[1], *Jaime H. Mejías*[3], *Shannon Mitchell*[4], *Diter von Wettstein*[1] *and Sachin Rustgi*[1,5]*

[1] Department of Crop and Soil Sciences, Washington State University, Pullman, WA, United States, [2] Agriaquaculture Nutritional Genomic Center, Temuco, Chile, [3] Centro Regional de Investigación Carillanca, Instituto de Investigaciones Agropecuarias INIA, Temuco, Chile, [4] Department of Biological Systems Engineering, Washington State University, Pullman, WA, United States, [5] Department of Plant and Environmental Sciences, Clemson University Pee Dee Research and Education Center, Florence, SC, United States

OPEN ACCESS

Edited by:
Michelle Lisa Colgrave, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia

Reviewed by:
Carmen Gianfrani, Institute of Biochemistry and Cell Biology (CNR), Italy
Ana Maria Calderon De La Barca, Centro de Investigación en Alimentación y Desarrollo (CIAD), Mexico

***Correspondence:**
Sachin Rustgi
srustgi@clemson.edu

Specialty section:
This article was submitted to Nutrition and Food Science Technology, a section of the journal Frontiers in Nutrition

Received: 14 October 2019
Accepted: 28 January 2020
Published: 18 February 2020

Citation:
Osorio CE, Wen N, Mejías JH, Mitchell S, von Wettstein D and Rustgi S (2020) Directed-Mutagenesis of Flavobacterium meningosepticum Prolyl-Oligopeptidase and a Glutamine-Specific Endopeptidase From Barley. Front. Nutr. 7:11. doi: 10.3389/fnut.2020.00011

Wheat gluten proteins are the known cause of celiac disease. The repetitive tracts of proline and glutamine residues in these proteins make them exceptionally resilient to digestion in the gastrointestinal tract. These indigested peptides trigger immune reactions in susceptible individuals, which could be either an allergic reaction or celiac disease. Gluten exclusion diet is the only approved remedy for such disorders. Recently, a combination of a glutamine specific endoprotease from barley (EP-B2), and a prolyl endopeptidase from *Flavobacterium meningosepticum* (Fm-PEP), when expressed in the wheat endosperm, were shown to reasonably detoxify immunogenic gluten peptides under simulated gastrointestinal conditions. However useful, these "glutenases" are limited in application due to their denaturation at high temperatures, which most of the food processes require. Variants of these enzymes from thermophilic organisms exist, but cannot be applied directly due to their optimum activity at temperatures higher than 37°C. Though, these enzymes can serve as a reference to guide the evolution of peptidases of mesophilic origin toward thermostability. Therefore, a sequence guided site-saturation mutagenesis approach was used here to introduce mutations in the genes encoding Fm-PEP and EP-B2. A thermostable variant of Fm-PEP capable of surviving temperatures up to 90°C and EP-B2 variant with a thermostability of up 60°C were identified using this approach. However, the level of thermostability achieved is not sufficient; the present study has provided evidence that the thermostability of glutenases can be improved. And this pilot study has paved the way for more detailed structural studies in the future to obtain variants of Fm-PEP and EP-B2 that can survive temperatures ∼100°C to allow their packing in grains and use of such grains in the food industry.

Keywords: glutenases, site-directed mutagenesis, thermostability, glutenin, gliadin, celiac disease

INTRODUCTION

Celiac disease with an autoimmune component affects about 1.4% of the global population (1). Currently, there is only one approved therapy for celiac disease, which is the lifelong gluten abstinence (2, 3). The effects of this prescription on individuals and families, makes it difficult to follow, due to cultural, social, technical, and financial concerns (4, 5) as well as problems associated with the use of the gluten-free commodities (6, 7). Therefore, alternative therapies are continuously being sought worldwide.

Gluten, the causal agent of celiac disease (8), is a complex mixture of polypeptides, which is also responsible for the unique technological properties to wheat (6, 9–11). Glutamine (Gln or Q, 35%) and proline (Pro or P, 15%) are the two major constituents of the gluten proteins, which give it its identity as prolamins (12). The iterative tracts of glutamine and proline-residues present in gluten proteins allow dense packing of nitrogen in grains, but also render gluten proteins highly resistant to gastric and pancreatic proteases thus producing a broad size range of Pro/Gln-rich peptides (8, 13). These indigestible peptides pass through the intestinal epithelium and reach the lamina propria where they get deamidated by the tissue transglutaminase 2 (tTG2), which increases their affinity to the human leukocyte antigen (HLA)-DQ2 or HLA-DQ8 (14, 15). Deamidated gluten peptide-DQ2 complex enhances the inflammatory gut mucosa response by eliciting an increase in the CD4+ T-helper 1 (Th 1) cell-mediated inflammation, which ultimately leads to the destruction of the intestinal microvilli (16, 17).

A strict gluten-free diet ameliorates the intestinal mucosa morphology in a large number of celiac patients. However, a prolonged reliance on such a diet leads to often ignored unwanted effects, such as poor gut health due to changes in microbial population of the gut (18) or increase in the body mass index due to an excess consumption of starch laden low-fiber content processed foods (6). Therefore, alternative methods are continually being sought. Among them, the processing methods include sourdough fermentation, use of gluten sequestering polymers or resins, food-grade enzymes of *Aspergillus* spp. (aspergillopepsin and dipeptidyl peptidase), microbial transglutaminase enzyme and flour derived from germinated or UV treated grains (6). Besides these processing methods, a large number of preventive methods, and intestinal barrier enhancing or immune targeted therapies are being developed and in various phases of clinical testing [cf. (19)]. Likewise, the use of "glutenases" and reduced-gluten wheat genotypes have also been subjected to testing (20–23). Among these approaches, the utilization of glutenases presents several advantages over the reduced-gluten wheat genotypes or other processing methods. Such as, with the use of glutenases, gluten contamination (at any level from farm to fork) can be handled without any extra effort. Additionally, the glutenases could be used in two ways as a food supplement ingested with or before each meal or by ectopically expressing these enzymes in the grains of plants producing gluten proteins. Also, glutenases do not lead to the production of bitter-tasting peptides often produced during the partial hydrolysis of gluten proteins via the food-grade enzymes (24, 25).

A large number of proteases of microbial, plant and synthetic origin have been proposed to be useful in reducing the content of immunogenic gluten peptides and epitopes (26–32). Most of these enzymes, such as pseudolysin (IasB) from *Pseudomonas aeruginosa* (33), nepenthesin from pitcher plants (31, 34), endopeptidase 40 from the soil actinomycete *Actinoallomurus* A8 (32), and an unknown enzyme from human salivary plaques (35), are at early stages of testing but seems to hold great promise. And a handful of these glutenases are already under advanced clinical trials (36). Monotherapies such as *A. niger* prolyl endopeptidase (AN-PEP), and a modified recombinant *Alicyclobacillus sendaiensis* endopeptidase (KumaMax, now Kuma030) (37, 38), and combination therapy (ALV003, now latiglutenase), which is a cocktail of a modified recombinant *Sphingomonas capsulate* prolyl endopeptidase (ALV002) and a barley endoprotease (ALV001), are among them. Specifically, latiglutenase is now under phase II clinical trials (36, 39). However, none of these therapies except AN-PEP with a tradename "Tolerase® G" are commercially available (40). Among these proposed treatments, the combination therapy, which relies on the action of two specifically selected peptidases with complementing properties (e.g., target specificity, substrate length, optimal pH, and site of action), offer specific advantages, i.e., capability to detoxify different gluten proteins in the diet, before they trigger an immune response in the gut.

Following the leads from the earlier research, our *in-silico* analysis (13), and *in vitro* gut simulation studies performed by (28), we expressed a combination of peptidases, a recombinant modified *Flavobacterium meningosepticum* prolyl endopeptidase (Fm-PEP) and a glutamine-specific endoprotease from barley (EP-B2) in the wheat endosperm (23). These two enzymes complement each other in their gluten processing properties, EP-B2 is a cysteine endopeptidase, which cuts at glutamine residues and prefers intact proteins as substrate. Whereas, Fm-PEP a serine endopeptidase, presents a substrate preference of 30 amino acids, cutting after proline residues. They also present complementary action in different portions of the digestive system, EP-B2 functions optimally in acidic pH ∼4 (in the stomach), and Fm-PEP prefers neutral pH (in the duodenum) (41). The initial *in vitro* studies cast some concern on pepsin sensitivity of the Fm-PEP, but later *in vivo* study in rats showed that the enzyme survives the gastric and brush border enzymes and shows up to 60% gastric activity (42).

Even though expressing glutenases in the wheat endosperm is an attractive approach, it poses a few technical challenges, such as the enzymes expressed in grains have to go through the harsh food processing conditions, specifically high temperatures, limiting the industrial application of this approach. When tested *in vitro*, the available thermostable variant of Fm-PEP (43) showed a stark decline in the activity at temperatures above 60°C and the same happens to EP-B2 at even lower temperature (above 50°C). However, under *in vivo* conditions, the propeptide of EP-B2 serves as an intramolecular chaperone, helping the enzyme refold to its native state after thermal denaturation. Also, the signal peptide-guided sequestration of both EP-B2

and Fm-PEP into protein bodies of the endosperm cells is expected to provide thermal stability to enzymes during the baking process with minimal effect on their catalytic properties. These assumptions, however, need to be tested in a real-life scale food processing experiments, which will be undertaken on the availability of the required amount of genetically stable seeds from the selected transformants.

An alternative to overcome this challenge and to ensure the enzyme stability under food processing conditions is to engineer enzymes for thermostability or to retain biological activity after being exposed to temperatures at or over 100°C, a temperature often used in industrial food processes. Therefore, this research was designed to set the basis for engineering thermostable variants of Fm-PEP and EP-B2 using a sequence guided mutagenesis approach, with a future objective to develop transgenic wheat lines expressing these enzymes in grains. Glutenases in such lines are expected to retain activity even after getting exposed to high temperatures and, upon consumption, the ability to detoxify immunogenic gluten peptides in the human gastrointestinal tract. The results of engineering glutenases and their biochemical characterization are presented in this manuscript.

MATERIALS AND METHODS

Materials

The plasmid containing barley EP-B2 in pET28 background was a gift from Dr. Chaitan Khosla of Stanford University. All reagents used in this study were analytical grade, until and unless notified and were purchased from Sigma.

Selection of Sites and Type of Changes to Induce Mutations in the Genes Encoding Fm-PEP and EP-B2

Two approaches were adopted to identify sites and types of changes to be induced in the genes encoding Fm-PEP and EP-B2. In the case of Fm-PEP, sequences of prolyl endopeptidases from thermophilic organisms were identified via BLASTP searches against NCBI non-redundant (nr) protein database, and the conserved sites in these sequences were compared with the corresponding sites in the porcine-PEP (NP_001004050.1), *Pyrococcus furiosus* (Pfu)-PEP (AAA73423.1), and Fm-PEP sequences. The clustering of sequences was performed using ClustalW (see **Figure S2**).

To identify sites and residues for mutagenesis in EP-B2, the sequence of a thermostable cysteine endoprotease, Ervatamin C, was used as a reference (44, 45). To identify corresponding sites the amino acid sequences of EP-B2 and Ervatamin C were aligned, and the three EP-B2 sites (Val34, Gly38, and Lys180) were marked to substitute, respectively with Ser, Ser, and Ala residues in the Ervatamin C sequence [**Table 1**; also see Wen, (46) for details].

TABLE 1 | List of amino acid residues selected for introducing substitutions in *Hordeum vulgare* cysteine endopeptidase B2 (EP-B2) sequence.

Organism	Enzyme	T_{max} (°C)	Amino acid locations		
			34	38	180
Ervatamia coronaria	Ervatamin C	70	$S_{(32)}$	$S_{(36)}$	$A_{(172)}$
Carica papaya	Papain	50	$V_{(32)}$	$G_{(36)}$	$K_{(174)}$
Hordeum vulgare	EP-B2	56	V	G	K

Amino acid locations are provided in accordance with the EP-B2 protein sequence. T_{max} = Temperature of maximum enzymatic activity.

Introduction of Selected Mutations in Fm-PEP and EP-B2

A codon-optimized version of Fm-PEP with a GC content of 61% was synthesized from GenScript, USA, and cloned into pUC57 vector. The gene was amplified from the plasmid (pUC57+Fmen) DNA to introduce point mutations, using specific primers (F: 5′-CGCCATATGAAGTACAACAAGCT-3′ and R: 5′-CGCGAGCTCCTACTTCAAACTCT-3′) flanked on either side by the *Nde*I and *Sac*I restriction sites to facilitate cloning. The PCR conditions used to amplify the gene fragment were as follows: initial melting at 98°C for 3 min; followed by 25 cycles at 98°C for 10 s, 63.7°C for 30 s, and 72°C for 3 min, and a final extension for 10 min at 72°C. Following PCR amplification, a 5 μl aliquot from the 25 μl reaction was loaded onto the 1% (w/v) agarose gel and electrophoresed for 60 min. After testing the product on the gel, 1 μl of it was ligated into pGEM®-T Easy vector following the manufacturer's instructions and transformed in *E. coli* DH5α cells. Positive colonies were selected by blue-white screening. The plasmid was isolated from positive colonies, and the presence of the desired plasmid was confirmed by restriction digestion with *Nde*I and *Sac*I. After electrophoresis, the insert was purified from the gel using Geneclean III kit (MP Biomedicals, USA) and ligated into the pET28b(+) vector also digested with the same restriction enzymes. The resultant plasmid was used to transform BL21(DE3) cells. The positive colonies containing the plasmid pET28b(+)(FmenWT) were selected and grown on liquid LB medium supplemented with 50 μg/ml kanamycin. Plasmid DNA was isolated using NucleoSpin Plasmid—plasmid Miniprep kit (Macherey Nagel, USA) following the manufacturer's instructions.

Primers were designed to introduce two mutations at a time in each blade of the Fm-PEP β-propeller domain [see above for the selection of target sites, **Table S1** for primer details, and Osorio, (47) for other pertinent details]. For this purpose, the gene was divided into four sections, and unique restriction sites were used to splice together desired gene fragments. Using this approach, different mutant combinations or haplotypes were created, and the assembled gene fragments were cloned in pET28b(+) backbone using the In-Fusion HD Cloning Kit (Clontech Laboratories, Mountain View, CA) (**Figure 1A**).

To create the thermostable variant of EP-B2, error-prone PCR was used following Uchiyama et al. (48). The gene was divided into three sections. Overlapping primers containing desired point mutations were designed to amplify each section [see Wen, (46)]. To obtain the complete gene sequence with desired modifications, five derived PCR products were mixed in different combinations to serve as the template in splicing by overlap extension reaction (**Figure 1B**). The first two-point mutations were close to each other, therefore, were introduced via

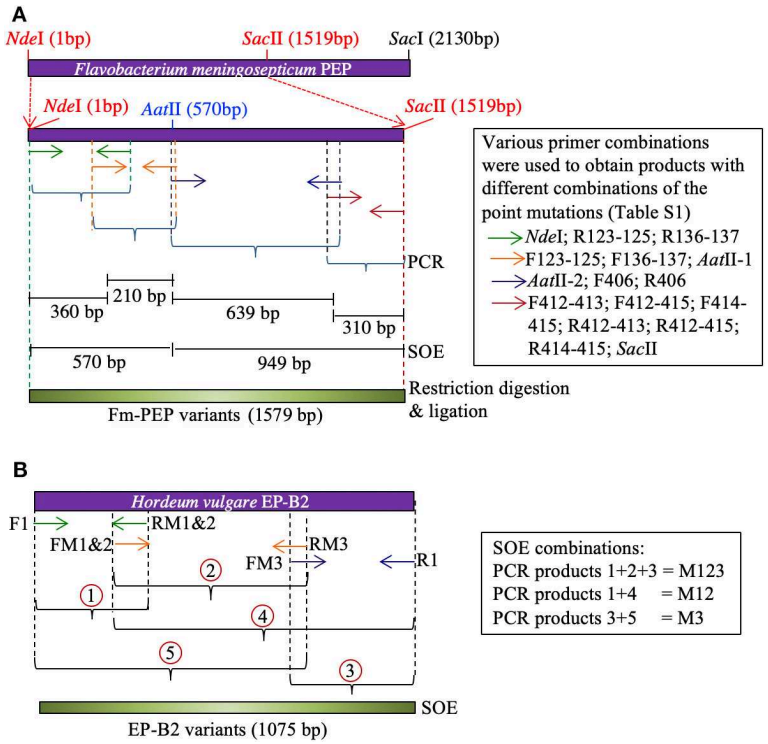


FIGURE 1 | Schematic representation of Splicing by Overlapping Extension (SOE)-PCR used as a strategy to introduce mutations in the genes encoding **(A)** *Flavobacterium meningosepticum* prolyl endoprotease, and **(B)** *Hordeum vulgare* cysteine endopeptidase B2. Corresponding primers are shown as same color arrows. For primer sequences, see **Table S1**.

a single primer. The single (M3, K180A), double (M1&2, V34S, and G38S), and triple (M1&2&3, V34S, G38S, and K180A) EP-B2 mutants thus obtained, were confirmed by Sanger sequencing of the corresponding clones.

Expression of Fm-PEP and EP-B2 Variants in *E. coli*

Colonies with the desired mutations were cultured into 5 ml LB medium with 50 μg/ml kanamycin at 37°C until OD_{600} 0.6 was reached. At that point, IPTG was added to the cultures at a final concentration of 0.25 mM and induced for another 14 h at 37°C under constant shaking (200 rpm). The cells were harvested by centrifugation at 10,000 g for 5 min, the supernatant discarded, and inclusion bodies were isolated using BugBuster™ protein extraction reagent (Novagen, USA), following the manufacturer's instructions.

The solubilization of the inclusion bodies was achieved following the protocol of Singh and Panda with minor modifications (49). Fifty microliter of inclusion body suspension at a concentration of 20 mg/ml, was solubilized in 500 μl of solubilization buffer (100 mM Tris, 2 M urea, pH 12.5). The suspension was incubated for 30 min with gentle shaking at room temperature and centrifuged at 14,000 rpm for 40 min. The supernatant was recovered, and protein concentration was measured using Bradford assay (BioRad, USA). The refolding of the extracted protein was accomplished by diluting the solubilized inclusion bodies to a final concentration of 50 μg/ml in refolding buffer pH 8.0 containing 50 mM Tris-HCl, 2 M urea, 5% sucrose, 10% glycerol, 0.5 mM EDTA, and 1 mM PMSF. Tubes were incubated overnight with gentle shaking at 4°C. The refolded sample was concentrated using Millipore concentrating devices following the manufacturer's instructions. For proEP-B2, the refolding and activation of the enzyme were accomplished as documented earlier [(41, 50); **Text S1** and **Figure S1**].

Evaluation of Thermostability of Fm-PEP and EP-B2 Variants

Initial screening of the mutants was performed to identify the best variants. Protein concentration was measured using the standard Bradford Protein Assay following the manufacturer's instructions (BioRad, USA). Absorbance was measured at 595 nm, and the concentration was calculated using bovine serum albumin (BSA; New England Biolabs, USA) as standard. Activity assays were performed in triplicate for each variant. Prolyl endopeptidase activity was evaluated using a synthetic peptide Z-Gly-Pro-pNA (Bachem, Torrance, CA) as described below: 100 μl aliquots of protein solution were incubated for 10 min at different temperatures starting from 60 to 90°C, with increments of 10°C. After heat treatment, the protein solution was added to 140 μl of PEP assay buffer (100 mM potassium phosphate, pH 7.0; 100 μg/ml BSA; 1 mM dithiothreitol; 0.2 mM Z-Gly-Pro-pNA to a final concentration of 0.02 μM) and incubated for 15 min at 30°C.

EP-B2 activity was analyzed using a synthetic peptide Z-Phe-Arg-pNA (Bachem, Torrance, CA). Protein concentration was measured and normalized as described above. For determination of mutant activity, a total volume of 35 µl of the proEP-B2 solution was heated for 10 min at temperatures ranging from 50 to 70°C, with increments of 2°C. Following heat-treatment, the proenzyme was activated to EP-B2, as described above. Peptide Z-Phe-Arg-pNA was added to a final concentration of 25 µM, and the enzyme/substrate mixture was incubated at room temperature.

In both cases (Fm-PEP and EP-B2), absorbance was measured at 410 nm every 5 min during a 5 h period. The mutant that showed enzyme-kinetics indicative of thermal stability was further analyzed. The kinetic parameters were calculated by measuring the initial velocity of the reaction, which was determined by the increase in absorbance at 410 nm. Initial velocities were plotted against substrate concentration, and K_M and k_{cat} values were calculated. The activity was calculated using the Beer-Lambert equation ($A = \varepsilon\ c\ l$), and the concentration of the product was calculated based on the extinction coefficient for pNa (8.8 mM).

A more detailed analysis was performed on the mutants that showed better performance in the initial screen. For such mutants, large-scale expression cultures in 60 ml LB medium were performed to obtain adequate quantities of the enzyme variants. After expression, the enzyme variants were retrieved in inclusion bodies, purified, and refolded. The kinetic parameters were calculated by measuring the initial velocity of the reaction, which was determined by the increase in absorbance at 410 nm. For making these calculations, the enzyme concentration was kept constant at 0.02 µM, and the substrate concentration ranged between 0.075 and 0.3 mM. Initial velocities were plotted against substrate concentration, and the slopes in each case were used to calculate K_M and K_{cat}.

Enzyme Performance Assay Using Wheat Gluten Proteins

After the characterization of enzymes using a synthetic substrate, activity was measured against wheat gluten standard procured from the National Institute of Standards & Technology (Canada). To study the effect of digestion on individual prolamin groups, 25 mg of gluten reference material was fractionated into gliadins and glutenins using the stepwise gluten extraction procedure described in Wen et al. (20). Digestion of gliadins and glutenins was performed stepwise, as described below. First, the gliadin and glutenin fractions at the final concentration of 25 mg/ml, were digested with pepsin (0.6 mg/ml) and pre-activated EP-B2 (0.375 mg/ml) under simulated gastric conditions (50 mM sodium acetate buffer, pH 4.5) for 60 min. The tube was incubated with gentle shaking (50 rpm) at 37°C. Second, pH of the solution was adjusted to 6.0 with 500 mM sodium phosphate buffer, and intestinal proteases, trypsin (0.375 mg/ml), chymotrypsin (0.375 mg/ml), elastase (0.075 mg/ml), and carboxypeptidase A (0.075 mg/ml) were added to the solution. Third, the prolyl endopeptidase variant (pre-exposed to 90°C for 10 min) was added to the solution at a concentration of 0.375 mg/ml, and the mixture was incubation at 37°C for 60 min with gentle shaking. Finally, the reaction was stopped by incubation of the tubes for 10 min at 100°C. Later, to clarify the solution, digests were centrifuged for 10 min at 9,300 g and filtered through 0.45 µm filters. Tricine-PAGE was used to analyze the digestion product using the protocol described in Schagger (51). After electrophoresis, each gel was subject to densitometric analysis using a Personal Densitometer SI, Model 375A (Molecular Dynamics), and the resulting images were analyzed using ImageJ software v 1.47.

Reverse-Phase High-Performance Liquid Chromatography (RP-HPLC)

The HPLC separations were performed using a C8 reversed-phase analytical column (Zorbax 300SB-C8, Agilent Technologies) with 5 µm particle size and 30 nm microporous silica diameter (250 mm length, 4.6 mm inner diameter) and a C18 reversed-phase analytical column (Eclipse Plus C18, Agilent Technologies) with 5 µm particle size, 150 mm length, and 4.6 mm inner diameter. Both column types were used on a 1200 Series Quaternary HPLC-System (Agilent Technologies) with a diode array UV-V detector. During the runs, the column temperature was maintained at 60°C. A linear elution gradient was implemented using two mobile solvents, the polar solvent A consisting of 0.1% trifluoroacetic acid (TFA) (vol/vol) in type I ultrapure water (18 MΩ·cm specific resistance), and the non-polar solvent B containing 0.1% TFA (vol/vol) and acetonitrile (ACN). Absorbance was monitored at a detection wavelength of 210 nm, and the flow rate was maintained at 1.0 ml min^{-1} on the C8 column and 0.5 mL min^{-1} on the C18 column. In the case of C8 column, the elution gradient conditions were selected as follows: for gliadins, a linear gradient from 20 to 60% B in 60 min, and for glutenins, from 0 to 24% B in 20 min followed by 24–60% B in 40 min. After each run, the column was cleared with 90% B for 3 min and equilibrated with the starting B concentration for 5 min. In the case of the C18 column, the elution gradient condition was: a linear gradient from 0 to 50% B in 30 min. After each run, the column was cleared by linearly decreasing solvent B to 0% in 4 min, and then the column was equilibrated with the starting B concentration for 10 min.

RESULTS AND DISCUSSION

Multiple Sequence Alignment and Phylogenetic Analysis of the Prolyl Endopeptidase Family

Prolyl endopeptidases can be found in archaeal, bacterial, and eukaryotic species (52). But for the purpose of this study, enzymes were selected for further analysis from the thermophilic organisms (**Table S2**). The BLASTP searches were performed against the NCBI nr database using the Pfu-PEP sequence as a query, and 16 PEP sequences from thermophilic organisms were identified. The range of survival temperatures for these thermophiles varied from 45 to 100°C. Among these thermophiles, members of the *Pyrococcus* family thrive at a temperature above 80°C and hence being classified as

hyperthermophile or extremophiles. The alignment of the protein sequence was achieved using ClustalW, and as expected, a significant similarity between the protein sequences was observed (**Figure S2**). In the catalytic domain (between residues 1–73 and 428–710), 49 residues were 100% conserved among these sequences. Whereas, in the β-propeller domain, only two residues were 100% conserved, which is not surprising, because, in the *Pyrococcus* family, the β-propeller was not reported to be involved in the filtering of peptides large than the optimal substrate size (53). Thus, due to lack of selection pressure, it was expected to evolve faster than other parts of the enzyme, which explains the observed levels of diversity (52). Considering the high levels of diversity in the β-propeller domain, it is very likely that the variations contributing to the differences in the thermostabilities of different PEPs also lie in this domain. The observed level of similarity ranged from 22.28 to 41.11% between pairwise comparisons of different enzymes with FmPEP. The lowest level of homology ranging from 22.28 to 27.76% was observed between the prolyl endopeptidases from *F. meningosepticum* and *Pyrococcus*/*Thermococcus* families. Interestingly, when the optimum growth temperatures for organisms used to obtain these sequences were checked, a decrease in growth temperature below 80°C was found associated with the increase in percentage homology in sequences. This trend persisted up to 41.11% sequence similarity, as was observed in the comparison between Fm-PEP with *Deinococcus radiodurans* PEP.

The conservation of residues at the substrate-binding site was also studied (**Table S3**). The catalytic triad and the specificity pocket S1 showed complete conservation. The similarity between residues decreases in the case of the specificity pocket 3 (S3), because of its overlapping location with the β-propeller domain. At this site conservation of the residues varied between 27.8 and 77.8%. These changes have an influence on the activity of the enzyme and its substrate-binding properties (52). An example of such variability is residue 252. In the case of Fm-PEP, this amino acid location is occupied by tyrosine, whereas, in the PEPs of the *Pyrococcus* family, the same position is held by phenylalanine. In the case of *Pyrococcus* PEP, this change in the amino acid residue was found to be associated with the low turnover of the enzyme under experimental conditions, which might also be the case for other archaeal PEPs carrying phenylalanine at amino acid (aa) location 252. Not all PEP sequences derived from the thermophilic organisms have phenylalanine at position 252 and were also documented to have high cleavage efficiencies. It has been reported that members of the *Pyrococcus* family have the ability to hydrolyze proteins like azocasein (23.6 kDa), which was rather surprising for a prolyl endopeptidase. These observations led to the conclusion that the β-propeller domain of *Pyrococcus* PEP favors it opening on exposure to high temperatures, which also contributes to its autoproteolytic properties (54, 55).

The multiple sequence alignment of 16 PEPs formed the basis of constructing a phylogenetic tree using the distance- and character-based methods (**Figure S3**). The high conservation between the PEPs derived from the hyperthermophilic organisms suggested a close common ancestry. It is possible to draw conclusions about the structure-function relationships of PEPs by looking at the phylogenic relationships of the analyzed thermophiles. Also, it is possible to assume that the adaptation to different environmental conditions might have favored the selection of specific changes in the sequence of the β-propeller domain. The same is also true for the observed conservation of residues specific for the catalytic activity. An example of this situation is the high sequence similarity observed between the Fm-PEP and the PEP sequences of the *Deinococcus* family, and the fact that both bacterial species are adapted to mild environmental conditions.

Site-Directed Mutagenesis and Analysis of Thermostability of Enzyme Variants

Flavobacterium meningosepticum Prolyl Endopeptidase (Fm-PEP)

Conserved amino acid residues from the β-propeller domain in the *Pyrococcus* family were selected for mutagenesis (**Table 2**). Nine sites were identified to introduce point mutations leading to amino acid substitutions in Fm-PEP, to create a library of 59 enzyme variants with differences in number and order of substitutions (**Table S4**).

Residues 123 and 125 were, respectively mutated from Arginine and Aspartic acid to Isoleucine and Tryptophan. In this case, the arginine to isoleucine substitution, favored interaction by van der Waals forces with the residues in proximity. On the other hand, the substitution of Aspartic acid to Tryptophan, a non-polar and hydrophobic amino acid, resulted in increased interactions by the formation of hydrogen bonds, aromatic stacking, and increase of van der Waals force.

The second set of substitutions was made at amino acid residues Proline and Asparagine located at positions 136 and 137. In this case, the first amino acid substitution of Proline to Glutamic acid, which can interact ionically, showed the formation of four hydrogen bonds and also van der Waals interactions. The amino acid residue at position 137 was substituted with Leucine. This change did not affect the flexibility of the peptide chain but enhance van der Waals interactions. Leucine does not form hydrogen bonds, because of the conformation of its side chains that are arranged in "L" shape in contrary to the "Y" shape of side chains in Isoleucine, which is often found at the core of protein folds.

The first mutation on the seventh blade was the change of residue 406 from Asparagine to Serine. In this case, the hydroxyl group of Serine conferred hydrogen bonding potential, and this amino acid substitution also increased the possibility of van der Waals interacts with close by residues.

The amino acid substitution at residue 412 from Tyrosine to Arginine increased the possibility of forming hydrogen bonds at seven potential sites and also increased the chances of ionic and van der Waals interactions. The next amino acid substitution introduced at residue 413 from Isoleucine to Leucine also increased the flexibility of interactions, especially by van der Waals forces.

The last couple of amino acid substitutions were introduced at positions 414 and 415. In the first case, Phenylalanine was replaced with Tyrosine, which potentially formed three hydrogen

TABLE 2 | List of amino acid residues selected for introducing sequence guided substitutions in *Flavobacterium meningosepticum* prolyl endopeptidase gene sequence.

Organism	Optimum growth temperature (°C)	Amino acid locations*								
		123	125	136	137	406	412	413	414	415
Sulfolobus tokodaii	80	L	Y	K	T	S	R	V	V	K
Metallosphaera sedula	75	L	R	N	I	S	T	I	S	R
Thermococcus barophilus	48–95	I	L	K	L	S	R	L	Y	R
Thermococcus sibiricus	60–84	I	V	T	L	S	R	L	Y	K
Thermococcus kodakarensis	86	I	W	A	L	S	R	L	Y	Q
Thermococcus onnurineus	80–90	I	W	R	L	S	R	L	Y	E
Thermococcus gammatolerans	88	I	W	E	L	S	R	L	Y	E
Pyrococcus yayanosii	98	I	W	K	L	S	R	L	Y	E
Pyrococcus furiosus	100	I	W	E	L	S	R	L	Y	E
Pyrococcus horikoshii	98	I	W	E	I	S	R	L	Y	E
Pyrococcus abyssi	102	V	W	E	L	S	R	I	Y	E
Aciduliprofundum boonei	70	V	N	D	L	S	R	L	Y	E
Sphaerobacter thermophiles	65	L	E	P	N	S	T	V	F	Q
Haladaptatus paucihalophilus	45	V	D	P	N	S	T	V	Y	R
Deinococcus radiodurans	Variable	T	D	P	N	S	R	P	Y	R
Deinococcus maricopensis	45	V	D	A	N	S	T	P	H	H
Flavobacterium meningosepticum	37	R	D	P	N	N	T	I	F	K
Natranaerobius thermophiles	57	F	E	P	N	T	T	I	L	R

**Conserved amino acid residues are highlighted in red font.*

bonds and interacted by aromatic stacking as well as van der Waals forces. And as of the last mutation, Lysine residue at position 415 was replaced by Glutamic acid.

After the successful cloning of the mutants in the expression vector, DNA sequencing was performed to assure that the mutations were successfully introduced in the gene sequence. Initial screening of mutants was performed using a synthetic substrate to evaluate their activity after heat shock. Out of the 59 evaluated mutants, five showed significantly higher activity than the control at 80°C (**Table S4**). Mutations stacked at the amino acid positions 123, 125, 136, and 137 showed a detrimental effect on the stability of the enzyme at 80°C. Albeit, a combination of mutations involving residues at 412–415 showed stability at higher temperatures likely due to the formation of the hydrogen bond, witnessed in the *in-silico* structural analysis of mutants (**Figure S4**).

Based on the initial thermal analysis, five mutants, namely Fme5, Fme6, Fme10, Fme16, and Fme18, were exposed to 90°C for 10 min. After heat-shock, enzyme activity was measured, and only three mutants (Fme5, Fme6, and Fme10) retained their ability to cleave the synthetic substrate. Therefore, these mutants were selected for enzyme kinetics analysis. After heat shock at 60 or 90°C, k_{cat}, K_M, and k_{cat}/K_M values were determined for each Fm-PEP variant (**Table 3**). A combination of mutations involving amino acid residues at 412, 413, 414, and 415 increased thermostability of the enzyme (**Table 3**).

TABLE 3 | Kinetic parameters of Fm-PEP variants determined using a synthetic substrate, Z-Gly-Pro-pNA.

Enzyme	60°C k_{cat}/min^{-1}	90°C k_{cat}/min^{-1}	60°C K_M (μM)	90°C K_M (μM)	60°C k_{cat}/K_M	90°C k_{cat}/K_M
Wild type[a]	60.3898	2.9004	0.2034	0.1135	296.9157	25.5503
Fme5[b]	14.7133	7.49034	0.09896	0.1113	148.6823	67.3192
Fme6[c]	18.5709	5.4641	0.09655	0.0861	192.3406	63.4818
Fme10[d]	4.8203	5.8314	0.0418	0.1419	115.3152	41.0920

Enzymatic activity was measured after a heat-shock at 60 and 90°C for 10 min. Absorbance at 410 nm was recorded. The Beer-Lambert equation ($A = \varepsilon\, c\, l$) and the concentration of the product was calculated based on the extinction coefficient for pNa (8.8 mM).

[a]Corresponds to wild type prolyl endopeptidase from Flavobacterium meningosepticum used as a negative control.

[b]Corresponds to variant Fme5 having mutations at residues 412 and 413.

[c]Corresponds to variant Fme6 having mutations at residues 414 and 415.

[d]Corresponds to variant Fme10 having mutations at residues 406, 412, 413, 414, and 415.

Hordeum vulgare Endoprotease B2 (EP-B2)

Mutations in the EP-B2 gene sequence were introduced as described earlier, and the point mutations were confirmed by DNA sequencing (**Figure 2**). SDS-PAGE was used to confirm the expression of EP-B2 in *E. coli* (strain BL21) and *in-vitro* enzyme activity assay using a synthetic substrate, Z-Phe-Arg-pNA to confirm refolding as well as activation. The effect of temperature on enzyme activity was also determined using synthetic substrate, Z-Phe-Arg-pNA via spectrophotometry (**Figure S5**). The heat-shock analysis of the EP-B2 variants (with one, two, and three mutations) suggested that the enzyme variant with a single substitution at amino acid location 180, dubbed M3 (K180A), retained activity up to 60°C (T_{max}). In comparing with the wild type, the mutant showed an increase in thermostability by 4°C. On the other hand, the T_{max} of M1&2 and M1&2&3 did not

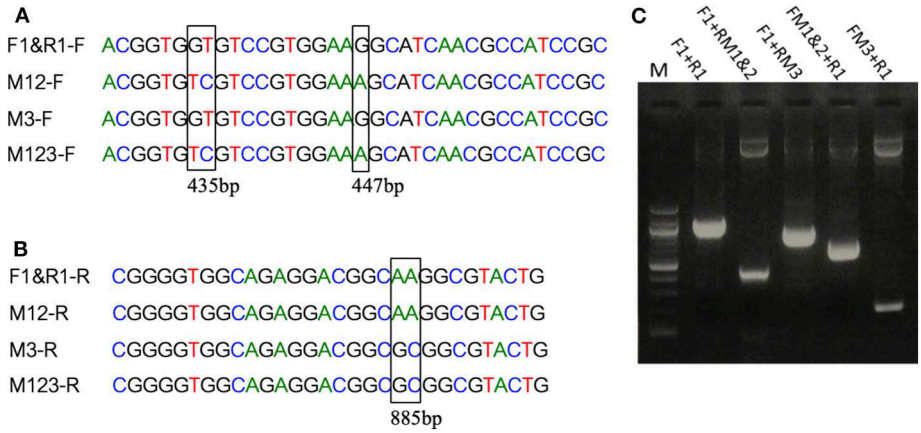


FIGURE 2 | Introduction of mutations in EP-B2 by error-prone PCR. **(A,B)** Results of DNA sequencing showing the location of the desired point mutations (marked by a rectangle in the sequence alignment). M1 = GT to TC transversions, M2 = G to A transition and M3 = A to G transition, M12 represents a combination of M1 and M2, and M123 a combination of M1, M2, and M3. F1&R1 represents the wildtype sequence. F and R in the sequence names designate DNA sequences in Forward (+) and Reverse (–) orientations. **(C)** Agarose gel analysis of the PCR products obtained using various combination of primers (for primer sequences, see **Table S1**). Different sets of primers were used in PCR to obtain different combinations of mutations and thus various EP-B2 variants.

show a noticeable improvement. In fact, the activities of the M1&2 (V34S and G38S) and M1&2&3 (V34S, G38S, and K180A) were slightly lower than the wild type. This could be explained because of amino acid substitutions affected protein structure, disturbing binding sites, and thus reducing catalytic activity. Although the T_{max} of M3 was higher than the control, the activity of the enzyme declined with the increase in temperature, as can be witnessed in **Figure 3** and **Table 4**. In sum, site-directed mutagenesis has the potential to increase EP-B2 thermostability. However, the 4°C increase in thermostability observed for the M3 (K180A) variant is not sufficient, and a further increase in it is needed, which could be achieved probably by the next round of directed evolution.

Testing the Activity of Thermostable Fm-PEP Variants Using Gliadins and Glutenins

To further affirm the activity of the thermostable Fm-PEP variants against the intended substrate after heat shock treatment, the enzyme was supplied with gluten protein fractions under simulated gastrointestinal conditions. For this purpose, the prolamins were extracted as described in materials and methods, and the resultant protein fractions were used as the substrate for the enzymatic digestion. Each fraction, gliadin and glutenin was subject to three treatments: (i) Gastric and pancreatic enzymes (pepsin, trypsin, chymotrypsin, elastase, and carboxypeptidase A); (ii) Gastric and pancreatic enzymes plus activated EP-B2; and (iii) Combination of gastric and pancreatic enzymes, activated EP-B2, and either of the three heat-treated Fm-PEP variants or its wild type version (negative control). And after each treatment, the digests were analyzed by RP-HPLC on the C18 column. Comparative analysis of the treatments comprising only gastric enzymes with the one with the addition of glutenases (**Figure 4A**) showed evidence of gliadin degradation,

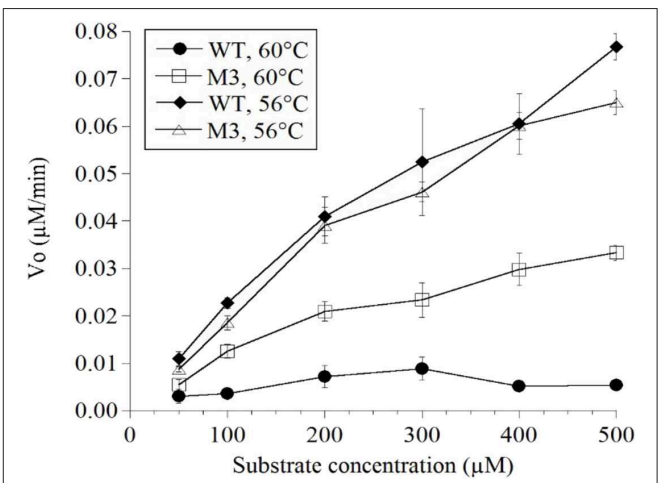


FIGURE 3 | Determination of the thermostability of wild type (WT) control and EP-B2 variant dubbed M3 (K180A) using a synthetic substrate, Z-Phe-Arg-pNA. Enzyme activity was measured at 56 and 60°C, and was compared with the wild type control. To monitor enzyme activity absorbance was recorded at 410 nm. The Beer-Lambert equation ($A = \varepsilon\, c\, l$) and the concentration of the product was calculated based on the extinction coefficient for pNa (8.8 mM). Concentration is expressed as μM/min.

where degraded peptides appeared between 18 and 24 min elution time on chromatograms, with only a small amount of peptides (potentially immunogenic hydrophobic large peptides) appeared after 30 min elution time. The combination of EP-B2 (wild type-control) and heat-treated Fm-PEP variant "Fme5" successfully detoxified the gluten reference material. It suggested that essentially, the enzymatic treatment with glutenases has detoxified the majority of the immunogenic peptides, based on the fact that large hydrophobic peptides elute late when the column conditions become more non-polar. The results of the

TABLE 4 | Kinetic parameters of EP-B2 variant "M3" and the wild type (WT) control was determined using a synthetic substrate, Z-Phe-Arg-pNA.

	M3, 56°C	M3, 60°C	WT, 56°C	WT, 60°C
V_{max} (μM/min)	0.236	0.0506	0.142	0.00746
K_M (μM)	1188	318.3	530.0	104.0

Enzymatic activity was measured after a heat-shock at 56 and 60°C for 10 min. Absorbance was recorded at 410 nm. The Beer-Lambert equation ($A = \varepsilon\ c\ l$) and the concentration of the product were calculated based on the extinction coefficient for pNa (8.8 mM).

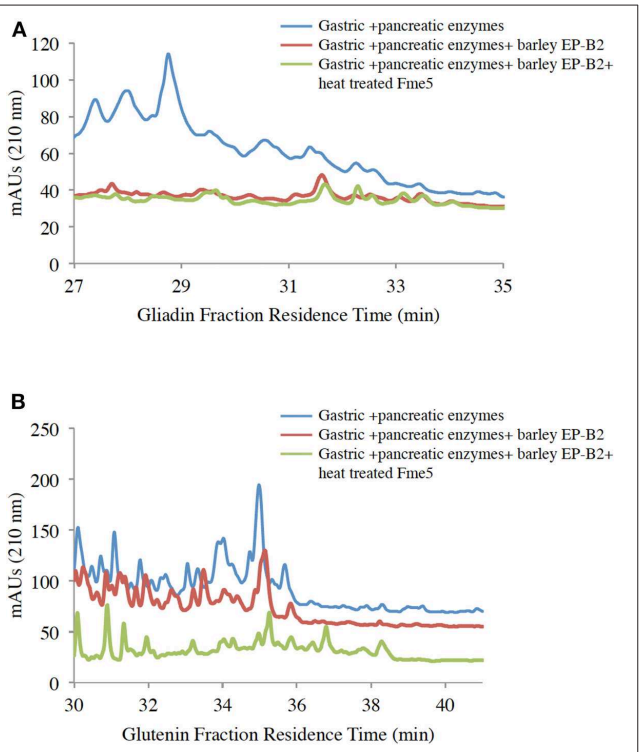


FIGURE 4 | RP-HPLC (reversed-phase high-performance liquid chromatography) analysis of wheat gliadin **(A)** and glutenin **(B)** fractions. Before HPLC, gliadins were sequentially treated with digestive enzymes (pepsin, trypsin, chymotrypsin, elastase, and carboxypeptidase **(A)**, EP-B2, and Fm-PEP variant "Fme5." Subsequently, pancreatic enzymes were added to the solution after pH adjustment. The control was wheat gliadins digested only with gastric and pancreatic enzymes (blue line), and the treatments were wheat gliadin digestion with gastric and pancreatic enzymes supplemented with (i) EP-B2 (red line), or (ii) EP-B2 and heat-treated (90°C for 10 min) Fm-PEP (green line).

analysis showed correspondence with the earlier studies (29), which showed gluten degradation to ∼28-mer peptides by the addition of EP-B2. It was also shown that the most immunogenic gluten peptides elute around 22 min (56) using a similar gradient of polar and non-polar solvents in RP-HPLC on the column type similar to the one used in the present study. These earlier studies further support our conclusion that the thermostable variants of Fm-PEP retain the capability to degrade immunogenic peptides after heat-treatment.

The glutenin fraction was also analyzed (**Figure 4B**) using RP-HPLC. Reduction in the content of immunogenic peptides appearing between 28 and 35 min in chromatograms of EP-B2 or EP-B2 plus Fm-PEP (variant) treated glutenin samples was observed. These observations were supported by the results of Gass et al. (30), who reported complete degradation of immunogenic peptides by sequential treatment of gluten samples with EP-B2 and *Myxococcus xanthus* PEP under simulated gastric conditions.

Densimetric analysis of digested gluten proteins resolved on tricine-PAGE gels also supported the observations made using RP-HPLC. Based on cumulative area estimations under each peak on the densitogram and using a BSA standard curve as a reference (57), amounts of loaded proteins were determined. Digestion with gastric and pancreatic enzymes resulted in a decrease in the amount of proteins from 344.54 to 60.69 μg/ml. Similarly, the addition of EP-B2 decreased the amount further to 48.08 μg/ml. The addition of native Fm-PEP did not result in any further decrease in the concentration (up to 48.08 μg/ml). Interestingly, the addition of prolyl endopeptidase variants numbered 5, 6, and 10 (Fme5, Fme6, and Fme10) resulted in a reduction of protein amount from 48.08 μg/ml to 25.54, 17.61, and 20.28 μg/ml, respectively (**Figure 5**, panel I). However, in the case of the glutenin fraction, no difference in terms of protein amount could be detected by the addition of prolyl endopeptidase in the fraction (**Figure 5**, panel II). These observations correspond well with the results of *in silico* analysis. Specifically in terms of the activity of endopeptidase on different gluten fractions, showing better cleavage of the gliadins than glutenins. It could be explained in terms of the size of the peptides produced after EP-B2 treatment, which makes them suitable for PEP-activity. In sum, these results suggested that Fm-PEP variants retain their activity after a heat-treatment at 90°C for 10 min and are capable of detoxifying the gluten protein to near-completion. This property will add value to the glutenases, as it will allow their expression in grain and use of such grains in the food industry without loss of enzyme activity due to thermal denaturation.

CONCLUSIONS

Since a combination of prolyl endopeptidase and barley cysteine endoprotease has been proven to be effective in detoxifying gluten proteins, efforts were made to increase their thermostability, to make them suitable for industrial applications. Site saturation mutagenesis effectively increased the thermostability of Fm-PEP, which should be sufficient to maintain the activity of the enzyme in the core of bread, where temperatures generally do not exceed 100°C. Given, accumulation of enzymes in protein storage bodies in endosperm cells and the use of "glutenase" expressing grains as such or cracked grains in bread is expected to provide further insulation from thermal denaturation. On the other hand, directed mutagenesis of EP-B2 resulted in the limited increase in the thermostability of the enzyme, however insufficient in combination with the inherent properties of the EP-B2 pro-peptide as a molecular chaperone and its accumulation in protein

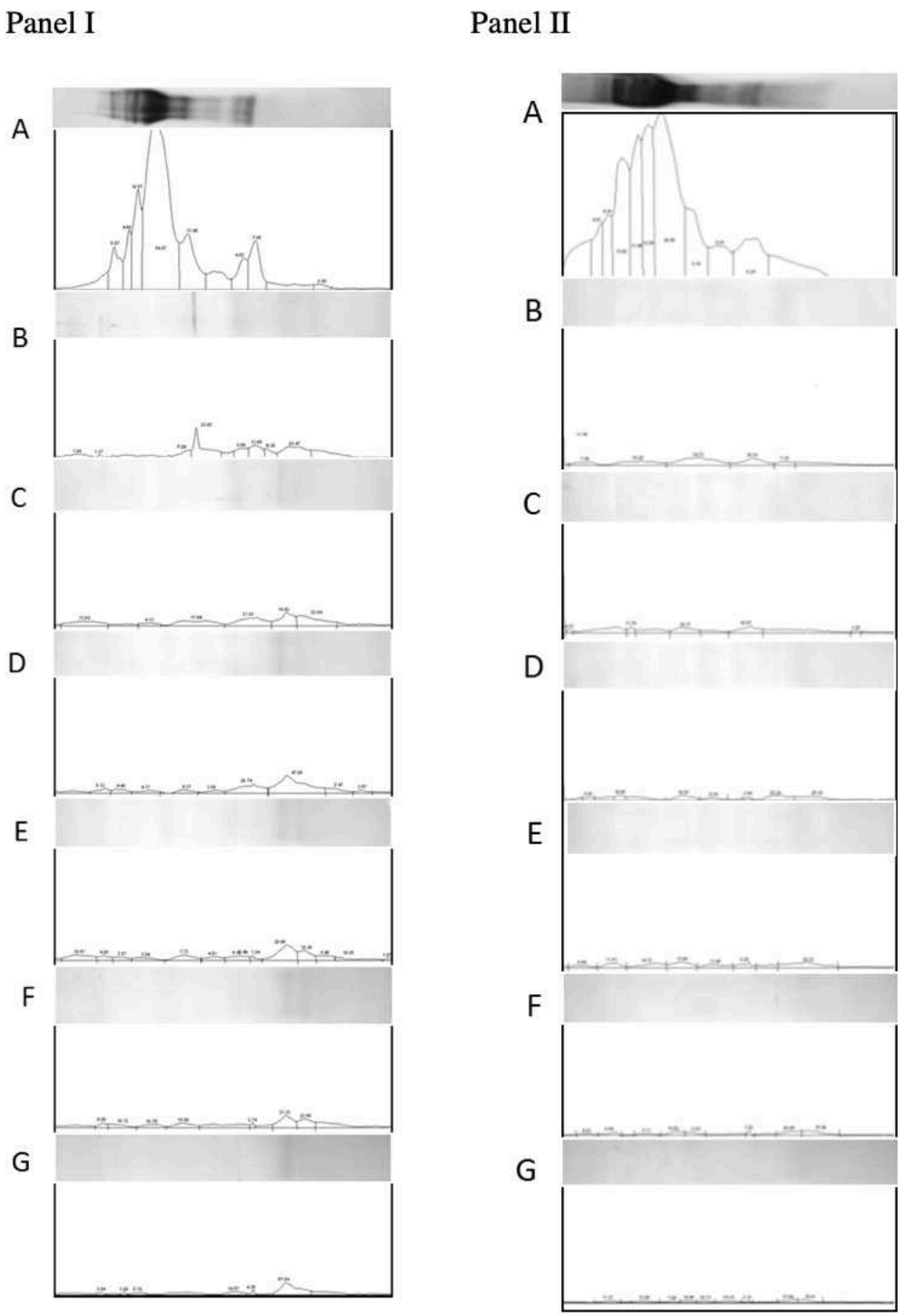


FIGURE 5 | Densitometric analysis of gliadin (panel I) and glutenin (panel II) fractions loaded on Tricine-PAGE. **(A)** Undigested gliadins/glutenins (negative control). **(B)** Gliadins/glutenins digested with gastric and pancreatic enzymes. **(C)** Gliadin/glutenin fraction digested with gastric-pancreatic enzymes in combination with barley cysteine endoprotease B2 (EP-B2). **(D)** Gliadin/glutenin fraction digested with gastric-pancreatic enzymes, EP-B2, and native *Flavobacterium meningosepticum* prolyl endopeptidase (Fm-PEP) (positive control). **(E)** Gliadin/glutenin fraction digested with gastric-pancreatic enzymes, EP-B2, and Fm-PEP variant, Fme5. **(F)** Gliadin/glutenin fraction digested with gastric-pancreatic enzymes, EP-B2, and Fm-PEP variant, Fme6. **(G)** Gliadin/glutenin fraction digested with gastric-pancreatic enzymes, EP-B2, and Fm-PEP variant, Fme10.

storage bodies is expected to provide further encapsulation and resistance from melting. Still, a further increase in the thermostability of EP-B2 is desirable to seek full advantage of the complimentary digestive properties of Fm-PEP and EP-B2 on gluten detoxification, especially in processed foods. Therefore, research is being conducted to mutagenize the gene encoding EP-B2 further. Once both thermostable enzymes have been developed, wheat grains expressing a combination of these glutenases could be produced, which will constitute an alternative for the treatment of celiac disease.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

CO, SR, and DW contributed the conception and design of the study. CO, NW, JM, SM, and SR performed the experiments. CO and SR wrote the first draft of the manuscript. SR edited the manuscript.

FUNDING

This work was supported by the NIH grants 1R01 GM080749-01A2, Clemson Faculty Succeeds grant 15-202-EQUIP-5701-430-1502211 and NIFA Hatch/Multi-state grant (S009).

ACKNOWLEDGMENTS

The authors would like to thank Dr. Rhoda Brew-Appiah for laboratory assistance. The research presented in this manuscript is part of the Ph.D. dissertations of the co-authors, CO and NW.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2020.00011/full#supplementary-material

REFERENCES

1. Singh P, Arora A, Strand TA, Leffler DA, Catassi C, Green PH, et al. Global prevalence of celiac disease: systematic review and meta-analysis. *Clin Gastroenterol Hepatol.* (2018) 16:823–36.e2. doi: 10.1016/j.cgh.2017.06.037
2. Verma A, Gatti S, Galeazzi T, Monachesi C, Padella L, Baldo G, et al. Gluten contamination in naturally or labeled gluten-free products marketed in Italy. *Nutrients.* (2017) 9:E115. doi: 10.3390/nu9020115
3. Howitt CA, Larkin PJ, Colgrave ML. Gluten reduction strategies for wheat and barley. *Cereal Food World.* (2018) 65:184–7. doi: 10.1094/CFW-63-5-0184
4. Rubio-Tapia A, Hill ID, Kelly CP, Calderwood AH, Murray JA. ACG clinical guidelines: diagnosis and management of celiac disease. *Am J Gastroenterol.* (2013) 108:656–76. doi: 10.1038/ajg.2013.79
5. Wolf RL, Lebwohl B, Lee AR, Zybert P, Reilly NR, Cadenhead J, et al. Hypervigilance to a gluten-free diet and decreased quality of life in teenagers and adults with Celiac disease. *Dig Dis Sci.* (2018) 63:1438–48. doi: 10.1007/s10620-018-4936-4
6. Rustgi S, Shewry P, Brouns F, Deleu L, Delcour JA. Wheat seed proteins: factor's influencing their content, composition, and technological properties, and strategies to reduce adverse reactions. *Compr Rev Food Sci Food Saf.* (2019) 18:1751–69. doi: 10.1111/1541-4337.12493
7. Cadenhead JW, Wolf RL, Lebwohl B, Lee AR, Zybert P, Reilly NR, et al. Diminished quality of life among adolescents with coeliac disease using maladaptive eating behaviours to manage a gluten-free diet: a cross-sectional, mixed-methods study. *J Hum Nutr Diet.* (2019) 32:311–20. doi: 10.1111/jhn.12638
8. Shan L, Molberg O, Parrot I, Hausch F, Filiz F, Gray GM, et al. Structural basis for gluten intolerance in celiac sprue. *Science.* (2002) 297:2275–9. doi: 10.1126/science.1074129
9. Tatham A, Buonocore F, Bertini L, Ronchi C, Wellner K, Belton P, et al. Structures and functional properties of wheat gluten proteins. *Cereal Foods World.* (1997) 42:663.
10. Shewry PR, Hey SJ. The contribution of wheat to human diet and health. *Food Energy Secur.* (2015) 4:178–202. doi: 10.1002/fes3.64
11. Shewry PR. What is gluten-why is it special? *Front Nutr.* (2019) 6:101. doi: 10.3389/fnut.2019.00101
12. Stern M, Ciclitira PJ, van Eckert R, Feighery C, Janssen FW, Mendez E, et al. Analysis and clinical effects of gluten in coeliac disease. *Eur J Gastroenterol Hepatol.* (2001) 13:741–7. doi: 10.1097/00042737-200106000-00023
13. Osorio C, Wen N, Gemini R, Zemetra R, von Wettstein D, Rustgi S. Targeted modification of wheat grain protein to reduce the content of celiac causing epitopes. *Funct Integr Genomics.* (2012) 12:417–38. doi: 10.1007/s10142-012-0287-y
14. Molberg O, McAdam SN, Korner R, Quarsten H, Kristiansen C, Madsen L, et al. Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. *Nat Med.* (1998) 4:713–7. doi: 10.1038/nm0698-713
15. Quarsten H, McAdam SN, Jensen T, Arentz-Hansen H, Molberg O, Lundin KEA, et al. Staining of celiac disease-relevant T cells by peptide-DQ2 multimers. *J Immunol.* (2001) 167:4861–8. doi: 10.4049/jimmunol.167.9.4861
16. Alaedini A, Green PH. Narrative review: celiac disease: understanding a complex autoimmune disorder. *Ann Intern Med.* (2005) 142:289–98. doi: 10.7326/0003-4819-142-4-200502150-00011
17. Brouns F, Van Rooy G, Shewry P, Rustgi S, Jonkers D. Adverse reactions to wheat or wheat components. *Compr Rev Food Sci Food Saf.* (2019) 18:1437–52. doi: 10.1111/1541-4337.12475
18. Chibbar R, Dieleman LA. The gut microbiota in celiac disease and probiotics. *Nutrients.* (2019) 11:E2375. doi: 10.3390/nu11102375
19. Tye-Din JA, Galipeau HJ, Agardh D. Celiac disease: a review of current concepts in pathogenesis, prevention, and novel therapies. *Front Pediatr.* (2018) 6:350. doi: 10.3389/fped.2018.00350
20. Wen S, Wen N, Pang J, Langen G, Brew-Appiah RA, Mejias JH, et al. Structural genes of wheat and barley 5-methylcytosine DNA glycosylases and their potential applications for human health. *Proc Natl Acad Sci USA.* (2012) 109:20543–8. doi: 10.1073/pnas.1217927109
21. Gil-Humanes J, Pistón F, Altamirano-Fortoul R, Real A, Comino I, Sousa C, et al. Reduced-Gliadin wheat bread: an alternative to the gluten-free diet for consumers suffering gluten-related pathologies. *PLoS ONE.* (2014) 9:e90898. doi: 10.1371/journal.pone.0090898
22. Rosell CM, Barro F, Sousa C, Mena MC. Cereals for developing gluten-free products and analytical tools for gluten detection. *J Cereal Sci.* (2014) 59:354–64. doi: 10.1016/j.jcs.2013.10.001
23. Osorio CE, Wen N, Mejias JH, Liu B, Reinbothe S, von Wettstein D, et al. Development of wheat genotypes expressing a glutamine-specific endoprotease from barley and a prolyl endopeptidase from *Flavobacterium meningosepticum* or *Pyrococcus furiosus* as a potential remedy to celiac disease. *Funct Integr Genomics.* (2019) 19:123–36. doi: 10.1007/s10142-018-0632-x

24. Liu B-Y, Zhu K-X, Guo X-N, Peng W, Zhou H-M. Changes in the enzyme-induced release of bitter peptides from wheat gluten hydrolysates. *RSC Adv.* (2016) 6:102249–57. doi: 10.1039/C6RA22155F
25. Liu B-Y, Zhu K-X, Peng W, Guo X-N, Zhou H-M. Effect of sequential hydrolysis with endo-and exo-peptidase on bitterness properties of wheat gluten hydrolysates. *RSC Adv.* (2016) 6:27659–68. doi: 10.1039/C5RA28171G
26. Piper JL, Gray GM, Khosla C. Effect of prolyl endopeptidase on digestive-resistant gliadin peptides *in vivo*. *J Pharmacol Exp Ther.* (2004) 311:213–9. doi: 10.1124/jpet.104.068429
27. Shan L, Martin T, Sollid LM, Gray GM, Khosla C. Comparative biochemical analysis of three bacterial prolyl endopeptidases: implications for coeliac sprue. *Biochem J.* (2004) 383:311–8. doi: 10.1042/BJ20040907
28. Gass J, Vora H, Bethune MT, Gray GM, Khosla C. Effect of barley endoprotease EP-B2 on gluten digestion in the intact rat. *J Pharmacol Exp Ther.* (2006) 318:1178–86. doi: 10.1124/jpet.106.104315
29. Siegel M, Bethune MT, Gass J, Ehren J, Xia J, Johannsen A, et al. Rational design of combination enzyme therapy for celiac sprue. *Chem Biol.* (2006) 13:649–58. doi: 10.1016/j.chembiol.2006.04.009
30. Gass J, Bethune MT, Siegel M, Spencer A, Khosla C. Combination enzyme therapy for gastric digestion of dietary gluten in patients with celiac sprue. *Gastroenterology.* (2007) 133:472–80. doi: 10.1053/j.gastro.2007.05.028
31. Rey M, Yang M, Lee L, Zhang Y, Sheff JG, Sensen CW, et al. Addressing proteolytic efficiency in enzymatic degradation therapy for celiac disease. *Sci Rep.* (2016) 6:30980. doi: 10.1038/srep30980
32. Cavaletti L, Taravella A, Carrano L, Carenzi G, Sigurtà A, Solinas N, et al. E40, a novel microbial protease efficiently detoxifying gluten proteins, for the dietary management of gluten intolerance. *Sci Rep.* (2019) 9:1–11. doi: 10.1038/s41598-019-48299-7
33. Wei G, Tian N, Valery AC, Zhong Y, Schuppan D, Helmerhorst EJ. Identification of pseudolysin (lasB) as an aciduric gluten-degrading enzyme with high therapeutic potential for Celiac disease. *Am J Gastroenterol.* (2015) 110:899–908. doi: 10.1038/ajg.2015.97
34. Lee L, Zhang Y, Ozar B, Sensen CW, Schriemer DC. Carnivorous nutrition in pitcher plants (*Nepenthes* spp.) via an unusual complement of endogenous enzymes. *J Proteome Res.* (2016) 15:3108–17. doi: 10.1021/acs.jproteome.6b00224
35. Helmerhorst EJ, Zamakhchari M, Schuppan D, Oppenheim FG. Discovery of a novel and rich source of gluten-degrading microbial enzymes in the oral cavity. *PLoS ONE.* (2010) 5:e13264. doi: 10.1371/journal.pone.0013264
36. Lähdeaho M-L, Kaukinen K, Laurila K, Vuotikka P, Koivurova O-P, Kärjä-Lahdensuu T, et al. Glutenase ALV003 attenuates gluten-induced mucosal injury in patients with celiac disease. *Gastroenterology.* (2014) 146:1649–58. doi: 10.1053/j.gastro.2014.02.031
37. Gordon SR, Stanley EJ, Wolf S, Toland A, Wu SJ, Hadidi D, et al. Computational design of an α-gliadin peptidase. *J Am Chem Soc.* (2012) 134:20513–20. doi: 10.1021/ja3094795
38. Wolf C, Siegel JB, Tinberg C, Camarca A, Gianfrani C, Paski S, et al. Engineering of Kuma030: a gliadin peptidase that rapidly degrades immunogenic gliadin peptides in gastric conditions. *J Am Chem Soc.* (2015) 137:13106–13. doi: 10.1021/jacs.5b08325
39. Syage JA, Murray JA, Green PHR, Khosla C. Latiglutenase improves symptoms in seropositive celiac disease patients while on a gluten-free diet. *Dig Dis Sci.* (2017) 62:2428–32. doi: 10.1007/s10620-017-4687-7
40. König J, Holster S, Bruins MJ, Brummer RJ. Aspergillus niger-derived enzyme degrades gluten in the stomach of Gluten-sensitive subjects. *Gastroenterology.* (2017) 152:S481. doi: 10.1016/S0016-5085(17)31796-1
41. Bethune MT, Strop P, Tang Y, Sollid LM, Khosla C. Heterologous expression, purification, refolding, and structural-functional characterization of EP-B2, a self-activating barley cysteine endoprotease. *Chem Biol.* (2006) 13:637–47. doi: 10.1016/j.chembiol.2006.04.008
42. Fuhrmann G, Leroux JC. *In vivo* fluorescence imaging of exogenous enzyme activity in the gastrointestinal tract. *Proc Natl Acad Sci USA.* (2011) 108:9032–7. doi: 10.1073/pnas.1100285108
43. Hamamatsu N, Aita T, Nomiya Y, Uchiyama H, Nakajima M, Husimi Y, et al. Biased mutation-assembling: an efficient method for rapid directed evolution through simultaneous mutation accumulation. *Protein Eng Des Sel.* (2005) 18:265–71. doi: 10.1093/protein/gzi028
44. Thakurta PG, Biswas S, Chakrabarti C, Sundd M, Jagannadham MV, Dattagupta JK. Structural basis of the unusual stability and substrate specificity of ervatamin C, a plant cysteine protease from *Ervatamia coronaria*. *Biochemistry.* (2004) 43:1532–40. doi: 10.1021/bi0357659
45. Choudhury D, Biswas S, Roy S, Dattagupta JK. Improving thermostability of papain through structure-based protein engineering. *Protein Eng Des Sel.* (2010) 23:457–67. doi: 10.1093/protein/gzq016
46. Wen N. *Celiac-safe wheat, reaching one objective by two approaches, gluten elimination by random and site-directed mutagenesis, and detoxification by ectopic expression of "glutenases"* (Dissertation). Washington State University, Pullman, Washington (2014).
47. Osorio CE. *Expression of glutenases in wheat (Triticum aestivum L.) endosperm and increase in the thermostability of prolyl endopeptidase from Flavobacterium meningosepticum as an alternative therapy for celiac patients* (Dissertation). Washington State University, Pullman, Washington (2013).
48. Uchiyama H, Inaoka T, Ohkuma-Soyejima T, Togame H, Shibanaka Y, Yoshimoto T, et al. Directed evolution to improve the thermostability of prolyl endopeptidase. *J Biochem.* (2000) 128:441–7. doi: 10.1093/oxfordjournals.jbchem.a022772
49. Singh SM, Panda AK. Solubilization and refolding of bacterial inclusion body proteins. *J Biosci Bioeng.* (2005) 99:303–10. doi: 10.1263/jbb.99.303
50. Vora H, McIntire J, Kumar P, Deshpande M, Khosla C. A scaleable manufacturing process for pro-EP-B2, a cysteine protease from barley indicated for celiac sprue. *Biotechnol Bioeng.* (2007) 98:177–85. doi: 10.1002/bit.21423
51. Schagger H. Tricine-SDS-PAGE. *Nat Protoc.* (2006) 1:16–22. doi: 10.1038/nprot.2006.4
52. Venalainen JI, Juvonen RO, Mannisto PT. Evolutionary relationships of the prolyl oligopeptidase family enzymes. *Eur J Biochem.* (2004) 271:2705–15. doi: 10.1111/j.1432-1033.2004.04199.x
53. Harris MN, Madura JD, Ming LJ, Harwood VJ. Kinetic and mechanistic studies of prolyl oligopeptidase from the hyperthermophile *Pyrococcus furiosus*. *J Biol Chem.* (2001) 276:19310–7. doi: 10.1074/jbc.M010489200
54. Szeltner Z, Renner V, Polgar L. The noncatalytic beta-propeller domain of prolyl oligopeptidase enhances the catalytic capability of the peptidase domain. *J Biol Chem.* (2000) 275:15000–5. doi: 10.1074/jbc.M000942200
55. Polgar L. The prolyl oligopeptidase family. *Cell Mol Life Sci.* (2002) 59:349–62. doi: 10.1007/s00018-002-8427-5
56. Gass J, Ehren J, Strohmeier G, Isaacs I, Khosla C. Fermentation, purification, formulation, and pharmacological evaluation of a prolyl endopeptidase from *Myxococcus xanthus*: implications for celiac sprue therapy. *Biotechnol Bioeng.* (2005) 92:674–84. doi: 10.1002/bit.20643
57. Mejías JH, Lu X, Osorio C, Ullman JL, von Wettstein D, Rustgi S. 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.* (2014) 6:1578–97. doi: 10.3390/nu6041578

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 February 2020
doi: 10.3389/fpls.2020.00020

Reducing the Immunogenic Potential of Wheat Flour: Silencing of Alpha Gliadin Genes in a U.S. Wheat Cultivar

Susan B. Altenbach[1], Han-Chang Chang[1], Matthew H. Rowe[1,2], Xuechen B. Yu[3,4], Annamaria Simon-Buss[1,5], Bradford W. Seabourn[6], Peter H. Green[3,7] and Armin Alaedini[3,4,7,8]**

[1] Western Regional Research Center, United States Department of Agriculture-Agricultural Research Service, Albany, CA, United States, [2] Takara Bio USA, Inc., Mountain View, CA, United States, [3] Department of Medicine, Columbia University, New York, NY, United States, [4] Institute of Human Nutrition, Columbia University, New York, NY, United States, [5] Hamburg School of Food Science, Institute of Food Chemistry, University of Hamburg, Hamburg, Germany, [6] Hard Winter Wheat Quality Laboratory, Center for Grain and Animal Health Research, United States Department of Agriculture-Agricultural Research Service, Manhattan, KS, United States, [7] Celiac Disease Center, Columbia University, New York, NY, United States, [8] Department of Medicine, New York Medical College, Valhalla, NY, United States

The alpha gliadins are a group of more than 20 proteins with very similar sequences that comprise about 15%–20% of the total flour protein and contribute to the functional properties of wheat flour dough. Some alpha gliadins also contain immunodominant epitopes that trigger celiac disease, a chronic autoimmune disease that affects approximately 1% of the worldwide population. In an attempt to reduce the immunogenic potential of wheat flour from the U.S. spring wheat cultivar Butte 86, RNA interference was used to silence a subset of alpha gliadin genes encoding proteins containing celiac disease epitopes. Two of the resulting transgenic lines were analyzed in detail by quantitative two-dimensional gel electrophoresis combined with tandem mass spectrometry. Although the RNA interference construct was designed to target only some alpha gliadin genes, all alpha gliadins were effectively silenced in the transgenic plants. In addition, some off-target silencing of high molecular weight glutenin subunits was detected in both transgenic lines. Compensatory effects were not observed within other gluten protein classes. Reactivities of IgG and IgA antibodies from a cohort of patients with celiac disease toward proteins from the transgenic lines were reduced significantly relative to the nontransgenic line. Both mixing properties and SDS sedimentation volumes suggested a decrease in dough strength in the transgenic lines when compared to the control. The data suggest that it will be difficult to selectively silence specific genes within families as complex as the wheat alpha gliadins. Nonetheless, it may be possible to reduce the immunogenic potential of the flour and still retain many of the functional properties essential for the utilization of wheat.

Keywords: alpha gliadins, celiac disease, gluten proteins, immunogenic potential, proteomics, wheat flour quality

OPEN ACCESS

Edited by:
Katharina Anne Scherf, Karlsruhe Institute of Technology (KIT), Germany

Reviewed by:
Marinus J. M. Smulders, Wageningen University & Research, Netherlands
Francisco Barro, Spanish National Research Council (CSIC), Spain
Alecia Marie Kiszonas, United States Department of Agriculture, United States

****Correspondence:***
Susan B. Altenbach
susan.altenbach@usda.gov
Armin Alaedini
aa819@columbia.edu

Specialty section:
This article was submitted to Plant Proteomics, a section of the journal Frontiers in Plant Science

Received: *15 August 2019*
Accepted: *10 January 2020*
Published: *25 February 2020*

Citation:
Altenbach SB, Chang H-C, Rowe MH, Yu XB, Simon-Buss A, Seabourn BW, Green PH and Alaedini A (2020) Reducing the Immunogenic Potential of Wheat Flour: Silencing of Alpha Gliadin Genes in a U.S. Wheat Cultivar. Front. Plant Sci. 11:20. doi: 10.3389/fpls.2020.00020

INTRODUCTION

The gluten proteins are a complex group of more than 50 proteins that have been intensively studied because of their important contributions to the commercial value of wheat. These proteins comprise 70%–80% of wheat flour protein, contain regions of very repetitive sequences with large proportions of glutamine (Q) and proline (P), and are responsible for the unique viscoelastic properties of the flour. The gluten proteins include glutenins, polymeric proteins that contribute elasticity to wheat flour dough, and gliadins, monomeric proteins that contribute extensibility to dough. The glutenins are composed of high molecular-weight glutenin subunits (HMW-GS) and low-molecular weight glutenin subunits (LMW-GS) that are linked by disulfide bonds whereas the gliadins consist of four distinct types of proteins, referred to as alpha, gamma, delta and omega gliadins (reviewed by Shewry, 2019). Most hexaploid wheat cultivars contain six or less HMW-GS genes. However, the numbers of genes within the complex gliadin and LMW-GS families were not known until the completion of a high-quality genome sequence from the reference wheat Chinese Spring (IWGSC, 2018) made it possible to assemble and annotate a complete set of gluten protein genes from a single hexaploid cultivar (Huo et al., 2018a; Huo et al., 2018b). In Chinese Spring, the sequences of 47 alpha gliadin, 14 gamma gliadin, five delta gliadin, 19 omega gliadin, and 17 LMW-GS genes were reported. Of these, 26 alpha, 11 gamma, two delta, and five omega gliadin, and 10 LMW-GS genes encode full-length proteins.

In addition to their role in end-use quality, the gluten proteins also trigger celiac disease (CD), a chronic autoimmune disease that affects 1.4% and 0.7% of the population worldwide, based on serology and biopsy assessments, respectively (Singh et al., 2018). CD occurs in genetically susceptible individuals that carry the human leukocyte antigen (HLA) genes DQ2 and/or DQ8 and results in damage to the lining of the intestine and malabsorption of nutrients that are manifested in a wide range of intestinal and extraintestinal symptoms (Koning, 2012). It is likely that the high glutamine and proline contents of the gluten proteins contribute to their immunogenic properties. The average Q + P content for gluten proteins in the different classes ranges from ~45% (delta gliadins) to 73% (omega gliadins). As a result, these proteins are highly resistant to proteolytic degradation within the gastrointestinal tract. The high Q + P contents of the gluten proteins also makes them good substrates for tissue transglutaminase, an enzyme in the small intestine that converts glutamine to negatively charged glutamate residues. Deamidation of gluten peptides increases their binding affinity for HLA-DQ2 and -DQ8 on antigen presenting cells, allowing them to be processed and presented to T-cells to trigger an inflammatory immune response.

Epitopes relevant to CD have been identified within all of the major classes of gluten proteins (Sollid et al., 2012). Five epitopes from alpha gliadins include the core sequences PFPQPQLPY, PYPQPQLPY, PQPQLPYPQ, FRPQQPYPQ, and QGSFQPSQQ. In some alpha gliadins, six epitopes overlap in a 33-mer protease-resistant peptide that has been found to be particularly toxic (Shan et al., 2002). Sixteen of the 26 alpha gliadins from Chinese Spring (62%) contain from one to eight CD epitopes. However, only one alpha gliadin encoded by the D genome contains the 33-mer toxic peptide. The greatest number of epitopes are found in proteins encoded by the D genome while nine of 11 alpha gliadins encoded by the B genome and one alpha gliadin encoded by the D genome do not contain any previously described epitopes (Huo et al., 2018b; Altenbach et al., 2020). Eight CD epitopes have been described in gamma gliadins, including PQQSFPQQQ, IQPQQPAQL, QQPQQPYPQ, SQPQQQFPQ, PQPQQQFPQ, PQPQQPFCQ, QQPFPQQPQ, and QQPQQPFPQ. All gamma gliadins from Chinese Spring contain from five to ten CD epitopes (Altenbach et al., 2020). Two epitopes, PFPQPQQPF and PQPQQPFPW, were identified in omega-1,2 gliadins, a subset of omega gliadins. All omega-1,2 gliadins from Chinese Spring contain these epitopes as well as multiple copies of the QQPQQPFPQ and QQPFPQQPQ gamma gliadin epitopes. Two epitopes have been described for LMW-GS, PFSQQQQPV and FSQQQQSPF. Seven of the ten LMW-GS in Chinese Spring contain from one to three of these epitopes. Finally, one epitope was identified in HMW-GS, QGYYPTSPQ (Sollid et al., 2012). In general, epitopes from alpha and omega gliadins are immunodominant (Tye-Din et al., 2010), possibly because these epitopes have a greater number of proline residues and may be more resistant to proteolytic digestion.

Currently, the only effective treatment for CD is a lifelong gluten-free diet. Thus, there is a critical need for new approaches to reduce the immunogenic potential of wheat flour. However, these studies are challenging because of the large numbers of different wheat cultivars that are grown around the world, the tremendous allelic variation in gluten protein genes among cultivars and the large number of immunogenic sequences in all of the major classes of gluten proteins. A number of studies have focused on identifying cultivars that are low in CD epitopes using DNA sequencing, quantitative protein analyses, antibody screening or targeted mass spectrometry methods (van den Broeck et al., 2010; Salentijn et al., 2013; van den Broeck et al., 2015; Prandi et al., 2016; Ribeiro et al., 2016; Malalgoda et al., 2018; Pilolli et al., 2019). Other studies have used gene silencing to reduce the amounts of immunogenic proteins in wheat flour. In a recent study, RNA interference (RNAi) was used to eliminate omega-1,2 gliadins from wheat flour (Altenbach et al., 2019). This was accomplished without notable effects on the levels of other gluten proteins in the flour. Flour from the resulting transgenic plants showed decreased reactivity to IgG and IgA antibodies from a cohort of CD patients as well as improved mixing properties relative to the nontransgenic control. A number of other studies have focused on the more

Abbreviations: AACCI, American Association of Cereal Chemists International; AAI, alpha-amylase inhibitors; CD, celiac disease; ELISA, enzyme-linked immunosorbent assay; ESTs, expressed sequence tags; HMW-GS, high molecular weight glutenin subunits; LMW-GS, low molecular weight glutenin subunits; MS/MS, tandem mass spectrometry; RNAi, RNA interference; 2-DE, two-dimensional gel electrophoresis.

complex family of alpha gliadins. Barro et al. (2015) and Becker et al. (2012) used RNAi to target all alpha gliadin genes. While the alpha gliadins were reduced significantly in both studies, there were numerous changes in the levels of other gluten proteins in the resulting transgenic lines. Sánchez-León et al. (2018) used genome editing to introduce mutations into a conserved region in the alpha gliadin genes. Alpha gliadins were reduced from 32%–82% in the resulting plants. However, reductions in alpha gliadins were accompanied by significant changes in the levels of most other types of gluten proteins. While off-target and compensatory effects on the proteome have been observed, little is known about how the proteome adjusts to significant reductions in proteins that normally comprise as much as 15%–20% of the total protein or the molecular mechanisms involved. In this study, the goal was to use RNA interference to silence only those alpha gliadin genes containing known CD epitopes with the hope that the immunoreactivity of the flour might be reduced with minimal effects on the proteome. The work highlights some of the challenges faced in experiments aimed at eliminating specific proteins within large families of gluten proteins with very similar and repetitive sequences.

MATERIAL AND METHODS

Plant Material

The U.S. hard red spring wheat *Triticum aestivum* cv. Butte 86 was used for all studies. All plant material was grown in a temperature-controlled greenhouse with daytime/nighttime temperatures of 24/17°C as described previously (Altenbach et al., 2003). Plants were supplied with water mixed with 0.6 g/l of Peters Professional 20-20-20 water-soluble fertilizer (Scotts-Sierra Horticultural Products Company, Marysville, OH) by a drip irrigation system.

RNAi Construct and Transformation of Plants

A 608-bp DNA fragment designed to target alpha gliadins was synthesized by GenScript (Piscataway, NJ) and cloned into the vector pUC57. The 608-bp fragment consisted of a 14-bp region that included a Hpa I restriction site, a 217-bp trigger in sense orientation, a 146-bp spacer region corresponding to an intron from a wheat starch synthase gene, a 217-bp trigger in antisense orientation, and a 14-bp spacer that included a Hpa I restriction site. This plasmid was digested with Hpa I (New England Biolabs, Ipswich, MA). Following purification, the fragment was ligated into the Hpa I site of the plasmid pJL10P5 between the promoter from the HMW-GS Dy10 gene and the terminator from the HMW-GS Dx5 gene as described in Altenbach and Allen (2011). The final construct, referred to as Bazooka-pJL10P5-#6, was verified by DNA sequencing. Bazooka-pJL10P5-#6 and the plasmid pAHC20 that facilitates selection of transgenic plants with phosphinothricin (Christensen and Quail, 1996) were used to transform Butte 86 wheat plants as described in detail in Altenbach and Allen (2011). Putative transgenic plants were identified by PCR analysis using primers described in Altenbach and Allen (2011). Initial screening of gliadin fractions from grain by SDS-PAGE was also described in Altenbach and Allen (2011). Lines in which alpha gliadins were significantly down-regulated were identified and homozygous plants were selected in subsequent generations.

Protein Extraction and Analysis by Two-Dimensional Gel Electrophoresis (2-DE)

Triplicate samples of grain from selected lines were milled into flour using a Quadrumat Senior experimental flour mill following AACCI Method 26.10.02 (AACC Int., 1988). Total proteins were extracted from the resulting flour, quantified using a modified Lowry assay and analyzed on triplicate 2-D gels using capillary tube gels with a pI range of 3 to 10 in the first dimension and NuPAGE 4%–12% BIS-Tris protein gels in the second dimension (Life Technologies, Carlsbad, CA) as described in detail in Dupont et al. (2011). Following staining with Coomassie G-250 (Sigma Aldrich, St. Louis, MO), the gels were digitized using a calibrated scanner. 2-D gels used for the analysis are shown in **Supplementary File 1**. Individual gel spots were aligned between gels and quantified using SameSpots Version 5.0 (Nonlinear Dynamics Limited, Newcastle upon Tyne, UK). Statistical analyses of spot volume data were conducted using the SameSpots software. Identifications of individual protein spots in the Butte 86 nontransgenic line were as reported in Dupont et al. (2011) or as determined in this study. Individual spots in transgenic lines were deemed to show significant changes from the nontransgenic if they had ANOVA values < 0.02 and had changes in average normalized spot volumes that were greater than 20%.

Identification of Proteins in 2-DE Spots by Tandem Mass Spectrometry (MS/MS)

Selected protein spots from the alpha gliadin region of 2-D gels of nontransgenic and transgenic lines were excised from triplicate gels, placed in 96-well plates and digested with either chymotrypsin, thermolysin, or trypsin using a DigestPro according to the directions of the manufacturer (INTAVIS Bioanalytical Instruments AG, Cologne, Germany). The resulting samples were then analyzed using an Orbitrap Elite mass spectrometer (Thermo Scientific, San Jose, CA, USA) as described in Vensel et al. (2014). Two search engines, Mascot (www.matrixscience.com) and XTandem! (https://www.thegpm.org/TANDEM/), were used to interrogate a database of 125,400 protein sequences. The database included Triticeae sequences downloaded from NCBI on 06-18-2018 plus Chinese Spring sequences reported by Huo et al. (2018a; 2018b); Butte 86 sequences from Dupont et al. (2011) and Altenbach et al. (2011); Xiaoyan 81 sequences from Wang et al. (2017); and common mass spectrometry contaminant sequences contained in the common Repository of Adventitious Proteins (cRAP) database (ftp://ftp.thegpm.org/fasta/cRAP/crap.fasta). Data from the two searches and three enzymes were compiled and further validated using Scaffold version 4.8.9 (http://www.proteomesoftware.com/) using a protein threshold of 99%, peptide threshold of 95% and 20 ppm mass error, and a

minimum of four peptides. The mass spectrometry data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) *via* the PRIDE partner repository (Perez-Riverol et al., 2019) with the dataset identifier PXD016930 and 10.6019/PXD016930. The protein that was assigned the greatest number of unique peptides was reported as the predominant protein for each spot. Proteins that were assigned at least half the number of unique peptides as the predominant protein are also reported for each spot along with the numbers of unique peptides, total spectra, and protein coverage for each. Summaries of MS/MS data from each spot are shown in **Supplementary Files 2–5**.

Assessment of Immune Reactivity by ELISA and 2-D Immunoblot Analysis

Serum samples from a cohort of patients with celiac disease were used to assess immune reactivity toward gluten proteins from the nontransgenic and transgenic wheat lines. The celiac disease patients included twenty with elevated levels of IgG antibody to gluten [15 female, 17 white race, mean (SD) age 42.9 (18.5) years] and twenty with elevated levels of IgA antibody to gluten [13 female, 19 white race, mean (SD) age 46.7 (17.3) years]. Positivity for IgG or IgA antibody reactivity to gluten was determined as described previously (Samaroo et al., 2010). All patients were biopsy proven, diagnosed with CD according to previously described criteria (Alaedini and Green, 2005), and on a gluten-containing diet. In addition, all patients were positive for antibody reactivity to transglutaminase 2, the most sensitive and specific serologic marker of CD, determined as previously described (Lau et al., 2013). Serum samples were obtained under institutional review board-approved protocols at Columbia University. This study was approved by the Institutional Review Board of Columbia University Medical Center. Serum samples were maintained at −80°C to maintain stability.

Levels of serum IgG and IgA antibody reactivity to gluten were measured separately by enzyme-linked immunosorbent assay (ELISA) as described in Altenbach et al. (2019). All serum samples were tested in duplicate. Absorbance values were corrected for nonspecific binding by subtraction of the mean absorbance of the associated uncoated wells and corrected values were normalized according to the mean value of the positive controls on each plate. The change in immune reactivity towards the transgenic wheat lines in comparison to the nontransgenic line, as determined by ELISA, was assessed by the Wilcoxon matched-pairs test. All P values were two-sided and differences were considered statistically significant at $P < 0.05$. Statistical analyses were performed with Prism 8 (GraphPad) software.

IgG and IgA antibody reactivity to gluten proteins was further analyzed by two-dimensional immunoblotting as described in detail in Altenbach et al. (2019).

Analysis of Flour End-Use Quality

End-use functionality tests were conducted at the USDA-ARS-HWWQL (Manhattan, KS) using standardized methods approved by American Association of Cereal Chemists International (AACCI). Flour protein content was determined by NIR using AACCI method 39-11.01 (AACC Int., 1985), mixing properties were determined on 10 g flour samples (14% mb) using a Mixograph (TMCO, National Mfg., Lincoln, NE) and AACCI Method 54-40.02 (AACC Int., 1995), and SDS sedimentation tests were done according to AACCI Method 56-60.01 (AACC Int., 1961). Averages and standard deviations from triplicate samples were calculated for each wheat line.

RESULTS

Design of the Trigger for the RNAi Construct

The 217 bp trigger for the RNAi construct consisted of three distinct fragments of 74, 65, and 78 bp that were based on sequences of 13 full-length alpha gliadin coding regions assembled from Butte 86 expressed sequence tags (ESTs) available at the time the study was initiated (Altenbach et al., 2010). All three target fragments encode a portion of the first nonrepetitive region of the alpha gliadin that lies between the two poly Q regions (**Figure 1**). Target 1, a 74-bp fragment with the sequence AAAGTACTTACCAGCTGGTGCAACAATTGTGT TGTCAGCAGCTGTGGCAGATCCCCGAGCAGTCGCG GTGCCAA, was a perfect match with alpha gliadins Bu-1, Bu-2, Bu-3, Bu-4, and Bu-10 (**Table 1**). These genes are likely from the D genome and encode proteins containing from three to eight CD epitopes. This fragment also had 25 bp of identity with Bu-11. Target 2, a 65-bp fragment with the sequence TTGCAAG AATTGTGTTGTCAGCACCTATGGCAGATCCCTG AGCAGTCGCAGTGCCAGGCCATCCA, was a perfect match with Bu-5 and Bu-14, likely from the A genome and encoding proteins containing two CD epitopes, while Target 3, a 78-bp fragment with the sequence AAGTATTGCAGCAAAGTAGTT ACCAAGTGTTGCAACAATTATGTTGTCAGCAG CTGCGGCTGATCCCCGAGCAGTCGC, was a perfect match with Bu-11 encoding a protein with one CD epitope. The 78-bp fragment also had 25 bp of identity with Bu-12 that encodes a protein without any epitopes. Seven alpha gliadin genes from Butte 86 had 19 or less bp of identity with Target 1 while 11 genes had 19 or less bp of identity with either Target 2 or Target 3. None of the targets had identities greater than 19 bp with Bu-8, Bu-13, Bu-23, or Bu-27, all of which are likely to be from the B genome in Butte 86 and encode proteins devoid of CD epitopes.

Because the collection of alpha gliadin sequences from Butte 86 is incomplete, the specificities of the target sequences also were assessed using the complete set of 26 full-length alpha gliadin genes that was recently reported from the reference wheat Chinese Spring (Huo et al., 2018b) (**Table 2**). Five alpha gliadins encoded by the D genome (CS-D4, CS-D5, CS-D6, CS-D8, CS-D9) were perfect matches with Target 1, four encoded by the A genome (CS-A4, CS-A5, CS-A9, CS-A10) were perfect matches with Target 2, and one encoded by the A genome (CS-A2) was a perfect match with Target 3. In addition, five genes had regions of identities between 25 and 56 bp with Target 1, one gene had a 53-bp region of identity with Target 2 and nine

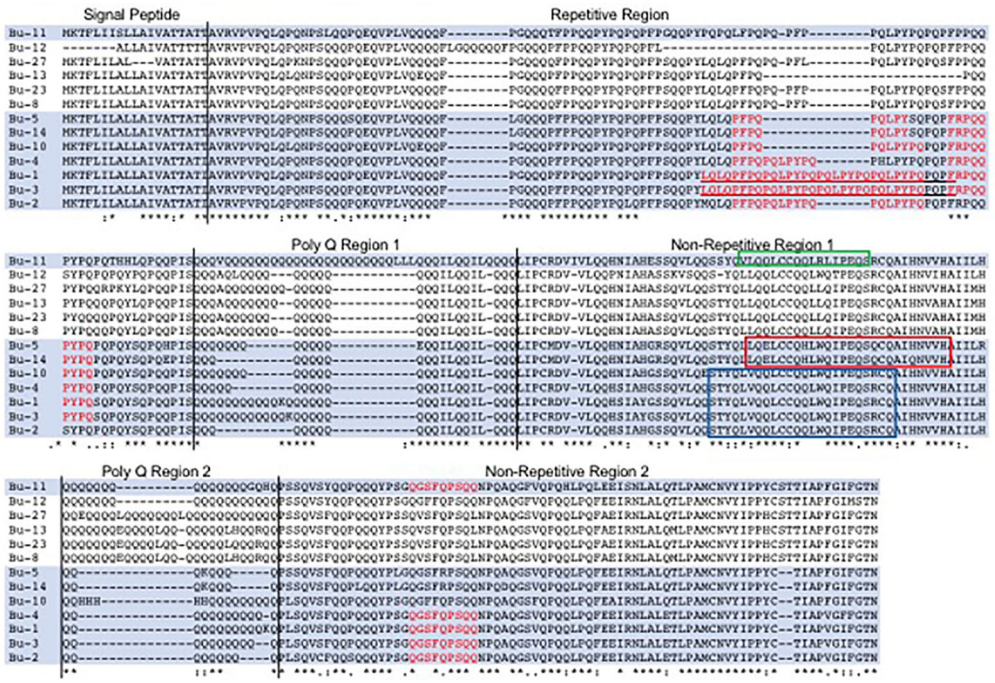


FIGURE 1 | Comparison of alpha gliadins from Butte 86 and selection of target regions for the RNAi construct. Celiac disease (CD) epitope sequences within the alpha gliadins are shown in red and the 33-mer toxic peptide is underlined. Proteins that were targeted in RNAi experiments are shaded. The 74, 65, and 78 bp fragments used in the RNAi construct correspond to regions of the proteins shown in the blue, red, and green boxes, respectively. Protein sequences were derived from genes assembled from Butte 86 expressed sequence tags reported in Altenbach et al., 2010.

TABLE 1 | Identities of target sequences used in trigger of RNAi construct to alpha gliadin genes from Butte 86.

Alpha gliadin gene	# CD epitopes in protein	Longest region of identity (bp)		
		Target 1[1]	Target 2[2]	Target 3[3]
Bu-1	8	74	16	17
Bu-2	4	74	16	17
Bu-3	8	74	16	17
Bu-4	4	74	16	17
Bu-5	2	17	65	<16
Bu-8	0	19	19	<16
Bu-10	3	74	16	17
Bu-11	1	25	<16	78
Bu-12	0	<16	<16	25
Bu-13	0	19	19	<16
Bu-14	2	16	65	<16
Bu-23	0	19	19	<16
Bu-27	0	19	19	<16

[1]*74 bp target sequence is AAAGTACTTACCAGCTGGTGCAACAATTGTGTTGTCAGCA GCTGTGGCAGATCCCCGAGCAGTCGCGGTGCCAA.*
[2]*65 bp target sequence is TTGCAAGAATTGTGTTGTCAGCACCTATGGCAGATCCCTG AGCAGTCGCAGTGCCAGGCCATCCA.*
[3]*78 bp target sequence is AAGTATTGCAGCAAAGTAGTTACCAAGTGTTGCAACAATT ATGTTGTCAGCAGCTGCGGCTGATCCCCGAGCAGTCGC.*

genes had regions of identities between 21 and 29 bp with Target 3. Only four of the 26 alpha gliadins in Chinese Spring did not have identities greater than 19 bp with any of the targeting regions. Three of these are from the B genome (CS-B7, CS-B8, CS-B9) and do not contain CD epitopes and one is from the A genome (CS-A1) and contains a single epitope. A BLASTn search also revealed that there were no regions of identity 16 bp or greater with any gamma, delta, or omega gliadins, LMW-GS, or HMW-GS from Chinese Spring, suggesting that the RNAi construct should target only alpha gliadins.

Analysis of Flour Proteins From Transgenic Lines

Following transformation of Butte 86 plants and initial DNA and protein analyses, two homozygous transgenic lines showing altered alpha gliadin profiles in SDS-PAGE were selected for detailed analysis by quantitative 2-DE. Total protein profiles are shown in **Figure 2** for transgenic lines SA35a-124j and SA39b-658-5, referred to as 124j and 658-5, respectively. While most alpha gliadins in the nontransgenic flour are found within the red box shown in Panel A, this region also contains some gamma gliadins, LMW-GS and the nongluten storage proteins called triticins. There are notable changes in this region of the gel for the two transgenic lines shown in panels B and C. As can be seen in the enlarged alpha gliadin regions in **Figure 3**, some spots present in the nontransgenic line are missing in the transgenic lines and some are significantly reduced. Additionally, in a number of cases the suppression of a major spot found in the nontransgenic line revealed the presence of several minor spots

TABLE 2 | Identities of target sequences used in trigger of RNAi construct to alpha gliadin genes from Chinese Spring.

Alpha gliadin gene	# CD epitopes in protein	Longest region of identity (bp)		
		Target 1[1]	Target 2[2]	Target 3[3]
CS-A1	1	<16	<16	<16
CS-A2	1	25	<16	78
CS-A4	2	17	65	<16
CS-A5	2	17	65	<16
CS-A6	2	17	62	<16
CS-A8	2	17	53	<16
CS-A9	1	17	65	<16
CS-A10	2	16	65	<16
CS-B3	1	32	<16	22
CS-B7	0	19	19	<16
CS-B8	0	19	19	<16
CS-B9	0	19	19	<16
CS-B11	0	<16	<16	21
CS-B14	0	<16	<16	25
CS-B15	0	<16	<16	25
CS-B16	0	<16	<16	25
CS-B17	0	<16	<16	25
CS-B18	0	<16	<16	25
CS-B25	1	56	16	17
CS-D1	1	45	<16	29
CS-D4	3	74	16	17
CS-D5	8	74	16	17
CS-D6	5	74	16	17
CS-D8	6	74	16	17
CS-D9	6	74	16	17
CS-D12	0	27	<16	23

[1]*74 bp target sequence is AAAGTACTTACCAGCTGGTGCAACAATTGTGTTGTCAG CAGCTGTGGCAGATCCCCGAGCAGTCGCGGTGCCAA.*
[2]*65 bp target sequence is TTGCAAGAATTGTGTTGTCAGCACCTATGGCAGATCCCT GAGCAGTCGCAGTGCCAGGCCATCCA.*
[3]*78 bp target sequence is AAGTATTGCAGCAAAGTAGTTACCAAGTGTTGCAACAA TTATGTTGTCAGCAGCTGCGGCTGATCCCCGAGCAGTCGC.*

in the transgenic lines. To investigate this further, 30 spots in the alpha gliadin region as well as two spots that lie outside of this region and were previously identified as alpha gliadins (spots 31, 32) were excised from triplicate 2-D gels of the nontransgenic Butte 86. Spots in corresponding positions in the transgenic lines as well as new spots uncovered in the transgenic lines were also excised from triplicate gels (**Figure 3**). Following digestion with either chymotrypsin, thermolysin, or trypsin, all spots were analyzed by MS/MS.

Not surprisingly, the identifications were complex (**Table 3**, **Supplementary Files 2–5**). In the nontransgenic line, the predominant proteins in 20 spots were alpha gliadins (10, 11, 13, 14, 15, 17–27, 29–32). For these spots, the MS sequence coverage ranged from 29 to 89% with an average of 63%. The predominant proteins in four spots were gamma gliadins (4, 5, 9, 12) while the predominant proteins in three spots were LMW-GS (6, 7, 8). Three of the alpha gliadin spots also contained other alpha gliadins (20, 24, 26) while five also contained gamma gliadins (10, 14, 17, 22, 24). In addition, one spot contained a delta gliadin mixed with an alpha gliadin (28), one contained an avenin-like protein (16), and three contained triticins (1, 2, 3), two of which were mixed with gamma gliadins (2, 3).

Twelve of the spots identified as alpha gliadins in the nontransgenic line were missing in transgenic line 124j (11, 13, 17, 20, 21, 23, 25, 26, 29, 30, 31, 32) while seven spots identified as alpha gliadins in the nontransgenic line were identified as gamma gliadins (10, 14, 15, 18, 19, 22, 24) in 124j. A delta gliadin was the only protein identified in spot 28 in the transgenic line. Of the ten other minor spots that were uncovered in the transgenic line, six were gamma gliadins (34, 37, 39–42), two were LMW-GS (33, 36), one was triticin (38) and one was glyceraldehye-3-phosphate dehydrogenase (35). No alpha gliadins were identified in 124j in the analysis.

Similarly, twelve spots identified as alpha gliadins in the nontransgenic line were missing in transgenic line 658-5 (11, 13, 17, 20, 21, 22, 25, 26, 27, 29, 31, 32), and the same seven spots that were identified as alpha gliadins in the nontransgenic line but gamma gliadins in 124j were also identified as gamma gliadins in 658-5 (10, 14, 15, 18, 19, 22, 24). Spot 28 in 658-5 contained only a delta gliadin and the identities of the minor spots uncovered in 658-5 were similar to those from 124j (**Table 3**).

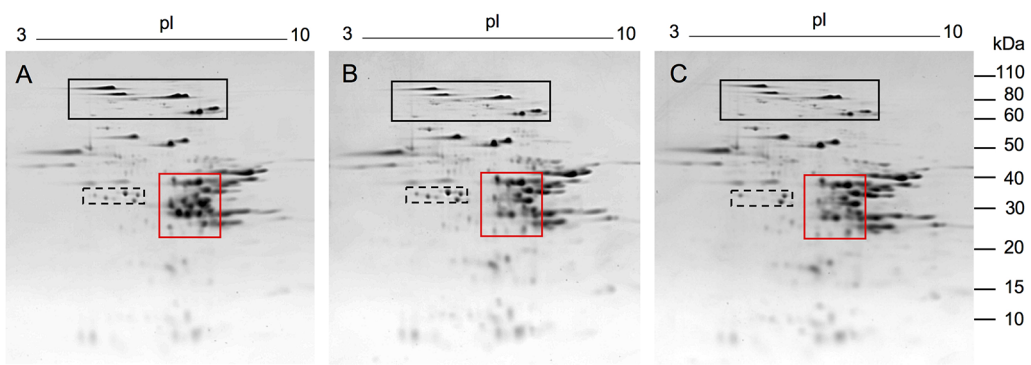


FIGURE 2 | 2-DE analysis of total flour proteins from the nontransgenic **(A)** and transgenic lines 124j **(B)** and 658-5 **(C)**. MW are indicated on the right and the pI range is shown above each gel. Red boxes show regions of gels containing alpha gliadins. These regions were enlarged in **Figure 3**. The black and black dashed boxes show regions of gels containing HMW-GS and serpins, respectively. Proteins were identified by MS/MS in Dupont et al. (2011).

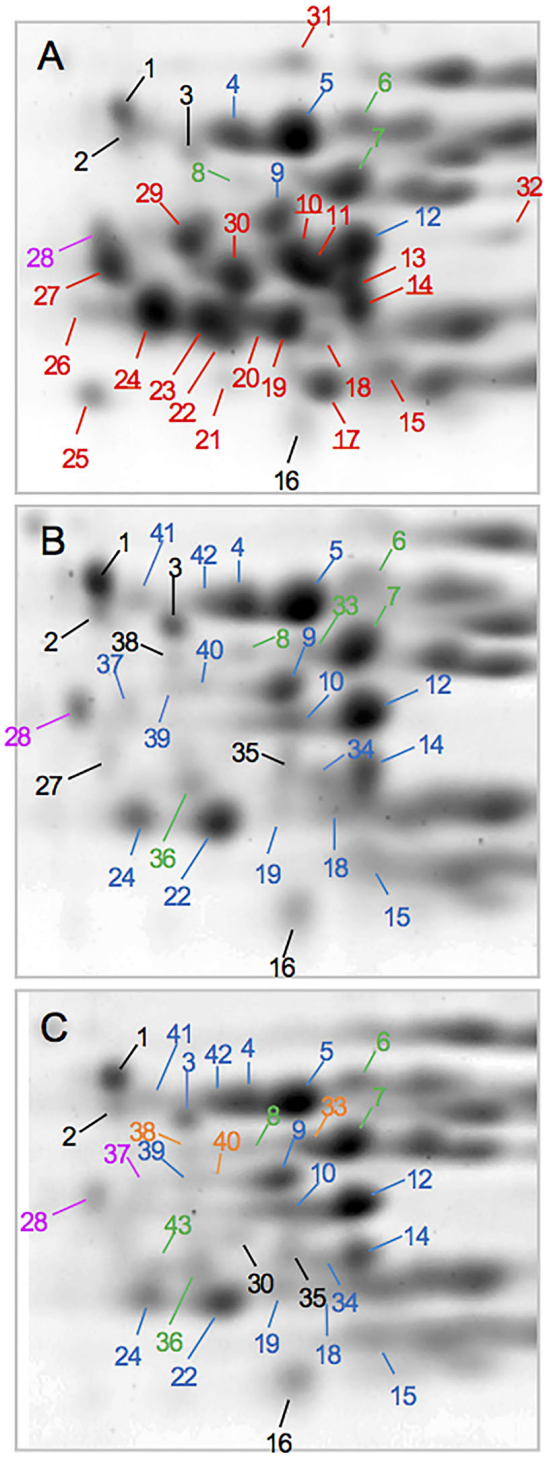


FIGURE 3 | Regions of 2-D gels containing alpha gliadins from the nontransgenic **(A)** and transgenic lines 124j **(B)** and 658-5 **(C)**. Spots in which the predominant protein was identified as an alpha gliadin are shown in red, gamma gliadin in blue, delta gliadin in magenta, and LMW-GS in green. Spots with underlined numbers in panel A contained a gamma gliadin in addition to an alpha gliadin. Spots shown in black were identified as nongluten proteins while those labeled in orange did not yield identifications.

Alpha gliadins were identified as one of several components of two minor spots (19, 30) in 658-5.

Quantitative Analysis of Proteins in Nontransgenic and Transgenic Lines

Most spots in which the predominant proteins were alpha gliadins in Butte 86 showed significantly reduced volumes in the transgenic lines (73% of alpha gliadin spots in 124j and 79% of alpha gliadin spots in 658-5) (**Supplementary Files 6, 7**). Decreases ranged from 25.2% to 83.9% with an average reduction of 56.6% in 124j, and 28.5% to 67.0% with an average reduction of 48.8% in 658-5. Surprisingly, a large percentage of the spots that were identified as HMW-GS also showed significantly reduced volumes (71% of HMW-GS spots in 124j and 76% of HMW-GS spots in 658-5), although the reductions were generally much smaller for the HMW-GS than the alpha gliadins. Changes in a few spots identified as omega gliadins (four of 16 spots in 124j and one of 15 in 658-5) and LMW-GS (three of 22 spots in 124j and four of 22 spots in 658-5) also were observed. In 124j, increases were observed in the volumes of spots containing a variety of nongluten proteins, including purinins, triticins, globulins, serpins and alpha amylase inhibitors (AAI) (**Supplementary File 6**) while decreases in a number of serpins were observed in 658-5 as well as increases among purinins, globulins and a few AAI (**Supplementary File 7**). It is notable that some of these proteins, including purinins, globulins, serpins and AAI may also be involved in wheat-related pathologies.

A number of adjustments were made to the normalized spot volume data for 2-DE spots containing either alpha or gamma gliadins (**Supplementary Files 6, 7**). First, for the five Butte 86 spots in which both alpha and gamma gliadins were identified (10, 14, 17, 22, 24), average spot volumes for Butte 86 were divided among the two protein types according to the percentage of unique peptides that were obtained for each type as detailed in **Supplementary Files 6, 7**. Second, in the transgenic lines, the average normalized volumes were assigned to gamma gliadins since gamma gliadins were the only proteins identified in these spots in these lines. Additionally, in cases where spots were identified as alpha gliadins in Butte 86 but as gamma gliadins in the transgenic lines (15, 18, 19, 22), the entire spot volume was assigned to alpha gliadins in Butte 86, but to gamma gliadins in the transgenic lines. Additionally, spot volume data for spot 28 was divided between alpha and delta gliadins in Butte 86, but assigned to delta gliadins for the transgenic lines and spot volume data for spots 27 and 30 were assigned to alpha gliadins for Butte 86, but to the nongluten protein group for lines 124j and 658-5, respectively, since the spots were identified as glyceraldehyde-3-phosphate dehydrogenase in the transgenic lines.

Overall, decreases in the amounts of alpha gliadins of 70.4% and 66.1% and decreases in the amounts of HMW-GS of 26% and 28.8% were observed in transgenic lines 124j and 658-5, respectively (**Table 4**). Among the HMW-GS, significant decreases were noted for all subunits except Ax2* with the greatest decreases noted for HMW-GS Dy10 (34.7% and 42.2% for 124j and 658-5, respectively) (**Supplementary Files 6, 7**). Serpins showed a 30.5% decrease in 658-5 but a slight increase in 124j. Small increases in the amounts of purinins were also noted

TABLE 3 | Predominant proteins identified by MS/MS in 2-DE protein spots from the alpha gliadin region of nontransgenic Butte 86 and transgenic lines 124j and 658-5. The positions of the spots are shown in **Figure 3**. MS data can be found in **Supplementary Files 2–5**.

Line	Spot #	Predominant Protein	Accession[1]	# Unique Peptides	# Spectra	% Coverage
Butte 86	1	triticin	ACB41345	7	14	14
Butte 86	2	triticin[2]	AAB27108	21	39	291
Butte 86	3	triticin[2]	EMS60011	20	34	26
Butte 86	4	gamma gliadin	Bu-Gamma5	49	121	68
Butte 86	5	gamma gliadin	Bu-Gamma5	79	181	84
Butte 86	6	LMW-GS (s-type)	CS-LMW-D1	50	109	75
Butte 86	7	LMW-GS (m-type)	Bu-LMW-7	87	210	85
Butte 86	8	LMW-GS (m-type)	AVY03606 (Bu-LMW7)	13	23	43
Butte 86	9	gamma gliadin	EMS45054 (Bu-Gamma1)	40	80	57
Butte 86	10	alpha gliadin[2]	X-Alpha18 (Bu-Alpha23)	30	67	52
Butte 86	11	alpha gliadin	Bu-Alpha23	67	145	77
Butte 86	12	gamma gliadin	Bu-Gamma2	61	165	87
Butte 86	13	alpha gliadin	SCW25764 (Bu-Alpha 1)	47	100	73
Butte 86	14	alpha gliadin[2]	AKC91252 (Bu-Alpha 1)	26	53	58
Butte 86	15	alpha gliadin	Bu-Alpha2	43	89	70
Butte 86	16	avenin-like protein	Bu-farinin-2	16	34	49
Butte 86	17	alpha gliadin[2]	Bu-Alpha5	44	120	81
Butte 86	18	alpha gliadin	SCW25769 (Bu-Alpha10)	32	53	59
Butte 86	19	alpha gliadin	Bu-Alpha10	35	73	57
Butte 86	20	alpha gliadin[3]	CS-Alpha-B11 (Bu-Alpha12)	45	118	71
Butte 86	21	alpha gliadin	Bu-Alpha5	5	9	29
Butte 86	22	alpha gliadin[2]	Bu-Alpha4	44	89	77
Butte 86	23	alpha gliadin	Bu-Alpha4	82	159	89
Butte 86	24	alpha gliadin[2,3]	AKC91122	49	122	76
Butte 86	25	alpha gliadin	Bu-Alpha14	14	22	47
Butte 86	26	alpha gliadin[3]	AKC91122	8	12	42
Butte 86	27	alpha gliadin	CS-Alpha-B16/B17	18	36	40
Butte 86	28	delta gliadin[3]	CS-delta-D1	11	22	33
Butte 86	29	alpha gliadin	CAY54134 (Bu-Alpha12)	51	120	74
Butte 86	30	alpha gliadin	Bu-Alpha3	83	165	86
Butte 86	31	alpha gliadin	CS-Alpha-B3	16	32	39
Butte 86	32	alpha gliadin	Bu-Alpha27	33	70	60
transgenic 124j	1	triticin	ABC41345	48	102	46
transgenic 124j	2	triticin	ABC41345	25	57	37
transgenic 124j	3	triticin[2]	EMS60011	40	78	36
transgenic 124j	4	gamma gliadin	Bu-Gamma5	37	80	72
transgenic 124j	5	gamma gliadin	Bu-Gamma5	77	181	81
transgenic 124j	6	LMW-GS (s-type)	CS-LMW-D1	44	80	79
transgenic 124j	7	LMW-GS (m-type)	ALN96387 (Bu-LMW-7)	52	118	71
transgenic 124j	8	LMW-GS (m-type)	Bu-LMW-7	16	27	39
transgenic 124j	9	gamma gliadin	Bu-Gamma1	29	59	52
transgenic 124j	10	gamma gliadin	Bu-Gamma2	40	91	81
transgenic 124j	12	gamma gliadin	Bu-Gamma2	49	139	83
transgenic 124j	14	gamma gliadin	Bu-Gamma3	26	52	59
transgenic 124j	15	gamma gliadin[2]	Bu-Gamma7	18	39	39
transgenic 124j	16	avenin-like protein	AEW43832	10	24	29
transgenic 124j	18	gamma gliadin[2]	Bu-Gamma4	19	36	57
transgenic 124j	19	gamma gliadin	ACI04093	15	26	34
transgenic 124j	22	gamma gliadin	CS-Gamma-D3	45	104	74
transgenic 124j	24	gamma gliadin	CS-Gamma-D3	40	76	71
transgenic 124j	27	glyceraldehyde-3-phosphate dehydrogenase[6]	ANW11922	17	32	40
transgenic 124j	28	delta gliadin	CS-delta-D1	30	59	59
transgenic 124j	33	LMW-GS (m-type)[2]	Bu-LMW-7	19	29	39
transgenic 124j	34	gamma gliadin[4]	Bu-Gamma3	16	31	37
transgenic 124j	35	glyceraldehyde-3-phosphate dehydrogenase[2]	ALE18233	14	24	36
transgenic 124j	36	LMW-GS (m-type)	CS-LMW-D7	35	109	59
transgenic 124j	37	gamma gliadin[5,6]	EMS45054	7	11	15
transgenic 124j	38	triticin	EMS60011	12	19	25
transgenic 124j	39	gamma gliadin	EMS45054	6	14	15
transgenic 124j	40	gamma gliadin	Bu-Gamma8	14	30	32
transgenic 124j	41	gamma gliadin	Bu-Gamma5	17	34	37
transgenic 124j	42	gamma gliadin	Bu-Gamma5	19	39	39

(Continued)

TABLE 3 | Continued

Line	Spot #	Predominant Protein	Accession[1]	# Unique Peptides	# Spectra	% Coverage
transgenic 658-5	1	triticin	ACB41345	28	69	43
transgenic 658-5	2	triticin	ACB41345	22	60	42
transgenic 658-5	3	gamma gliadin[6]	Bu-Gamma5	16	27	33
transgenic 658-5	4	gamma gliadin	Bu-Gamma5	26	48	52
transgenic 658-5	5	gamma gliadin	Bu-Gamma5	47	140	69
transgenic 658-5	6	LMW-GS (s-type)	CS-LMW-D1	23	40	61
transgenic 658-5	7	LMW-GS (m-type)	Bu-LMW7	46	115	66
transgenic 658-5	8	LMW-GS (m-type)	Bu-LMW7	21	36	47
transgenic 658-5	9	gamma gliadin	EMS45054 (Bu-Gamma1)	17	33	44
transgenic 658-5	10	gamma gliadin	AGO17694 (Bu-Gamma2)	22	44	60
transgenic 658-5	12	gamma gliadin	AAF42989 (Bu-Gamma2)	42	86	66
transgenic 658-5	14	gamma gliadin	Bu-Gamma3	18	30	54
transgenic 658-5	15	gamma gliadin	ACJ03439 (Bu-Gamma7)	12	27	41
transgenic 658-5	16	avenin-like protein	AEW43832	20	47	57
transgenic 658-5	18	gamma gliadin[2]	ATD83912	20	37	46
transgenic 658-5	19	gamma gliadin[2,3,6]	ACl04093	7	22	24
transgenic 658-5	22	gamma gliadin	BAN29066	40	94	77
transgenic 658-5	24	gamma gliadin	CS-Gamma-D3	18	32	51
transgenic 658-5	28	delta gliadin	CS-delta-D1	14	27	34
transgenic 658-5	30	glyceraldehyde-3-phosphate dehydrogenase[2,3,6]	ANW11921	8	15	30
transgenic 658-5	33	no ID				
transgenic 658-5	34	gamma gliadin	Bu-gamma3	7	14	24
transgenic 658-5	35	glyceraldehyde-3-phosphate dehydrogenase	ANW11922	16	32	47
transgenic 658-5	36	LMW-GS (m-type)[6]	AFL55408	7	18	23
transgenic 658-5	37	delta gliadin[2]	CS-delta-D1	6	14	24
transgenic 658-5	38	no ID				
transgenic 658-5	39	gamma gliadin[6]	Bu-Gamma8	5	11	16
transgenic 658-5	40	no ID				
transgenic 658-5	41	gamma gliadin[6]	Bu-Gamma5	11	22	35
transgenic 658-5	42	gamma gliadin	Bu-Gamma5	27	59	60
transgenic 658-5	43	LMW-GS (m-type)	CS-LMW-D7	5	9	18

[1]*Protein accessions from Butte 86 begin with Bu and are reported in Dupont et al. (2011), proteins from Chinese Spring begin with CS and are reported in Huo et al. (2018a; 2018b), and proteins from Xioayan 81 begin with X and are reported in Wang et al. (2018). All other accessions are from NCBI. Proteins from Butte 86 that are very similar to the protein identified by Scaffold are shown in parentheses.*

[2]*Spot also contains gamma gliadin.*

[3]*Spot also contains other alpha gliadins.*

[4]*Spot also contains LMW-GS.*

[5]*Spot also contains delta gliadin.*

[6]*Spot also contains non-gluten protein.*

TABLE 4 | Changes in amounts of different classes of flour proteins in transgenic lines 124j and 658-5 relative to the nontransgenic.

	% Change	
	SA35-124j	**SA39b-658-5**
alpha gliadins	–70.4	–66.1
gamma gliadins	15.0	1.7
omega gliadins	8.1	–3.3
delta gliadins	26.8	4.4
HMW-GS	–26.0	–28.8
LMW-GS	–2.6	–4.4
purinins	32.0	26.4
farinins	–0.3	–0.8
triticins	9.2	–2.3
globulins	21.9	15.2
serpins	21.9	–30.5
AAI	11.8	3.6
other nongluten proteins	21.9	27.0

in both transgenic lines as well as small increases in the amounts of some of the other nongluten proteins.

The ratio of glutenin to gliadin in the nontransgenic line was 0.94 to 0.96 while that of both transgenics was slightly higher, 1.04 for 124j and 1.06 for 658-5. The ratio of HMW-GS to LMW-GS was 0.64 and 0.67 in Butte 86, but 0.49 and 0.50 in 124j and 658-5, respectively (**Supplementary Files 5** and **6**).

Immunogenic Potential of Transgenic Lines

The immunogenic potential of the transgenic lines relative to the nontransgenic line was assessed by comparing the reactivity of antibodies from patients with biopsy-confirmed cases of CD towards flour proteins from the nontransgenic and transgenic lines. Levels of serum IgG and IgA reactivity were significantly reduced for the transgenic lines when compared to the nontransgenic line as determined by ELISA ($p < 0.0001$ for all

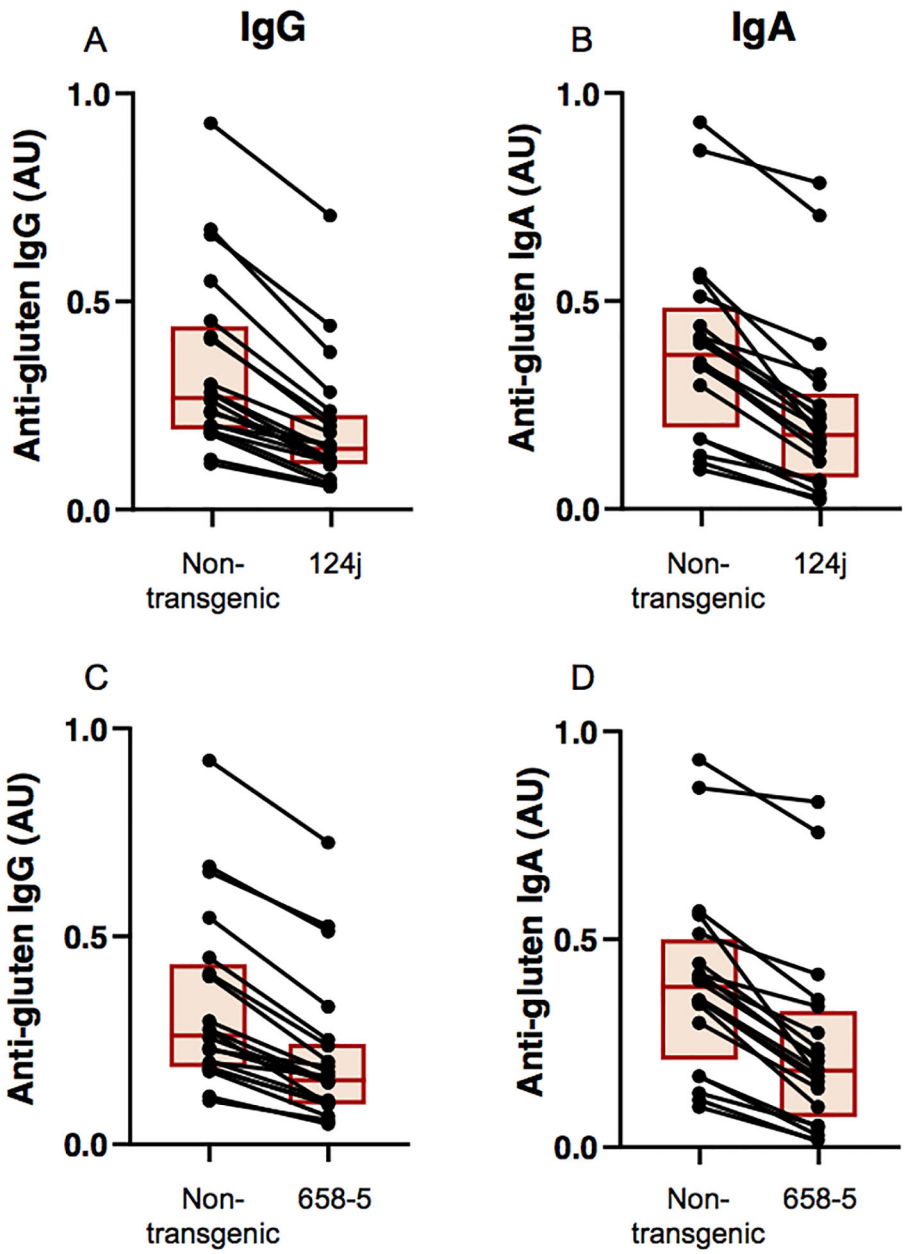


FIGURE 4 | Measurement of celiac disease antibody reactivity to nontransgenic and transgenic wheat gluten proteins. Levels of antibody reactivity towards gluten proteins from nontransgenic and transgenic 124j **(A**, **B)** and 658-5 **(C**, **D)** plants are shown for each of the 20 antigluten IgG-positive **(A**, **C)** and 20 antigluten IgA-positive **(B**, **D)** celiac disease patients, as determined by ELISA. Each individual is represented by a dot and the two points corresponding to the same individual are connected by a line. Each box indicates the 25th–75th percentiles of distribution, with the horizontal line inside the box representing the median.

comparisons) (**Figure 4**). All patients in the study had lower IgG and IgA reactivities to the transgenic lines than to the nontransgenic line, although differences were small for some patients. Reductions were similar for both transgenic lines.

The molecular specificity of immune reactivity to gluten proteins in the transgenic lines was examined by two-dimensional immunoblotting (**Figure 5**). The observed decrease in levels of IgG and IgA antigluten antibodies to transgenic lines as determined by ELISA was confirmed to be due to a reduction in antibody binding to alpha gliadins. Generally, IgG and IgA antibodies from patients reacted with a number of proteins in addition to the alpha gliadins and the profiles of reactivity varied among patients. For the representative case shown in **Figure 5A**, IgG antibodies from one patient exhibited reactivity with alpha gliadins, LMW-GS, serpins and purinins in the nontransgenic line, while in another

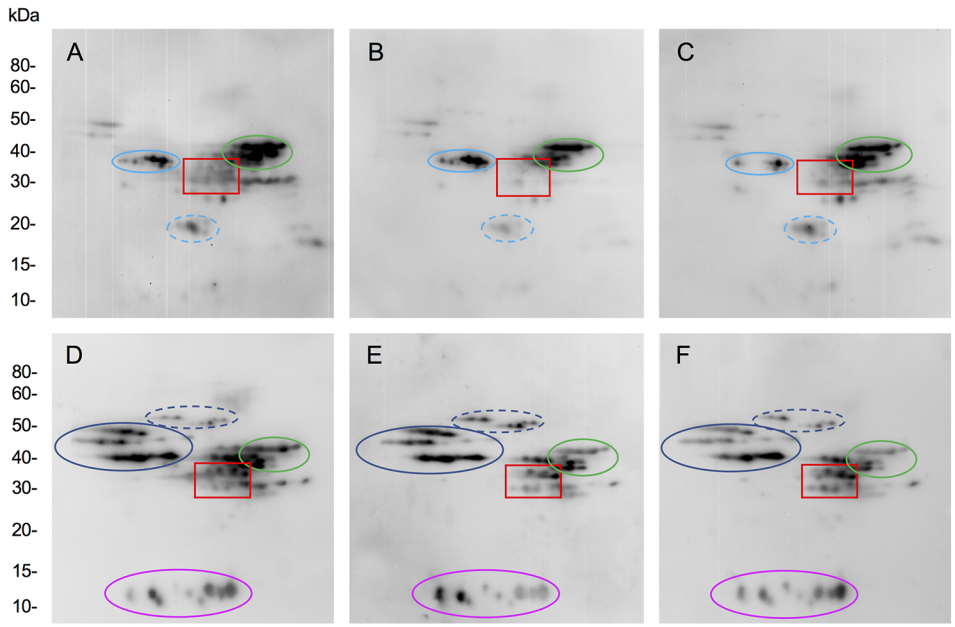


FIGURE 5 | Assessment of the molecular specificity of immune reactivity towards nontransgenic and transgenic wheat lines. Immunoblots show IgG **(A–C)** and IgA **(D–F)** antibody reactivity in representative celiac disease patients towards two dimensionally separated total flour proteins from nontransgenic **(A**, **D)** and transgenic plants 124j **(B**, **E)** and 658-5 **(C**, **F)**. Red boxes show the alpha gliadin regions. In panels **(A–C)**, the positions of serpins, low-molecular weight glutenin subunits (LMW-GS) and purinins are shown in solid light blue, green, and dashed light blue ovals. In panels **(D–F)**, the positions of omega-1,2 gliadins, omega-5 gliadins, LMW-GS, and AAI are shown in solid blue, dashed blue, green, and magenta ovals. MW are indicated on the left.

case shown in **Figure 5D**, IgA antibodies showed reactivity with alpha gliadins, omega-1,2 gliadins, some LMW-GS and gamma gliadins, omega-5 gliadins and AAI proteins. In both cases, the overall observed reduction in IgG and IgA antibody reactivity toward the transgenic lines was attributable to a reduction in reactivity to alpha gliadin proteins.

End-Use Quality Analysis of Transgenic Lines

Sufficient quantities of grain from the nontransgenic and transgenic lines were produced in the greenhouse for end-use quality testing using methods that are commonly utilized to assess breeding lines in the U.S. The average kernel weight of transgenic line 124j was similar to that of the nontransgenic control, 42.7 mg +/−0.7 versus 42.8 mg +/−2.2 while that of 658-5 was 24.6% less, 32.3 mg +/−1.6. Nonetheless, grain protein contents (%) were similar in the transgenic lines and the nontransgenic control while flour protein contents (%) were somewhat less than the control in the two transgenic lines (**Table 5**). The overall shapes of the mixing curves generated with a 10-g mixograph were similarly poor in both the transgenic and the nontransgenic lines. However, both transgenic lines had shorter mix times and peak heights than the nontransgenic control (**Figure 6**, **Table 5**). Water absorption decreased in flour from both transgenic lines relative to the control while mixing tolerance was poor in all lines. The SDS sedimentation volumes for the two transgenic lines were 69.2% and 59.6% less than that of the nontransgenic line, 21.1 ml/g for 124j and 27.7 ml/g for 658-5 as opposed to 68.6 ml/g for Butte 86 (**Table 5**).

TABLE 5 | End-use quality data from nontransgenic Butte 86 and transgenic lines 124j and 658-5.

Line	Grain protein (%)	Flour protein[1] (%)	Water Absorption[1] (%)	Mix time (min)	Mix tolerance[2]	SDS sedimentation volume (ml/g)
Butte 86[3]	19.7 (0.2)	17.4 (0.5)	73.2 (1.3)	2.0 (0.4)	0.3 (0.6)	68.6 (2.17)
124j[3]	19.2 (0.7)	15.3 (0.6)	59.8 (0.8)	0.7 (0.1)	0.0 (0.0)	21.1 (2.71)
658-5[3]	19.2 (0.1)	15.5 (0.1)	64.7 (1.7)	1.1 (0.1)	0.3 (0.6)	27.7 (3.55)

[1]*Based on 14% moisture.*
[2]*Recorded on a 0–6 scale with 6 having the greatest tolerance.*
[3]*Averages and (standard deviations) from flour samples from three biological replicates are reported.*

DISCUSSION

Given the large numbers of alpha gliadin genes as well as the high similarities and repetitiveness of their sequences, it can be difficult to find regions that are unique for specific genes. This is even more challenging when not all of the gene sequences from the cultivar of interest are known. In this study, only 13 alpha gliadin gene sequences were available from Butte 86 when the RNAi construct was designed. Based on the finding that Chinese Spring contains 26 genes encoding full-length alpha gliadins, it is likely that the available sequences account for only one half of the total number of gene sequences expressed in this cultivar.

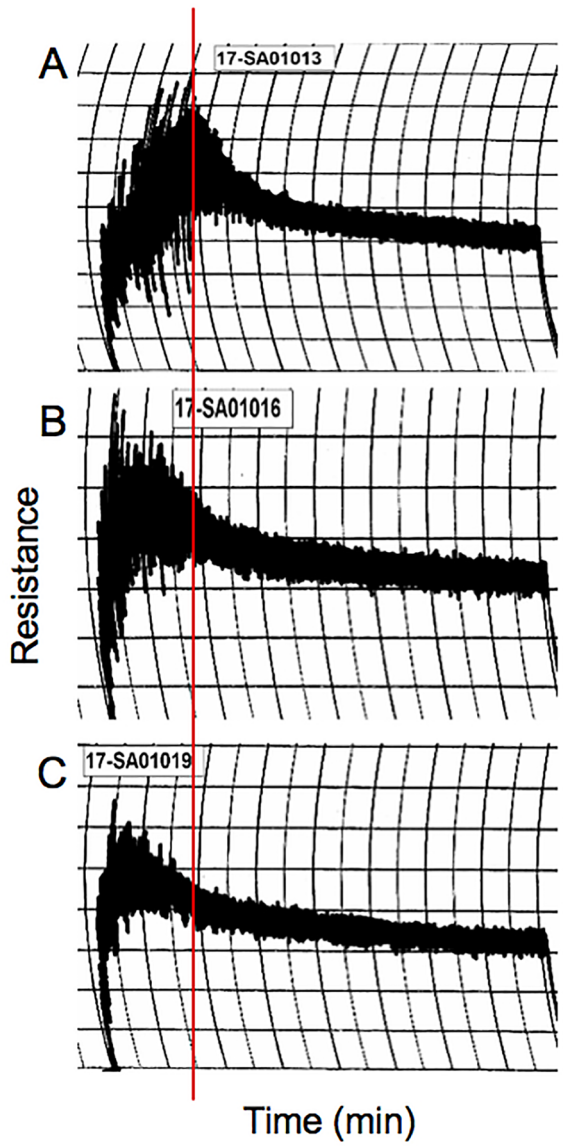


FIGURE 6 | 10 g mixograph curves produced using flour from nontransgenic **(A)** and transgenic lines 124j **(B)** and 658-5 **(C)**. The red line shows the position of the peak mixing time in the nontransgenic line.

Nonetheless, we focused on the regions of the 13 genes encoding the first nonrepetitive portion of the proteins with the aim of designing an RNAi construct that would target only those alpha gliadins containing CD epitopes. The average percentage of Q+P in this region is much lower than that of the N-terminal and repetitive regions that were used as targets in the RNAi constructs of Barro et al. (2015) and Becker et al. (2012). In fact, among the 26 alpha gliadins from Chinese Spring, the average percentage of Q+P in this region was 24.9% as opposed to 69.2% for the N-terminal and repetitive regions. Thus, it might be expected that the construct would be less likely to silence nontargeted genes or genes in other gluten protein families. To further decrease the likelihood of off-target effects, we also selected a trigger for the RNAi construct that consisted of fragments of contiguous sequence that were less than 78 bp in length. In comparison, the triggers in the Barro et al. (2015) and Becker et al. (2012) constructs contained 377 and 313 bp, respectively. Despite these efforts, the desired specificity was not achieved in our experiments. Rather, all alpha gliadin genes were silenced effectively in the transgenic plants, demonstrating that regions of identity less than 20 bp between the trigger and the target gene can result in silencing of gluten protein genes. This is consistent with reports that triggers with as few as 14 nucleotides of contiguous sequence complementarity sometimes result in suppression (Senthil-Kumar and Mysore, 2011).

Even more surprising is the partial suppression of HMW-GS genes that was observed in the transgenic lines given that there is little identity between the trigger region and HMW-GS sequences. However, the RNAi construct did include the promoter and 5' untranslated region from the Dy10 HMW-GS gene and the 3' untranslated region and terminator from the Dx5 HMW-GS gene. It is thus possible that the partial reduction of HMW-GS is due to cosuppression of the HMW-GS genes as a result of homology dependent gene silencing. Silencing of endogenous genes has been reported when HMW-GS transgenes were introduced into transgenic plants (Alvarez et al., 2000). However, it should be noted that the same HMW-GS promoter and terminator regions were included in RNAi constructs that targeted the omega-5 and omega-1,2 gliadin genes in other studies. Partial decreases in HMW-GS were observed in one of four transgenic lines in which the omega-5 gliadins were suppressed (Altenbach et al., 2014a), but not in lines in which the omega-1,2 gliadins were suppressed (Altenbach et al., 2019), suggesting that the copy number and/or site of insertion may also be important. Nonetheless, as a caution it may be wise to use avoid using regulatory regions derived from wheat gluten protein genes in future studies.

Surprisingly, there was little change in the levels of other gluten proteins in the transgenic lines as determined by quantitative 2-DE. Rather, the reductions in alpha gliadins and HMW-GS were compensated partially by increases in a number of nongluten proteins. Thus, it is possible to eliminate an entire group of gluten proteins without compensatory effects on other gluten protein classes. In contrast, Barro et al. (2015) stated that decreases in alpha gliadins were offset mostly by increases in HMW-GS while Becker et al. (2012) reported that lines that had the largest decreases in alpha gliadins showed increases in omega and gamma gliadins, HMW-GS, and albumins/globulins. Becker et al. (2012) also observed off-target suppression of LMW-GS in some of their lines. They hypothesized that some of the reductions in LMW-GS might be due to the suppression of alpha gliadins that contain an extra cysteine and thus are linked into the glutenin polymer. However, no evidence was provided to support this notion. Undoubtedly, both off-target and compensatory effects of RNA interference are complicated, particularly among gene families as complex as the gliadins and glutenins, and therefore require further study. Indeed, one study in which omega-5 gliadins genes were silenced by RNAi yielded transgenic lines in which there were minimal off-target or compensatory effects on the proteome as well as lines that showed notable changes in protein groups other than

those targeted (Altenbach et al., 2014a). This study also employed quantitative 2-DE.

Both transgenic lines exhibited a significant reduction in binding to IgG and IgA antibodies from CD patients in comparison to the nontransgenic wheat, suggesting a decreased immunogenic potential. The reductions were similar in the two lines, as expected from their similarities in protein composition. However, the reductions were less substantive than what was observed in a previous study with transgenic plants missing the omega-1,2 gliadins (Altenbach et al., 2019), demonstrating the complexity and broad range of immunogenic gluten proteins in the context of celiac disease. Clearly, additional studies that address potential T cell reactivity of flour proteins from the transgenic lines are warranted. Considering the fact that the alpha gliadins contain known T cell epitopes and that most of these epitopes are located within the B cell epitope sequences, the data here suggest that T cell reactivity to the transgenic lines would also be diminished to a similar extent.

With regards to end-use quality, the decreases in mix time in the transgenic lines relative to nontransgenic Butte 86 suggest that the altered protein compositions may have a negative effect on the mixing properties of the flour. However, given the short mix time and tolerance observed in the control, it is difficult determine the significance of the effects. Nonetheless, reductions in SDS sedimentation volumes in the transgenic lines suggest that one effect of the gene silencing was a decrease in gluten strength. This is consistent with the decrease in the ratio of HMW-GS to LMW-GS that was observed in the transgenic lines and is further supported by the observed decreases in water absorption in the transgenic lines. While the most notable difference between the nontransgenic and the transgenic lines was the absence of alpha gliadins, it must be kept in mind that the HMW-GS also showed a small decrease. Because the HMW-GS have a major effect on the functional properties of the flour (Shewry et al., 2003), even small decreases in these proteins could confound the interpretation of the quality data. In other transgenic studies, the effects of the alpha gliadins on flour functional properties were inconclusive. Becker et al. (2012) performed small-scale rheology tests on a mixture of flour from two transgenic lines because of the small amounts of transgenic material that were available and concluded that flour from the transgenic lines did not differ from controls in dough rheology. However, gluten from the transgenic lines had a higher maximum resistance to extension and a lower extensibility than gluten from the control. Barro et al. (2015) examined only SDS sedimentation volumes, but it was not possible to determine whether the increased values observed in their transgenic lines were due to the decrease in alpha gliadins or alterations in the levels of other proteins. In comparison, when omega-5 gliadins or omega-1,2 gliadins were down-regulated in transgenic plants, there was an increase in both mix time and mix tolerance, suggesting that the mixing properties of the flour were improved, and SDS sedimentation volumes were similar or slightly increased (Altenbach et al., 2014b; Altenbach et al., 2019). In the future, it may be interesting to cross transgenic plants in which the alpha gliadins have been eliminated with ones in which the omega-1,2 gliadins have been eliminated and assess both flour quality and IgG and IgA antibody reactivities of the resulting lines.

An important question is whether it will be feasible to target only those alpha gliadin genes encoding proteins with CD epitopes or, alternately, a subset of genes encoding proteins with the greatest numbers of epitopes. To achieve this goal, it will be important first to obtain all of the alpha gliadin gene sequences from the cultivar of interest. The availability of a reference genome sequence from Chinese Spring makes it possible to design gene capture methods to obtain complete sets of gluten protein genes from different cultivars. This might be accomplished using a capture system that includes baits for all high-confidence exons from the International Wheat Genome Sequencing Consortium (IWGSC) genome assembly of Chinese Spring that is commercially available from Arbor Biosciences (Ann Arbor, MI). Alternately, baits for the capture system may be specific for genomic regions encoding the major gluten proteins in Chinese Spring as annotated by Huo et al. (2018a; 2018b) or based on the sequences of all known gluten genes from various Triticeae species and cultivars (Jouanin et al., 2019). However, even with complete sequence information it will be very challenging to identify regions that can be used as triggers in RNAi constructs given the similarities in the sequences of the different genes. Genome editing using CRISPR/Cas9 is an alternate approach that promises greater specificity since it requires only 20 bp of identical sequence for the guide RNAs that determine the sites of the mutations introduced into target genes. In addition, genome editing approaches potentially could be used to alter specific epitope sequences within alpha gliadin genes, as suggested by Ruiz-Carnicer et al. (2019). However, off-target mutations have also been reported in CRISPR/Cas9 edited plants (Endo et al., 2015). Additionally, genome editing is not simple in a family as complex as the alpha gliadins. The method can create indels of various sizes in both expressed genes and pseudogenes that either eliminate proteins or introduce new protein variants. And, when multiple genes are present in tandem in the genome, as in the case of the alpha gliadins, one or more genes may be deleted. Without a doubt, the challenges are many to achieve the long-term goal of reducing the immunogenic potential of wheat. But insight into both the complement of proteins in wheat flour and the roles that different groups of wheat gluten proteins play in determining the functional properties of the flour should make it easier to do so while retaining the unique viscoelastic properties of the flour.

DATA AVAILABILITY STATEMENT

The mass spectrometry data was uploaded to ProteomeXchange *via* the PRIDE database. The dataset identifier is PXD016930 and 10.6019/PXD016930.

AUTHOR CONTRIBUTIONS

SA designed the study, analyzed the data, and wrote the manuscript. AA contributed to study design, assay protocol development, data analysis, and writing of the manuscript. MR was responsible for

designing the RNAi construct, transforming the plants, and identifying transgenic lines. H-CC conducted 2-DE analyses. H-CC and AS-B were responsible for MS/MS analyses. XY was responsible for immunoassay experiments and interpretation of data. BS was responsible for end-use quality testing and interpretation of results. PG was responsible for subject recruitment and clinical characterization of patients. All authors contributed to editing of the manuscript and approved the manuscript.

FUNDING SOURCES

This work was supported by the United States Department of Agriculture, Agricultural Research Service CRIS 2030-21430-014-00D (to SBA). Additional support was provided by the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant Number UL1 TR000040 (to AA).

ACKNOWLEDGMENTS

The authors thank Anita Cheung, Trenton Owens, and Julian DeLellis-Mitch for excellent technical assistance and Leslie Harden and Melissa Erickson-Beltran for expert advice and assistance with mass spectrometry work. Mention of a specific product by the USDA does not constitute an endorsement and does not imply a recommendation over other suitable products. USDA is an equal opportunity provider and employer.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00020/full#supplementary-material

REFERENCES

AACC International. (1961). Approved methods of analysis, 11th Ed. Method 56-60.01. Sedimentation test for flour. Approved April 13, 1961. AACC International, St. Paul, MN, U.S.A. doi: 10.1094/AACCIntMethod-56-60.01

AACC International. (1985). Approved methods of analysis, 11th Ed. Method 39-11.01. Near-infrared reflectance method for protein determination in wheat flour. Approved September 26, 1985. AACC International, St. Paul, MN, U.S.A. doi: 10.1094/AACCIntMethod-39-11.01

AACC International. (1988). Approved methods of analysis, 11th Ed. Method 26-10.02. Experimental milling: introduction, equipment, sample preparation, and tempering. Approved October 12, 1988. AACC International, St. Paul, MN, U.S.A. doi: 10.1094/AACCIntMethod-26-10.02

AACC International. (1995). Approved methods of analysis, 11th Ed. Method 54-40.02. Mixograph method. Approved November 8, 1995. AACC International, St. Paul, MN, U.S.A. doi: 10.1094/AACCIntMethod-54-40.02

Alaedini, A., and Green, P. H. (2005). Narrative review: celiac disease: understanding a complex autoimmune disorder. *Ann. Intern. Med.* 142, 289–298. doi: 10.7326/0003-4819-142-4-200502150-00011

Altenbach, S. B., and Allen, P. V. (2011). Transformation of the US bread wheat 'Butte 86' and silencing of omega-5 gliadin genes. *GM Crops* 2, 67–74. doi: 10.4161/gmcr.2.1.15884

Altenbach, S. B., Dupont, F. M., Kothari, K. M., Chan, R., Johnson, E. L., and Lieu, D. (2003). Temperature, water and fertilizer influence the timing of key events during grain development in a US spring wheat. *J. Cereal Sci.* 37, 9–20. doi: 10.1006/jcrs.2002.0483

Altenbach, S. B., Vensel, W. H., and DuPont, F. M. (2010). Integration of transcriptomic and proteomic data from a single wheat cultivar provides new tools for understanding the roles of individual alpha gliadin proteins in flour quality and celiac disease. *J. Cereal Sci.* 52, 143–151. doi: 10.1016/j.jcs.2010.04.006

Altenbach, S. B., Vensel, W. H., and Dupont, F. M. (2011). The spectrum of low molecular weight alpha-amylase/protease inhibitor genes expressed in the US bread wheat cultivar Butte 86. *BMC Res. Notes* 4, 242. doi: 10.1186/1756-0500-4-242

Altenbach, S. B., Tanaka, C. K., and Allen, P. V. (2014a). Quantitative proteomic analysis of wheat grain proteins reveals differential effects of silencing of omega-5 gliadin genes in transgenic lines. *J. Cereal Sci.* 59, 118–125. doi: 10.1016/j.jcs.2013.11.008

Altenbach, S. B., Tanaka, C. K., and Seabourn, B. W. (2014b). Silencing of omega-5 gliadins in transgenic wheat eliminates a major source of environmental variability and improves dough mixing properties of flour. *BMC Plant Biol.* 14, 1. doi: 10.1186/s12870-014-0393-1

Altenbach, S. B., Chang, H.-C., Simon-Buss, A., Mohr, T., Huo, N., and Gu, Y. Q. (2020). Exploiting the reference genome sequence of hexaploid wheat: a proteomic study of flour proteins from the cultivar Chinese Spring. *Func. Int. Genomics*. 20, 1–16. doi: 10.1007/s10142-019-00694-z

Altenbach, S. B., Chang, H.-C., Yu, X. B., Seabourn, B. W., Green, P. H., and Alaedini, A. (2019). Elimination of omega-1,2 gliadins from bread wheat (*Triticum aestivum*) flour: effects on immunogenic potential and end-use quality. *Front. Plant Sci.* 10, 580. doi: 10.3389/fpls.2019.00580

Alvarez, M. L., Guelman, S., Halford, N. G., Lustig, S., Reggiardo, M. I., Ryabushkina, N., et al. (2000). Silencing of HMW glutenins in transgenic wheat expressing extra HMW subunits. *Theor. Appl. Genet.* 100, 319–327. doi: 10.1007/s001220050042

Barro, F., Iehisa, J. C. M., Giménez, M. J., García-Molina, M. D., Ozuna, C. V. Comino, I., et al. (2015). Targeting of prolamins by RNAi in bread wheat: effectiveness of seven silencing-fragment combinations for obtaining lines devoid of coeliac disease epitopes from highly immunogenic gliadins. *Plant Biotech. J.* 14, 986–996. doi: 10.1111/pbi.12455

Becker, D., Wieser, H., Koehler, P., Folck, A., Mühling, K. H., and Zörb, C. (2012). Protein composition and techno-functional properties of transgenic wheat with reduced α-gliadin content obtained by RNA interference. *J. Appl. Bot. Food Qual.* 85, 23–33.

Christensen, A. H., and Quail, P. H. (1996). Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Res.* 5, 213–218. doi: 10.1007/BF01969712

Dupont, F. M., Vensel, W. H., Tanaka, C. K., Hurkman, W. J., and Altenbach, S. B. (2011). Deciphering the complexities of the wheat flour proteome using quantitative two-dimensional electrophoresis, three proteases and tandem mass spectrometry. *Proteome Sci.* 9, 10. doi: 10.1186/1477-5956-9-10

Endo, M., Mikami, M., and Toki, S. (2015). Multigene knockout utilizing off-target mutations of the CRISPR/Cas9 system in rice. *Plant Cell Physiol.* 56, 41–47. doi: 10.1093/pcp/pcu154

Huo, N., Zhang, S., Zhu, T., Dong, L., Mohr, T., Hu, T., et al. (2018a). Gene duplication and evolution dynamics in the homeologous regions harboring multiple prolamin and resistance gene families in hexaploid wheat. *Front. Plant Sci.* 9, 673. doi: 10.3389/fpls.2018.00673

Huo, N., Zhu, T., Altenbach, S., Dong, L., Wang, Y., Mohr, T., et al. (2018b). Dynamic evolution of α-gliadin prolamin gene family in homeologous genomes of hexaploid wheat. *Sci. Rep.* 8, 5181. doi: 10.1038/s41598-018-23570-5

International Wheat Genome Sequencing Consortium (IWGSC). (2018). Shifting the limits in wheat research and breeding through a fully annotated and anchored reference genome sequence. *Science* 361, eaar7191. doi: 10.1126/science.aar7191

Jouanin, A., Borm, T., Boyd, L. A., Cockram, J., Leigh, F., Santos, B. A. C. M., et al. (2019). Development of the GlutEnSeq capture system for sequencing gluten gene families in hexaploid bread wheat with deletions of mutations induced by γ-irradiation or CRISPR/Cas9. *J. Cereal Sci.* 88, 157–166. doi: 10.1016/j.jcs.2019.04.008

Koning, F. (2012). Celiac disease: quantity matters. *Semin. Immunopathol.* 34, 541–549. doi: 10.1007/s00281-012-0321-0

Lau, N. M., Green, P. H. R., Taylor, A. K., Hellberg, D., Ajamian, M., Tan, C. Z., et al. (2013). Markers of celiac disease and gluten sensitivity in children with autism. *PloS One* 8 (6), e66155. doi: 10.1371/journal.pone.0066155

Malalgoda, M., Meinhardt, S. W., and Simsek, S. (2018). Detection and quantitation of immunogenic epitopes related to celiac disease in historical and modern hard red spring wheat cultivars. *Food Chem.* 264, 101–107. doi: 10.1016/j.foodchem.2018.04.131

Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J., et al. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* 47 (D1), D442–D450. doi: 10.1093/nar/gky1106

Pilolli, R., Gadaleta, A., Mamone, G., Nigro, D., DeAngelis, E., Montemurro, N., et al. (2019). Scouting for naturally low-toxicity wheat genotypes by a multidisciplinary approach. *Sci. Rep.* 9, 1646. doi: 10.1038/s41598-018-36845-8

Prandi, B., Tedeschi, T., Folloni, S., Galverna, G., and Sforza, S. (2016). Peptides from gluten digestion: a comparison between old and modern wheat varieties. *Food Res. Int.* 91, 92–102. doi: 10.1016/j.foodres.2016.11.034

Ribeiro, M., Rodriquez-Quijano, M., Nunes, F. M., Carrillo, J. M., Branlard, G., and Igrejas, G. (2016). New insights into wheat toxicity: breeding did not seem to contribute to a prevalence of potential celiac disease's immunostimulatory epitopes. *Food Chem.* 213, 8–18. doi: 10.1016/j.foodchem.2016.06.043

Ruiz-Carnicer, A., Comino, I., Segura, V., Ozuna, C. V., Moreno, M. D. L., López-Casado, M. A., et al. (2019). Celiac immunogenic potential of α-gliadin epitope variants from *Triticum* and *Aegilop* species. *Nutrients* 11, 220. doi: 10.3390/nu11020220

Sánchez-Léon, S., Gil-Humanes, J., Ozuna, C. V., Giménez, M. J., Sousa, C., Voytas, D., et al. (2018). Low-gluten, nontransgenic wheat engineered with CRISPR/Cas9. *Plant Biotech. J.* 16, 902–910. doi: 10.1111/pbi.12837

Salentijn, E. M. J., Esselink, D. G., Goryunova, S. V., van der Meer, I. M., Gliissen, L. J. W. J., and Smulders, M. J. M. (2013). Quantitative and qualitative differences in celiac disease epitopes among durum wheat varieties identified through deep RNA-amplicon sequencing. *BMC Genomics* 14, 905. doi: 10.1186/1471-2164-14-905

Samaroo, D., Dickerson, F., Kasarda, D. D., Green, P. H. R., Briani, C., Yolken, R. H., et al. (2010). Novel immune response to gluten in individuals with schizophrenia. *Schizophr. Res.* 118, 248–255. doi: 10.1016/j.schres.2009.08.009

Senthil-Kumar, M., and Mysore, K. S. (2011). "Caveat of RNAi in plants: the off-target effect," in *RNAi and plant gene function analysis, Methods Mol. Biol.*, vol. 744 . Eds. H. Kodama and A. Komamine, (Humana Press) 13–25. doi: 10.1007/978-1-61779-123-9_2

Shan, L., Molberg, O., Parrot, I., Hausch, F., Filia, F., Gray, G. M., et al. (2002). Structural basis for gluten intolerance in celiac sprue. *Science* 297, 2275–2279. doi: 10.1126/science.1074129

Shewry, P. R., Halford, N. G., Tatham, A. S., Popineau, Y., Lafiandra, D., and Belton, P. S. (2003). The high molecular weight subunits of wheat glutenin and their role in determining wheat processing properties. *Adv. Food Nutr. Res.* 45, 219–302. doi: 10.1016/S1043-4526(03)45006-7

Shewry, P. R. (2019). What is gluten-Why is it special? *Front. Nutrition.* 6, 101. doi: 10.3389/fnut.2019.00101

Singh, P., Arora, A., Strand, T. A., Leffler, D. A., Catassi, C., Green, P. H., et al. (2018). Global prevalence of celiac disease: systematic review and meta-analysis. *Clin. Gastro. Hepat.* 16, 823–836. doi: 10.1016/j.cgh.2017.06.037

Sollid, L. M., Qiao, S.-W., Anderson, R. P., Gianfrani, C., and Konig, F. (2012). Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenetics* 64, 455–460. doi: 10.1007/s00251-012-0599-z

Tye-Din, J. A., Stewart, J. A., Dromey, J. A., Beissbarth, T., van Heel., D. A., Tatham, A., et al. (2010). Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Sci. Transl. Med.* 4, 41ra51. doi: 10.1126/scitranslmed.3001012

van den Broeck, H. C., de Jong, H. C., Salentijn, E. M. J., Dekking, L., Bosch, D., Hamer, R. J., et al. (2010). Presence of celiac disease epitopes in modern and old hexaploid wheat varieties: wheat breeding may have contributed to increased prevalence of celiac disease. *Theor. Appl. Genet.* 121, 1527–1539. doi: 10.1007/s00122-010-1408-4

van den Broeck, H. C., Cordewener, J. H. G., Nessen, M. A., America, A. H. P., and van der Meer, I. M. (2015). Label free targeted detection and quantification of celiac disease immunogenic epitopes by mass spectrometry. *J. Chrom.* 1392, 60–71. doi: 10.1016/j.chroma.2015.02.070

Vensel, W. H., Tanaka, C. K., and Altenbach, S. B. (2014). Protein composition of wheat gluten polymer fractions determined by quantitative two-dimensional gel electrophoresis and tandem mass spectrometry. *Proteome Sci.* 12, 8. doi: 10.1186/1477-5956-12-8

Wang, D. W., Li, D., Wang, J., Zhao, Y., Wang, Z., Yue, G., et al. (2017). Genome-wide analysis of complex wheat gliadins, the dominant carriers of celiac disease epitopes. *Sci. Rep.* 7, 44609. doi: 10.1038/srep44609

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Nutrition

REVIEW
published: 17 March 2020
doi: 10.3389/fnut.2020.00029

Recent Progress and Recommendations on Celiac Disease From the Working Group on Prolamin Analysis and Toxicity

*Katharina A. Scherf[1], Carlo Catassi[2], Fernando Chirdo[3], Paul J. Ciclitira[4], Conleth Feighery[5], Carmen Gianfrani[6], Frits Koning[7], Knut E. A. Lundin[8], Detlef Schuppan[9], Marinus J. M. Smulders[10], Olivier Tranquet[11], Riccardo Troncone[12] and Peter Koehler[13]**

[1] Department of Bioactive and Functional Food Chemistry, Institute of Applied Biosciences, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, [2] Department of Pediatrics, Polytechnic University of Marche, Ancona, Italy, [3] Instituto de Estudios Inmunologicos y Fisiopatologicos- IIFP (UNLP-CONICET), Universidad Nacional de La Plata, La Plata, Argentina, [4] Norwich Medical School, University of East Anglia, Norwich, United Kingdom, [5] St. James's Hospital, University of Dublin, Dublin, Ireland, [6] Institute of Biochemistry and Cell Biology, Italian National Council of Research, Naples, Italy, [7] Leiden University Medical Centre, Leiden, Netherlands, [8] Department of Gastroenterology, Oslo University Hospital Rikshospitalet and Stiftelsen KG Jebsen Coeliac Disease Research Centre, University of Oslo, Oslo, Norway, [9] Institute for Translational Medicine, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany, [10] Plant Breeding, Wageningen University & Research, Wageningen, Netherlands, [11] INRAE, UR BIA, Nantes, France, [12] European Laboratory for the Investigation of Food Induced Diseases (ELFID), Department of Medical Translational Sciences, University Federico II, Naples, Italy, [13] biotask AG, Esslingen, Germany*

OPEN ACCESS

Edited by:
Jun Lu,
Auckland University of Technology, New Zealand

Reviewed by:
Sachin Rustgi,
Clemson University, United States
Govind K. Makharia,
All India Institute of Medical Sciences, India
Emanuele Zannini,
University College Cork, Ireland
Luis Rodrigo,
Central University Hospital of Asturias, Spain

***Correspondence:**
Peter Koehler
peter.koehler@biotask.de

Specialty section:
This article was submitted to Nutrition and Food Science Technology, a section of the journal Frontiers in Nutrition

Received: *15 December 2019*
Accepted: *26 February 2020*
Published: *17 March 2020*

Citation:
Scherf KA, Catassi C, Chirdo F, Ciclitira PJ, Feighery C, Gianfrani C, Koning F, Lundin KEA, Schuppan D, Smulders MJM, Tranquet O, Troncone R and Koehler P (2020) Recent Progress and Recommendations on Celiac Disease From the Working Group on Prolamin Analysis and Toxicity. Front. Nutr. 7:29. doi: 10.3389/fnut.2020.00029

Celiac disease (CD) affects a growing number of individuals worldwide. To elucidate the causes for this increase, future multidisciplinary collaboration is key to understanding the interactions between immunoreactive components in gluten-containing cereals and the human gastrointestinal tract and immune system and to devise strategies for CD prevention and treatment beyond the gluten-free diet. During the last meetings, the Working Group on Prolamin Analysis and Toxicity (Prolamin Working Group, PWG) discussed recent progress in the field together with key stakeholders from celiac disease societies, academia, industry and regulatory bodies. Based on the current state of knowledge, this perspective from the PWG members provides recommendations regarding clinical, analytical and legal aspects of CD. The selected key topics that require future multidisciplinary collaborative efforts in the clinical field are to collect robust data on the increasing prevalence of CD, to evaluate what is special about gluten-specific T cells, to study their kinetics and transcriptomics and to put some attention to the identification of the environmental agents that facilitate the breaking of tolerance to gluten. In the field of gluten analysis, the key topics are the precise assessment of gluten immunoreactive components in wheat, rye and barley to understand how these are affected by genetic and environmental factors, the comparison of different methods for compliance monitoring of gluten-free products and the development of improved reference materials for gluten analysis.

Keywords: barley, celiac disease, gluten, gluten-free diet, Prolamin Working Group, rye, wheat

INTRODUCTION

About 60% of agricultural land worldwide is used to grow cereals, with maize (1,135 × 10^6 metric tons), wheat (772 × 10^6 metric tons) and paddy rice (770 × 10^6 metric tons) as major crops in terms of global production (data from 2017, FAOSTAT[1]). As an important source of dietary nutrients such as carbohydrates, proteins, dietary fiber, vitamins and minerals, wheat is an essential cornerstone for food security. However, the consumption of products made of wheat, and the closely related cereals rye and barley, may cause adverse reactions such as celiac disease (CD), non-celiac gluten/wheat sensitivity (NCGS) and wheat allergy. With increasing evidence from epidemiological studies pointing to a large number of affected individuals in many countries around the world, there is a strong need to understand the fundamental interactions between immunoreactive components in gluten-containing cereals and the human gastrointestinal tract and immune system in order to develop strategies for disease prevention and treatment beyond the gluten-free diet (GFD). The term "gluten" includes the closely related storage proteins of wheat (gliadins and glutenins), rye (secalins), barley (hordeins), and oats (avenins). The part of gluten soluble in aqueous alcohols has been termed prolamins and the insoluble part glutelins.

CD is defined as a lifelong small intestinal immune enteropathy with autoimmune features caused by ingestion of gluten from wheat, rye and barley in subjects with a dominant and necessary genetic predisposition [human leukocyte antigen (HLA)-DQ2 or -DQ8] (1). The main known environmental factor responsible for CD is the consumption of gluten, but there still needs to be a largely unknown factor as initial trigger of the disease. Certain viruses and bacteria are prime suspects, and the idea is that virus infection can prime the immune system in susceptible individuals so that not only the virus is recognized and defeated but the intestinal immune system also misinterprets gluten as "dangerous" (*vide infra*). Patients develop characteristic mucosal (usually IgA) antibodies to the autoantigen tissue transglutaminase [TG2, (2)]. TG2 can deamidate gluten peptides, which improves their presentation by HLA-DQ2 or -DQ8 on antigen-presenting cells of the intestinal mucosa, and this increases their T-cell stimulatory potential (3, 4). While such gluten-specific T-cell responses are characteristic for CD, it is unclear which events cause the loss of mucosal tolerance to food antigens in CD. Many studies now imply a role for additional environmental agents, including the exposure to (intestinal) viruses and bacteria. CD is a systemic disorder that predominantly manifests itself in the mucosa of the upper small intestine (duodenum, proximal jejunum) and is characterized by villous atrophy and crypt hyperplasia, which can vary from mild partial damage to a total absence of villi. As a clinical chameleon (5), CD presents in symptomatic, asymptomatic, potential and refractory forms and can occur at any age. Notably, CD often also presents with a wide variety of extra-intestinal symptoms, including associated autoimmune diseases (6–9). The only effective treatment so far is a GFD that essentially relies on the consumption of naturally gluten-free foods such as animal-based products, fruits, vegetables, legumes and nuts as well as dietary gluten-free products that may not contain more than 20 mg/kg of gluten according to Codex Alimentarius (Codex Standard 118-1979[2]). There are several ongoing attempts to develop non-dietary treatments of the disease—this is briefly discussed later.

Founded in 1985 by Professor Wim Hekkens, University of Leiden, The Netherlands, the Working Group on Prolamin Analysis and Toxicity (Prolamin Working Group, PWG) coordinates multidisciplinary research efforts primarily related to CD. The PWG currently has 13 executive members all of whom are renowned experts in the fields of pediatric and adult gastroenterology, immunology, biochemistry, plant science, food chemistry, and gluten analysis. Building upon this unique multidisciplinary knowledgebase, the PWG has made important achievements both in clinical research into CD and in improving food safety for CD patients by advancing analytical methods for gluten detection.

Some of the highlighted clinical research work of the PWG include the assessment of the safety of oats in the GFD (10–12), the establishment of 10 mg of gluten intake per day as the safe gluten threshold for the vast majority of CD patients (13), the search for wheat species with a reduced content of immunogenic sequences for disease prevention (14, 15), and the study of the signals for T- and B-cell recruitment into the *lamina propria* and epithelial compartment (16).

Having been granted observer status at Codex Alimentarius in 1999, the PWG plays a leading role in the development of enzyme-linked immunosorbent assays (ELISA) for gluten analysis (17) and the validation of such methods in collaboration with the Cereals & Grains Association [formerly known as AACC International; (18–20)], and AOAC International (21). It also produced the only well-characterized reference material, the so-called PWG-gliadin (22) that is used to calibrate a variety of gluten analytical methods and is available in 100 mg batches from the Association of Cereal Research (Arbeitsgemeinschaft Getreideforschung e.V., Detmold, Germany).

During its annual meetings, the PWG regularly unites a select group of about 60 international stakeholders including researchers, celiac disease societies, regulatory bodies, manufacturers of gluten-free foods and raw materials, and manufacturers of test systems for gluten analysis in foods. This paper will report the recent progress and recommendations that were presented and discussed during the last PWG meetings.

UPDATE ON CLINICAL ASPECTS OF CD

The Epidemiology of CD

In several countries the epidemiology of CD has been intensively investigated during these last decades (23, 24). In these studies, the incidence of CD is calculated by counting the number of new CD diagnoses in a population over a given period of time,

[1]Food and Agriculture Organization of the United Nations (accessed December 11, 2019).

[2]Codex standard for foods for special dietary use for persons intolerant to gluten. Codex Alimentarius Commission; revision 2008, amendment 2015.

usually 1 year. On the other hand, the overall prevalence of CD is determined through mass CD screening of general population samples. The screening algorithm usually consists of serological tests like IgA class anti-transglutaminase (TG2) antibodies. In some of the studies, positive serology is backed up by gastroscopy with duodenal biopsies for final confirmation of CD on an individual basis. Taken together, these studies have shown that there have been substantial increases in prevalence and incidence over the last two decades (24).

Prevalence of CD on a Worldwide Basis

According to a recent meta-analysis, the pooled worldwide prevalence of CD autoimmunity is 1.4% (95% confidence interval, CI: 1.1–1.7%), based on positive results from tests for IgA anti-TG2 and/or anti-endomysial antibodies (so-called seroprevalence). This study found that the pooled global prevalence of biopsy-confirmed CD is 0.7% (95% CI: 0.5–0.9%) with wide regional variations. CD prevalence is 0.4% in South America, 0.5% in Africa and North America, 0.6% in Asia, and 0.8% in Europe and Oceania; it is higher in female vs. male individuals (0.6 vs. 0.4%; $p < 0.001$), and significantly greater in children than adults (0.9 vs. 0.5%) (25). It should however be noted that including only biopsy-confirmed CD cases tends to underestimate the true CD prevalence (as it seems to be the case for North America) since cases of potential CD (CD serology positive with normal/nearly normal intestinal mucosa at the small intestinal biopsy) are excluded from the prevalence calculation. In some European countries, e.g., Sweden, Finland, and Italy, data indeed show a significantly higher overall CD prevalence (1.6–2.3%) (26, 27). Generally speaking, the prevalence of CD is directly related to the population prevalence of HLA-DQ2 or -DQ8 (30–40% in most Western countries) and to the average level of wheat consumed per capita, as shown by data from India: CD is much more common in the Northern part of the country where wheat is the staple food (CD prevalence = 1.2), than in the Southern part with both a lower prevalence of HLA-DQ2/DQ8 and a lower wheat consumption (CD prevalence = 0.13%) (28).

The Concept of the Celiac Iceberg

CD screening studies have clearly shown that the percentage of cases that are diagnosed clinically (the visible part of the iceberg) is much smaller than the overall CD prevalence (the submerged CD iceberg). The clinical severity of those detected in regular clinical care and those detected by screening, does not, however, seem to differ (29). In countries showing a high level of awareness of the CD clinical spectrum, still 50–75% of cases remain undiagnosed and are therefore exposed to the risks of long-term complications. In some countries, such as India and China, the visible CD iceberg is <5% of the overall "mountain of ice." How to increase the CD diagnostic rate (e.g., via mass-screening or case-finding) is still a matter of debate (30, 31).

Is CD Prevalence Increasing Over Time?

Studies from several countries, particularly the US, Finland and Italy, suggest that the overall CD prevalence is increasing over time. For instance, the analysis of "old" sera samples taken at two different time-points (15 years apart), coupled with recent population screening data suggested that CD prevalence increased 5-fold in the US during a 50-year period beginning from 1948 to 1954 (32). The environmental factor/s responsible for this huge increase are still unclear (33). A recent study in Denmark showed that the prevalence of diagnosed CD has doubled every decade from 1986 to 2016, the female/male ratio has increased, and also the prevalence of autoimmune comorbidity in 2016 was three times higher among CD patients compared with the general Danish population (34).

Risk Factors

A number of prospective studies have been performed to identify risk factors for CD. They were focused on the genetic factors predisposing to the disease (in this context the dose of HLA-DQ2 seems to play the most important role) and on environmental factors that increase the risk of developing celiac autoimmunity and then mucosal damage.

Infant Feeding

Amongst those studies, two (PreventCD and CeliPrev) have carried out an intervention based on the timing of gluten introduction in infants. Other observational studies have assessed the relationship between infant feeding practices and the risk of CD (Generation R, Norwegian Mother and Child Cohort Study, BabiDiab, TEDDY). In general, prospective studies have not been able to confirm the previous findings that both age of gluten introduction and breastfeeding influence CD risk (35, 36). Recent epidemiologic studies reported a positive correlation between the incidence of CD cases and the amount of gluten in the diet within the early years of life (37–39), but there are also reports questioning this relationship (40–42). Further immunological and multicentre studies are mandatory to assess whether a reduced gluten exposure in early life may protect from CD onset in predisposed individuals.

Early Events

A large registry-based cohort study that included over 1.5 million children from Denmark and Norway found no association between the mode of delivery (cesarean section vs. vaginal birth) and the risk of diagnosed CD (43). Data collected from the same cohort indicated that exposure to systemic antibiotics in the first year of life was positively associated with diagnosed CD, with a dose-dependent relation between an increasing number of dispensed antibiotics and CD risk (44). However, a recent systematic review of two studies on prenatal and three studies on postnatal antibiotic exposure reported contradictory results and thus rather excluded an association between antibiotic use and the risk of developing CD (45), as already suggested by the TEDDY study (46).

Infections

Longitudinal prospective studies have suggested an association between frequent rotavirus infection and an increased risk of CD (47). A protective effect of rotavirus vaccination has also been

reported (48). Both reovirus and norovirus (49) have been shown to be able to break oral tolerance in murine models and there is evidence for the role of reovirus in the pathogenesis of CD (50). In addition, infections with enterovirus A and B, especially with high titer and long duration, during early childhood were associated with later CD, whereas adenovirus infections were unlikely to contribute to CD onset (51). Interestingly, also the occurrence of acute respiratory infections seems to play a role (52).

Microbiota

Microbiota has been hypothesized to influence the risk of developing CD. Studies on active CD patients have suggested that microbiota from CD patients may harbor more pathogenic or proinflammatory bacteria (53, 54). However, in such studies on active CD patients, it cannot be stated whether dysbiosis is a risk factor for CD or a consequence of mucosal damage and inflammation. In infants carrying the high risk genotype a reduced number of Bifidobacterium (*B. longus*) was found. Early alterations of the proportions of Firmicutes were noted in children who later progressed to CD (55). However, in another study no statistically significant differences in the fecal microbiota composition were found between children who later developed CD and the control children without disease or associated autoantibodies (56). Mouse experimental studies, including fecal transplants from patients, demonstrated a protective effect of certain lactobacilli that are able to degrade immunogenic gluten peptides, thus alleviating small intestinal damage (57). Microbe-host interactions were recently identified as relevant factors in the development of food sensitivities. Duodenal biopsies from CD patients displayed increased proteolytic activity due to higher abundance of Proteobacteria that express gluten-degrading enzymes such as elastase (LasB). This resulted in the activation of an elastase-dependent, but gluten-independent inflammatory response mediated by the protease-activated receptor 2 pathway. In the presence of CD risk genes, a synergistic effect between elastase and gluten was demonstrated, thus highlighting the importance of microbiota in modulating the host immune response (58).

Natural History

The natural history of CD proceeds based on the prominent genetic risk (HLA-DQ2/DQ8) and exposure to environmental risk factors finally to break the oral tolerance to gluten. The seroconversion with the appearance of anti-TG2 autoantibodies (CD autoimmunity) is considered to be a sign of activation of anti-gluten adaptive immunity, being sustained by gluten-specific T cells. However, not all subjects in this stage appear to progress further to villous atrophy and consequent malabsorption (59). Some may remain at this stage with no histological damage or very mild lesions. This condition has been dubbed "potential CD." In a subset, anti-TG2 antibodies may fluctuate or even disappear (60, 61). Recently, factors predicting such evolution have been suggested, such as increased density of intraepithelial γδ T cells, small intestinal mucosal deposits of anti-TG2 antibodies, and HLA dose (62).

Clinical Gluten Challenge and the Adaptive Response to Gluten

For decades, the roles of the adaptive and the innate immune system as the key players in CD immunopathogenesis have been discussed. The very clear genetic association primarily to HLA-DQ2.5 and to a lesser degree to HLA-DQ2.2 and HLA-DQ8, and the finding that these HLA molecules present gluten to *lamina propria* $CD4^+$ T cells all argue for a prominent role of the adaptive immune system (63–66). In addition, there is a very strong HLA-DQ2 gene dose effect which correlates with stronger gluten-specific T-cell responses in individuals homozygous for HLA-DQ2.5 compared to heterozygotes (67). Very recent studies demonstrated that the DQ2.5 genes are more expressed than non-CD associated alleles in antigen-presenting cells heterozygous for DQ2.5. This differential expression of CD risk genes affects the level of the encoded DQ2.5 molecules on the cell surface and the strength of gluten-specific $CD4^+$ T-cell response (68–70). According to these findings, the magnitude of the T-cell response appears more prominent dependent on the amount of gluten and less on the DQ2.5 gene doses. Although innate effects of gluten also may be important, these are generally only found in patients with CD and not in healthy individuals. *Lamina propria* $CD4^+$ T cells recognize certain peptides from the gluten protein types α/β-, γ-, ω-gliadins, and high-molecular-weight glutenin subunits (HMW-GS) where the common denominator of this peptide recognition is that the peptides are deamidated by the enzyme TG2 (2, 3, 71, 72), although some sequences do not need deamidation to be CD-active. The set of peptides presented by any given of the two CD-associated HLA molecules share common features where the charged amino acids of the gluten peptides fit into pockets of HLA-DQ2/DQ8 molecules (9, 73–75). The $CD4^+$ T cells preferentially recognize these, partly deamidated peptides that cluster in proline- and glutamine-rich stretches of the gluten proteins. However, the peptide sequence of only ≈50% of the total number of such *lamina propria* T cells can be accounted for. Although often termed an "autoimmune" disease, this is mainly related to the production of autoantibodies to TG2, as hallmark of CD, while T cells recognizing TG2 do not appear to play a role.

The gluten-specific T cells can be demonstrated by *in vitro* culture of biopsies from CD patients, as first shown in the early 1990's (64, 65). They can also be demonstrated by direct staining using so-called HLA-DQ:gluten peptide tetramers; i.e., tetramers of HLA-DQ molecules with gluten peptides bound (76). It was shown that a short gluten challenge will mobilize gluten-specific T cells into the peripheral blood (77, 78), and such cells can be quantified by ELISpot or by HLA-DQ:gluten tetramers (79). Such T cells express markers for gut-homing, but if they actually are mobilized from the intestine remains uncertain. The procedure can be used for diagnostic purposes after gluten challenge (80) and may perform better than a 2-week gluten challenge followed by upper endoscopy with biopsy (81). Furthermore,

employing an improved methodology, such HLA-DQ:gluten tetramer$^+$ cells can be detected without gluten challenge and clearly distinguish CD patients and healthy individuals (82, 83). When the HLA tetramer technology was coupled with the CyTof technology, it was found that these cells carry a surprisingly rare phenotype with a profile suggesting that they may help plasma cells. Importantly, they are similar in profile to disease-relevant T cells in other autoimmune diseases, where the antigen specificity is unknown (84). Most recently, Zühlke et al. have demonstrated that the HLA-DQ:gluten tetramer$^+$ cells after gluten challenge show interesting features: (1) the kinetics of appearance peaks between day six and eight, (2) there are large inter-individual differences in numbers of cells, (3) even a one-day, single challenge with gluten mobilizes detectable cells, (4) although the numbers of cells vary between individuals, expression of the activation marker CD38 on the HLA-DQ:gluten tetramer$^+$ cells is a very specific and sensitive parameter (85). Thus, the HLA-DQ:gluten tetramers may be developed as a powerful diagnostic tool for CD but are not yet available outside the research setting and are not approved by any guidelines (86).

The importance of the adaptive immune system has recently been strongly supported by finding of the *bona fide* T cell cytokine IL-2 as fast as 4 h after ingestion of gluten. This was first shown after intradermal injection of gluten peptides but the same is seen after peroral gluten exposure (87). No such response is seen after gluten intake by non-celiac, gluten-free subjects (88). It is conceivable that either the HLA-DQ:gluten tetramers or the IL-2 response can be used as surrogate markers for testing of therapies for CD; this is the focus of ongoing research.

Compliance With the Gluten-Free Diet

Although a strict GFD remains the only effective treatment for CD, the rate of compliance is far from 100%. Adherence to the diet is higher in children and, in general, in those who have received diagnosis in early childhood (89). Socioeconomic factors, sex, access to health care facilities also influence the level of compliance. Adherence to the GFD is not easy to assess, clinical improvement not being a valid criterion. Periodic interviews conducted by dieticians could monitor compliance, with structured short, validated, dietary questionnaires being an alternative to consultations with a dietician (90). Anti-TG2 serology is in clinical practice the most used current method. In fact, antibody titers decrease after a few weeks on a strict GFD, but sometimes it can take longer, particularly if high titers are present at diagnosis. In any case antibody measurement cannot reveal minor dietary transgressions (91). The best way remains performing duodenal biopsies, but this is invasive and should be reserved to cases with no clinical improvement or no decrease of serological titers. More recently detection of gluten immunogenic peptides (GIP) in feces and urine has been proposed as new biomarker to detect gluten intake and verify GFD compliance in CD patients. Their determination is non-invasive and relatively simple, but shows poor correlation with antibody levels or with the response to dietary questionnaires (92).

UPDATE ON GLUTEN DIGESTIBILITY AND DEVELOPMENT OF NON-DIETARY TREATMENT

Gluten Digestibility Influences Its Stimulatory Properties

Gluten, the trigger factor of CD, is composed of hundreds of monomeric, oligomeric, and polymeric proteins, these latter interlinked by disulfide bonds (93). The unique amino acid composition of gluten proteins, enriched in glutamine and proline residues, makes this important dietary component highly resistant to gastrointestinal digestion (94). The inability of gastric and pancreatic proteases, as well as of the brush border membrane endopeptidases, to cleave proline-glutamine bonds throughout the gluten protein sequences leads to peptide fragments of different lengths that escape proteolytic degradation (95). These gluten peptides retain a marked immunogenic potential, as they pass across the small intestinal epithelial barrier and may trigger an adverse immune response in genetically susceptible individuals (96). It has been demonstrated that in patients with CD some long gluten peptides are site-specifically deamidated by TG2, bound to HLA class II molecules of antigen-presenting cells, and stimulate a specific immune response mediated by CD4$^+$ T cells. These mucosal T cells proliferate and release several inflammatory cytokines, such as interferon-γ and interleukin-21 with a key role in activating the injurious process of villous atrophy (97). Recently, the nomenclature of the existing CD-relevant gluten epitopes recognized by CD4$^+$ T cells has been updated (98). However, it has to be emphasized that the pool of CD-active sequences is far from being complete to date as the epitopes recognized by many T cells are not known. The clinical importance of these sequences is, however, uncertain.

A "new" class of poorly digestible proteins in wheat, the amylase/trypsin-inhibitors (ATI), has received a lot of attention recently, but their potential role in the pathogenesis of CD needs further investigation (99–101).

Gluten Degradation as a Treatment

The current therapy for CD patients is the lifelong withdrawal of gluten from the diet. Although, the GFD is efficacious in the great majority of patients, with the restoration of mucosa villous morphology and function, many young patients are poorly compliant, so that the identification of an alternative treatment would be beneficial for those patients for whom the GFD fails or is impracticable (102). Recently, great efforts were made to identify a pharmacological therapy that could be used to replace or support the GFD for treatment of CD patients (103). Currently, several proteolytic enzymes of microbial or plant origins have demonstrated a high efficiency to quickly degrade gluten proteins at very low pH, as occurring in gastric conditions (104–107). These glutenases, thanks to their efficacy in cleaving the proline- and glutamine-rich gluten sequences are promising drugs to abolish the immunogenic potential of dietary gluten. Both *in vitro* and pre-clinical studies have shown that the glutenase treatment results in a marked reduction of the amount of gluten epitopes in wheat-containing food. The

possibility of preventing that gluten immunogenic peptides reach the duodenal mucosa strongly suggests the possible use of glutenases in oral enzymatic treatment for CD (108). AN-PEP, a prolyl endopeptidase from *Aspergillus niger*, even though it was not intended to replace a GFD, was effective as a digestive aid protecting against the unintentional intake of gluten (109), or when consuming food which may contain small amounts of gluten, e.g., beer. A recent study demonstrated that the endopeptidase E40 from *Actinoallomurus* A8 is a fast-acting and strongly efficient glutenase, and thus a candidate as enzyme adjuvant to a GFD for the dietary management of CD (104). Glutenases can also be induced in wheat by germination but the activity is not high enough to be useful as an oral food supplement. However, this strategy can, for example, be used to eliminate residual gluten from food such as beer (110). Special wheat lines were developed recently to express the barley endoprotease B2 combined with a prolyl endopeptidase from *Flavobacterium meningosepticum* or *Pyrococcus furiosus* that significantly reduced the amount of indigestible gluten peptides (111). Sequence guided site-saturation mutagenesis was used to enhance the thermostability of these enzymes and allow their use in heat-treated cereal products (112).

Development of Other Non-dietary Treatment Options

As repeatedly stated in this paper, the GFD is a well-established and effective treatment for CD, at least as long as the patient is fully compliant. Here it is also important to emphasize that a GFD is inherently associated with nutritional deficiencies and not recommended except in the treatment of gluten-related disorders (113). It may be noted that no randomized, controlled trials have been performed to prove the real effects of the treatment, but this is not unusual in medicine. CD patients themselves express huge interest in non-dietary treatments like drugs or vaccines (114). Attractive options include sequestering of gluten within the lumen, luminal digestion of gluten by exogenous enzymes, interfering with mucosal integrity (tight junctions), inhibition of TG2, inhibition of antigen presentation, immune skewing and re-establishment of oral tolerance or clonal deletion, to mention a few (115–119). A plethora of Phase 1 studies, a small handful of Phase 2 and a single Phase 3 study is ongoing at the moment, but no drugs have reached the market. Almost all of these studies are based on preclinical studies *in vitro* and *ex vivo*. This research is hampered by the lack of good line research opportunity as the immune reaction to gluten is dependent on intact mucosal interaction. It is also hampered by lack of good animal (mouse) models for CD, although there are several mouse models for immune reaction to gluten (120, 121). Recently, a mouse model was developed that reproduces the overexpression of interleukin-15 (IL-15) in the gut epithelium and lamina propria, expresses the predisposing HLA-DQ8 molecule, and develops villous atrophy after ingestion of gluten (122). At any rate, it can be foreseen that such non-dietary options will come to the market, either as add-on therapy to the GFD, as rescue therapy after incidental gluten exposure or as replacement of the GFD.

UPDATE ON GLUTEN COMPOSITION OF WHEAT AND METHODS FOR MODIFICATION

Gluten Content and Composition in Wheat Species and Cultivars

The availability of the first annotated reference sequence for the hexaploid bread wheat genome containing 107,891 high-confidence gene models (123) recently allowed the establishment of a genome reference map for immunostimulatory wheat proteins (124). One of the hypotheses being discussed to explain the increasing prevalence of CD is that the protein composition of wheat may have changed over the past decades due to breeding and agronomic practices. The main goals of wheat breeding are increased yield, improved resistance against plant diseases, pests, and climatic stress, more efficient use of fertilizers as well as increased protein content. The protein content represents one of the key quality aspects that significantly influences the bread wheat quotation worldwide. With CD being determined by gluten as major and necessary environmental risk factor and HLA-DQ2/DQ8 as genetic risk factors, one might envisage to be able to predict the prevalence of CD in different countries. A systematic worldwide compilation of this data revealed that those two factors are clearly required for the development of CD, but not suitable to predict the prevalence. There was no correlation between CD prevalence, the levels of wheat consumption and the frequencies of HLA-DQ2/DQ8 or the combination of both. This rather surprising result was primarily due to several outlier populations in regions such as north-western India, northern Africa, Mexico, Finland, and Russia. For example, the prevalence of CD in Finland is among the highest worldwide (2.4%), whereas that of neighboring Karelia (Russia) is very low (0.2%), although both regions share similar levels of wheat consumption and frequencies of HLA-DQ2/8 (125). Within the United States, CD was 5.4-fold more common among individuals who lived at latitudes of 40° North or more than among individuals who lived at latitudes below 35° North, independent of race or ethnicity, socioeconomic status, and body mass index (126). This discrepancy can only be explained by further environmental factors that cause a loss of tolerance to dietary gluten and initiate CD (23, 24). Although there was no clear trend toward higher protein or gluten contents since the 1950s (127), the selection criteria for breeding might have resulted in a higher immunostimulatory potential of wheat (128). Several studies have explored the protein composition of different wheat species and cultivars of hexaploid bread wheat (*Triticum aestivum* subsp. *aestivum*) and spelt (*T. aestivum* subsp. *spelta*), tetraploid durum wheat (*T. turgidum* subsp. *durum*), and emmer (*T. turgidum* subsp. *dicoccum*) as well as diploid einkorn (*T. monococcum*) with respect to their content of potentially immunostimulatory proteins (15, 129–131).

There is evidence for changes in gluten protein composition (132–134) with decreasing contents of gliadins and total gluten, but increasing contents of glutenins from diploid to tetraploid and hexaploid wheats. However, within bread wheat or durum wheat, there were no clear differences in the contents of selected

CD-active epitopes between modern cultivars and landraces not subjected to breeding (135–138). All studies consistently reported a significant effect of environmental conditions on the expression of CD-immunogenic peptides, with e.g., low or high cultivation temperatures affecting the expression of immunostimulatory proteins in different ways (124). This high variability in protein composition regulated by mechanisms in the wheat plant that are still incompletely understood complicates the search for specific cultivars that naturally express low amounts of CD-immunogenic peptides independent of the environmental conditions.

Recent studies have shown that gluten proteins of several einkorn (*T. monococcum*) landraces have a reduced capability of activating the mucosal innate immune cells and inducing enterocyte apoptosis (14). Other studies have attributed the reduced immunotoxicity to the presence of protective sequences (139). A recent study demonstrated that wheat gluten from einkorn is extensively degraded by the gastrointestinal protease cocktail, including endopeptidases of the villous brush border membrane (15). This results in the release of a reduced amount of peptides that can activate pathogenic $CD4^+$ T cells in the gut mucosa. Altogether, these findings are relevant from the perspective of disease prevention, taking into account that the incidence of CD is much higher (≈10%) in first-degree relatives of CD patients carrying the high-risk HLA-DQ genes (36). However, the evidence cited in support of *T. monococcum* is still insufficient, and it is not recommended to include this wheat species into the diet of CD patients.

Removing Gluten From Wheat

It is likely, that no existing wheat species or variety is completely safe for use by CD patients, as they all contain far more than 20 mg/kg of gluten. Can we specifically develop wheat varieties that are CD-safe? In barley, an ultra-low gluten variety has been developed that is safe for CD patients, as it contains <20 mg/kg of gluten. It was developed by combining existing induced recessive mutations through breeding (140). This strategy is very difficult to implement in bread wheat as it has three genomes and one does not want to remove all gluten genes (141), but other approaches may be successful (142, 143). The general applicability of these breeding approaches needs to be discussed, since only gluten endows the dough with the desired properties for bread making.

Using RNA interference (RNAi), several groups have shown that it is possible to strongly reduce the expression of α-gliadins (144), γ-gliadins (145), ω-gliadins (146, 147), or all of them (148) in wheat. In the latter study, the α- and ω-gliadins were downregulated to the extent that no CD epitopes could be detected using LC-MS/MS (148). Unfortunately, these lines are genetically modified (GM) as the RNAi construct must remain present. As no GM wheat has been commercially introduced anywhere in the world, it is unlikely that these lines will reach the market shortly.

A recent alternative technology is to use gene editing with CRISPR/Cas9 to delete gliadin genes in order to produce gluten-free wheat and/or to edit the epitopes in them to generate wheat with safe gluten. Sánchez-León et al. (149) targeted two conserved sites adjacent to the epitope-containing region in the α-gliadin genes. Up to 35 of the 45 α-gliadin genes were mutated in a single line, with small or larger deletions around the target sites. This line showed a 85% reduction of the R5 and G12 ELISA signals. Jouanin et al. (150) simultaneously edited multiple sites in α- and γ-gliadins with a single construct. Although the lines produced in these pilot studies are not yet safe, they demonstrate the power of gene editing for effectively modifying tens of genes of multiple gene families in a polyploid species at once. The Cas9 construct used to generate the edits is removed afterwards by crossing, leaving only mutations that are identical to what can occur naturally. In most of the world, the resulting plants are not considered as GM, with the exception of the EU (151).

UPDATE ON THE USE OF ELISA FOR GLUTEN ANALYSIS

Advances in Compliance Monitoring of Gluten-Free Products

Apart from evolving proteomics-based detection methods, enzyme-linked immunosorbent assays (ELISAs) are most commonly used to monitor the compliance of gluten-free products to the regulatory threshold of 20 mg/kg of gluten (Codex Standard 118-1979[2]). There are more than 20 ELISA test kits on the market that use different principles (sandwich vs. competitive), extraction procedures, reference materials for calibration and various polyclonal or monoclonal antibodies such as the Skerritt (401.21) (152), R5 (153), G12 (154), and α20 (155). A wide variation in the reported measurement results between different commercial kits were observed in several studies (156–158). These discrepancies were mainly attributed to the use of different reference materials and to the fact that the used antibodies target only a fraction of gluten components whose proportions may vary according to the contamination source. The ELISA R5 Mendez Method is currently laid down as a Codex type I method for gluten determination in foods and, therefore, continues to be the most widely used assay. However, the R5 ELISA has two disadvantages; first it overestimates gluten from rye and barley when calibrated to gliadins and secondly, it does not detect glutelins adequately and the gluten content is calculated by multiplying the prolamin content detected by a factor of two. To address these limitations, a new sandwich ELISA based on four different monoclonal antibodies was developed that detects prolamins from wheat, rye and barley as well as HMW-GS, HMW-secalins from rye and low-molecular-weight (LMW)-GS from wheat. The performance of the test kit was recently validated for the quantitative analysis of wheat, rye and barley gluten in oat and oat products by an international collaborative study with 19 laboratories. The results of the study showed recoveries ranging from 99 to 137% for wheat, rye and barley when analyzing defined validation materials (159, 160) and relative reproducibility standard deviations from 10 to 53% for samples containing 10 mg/kg of gluten or higher. Following review by the AOAC Expert Review Panel for Gluten Assays, the method was adopted as AOAC *Official Method* First Action 2018.15 (21).

A Five Cultivar Wheat Blend Is Suitable for New Reference Material Production

Reference materials that are representative of the target analyte are essential prerequisites for calibrating and assuring the accuracy of analytical methods. They form the basis for method establishment and validation, proficiency tests, and verification of the comparability between different methods and laboratories (161–163). The use of appropriate reference materials was recently shown to efficiently reduce the disparity of gluten analysis between different commercial kits (157). A variety of different reference materials are used in ELISA test kits for gluten analysis including wheat or gluten preparations, some of which are proprietary to the respective kit manufacturer with little information on their exact composition related to immunoreactive sequences. Food matrix and food processing are also known to influence analytical results, which is why incurred materials are recommended to reflect the properties of actual food samples as closely as possible.

The best characterized reference material available for gluten analysis is the so-called PWG-gliadin that was developed by our group of experts. PWG-gliadin constitutes the purified gliadin fraction extracted from a mixture of the 28 most common European wheat cultivars, as of 1999 (22). It is homogeneous, completely soluble in 60% ethanol, representative for European wheat, regularly monitored for stability and widely used to calibrate ELISA test kits and other methods for gluten analysis (164, 165). However, as its supply is limited, efforts to develop new reference and incurred materials are currently underway. A collection of wheat cultivars from different countries was characterized for gluten protein composition and ELISA response to establish selection criteria and identify cultivars that are as representative as possible for the multitude of cultivars grown worldwide. A blend of the selected five cultivars from Asia (Yumai-34), Australia (Yitpi), Europe (Akteur, Mv Magvas), and North America (Carberry) was further characterized and appears to be suitable for further reference material development (166, 167).

CODEX ALIMENTARIUS AND GLUTEN QUANTITATION BY ELISA

Codex analytical methods are being revised every 10 years and revision of the R5 Mendez ELISA was due in 2018. Thus, it is time to discuss how to handle ELISAs for gluten quantitation regarding approval by the Codex Alimentarius. Based on the matrices used for validation, the R5 method has been recommended for gluten quantitation in maize matrices and the G12 method for the analysis of rice matrices. Both methods fulfill the performance requirements for gluten analysis set by the Codex Standard 118-1979[2], i.e., a limit of quantitation of 10 mg gluten/kg or less and the detection of CD-active epitopes. The recently developed Total Gluten ELISA covers both gluten fractions and, thus, measures the gluten content (21). This is an important step forward compared to the R5 and G12 ELISAs that measure the prolamin content and this is then converted into the gluten content by multiplication with the factor of two.

Concerns

A major concern, in particular of celiac societies, CD patients, food producers, and national food control laboratories is the unclear situation, if several ELISA methods for gluten quantitation were endorsed. It can be assumed that each analytical laboratory would use one ELISA as the default method for gluten quantitation. Consequently, it would be unclear, if a value obtained by one laboratory with one kit would be comparable to the value provided by a different laboratory with a different kit. This would lead to the question of how to handle conflicting results from different laboratories. In general, the possibility of having two type I methods has to be questioned, because the definition of a type I method as "the only method" should exclude approval of a second type I method. On the other hand, if a proprietary method fulfilling the performance criteria of the Codex is on the market and has been approved by suitable collaborative studies, it should not be excluded due to the fact that another method has already been endorsed by the Codex Alimentarius. Both the R5 and G12 sandwich ELISAs have been compared in a number of scientific studies. In summary, the results of these studies strongly suggest that these methods do not yield comparable results. Typical examples are papers published by Bugyi et al. (156), Bruins Slot et al. (168), and Scherf (158).

Position of the PWG

Therefore, the PWG suggests that ELISA methods should be approved using a combination of

(1) information on the method that has been used to provide the analytical value,
(2) strict performance criteria and
(3) a pre-defined set of maximum five matrices.

This would be similar to the Standard Method Performance Requirements (SMPR) published by AOAC International for allergen-containing commodities such as whole egg, milk, peanut, and hazelnut (169).

Performance Criteria

Performance criteria include the correct setup and statistical evaluation of validation studies (161, 170, 171) as well as the fulfillment of the requirements for standard method performance (172). The minimal performance requirements set in AOAC SMPR 2016.002 (169) for whole egg, milk, peanut and hazelnut can be adapted to gluten. Possible performance criteria are summarized in **Table 1**. AOAC suggests low recovery rates of 60%, but the PWG feels that, in general, the recovery range should be between 80 and 120 %, which is in line with Abbott et al. (161). With respect to LOD and LOQ, the three ELISAs under consideration perform well and meet the requirements (18–21).

Matrices

Matrices should not be based on botanical origin (e.g., rice- or maize-based), but on constituents that most likely affect the interaction of the antibodies with the gluten antigens. Possible matrices should be categorized into protein-based, starch-based, fat-based, polyphenol-rich, and fiber-rich foods (173). **Table 2** suggests categories and examples for foods from each category.

TABLE 1 | Method performance requirements for gluten ELISAs.

Parameter	Value/Range
Analytical range (mg/kg)	5–100
Limit of detection (LOD) (mg/kg)	3
Limit of quantitation (LOQ) (mg/kg)	10
Recovery (%)	80–120
Repeatability relative standard deviation (RSD_r) (%)	20
Reproducibility relative standard deviation (RSD_R) (%)	30

TABLE 2 | Suggested matrix categories and examples for foods from each category.

Category	Examples
Protein-based	Coated meat, sausage, protein isolate/concentrate
Starch-based	Starch, baked goods, sauce
Fat-based	Cookie, cake, ice cream
Polyphenol-rich	Chocolate cake, cocoa powder, beer
Fiber-rich	Cereal bran, breakfast cereals, legume seed flours

Examples are limited to three per category to keep the number of required analyses in validation studies in a range that can be handled. Kit manufacturers are encouraged to agree on a set of matrices which should be comparatively analyzed using their methods. In case of conflicting R5/G12 results, in particular in the concentration range of the 20 mg/kg threshold, the higher concentration value should be considered relevant in the interest of the celiac consumers. In future analyses, it should then be avoided having to do two ELISAs. For any analysis value, the ELISA method that was used should be indicated alongside the results.

RECOMMENDATIONS FOR FUTURE RESEARCH AND ACTIONS

Clinical Aspects of CD

Based on the most recent findings regarding epidemiological and clinical aspects of CD as discussed above, the PWG recommends the following priority research areas. With the recognition of CD just beginning to emerge, particularly in sub-Saharan Africa and Eastern Asia, more data needs to be collected in order to make a robust estimate of the prevalence of CD in these parts of the world. Recent epidemiological findings from Denmark, Finland, Italy and the US suggest an increasing prevalence of CD over time, but the reasons for this still remain unknown. Among the factors being suggested are genetic and epigenetic as well as environmental factors of which infectious agents are most likely to play a role. In order to assess the specific contributions of these factors toward increasing the risk of CD development, the natural history of the disease needs to be understood in more detail. Currently, there are still gaps in our knowledge on how the disease proceeds from the genetic risk combined with exposure to environmental risk factors in the initial loss of oral tolerance to gluten. Even then, some individuals remain at this stage of activated anti-gluten adaptive immunity with no or very mild histological abnormalities, whereas others progress to full-blown villous atrophy. It will be critical to clarify the role of microbiome/virome changes and infections in the period preceding the development of CD and to find markers (epigenetic changes, genetic expression, metabolome alterations, T-cell markers) that predict the development of the disease at the earliest stage possible. Related to this, it will be equally important to identify the factors controlling evolution from CD autoimmunity to mucosal damage and the biomarkers predictive of such evolution to enable the identification of preventive measures and non-dietary treatments. Finally, despite increased awareness of CD, diagnostic delays are still common and the appropriate policy to be implemented to improve the diagnostic rate needs to be determined.

Analytical Aspects of Gluten

For reasons of better handling and long-term stability compared to flours as evidenced by the excellent properties of PWG-gliadin since its production almost 20 years ago, we continue to support the use of isolates as reference materials. Because PWG-gliadin only constitutes the alcohol-soluble fraction of wheat gluten, we aim to provide prolamin and total gluten isolates from wheat, rye and barley flours, respectively. The first steps will be to establish a suitable protocol to extract all relevant immunoreactive gluten proteins from the flours, characterize the exact composition of the isolates and ensure homogeneity and solubility. Research efforts to identify representative rye and barley cultivars have just started as well as fundamental studies on suitable extraction protocols. We recommend using the same reference material for calibration of analytical methods for better comparability and reproducibility of results.

The PWG acknowledges that more than one ELISA method for the analysis of gluten in foods are currently used and that the results of these methods are not comparable. The group does not support the policy of the Codex Alimentarius to allow approval of more than one type 1 method, because this is in disagreement to the definition of a type 1 method. The Codex Alimentarius should decide soon how to proceed, because several methods are currently already available that fulfill all performance criteria such as the R5, G12, and Total Gluten ELISAs. The PWG suggests that performance data of these ELISAs obtained with identical or at least comparable matrices should be compared. If existing data is not sufficient, comparative studies need to be carried out on the set of foods suggested in this paper. This could result in a kind of guidebook suggesting specific ELISAs for specific foods.

AUTHOR CONTRIBUTIONS

KS, CC, CG, KL, MS, RT, and PK wrote the first draft of the manuscript. FC, CF, FK, DS, and OT contributed to revising and editing the manuscript. All authors read and approved the final manuscript.

FUNDING

Open Access publication of this paper was funded by Arbeitsgemeinschaft Getreideforschung e.V. (Association of Cereal Research), Detmold, Germany.

ACKNOWLEDGMENTS

The PWG was grateful to all regular attendees of the annual meetings for their active contributions as presenters as well as during the discussions.

REFERENCES

1. Ludvigsson JF, Leffler DA, Bai JA, Biagi F, Fasano A, Green PHR, et al. The Oslo definitions for coeliac disease and related terms. *Gut*. (2013) 62:43–52. doi: 10.1136/gutjnl-2011-301346
2. Dieterich W, Ehnis T, Bauer M, Donner P, Volta U, Riecken EO, et al. Identification of tissue transglutaminase as the autoantigen of celiac disease. *Nat Med*. (1997) 3:797–801. doi: 10.1038/nm0797-797
3. Molberg O, McAdam SN, Korner R, Quarsten H, Kristiansen C, Madsen L, et al. Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. *Nat Med*. (1998) 4:713–7. doi: 10.1038/nm0698-713
4. Van de Wal Y, Kooy YMC, van Veelen PA, Pena SA, Mearin LM, Molberg ∅, et al. Small intestinal T cells of celiac disease patients recognize a natural pepsin fragment of gliadin. *Proc Natl Acad Sci USA*. (1998) 95:1050–1054. doi: 10.1073/pnas.95.17.10050
5. Fasano A. Celiac disease – how to handle a clinical chameleon. *N Engl J Med*. (2003) 348:2568–70. doi: 10.1056/NEJMe030050
6. Fasano A, Berti I, Gerarduzzi T, Not T, Colletti RB, Drago S, et al. Prevalence of celiac cisease in at-risk and not-at-risk groups in the United States: A large multicenter study. *Arch Intern. Med.* (2003) 163:286–292. doi: 10.1001/archinte.163.3.286
7. Kahaly GJ, Frommer L, Schuppan D. Celiac disease and endocrine autoimmunity – the genetic link. *Autoimmun Rev*. (2018) 17:1169–75. doi: 10.1016/j.autrev.2018.05.013
8. Lundin KE, Sollid LM. Advances in coeliac disease. *Curr Opin Gastroenterol*. (2014) 30:154–62. doi: 10.1097/MOG.0000000000000041
9. Schuppan D, Junker Y, Barisani D. Celiac disease: from pathogenesis to novel therapies. *Gastroenterology*. (2009) 137:1912–33. doi: 10.1053/j.gastro.2009.09.008
10. Peräaho M, Kaukinen K, Mustalahti K, Vuolteenaho N, Mäki M, Laippala P, et al. Effect of an oats-containing gluten-free diet on symptoms and quality of life in coeliac disease. *A randomized study Scand J Gastroenterol*. (2004) 39:27–31. doi: 10.1080/00365520310007783
11. Srinivasan U, Leonard N, Jones E, Kasarda DD, Weir DG, O'Farrelly C, et al. Absence of oats toxicity in adult coeliac disease. *BMJ*. (1996) 313:1300–1. doi: 10.1136/bmj.313.7068.1300
12. Srinivasan U, Jones E, Carolan J, Feighery C. Immunohistochemical analysis of coeliac mucosa following ingestion of oats. *Clin Exp Immunol*. (2006) 144:197–203. doi: 10.1111/j.1365-2249.2006.03052.x
13. Catassi C, Fabiani E, Iacono G, D'Agate C, Francavilla R, Biagi F, et al. A prospective, double-blind, placebo-controlled trial to establish a safe gluten threshold for patients with celiac disease. *Am J Clin Nutr*. (2007) 85:160–6. doi: 10.1093/ajcn/85.1.160
14. Gianfrani C, Maglio M, Rotondi Aufiero V, Camarca A, Vocca I, Iaquinto G, et al. Immunogenicity of monococcum wheat in celiac patients. *Am J Clin Nutr*. (2012) 96:1339–45. doi: 10.3945/ajcn.112.040485
15. Gianfrani C, Camarca A, Mazzarella G, Di Stasio L, Rotondi Aufiero V, Giardullo N, et al. Extensive *in vitro* gastrointestinal digestion markedly reduces the immune-toxicity of *Triticum monococcum* wheat: implication for celiac disease. *Mol Nutr Food Res*. (2015) 59:1844–54. doi: 10.1002/mnfr.201500126
16. Bondar C, Araya RE, Guzman L, Rua EC, Chopita N, Chirdo FG. Role of CXCR3/CXCL10 axis in immune cell recruitment into the small intestine in celiac disease. *PLoS ONE*. (2014) 9:e89068. doi: 10.1371/journal.pone.0089068
17. García E, Llorente M, Hernando A, Kieffer R, Wieser H, Méndez E. Development of a general procedure for complete extraction of gliadins for heat processed and unheated foods. *Eur J Gastroenterol Hepatol*. (2005) 17:529–39. doi: 10.1097/00042737-200505000-00010
18. Don C, Halbmayr-Jech E, Rogers A, Koehler P. AACCI approved methods technical committee report: collaborative study on the immunochemical quantitation of intact gluten in rice flour and rice-based products using G12 sandwich ELISA. *Cereal Foods World*. (2014) 59:187–93. doi: 10.1094/CFW-59-4-0605
19. Koehler P, Schwalb T, Immer U, Lacorn M, Wehling P, Don C. AACCI approved methods technical committee report: collaborative study on the immunochemical determination of intact gluten using an R5 sandwich ELISA. *Cereal Foods World*. (2013) 58:36–40. doi: 10.1094/CFW-58-1-0036
20. Koehler P, Schwalb T, Immer U, Lacorn M, Wehling P, Don C. AACCI approved methods technical committee report: collaborative study on the immunochemical determination of partially hydrolyzed gluten using an R5 competitive ELISA. *Cereal Foods World*. (2013) 58:154–8. doi: 10.1094/CFW-58-3-0402
21. Lacorn M, Weiss T, Wehling P, Arlinghaus M, Scherf K. Quantification of wheat, rye, and barley gluten in oat and oats products by ELISA RIDASCREEN® Total Gluten: Collaborative study, First Action 2018.15. *J. AOAC Int*. (2019) 102:1535–43. doi: 10.5740/jaoacint.19-0094
22. Van Eckert R, Berghofer E, Ciclitira PJ, Chirdo F, Denery-Papini S, Ellis HJ, et al. Towards a new gliadin reference material - isolation and characterisation. *J Cereal Sci*. (2006) 43:331–41. doi: 10.1016/j.jcs.2005.12.009
23. Abadie V, Sollid LM, Barreiro LB, Jabri B. Integration of genetic and immunological insights into a model of celiac disease pathogenesis. *Annu Rev Immunol*. (2011) 29:493–525. doi: 10.1146/annurev-immunol-040210-092915
24. Ludvigsson JF, Murray JA. Epidemiology of celiac disease. *Gastroenterol Clin North Am*. (2019) 48:1–18. doi: 10.1016/j.gtc.2018.09.004
25. Singh P, Arora A, Strand TA, Leffler DA, Catassi C, Green PH, et al. Global prevalence of celiac disease: systematic review and meta-analysis. *Clin Gastroenterol Hepatol*. (2018) 16:823–36.e2. doi: 10.1016/j.cgh.2017.06.037
26. Gatti S, Lionetti E, Balanzoni L, Verma AK, Galeazzi T, Gesuita R, et al. Increased prevalence of celiac disease in school-age children in Italy. *Clin Gastroenterol Hepatol*. (2019) 18:596–603. doi: 10.1016/j.cgh.2019.06.013
27. Ivarsson A, Myléus A, Norström F, van der Pals M, Rosén A, Högberg L, et al. Prevalence of childhood celiac disease and changes in infant feeding. *Pediatrics*. (2013) 131:e687–94. doi: 10.1542/peds.2012-1015
28. Ramakrishna BS, Makharia GK, Chetri K, Dutta S, Mathur P, Ahuja V, et al. Prevalence of adult celiac disease in India: regional variations and associations. *Am J Gastroenterol*. (2016) 111:115–23. doi: 10.1038/ajg.2015.39
29. Kivelä L, Kaukinen K, Huhtala H, Lähdeaho ML, Mäki M, Kurppa K. At-risk screened children with celiac disease are comparable in disease severity and dietary adherence to those found because of clinical suspicion: a large cohort study. *J Peds*. (2017) 183:115–21.e2. doi: 10.1016/j.jpeds.2016.12.077
30. Catassi C, Lionetti E. Case finding for celiac disease is ok, but is it enough? *J Pediatr Gastroenterol Nutr*. (2013) 57:415–7. doi: 10.1097/MPG.0b013e3182a45676
31. Hujoel IA, Van Dyke CT, Brantner T, Larson J, King KS, Sharma A, et al. Natural history and clinical detection of undiagnosed coeliac disease in a North American community. *Aliment Pharmacol Ther*. (2018) 47:1358–66. doi: 10.1111/apt.14625
32. Rubio-Tapia A, Kyle RA, Kaplan EL, Johnson DR, Page W, and Erdtmann F, et al. Increased prevalence and mortality in undiagnosed celiac disease. *Gastroenterology*. (2009) 137:88–93. doi: 10.1053/j.gastro.2009.03.059
33. Catassi C, Kryszak D, Bhatti B, Sturgeon C, Helzlsouer K, Clipp SL, et al. Natural history of celiac disease autoimmunity in a USA cohort followed since 1974. *Ann Med*. (2010) 42:530–8. doi: 10.3109/07853890.2010.514285
34. Grode L, Bech BH, Jensen TM, Humaidan P, Agerholm IE, Plana-Ripoll O, et al. Prevalence, incidence, and autoimmune

comorbidities of celiac disease: a nation-wide, population-based study in Denmark from 1977 to 2016. *Eur J Gastroeneterol Hepatol*. (2018) 30:83–91. doi: 10.1097/MEG.0000000000000992
35. Lionetti E, Castellaneta S, Francavilla R, Pulvirenti A, Tonutti E, Amarri S, et al. Introduction of gluten, HLA status, and the risk of celiac disease in children. *N Engl J Med*. (2014) 371:1295–303. doi: 10.1056/NEJMoa1400697
36. Vriezinga SL, Auricchio R, Bravi E, Castillejo G, Chmielewska A, Crespo Escobar P, et al. Randomized feeding intervention in infants at high risk for celiac disease. *N Engl J Med*. (2014) 371:1304–15. doi: 10.1056/NEJMoa1404172
37. Aronsson CA, Lee HS, Liu E, Uusitalo U, Hummel S, Yang J, et al. Age at gluten introduction and risk of celiac disease. *Pediatrics*. (2015) 135:239–45. doi: 10.1542/peds.2014-1787
38. Størdal K, White RA, Eggesbø M. Early feeding and risk of celiac disease in a prospective birth cohort. *Pediatrics*. (2013) 132:e1202–1209. doi: 10.1542/peds.2013-1752
39. Aronsson CA, Lee HS, Hård af Segerstad EM, Uusitalo U, Yang J, Koletzko S, et al. Association of gluten intake duuring the first 5 years of life with incidence of celiac disease autoimmunity and celiac disease among children at increased risk. *JAMA*. (2019) 322:514–23. doi: 10.1001/jama.2019.10329
40. Crespo-Escobar P, Mearin ML, Hervás D, Auricchio R, Castillejo G, Gyimesi J, et al. The role of gluten consumption at an early age in celiac disease development: a further analysis of the prospective PreventCD cohort study. *Am J Clin Nutr*. (2017) 105:890–6. doi: 10.3945/ajcn.116.144352
41. Leonard MM, Fasano A. Gluten and celiac disease risk: Is it just a matter of quantity? *JAMA*. (2019) 322:510–1. doi: 10.1001/jama.2019.9678
42. Lund-Blix NA, Mårild K, Tapia G, Norris JM, Stene LC, Størdal K. Gluten intake in early childhood and risk of celiac disease in childhood: a nationwide cohort study. *Am J Gastroenterol*. (2019) 114:1299–306. doi: 10.14309/ajg.0000000000000331
43. Dydensborg Sander S, Hansen AV, Størdal K, Andersen AN, Murray JA, Husby S. Mode of delivery is not associated with celiac disease. *Clin Epidemiol*. (2018) 10:323–32. doi: 10.2147/CLEP.S152168
44. Dydensborg Sander S, Nybo Andersen AM, Murray JA, Karlstad Ø, Husby S, Størdal K. Association between antibiotics in the first year of life and celiac disease. *Gastroenterology*. (2019) 156:2217–29. doi: 10.1053/j.gastro.2019.02.039
45. Kołodziej M, Patro-Gołab B, Gieruszczak-Białek D, Skórka A, Piescik-Lech M, Baron R, et al. Association between early life (prenatal and postnatal) antibiotic administration and coeliac disease: a systematic review. *Arch Dis Child*. (2019) 104:1083–89. doi: 10.1136/archdischild-2019-317174
46. Kemppainen KM, Vehik K, Lynch KF, Larsson HE, Canepa RJ, Simell V, et al. Association between early-life antibiotic use and the risk of islet or celiac disease autoimmunity. *JAMA Pediatr*. (2017) 171:1217–55. doi: 10.1001/jamapediatrics.2017.2905
47. Stene LC, Honeyman MC, Hoffenberg EJ, Haas JE, Sokol RJ, Emery L, et al. Rotavirus infection frequency and risk of celiac disease autoimmunity in early childhood: a longitudinal study. *Am J Gastroenterol*. (2006) 101:2333–40. doi: 10.1111/j.1572-0241.2006.00741.x
48. Kemppainen KM, Lynch KF, Liu E, Lönnrot M, Simell V, Briese T, et al. Factors that increase risk of celiac disease autoimmunity after a gastrointestinal infection in early life. *Clin Gastroenterol Hepatol*. (2017) 15:694–702.e5. doi: 10.1016/j.cgh.2016.10.033
49. Bouziat R, Biering SB, Kouame E, Sangani KA, Kang S, Ernest JD, et al. Murine norovirus infection induces T_H1 inflammatory responses to dietary antigens. *Cell Host Microbe*. (2018) 24:677–88.e5. doi: 10.1016/j.chom.2018.10.004
50. Bouziat R, Hinterleitner R, Brown JJ, Stencel-Baerenwald JE, Ikizler M, Mayassi T, et al. Reovirus infection triggers inflammatory responses to dietary antigens and development of celiac disease. *Science*. (2017) 356:44–50. doi: 10.1126/science.aah5298
51. Kahrs CR, Chuda K, Tapia G, Stene LC, Mårild K, Rasmussen T, et al. Enterovirus as trigger of coeliac disease: nested case-control study within prospective birth cohort. *BMJ*. (2019) 364:l231. doi: 10.1136/bmj.l231
52. Auricchio R, Cielo D, de Falco R, Galatola M, Bruno V, Malamisura B, et al. Respiratory infections and the risk of celiac disease. *Pediatrics*. (2017) 140:e20164102. doi: 10.1542/peds.2016-4102
53. Pozo-Rubio T, Olivares M, Nova E, De Palma G, Mujico JR, Ferrer MD, et al. Immune development and intestinal microbiota in celiac disease. *Clin Dev Immunol*. (2012) 2012:654143. doi: 10.1155/2012/654143
54. Sánchez E, Donat E, Ribes-Koninckx C, Fernández-Murga ML, Sanz Y. Duodenal-mucosal bacteria associated with celiac disease in children. *Appl Environ Microbiol*. (2013) 79:5472–9. doi: 10.1128/AEM.00869-13
55. Olivares M, Walker AW, Capilla A, Benítez-Páez A, Palau F, Parkhill J, et al. Gut microbiota trajectory in early life may predict development of celiac disease. *Microbiome*. (2018) 6:36. doi: 10.1186/s40168-018-0415-6
56. Rintala A, Riikonen I, Toivonen A, Pietilä S, Munukka E, Pursiheimo JP, et al. Early fecal microbiota composition in children who later develop celiac disease and associated autoimmunity. *Scand J Gastroenterol*. (2018) 53:403–9. doi: 10.1080/00365521.2018.1444788
57. Caminero A, Galipeau HJ, McCarville JL, Johnston CW, Bernier SP, Russell AK, et al. Duodenal bacteria from patients with celiac disease and healthy subjects distinctly affect gluten breakdown and immunogenicity. *Gastroenterology*. (2016) 151:670–83. doi: 10.1053/j.gastro.2016.06.041
58. Caminero A, McCarville JL, Galipeau HJ, Deraison C, Bernier SP, Constante M, et al. Duodenal bacterial proteolytic activity determines sensitivity to dietary antigen through protease-activated receptor-2. *Nat Commun*. (2019) 10:1198. doi: 10.1038/s41467-019-09037-9
59. Kurppa K, Collin P, Viljamaa M, Haimila K, Saavalainen P, Partanen J, et al. Diagnosing mild enteropathy celiac disease: a randomized, controlled clinical study. *Gastroenterology*. (2009) 136:816–23. doi: 10.1053/j.gastro.2008.11.040
60. Mäki M, Mustalahti K, Kokkonen J, Kulmala P, Haapalahti M, Karttunen T, et al. Prevalence of celiac disease among children in Finland. *N Engl J Med*. (2003) 348:2517–24. doi: 10.1056/NEJMoa021687
61. Salmi TT, Hervonen K, Kurppa K, Collin P, Kaukinen K, Reunala T. Celiac disease evolving into dermatitis herpetiformis in patients adhering to normal or gluten-free diet. *Scand J Gastroenterol*. (2015) 50:387–92. doi: 10.3109/00365521.2014.974204
62. Auricchio R, Mandile R, Del Vecchio MR, Scapaticci S, Galatola M, Maglio M, et al. Progression of celiac disease in children with antibodies against tissue transglutaminase and normal duodenal architecture. *Gastroenterology*. (2019) 157:413–20. doi: 10.1053/j.gastro.2019.04.004
63. Sollid LM, Markussen G, Ek J, Gjerde H, Vartdal F, Thorsby E, et al. Evidence for a primary association of celiac disease to a particular HLA-DQ alpha/beta heterodimer. *J Exp Med*. (1989) 169:345–50. doi: 10.1084/jem.169.1.345
64. Lundin KE, Scott H, Hansen T, Paulsen G, Halstensen TS, Fausa O, et al. Gliadin-specific, HLA-DQ(alpha 1^*0501, beta 1^*0201) restricted T cells isolated from the small intestinal mucosa of celiac disease patients. *J Exp Med*. (1993) 178:187–96. doi: 10.1084/jem.178.1.187
65. Lundin KE, Scott H, Fausa O, Thorsby E, Sollid LM. T cells from the small intestinal mucosa of a DR4, DQ7/DR4, DQ8 celiac disease patient preferentially recognize gliadin when presented by DQ8. *Hum Immunol*. (1994) 41:285–91. doi: 10.1016/0198-8859(94)90047-7
66. Bodd M, Kim CY, Lundin KE, Sollid LM. T-cell response to gluten in patients with HLA-DQ2.2 reveals requirement of peptide-MHC stability in celiac disease. *Gastroenterology*. (2012) 142:552–61. doi: 10.1053/j.gastro.2011.11.021
67. Vader W, Stepniak D, Kooy Y, Mearin L, Thompson A, van Rood JJ, et al. The HLA-DQ2 gene dose effect in celiac disease is directly related to the magnitude and breadth of gluten-specific T cell responses. *Proc Natl Acad Sci USA*. (2003) 100:12390–5. doi: 10.1073/pnas.2135229100
68. Pisapia L, Camarca A, Picascia S, Bassi V, Barba P, Del Pozzo G, et al. HLA-DQ2.5 genes associated with celiac disease risk are preferentially expressed with respect to non-predisposing HLA genes: Implication for anti-gluten T cell response. *J Autoimmun*. (2016) 70:63–72. doi: 10.1016/j.jaut.2016.03.016
69. Gianfrani C, Pisapia L, Picascia S, Strazzullo M, Del Pozzo G. Expression level of risk genes of MHC class II is a susceptibility factor for autoimmunity: new insights. *J Autoimmun*. (2018) 89:1–10. doi: 10.1016/j.jaut.2017.12.016
70. Farina F, Picascia S, Pisapia L, Barba P, Vitale S, Franzese A, et al. HLA-DQA1 and HLA-DQB1 alleles, conferring susceptibility to celiac disease and type 1 diabetes, are more expressed than non-predisposing alleles and are coordinately regulated. *Cells*. (2019) 8:E751. doi: 10.3390/cells8070751
71. Arentz-Hansen H, McAdam SN, Molberg O, Fleckenstein B, Lundin KEA, Jørgensen TJD, et al. Celiac lesion T cells recognize epitopes that cluster

in regions of gliadins rich in proline residues. *Gastroenterology*. (2002) 123:803–9. doi: 10.1053/gast.2002.35381
72. Van de Wal Y, Kooy Y, van Veelen P, Pena S, Mearin L, Papadopoulos G, et al. Selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity. *J Immunol*. (1998) 161:1585–8.
73. Petersen J, Montserrat V, Mujico JR, Loh KL, Beringer DX, van Lummel M, et al. T-cell receptor recognition of HLA-DQ2–gliadin complexes associated with celiac disease. *Nat Struct Mol Biol*. (2014) 21:480–8. doi: 10.1038/nsmb.2817
74. Petersen J, Kooy-Winkelaar Y, Loh KL, Tran M, van Berrgen J, Koning F, et al. Diverse T cell receptor gene usage in HLA-DQ8-associated celiac disease converges into a consensus binding solution. *Structure*. (2016) 24:1643–57. doi: 10.1016/j.str.2016.07.010
75. Sollid LM, Qiao SW, Anderson RP, Gianfrani C, Koning F. Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenetics*. (2012) 64:455–60. doi: 10.1007/s00251-012-0599-z
76. Quarsten H, McAdam SN, Jensen T, Arentz-Hansen H, Molberg O, Lundin KEA, et al. Staining of celiac disease-relevant T cells by peptide-DQ2 multimers. *J Immunol*. (2001) 167:4861–8. doi: 10.4049/jimmunol.167.9.4861
77. Anderson RP, Degano P, Godkin AJ, Jewell DP, Hill AVS. *In vivo* antigen challenge in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope. *Nat Med*. (2000) 6:337–42. doi: 10.1038/73200
78. Anderson RP, van Heel DA, Tye-Din JA, Barnardo M, Salio M, Jewell DP, et al. T cells in peripheral blood after gluten challenge in coeliac disease. *Gut*. (2005) 54:1217–723. doi: 10.1136/gut.2004.059998
79. Raki M, Fallang LE, Brottveit M, Bergseng E, Quarsten H, Lundin KEA, et al. Tetramer visualization of gut-homing gluten-specific T cells in the peripheral blood of celiac disease patients. *Proc Natl Acad Sci USA*. (2007) 104:2831–6. doi: 10.1073/pnas.0608610104
80. Brottveit M, Raki M, Bergseng E, Fallang L-E, Simonsen B, Lovik A, et al. Assessing possible celiac disease by an HLA-DQ2-gliadin tetramer test. *Am J Gastroenterol*. (2011) 106:1318–24. doi: 10.1038/ajg.2011.23
81. Sarna VK, Lundin KEA, Morkrid L, Qiao S-W, Sollid LM, Christophersen A, et al. HLA-DQ-Gluten tetramer blood test accurately identifies patients with and without celiac disease in absence of gluten consumption. *Gastroenterology*. (2018) 154:886–96.e6. doi: 10.1053/j.gastro.2017.11.006
82. Christophersen A, Raki M, Bergseng E, Lundin KEA, Jahnsen J, Sollid LM, et al. Tetramer-visualized gluten-specific CD4+ T cells in blood as a potential diagnostic marker for coeliac disease without oral gluten challenge. *U Eur Gastroenterol J*. (2014) 2:268–78. doi: 10.1177/2050640614540154
83. Sarna VK, Skodje GI, Reims HM, Risnes LF, Dahal-Koirala S, Sollid LM, et al. HLA-DQ:gluten tetramer test in blood gives better detection of coeliac patients than biopsy after 14-day gluten challenge. *Gut*. (2018) 67:1606–13. doi: 10.1136/gutjnl-2017-314461
84. Christophersen A, Lund EG, Snir O, Solà E, Kanduri C, Dahal-Koirala S, et al. Distinct phenotype of $CD4^+$ T cells driving celiac disease identified in multiple autoimmune conditions. *Nat Med*. (2019) 25:734–7. doi: 10.1038/s41591-019-0403-9
85. Zühlke S, Fremgaard Risnes L, Dahal-Koirala S, Christophersen A, Sollid LM, Lundin KEA. CD38 expression on gluten-specific T cells is a robust marker of gluten re-exposure in coeliac disease. *U Eur Gastroent J*. (2019) 7:1337–44. doi: 10.1177/2050640619874183
86. Al-Toma A, Volta U, Auricchio R, Castillejo G, Sanders DS, Cellier C, et al. European Society for the Study of Coeliac Disease (ESsCD) guideline for coeliac disease and other gluten-related disorders. *United Eur Gastroenterol J*. (2019) 7:583–613. doi: 10.1177/2050640619844125
87. Goel G, Tye-Din JA, Qiao SW, Russell AK, Mayassi T, Ciszewski C, et al. Cytokine release and gastrointestinal symptoms after gluten challenge in celiac disease. *Sci Adv*. (2019) 5:eaaw7756. doi: 10.1126/sciadv.aaw7756
88. Tye-Din JA, Skodje GI, Sarna VK, Dzuris JL, Russell AK, Goel G, et al. Cytokine release after gluten ingestion differentiates coeliac disease from self-reported gluten sensitivity. *United Eur. Gastroenterol. J*. (2019) 8:108–18. doi: 10.1177/2050640619874173
89. Penagini F, Dilillo D, Meneghin F, Mameli C, Fabiano V, Zuccotti GV. Gluten-free diet in children: an approach to a nutritionally adequate and balanced diet. *Nutrients*. (2013) 5:4553–65. doi: 10.3390/nu5114553
90. Leffler DA, Dennis M, Edwards George JB, Jamma S, Magge S, Cook EF, et al. A simple validated gluten-free diet adherence survey for adults with celiac disease. *Clin Gastroenterol Hepatol*. (2009) 7:530–6. doi: 10.1016/j.cgh.2008.12.032
91. Troncone R, Mayer M, Spagnuolo F, Maiuri L, Greco L. Endomysial antibodies as unreliable markers for slight dietary transgressions in adolescents with celiac disease. *J Pediatr Gastroenterol Nutr*. (1995) 21:69–72. doi: 10.1097/00005176-199507000-00012
92. Comino I, Fernández-Bañares F, Esteve M, Ortigosa L, Castillejo G, Fambuena B, et al. Fecal gluten peptides reveal limitations of serological tests and food questionnaires for monitoring gluten-free diet in celiac disease patients. *Am J Gastroenterol*. (2016) 111:1456–65. doi: 10.1038/ajg.2016.439
93. Keck B, Köhler P, Wieser H. Disulphide bonds in wheat gluten: cystine peptides derived from gluten proteins following peptic and thermolytic digestion. *Z Lebensm Unters Forsch*. (1995) 200:432–9. doi: 10.1007/BF01193253
94. Wieser H. Chemistry of gluten proteins. *Food Microbiol*. (2007) 24:115–9. doi: 10.1016/j.fm.2006.07.004
95. Shan L, Molberg O, Parrot I, Hausch F, Filiz F, Gray GM, et al. Structural basis for gluten intolerance in celiac sprue. *Science*. (2002) 297:2275–9. doi: 10.1126/science.1074129
96. Matysiak-Budnik T, Moura IC, Arcos-Fajardo M, Lebreton C, Ménard S, Candalh C, et al. Secretory IgA mediates retrotranscytosis of intact gliadin peptides via the transferrin receptor in celiac disease. *J Exp Med*. (2008) 205:143–54. doi: 10.1084/jem.20071204
97. Gianfrani C, Auricchio S, Troncone R. Adaptive and innate immune responses in coeliac disease. *Immunol Lett*. (2005) 99:141–5. doi: 10.1016/j.imlet.2005.02.017
98. Sollid LM, Tye-Din JA, Qiao SW, Anderson RP, Gianfrani C, Koning F. Update 2020: nomenclature and listing of celiac disease–relevant gluten epitopes recognized by CD4+ T cells. *Immunogenetics*. (2020) 72:85–8. doi: 10.1007/s00251-019-01141-w
99. Junker Y, Zeissig S, Kim SJ, Barisani D, Wieser H, Leffler DA, et al. Wheat amylase trypsin inhibitors drive intestinal inflammation via activation of toll-like receptor 4. *J Exp Med*. (2012) 209:2395–408. doi: 10.1084/jem.20102660
100. Brouns F, van Rooy G, Shewry P, Rustgi S, Jonkers D. Adverse reactions to wheat or wheat components. *Compr Rev Food Sci Food Saf*. (2019) 18:1437–52. doi: 10.1111/1541-4337.12475
101. Rustgi S, Shewry P, Brouns F, Deleu LJ, Delcour JA. Wheat seed proteins: factors influencing their content, composition, and technological properties, and strategies to reduce adverse reactions. *Compr Rev Food Sci Food Saf*. (2019) 18:1751–69. doi: 10.1111/1541-4337.12493
102. Sollid LM, Lundin KEA. Diagnosis and treatment of celiac disease. *Mucosal Immunology*. (2009) 2:3–7. doi: 10.1038/mi.2008.74
103. Tye-Din JA, Galipeau HJ, Agardh D. Celiac disease: a review of current concepts in pathogenesis, prevention, and novel therapies. *Front Pediatr*. (2018) 6:350. doi: 10.3389/fped.2018.00350
104. Cavaletti L, Taravella A, Carrano L, Carenzi G, Sigurta G, Solinas N, et al. E40, a novel microbial protease efficiently detoxifying gluten proteins, for the dietary management of gluten intolerance. *Sci Rep*. (2019) 9:13147. doi: 10.1038/s41598-019-48299-7
105. Gass J, Bethune MT, Siegel M, Spencer A, Khosla C. Combination enzyme therapy for gastric digestion of dietary gluten in patients with celiac sprue. *Gastroenterology*. (2007) 133:472–80. doi: 10.1053/j.gastro.2007.05.028
106. Mitea C, Havenaar R, Drijfhout JW, Edens L, Dekking L, Koning F. Efficient degradation of gluten by a prolyl endoprotease in a gastrointestinal model: Implications for coeliac disease. *Gut*. (2008) 57:25–32. doi: 10.1136/gut.2006.111609
107. Wolf C, Siegel JB, Tinberg C, Camarca A, Gianfrani C, Paski S, et al. Engineering of Kuma030: A gliadin peptidase that rapidly degrades immunogenic gliadin peptides in gastric conditions. *J Am Chem Soc*. (2015) 137:13106–13. doi: 10.1021/jacs.5b08325
108. Bethune MT, Khosla C. Oral enzyme therapy for celiac sprue. *Methods Enzymol*. (2012) 502:241–71. doi: 10.1016/B978-0-12-416039-2.00013-6
109. König J, Holster S, Bruins MJ, Brummer RJ. Randomized clinical trial: effective gluten degradation by *Aspergillus niger*-derived enzyme in a complex meal setting. *Sci Rep*. (2017) 7:13100. doi: 10.1038/s41598-017-13587-7

110. Knorr V, Wieser H, Koehler P. Production of gluten-free beer by peptidase treatment. *Eur Food Res Technol*. (2016) 242:1129–40. doi: 10.1007/s00217-015-2617-5
111. Osorio CE, Wen N, Mejías JH, Liu B, Reinbothe S, von Wettstein D, et al. Development of wheat genotypes expressing a glutamine-specific endoprotease from barley and a prolyl endopeptidase from *Flavobacterium meningosepticum* or *Pyrococcus furiosus* as a potential remedy to celiac disease. *Funct Integr Genomics*. (2019) 19:123–36. doi: 10.1007/s10142-018-0632-x
112. Osorio CE, Wen N, Mejías JH, Mitchell S, von Wettstein D, Rustgi S. Directed-mutagenesis of Flavobacterium meningosepticum prolyl-oligopeptidase and a glutamine-specific endopeptidase from barley. *Front Nutr*. (2020) 7:11. doi: 10.3389/fnut.2020.00011
113. Dennis M, Lee AR, McCarthy T. Nutritional considerations of the gluten-free diet. *Gastroenterol Clin North Am*. (2019) 48:53–72. doi: 10.1016/j.gtc.2018.09.002
114. Rabinowitz LG, Zylberberg HM, Levinovitz A, Stockwell MS, Green PHR, Lebwohl B. Skepticism regarding vaccine and gluten-free food safety among patients with celiac disease and non-celiac gluten sensitivity. *Dig Dis Sci*. (2018) 63:1158–64. doi: 10.1007/s10620-017-4879-1
115. Sollid LM, Khosla C. Novel therapies for coeliac disease. *J Intern Med*. (2011) 269:604–13. doi: 10.1111/j.1365-2796.2011.02376.x
116. Kurppa K, Hietikko M, Sulic A-M, Kaukinen K, Lindfors K. Current status of drugs in development for celiac disease. *Expert Opin Invest Drugs*. (2014) 23:1079–91. doi: 10.1517/13543784.2014.916274
117. Lindfors K, Ciacci C, Kurppa K, Lundin KEA, Makharia GK, Mearin ML, et al. Coeliac disease. *Nat Rev Dis Primers*. (2019) 5:3. doi: 10.1038/s41572-018-0054-z
118. Serena G, Kelly CP, Fasano A. Nondietary therapies for celiac disease. *Gastroenterol Clin North Am*. (2019) 48:145–63. doi: 10.1016/j.gtc.2018.09.011
119. Christophersen A, Risnes LF, Dahal-Koirala S, Sollid LM. Therapeutic and diagnostic implications of T cell sarring in celiac disease and beyond. *Trends Mol Med*. (2019) 25:836–52. doi: 10.1016/j.molmed.2019.05.009
120. Korneychuk N, Meresse B, Cerf-Bensussan N. Lessons from rodent models in celiac disease. *Mucosal Immunol*. (2015) 8:18–28. doi: 10.1038/mi.2014.102
121. Freitag TL, Podojil JR, Pearson RM, Fokta FJ, Sahl C, Messing M, et al. Gliadin nanoparticles induce immune tolerance to gliadin in mouse models of celiac disease. *Gastroenterology*. (2020). doi: 10.1053/j.gastro.2020.01.045
122. Abadie V, Kim SM, Lejeune T, Palanski BA, Ernest JD, Tastet O, et al. IL-15, gluten and HLA-DQ8 drive tissue destruction in coeliac disease. *Nature*. (2020) 578:600–4. doi: 10.1038/s41586-020-2003-8
123. International Wheat Genome Sequencing Consortium (IWGSC), Appels R, Eversole K, Stein N, Feuillet C, Keller B, et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*. (2018) 361:eaar7191. doi: 10.1126/science.aar7191
124. Juhasz A, Belova T, Florides CG, Maulis C, Fischer I, Gell G, et al. Genome mapping of seed-borne allergens and immunoresponsive proteins in wheat. *Sci Adv*. (2018) 4:eaar8602. doi: 10.1126/sciadv.aar8602
125. Kondrashova A, Mustalahti K, Kaukinen K, Viskari H, Volodicheva V, Haapala AM, et al. Lower economic status and inferior hygienic environment may protect against celiac disease. *Ann Med*. (2008) 40:223–1. doi: 10.1080/07853890701678689
126. Unalp-Arida A, Ruhl CE, Choung RS, Brantner TL, Murray JA. Lower prevalence of celiac disease and gluten-related disorders in persons living in Southern vs Northern latitudes of the United States. *Gastroenterology*. (2017) 152:1922–32.e2. doi: 10.1053/j.gastro.2017.02.012
127. Kasarda DD. Can an increase in celiac disease be attributed to an increase in the gluten content of wheat as a consequence of wheat breeding? *J Agric Food Chem*. (2013) 61:1155–9. doi: 10.1021/jf305122s
128. Van den Broeck HC, De Jong HC, Salentijn EMJ, Dekking L, Bosch D, Hamer RJ, et al. Presence of celiac disease epitopes in modern and old hexaploid wheat varieties. Wheat breeding may have contributed to increased prevalence of celiac disease. *Theor Appl Genet*. (2010) 121:1527–39. doi: 10.1007/s00122-010-1408-4
129. Molberg O, Uhlen AK, Jensen T, Flaete NS, Fleckenstein B, Arentz-Hansen H, et al. Mapping of gluten T-cell epitopes in bread wheat ancestors: Implications for celiac disease. *Gastroenterology*. (2005) 128:393–401. doi: 10.1053/j.gastro.2004.11.003
130. Spaenij-Dekking L, Kooy-Winkelaar Y, van Veelen P, Drijfhout JW, Jonker H, van Soest L, et al. Natural variation in toxicity of wheat: potential for selection of nontoxic varieties for celiac disease patients. *Gastroenterology*. (2005) 129:797–806. doi: 10.1053/j.gastro.2005.06.017
131. Van den Broeck H, Hongbing C, Lacaze X, Dusautoir JC, Gilissen L, Smulders M, et al. In search of tetraploid wheat accessions reduced in celiac disease-related gluten epitopes. *Mol Biosyst*. (2010) 6:2206–13. doi: 10.1039/c0mb00046a
132. Geisslitz S, Wieser H, Scherf KA, Koehler P. Gluten protein composition and aggregation properties as predictors for bread volume of common wheat, spelt, durum wheat, emmer and einkorn. *J Cereal Sci*. (2018) 83:204–12. doi: 10.1016/j.jcs.2018.08.012
133. Geisslitz S, Longin CFH, Scherf KA, Koehler P. Comparative study on gluten protein composition of ancient (einkorn, emmer and spelt) and modern wheat species (durum and common wheat). *Foods*. (2019) 8:409. doi: 10.3390/foods8090409
134. Ozuna CV, Barro F. Characterization of gluten proteins and celiac disease-related immunogenic epitopes in the *Triticeae*: cereal domestication and breeding contributed to decrease the content of gliadins and gluten. *Mol Breed*. (2018) 38:22. doi: 10.1007/s11032-018-0779-0
135. Boukid F, Prandi B, Sforza S, Sayar R, Seo YW, Mejri M, et al. Understanding the effects of genotype, growing year, and breeding on Tunisian durum wheat allergenicity. 2. The celiac disease case. *J Agric Food Chem*. (2017) 65:5837–46. doi: 10.1021/acs.jafc.7b02041
136. Malalgoda M, Meinhardt SW, Simsek S. Detection and quantitation of immunogenic epitopes related to celiac disease in historical and modern hard red spring wheat cultivars. *Food Chem*. (2018) 264:101–7. doi: 10.1016/j.foodchem.2018.04.131
137. Ribeiro M, Rodriguez-Quijano M, Nunes FM, Carrillo JM, Branlard G, Igrejas G. New insights into wheat toxicity: breeding did not seem to contribute to a prevalence of potential celiac disease's immunostimulatory epitopes. *Food Chem*. (2016) 213:8–18. doi: 10.1016/j.foodchem.2016.06.043
138. Schalk K, Lang C, Wieser H, Koehler P, Scherf KA. Quantitation of the immunodominant 33-mer peptide from α-gliadin in wheat flours by liquid chromatography tandem mass spectrometry. *Sci Rep*. (2017) 7:45092. doi: 10.1038/srep45092
139. Iacomino G, Di Stasio L, Fierro O, Picariello G, Venezia A, Gazza L, et al. Protective effects of ID331 *Triticum monococcum* gliadin on *in vitro* models of the intestinal epithelium. *Food Chem*. (2016) 212:537–42. doi: 10.1016/j.foodchem.2016.06.014
140. Tanner GJ, Blundell MJ, Colgrave ML, Howitt CA. Creation of the first ultra-low gluten barley (*Hordeum vulgare* L.) for coeliac and gluten-intolerant populations. *Plant Biotechnol J*. (2016) 14:1139–50. doi: 10.1111/pbi.12482
141. Jouanin A, Gilissen LJWJ, Boyd LA, Cockram J, Leigh FJ, Wallington EJ, et al. Food processing and breeding strategies for coeliac-safe and healthy wheat products. *Food Res Int*. (2018) 110:11–21. doi: 10.1016/j.foodres.2017.04.025
142. Shewry PR, Tatham AS. Improving wheat to remove coeliac epitopes but retain functionality. *J Cereal Sci*. (2016) 67:12–21. doi: 10.1016/j.jcs.2015.06.005
143. García-Molina MD, Giménez MJ, Sánchez-León S, Barro F. Gluten free wheat: are we there? *Nutrients*. (2019) 11:487. doi: 10.3390/nu11030487
144. Becker D, Wieser H, Koehler P, Folck A, Mühling K-H, Zörb C. Protein composition and techno-functional properties of transgenic wheat with reduced alpha-gliadin content obtained by RNA interference. *J Appl Bot Food Qual*. (2012) 85:23–33.
145. Gil-Humanes J, Pistón F, Hernando A, Alvarez JB, Shewry PR, Barro F. Silencing of γ-gliadins by RNA interference (RNAi) in bread wheat. *J Cereal Sci*. (2008) 48:565–8. doi: 10.1016/j.jcs.2008.03.005
146. Altenbach SB, Tanaka CK, Bradford W, Seabourn BW. Silencing of omega-5 gliadins in transgenic wheat eliminates a major source of environmental variability and improves dough mixing properties of flour. *BMC Plant Biol*. (2014) 14:1. doi: 10.1186/s12870-014-0393-1
147. Altenbach SB, Chang H-C, Yu XB, Seabourn BW, Green PH, Alaedini A. Elimination of omega-1,2 gliadins from bread wheat (*Triticum aestivum*) flour: effects on immunogenic potential and end-use quality. *Front Plant Sci*. (2019) 10:580. doi: 10.3389/fpls.2019.00580

148. Barro F, Iehisa JCM, Giménez MJ, García-Molina MD, Ozuna CV, Comino I, et al. Targeting of prolamins by RNAi in bread wheat: effectiveness of seven silencing-fragment combinations for obtaining lines devoid of coeliac disease epitopes from highly immunogenic gliadins. *Plant Biotechnol J*. (2016) 14:986–996. doi: 10.1111/pbi.12455
149. Sánchez-León S, Gil-Humanes J, Ozuna CV, Giménez MJ, Sousa C, Voytas DF, et al. Low-gluten, non-transgenic wheat engineered with CRISPR/Cas9. *Plant Biotechnol J*. (2018) 16:902–10. doi: 10.1111/pbi.12837
150. Jouanin A, Schaart JG, Boyd LA, Cockram J, Leigh FJ, Bates R, et al. Outlook for coeliac disease patients: towards bread wheat with hypoimmunogenic gluten by gene editing of α- and γ-gliadin gene families. *BMC Plant Biol*. (2019) 19:333. doi: 10.1186/s12870-019-1889-5
151. Jouanin A, Boyd LA, Visser RGF, Smulders MJM. Development of wheat with hypoimmunogenic gluten obstructed by the gene editing policy in Europe. *Front Plant Sci*. (2018) 9:1523. doi: 10.3389/fpls.2018.01523
152. Skerritt JH, Hill AS. Monoclonal antibody sandwich enzyme immunoassays for determination of gluten in foods. *J Agric Food Chem*. (1990) 38:1771–8. doi: 10.1021/jf00098a029
153. Valdés I, García E, Llorente M, Méndez E. Innovative approach to low-level gluten determination in foods using a novel sandwich enzyme-linked immunosorbent assay protocol. *Eur J Gastroenterol Hepatol*. (2003) 15:465–74. doi: 10.1097/01.meg.0000059119.41030.df
154. Morón B, Bethune MT, Comino I, Manyani H, Ferragud M, López MC, et al. Toward the assessment of food toxicity for celiac patients: characterization of monoclonal antibodies to a main immunogenic gluten peptide. *PLoS ONE*. (2008) 3:e2294. doi: 10.1371/journal.pone.0002294
155. Mitea C, Kooy-Winkelaar Y, van Veelen P, de Ru A, Drijfhout JW, Koning F, et al. Fine specificity of monoclonal antibodies against celiac disease-inducing peptides in the gluteom. *Am J Clin Nutr*. (2008) 88:1057–66. doi: 10.1093/ajcn/88.4.1057
156. Bugyi Z, Török K, Hajas L, Adonyi Z, Popping B, Tömösközi S. Comparative study of commercially available gluten ELISA kits using an incurred reference material. *Qual Assur Saf Crops Foods*. (2013) 5:79–87. doi: 10.3920/QAS2012.0174
157. Rzychon M, Brohée M, Cordeiro F, Haraszi R, Ulberth F, O'Connor G. The feasibility of harmonizing gluten ELISA measurements. *Food Chem*. (2017) 234:144–54. doi: 10.1016/j.foodchem.2017.04.092
158. Scherf KA. Gluten analysis of wheat starches with seven commercial ELISA test kits - up to six different values. *Food Anal Methods*. (2017) 10:234–46. doi: 10.1007/s12161-016-0573-8
159. Boison J, Allred L, Almy D, Anderson L, Baumert J, Bhandari S, et al. Standard method performance requirements (SMPRs®) 2017.021: quantitation of wheat, rye, and barley gluten in oats. *J AOAC Int*. (2018) 101:1238–42. doi: 10.5740/jaoacint.SMPR2017.021
160. Wehling P, Scherf KA. Preparation of validation materials for estimating gluten recovery by ELISA according to SMPR 2017.021. *J AOAC Int*. (2019) 103:1–6. doi: 10.5740/jaoacint.19-0081
161. Abbott M, Hayward S, Ross W, Godefroy SB, Ulberth F, Van Hengel AJ, et al. Validation procedures for quantitative food allergen ELISA methods: community guidance and best practices. J. *AOAC Int*. (2010) 93:442–50. doi: 10.1093/jaoac/93.2.442
162. Diaz-Amigo C, Popping B. Accuracy of ELISA detection methods for gluten and reference materials: a realistic assessment. *J Agric Food Chem*. (2013) 61:5681–8. doi: 10.1021/jf3046736
163. Scherf KA, Poms RE. Recent developments in analytical methods for tracing gluten. *J Cereal Sci*. (2016) 67:112–22. doi: 10.1016/j.jcs.2015.08.006
164. Lexhaller B, Tompos C, Scherf KA. Fundamental study on reactivities of gluten protein types from wheat, rye and barley with five sandwich ELISA test kits. *Food Chem*. (2017) 237:320–30. doi: 10.1016/j.foodchem.2017.05.121
165. Schalk K, Lexhaller B, Koehler P, Scherf KA. Isolation and characterization of gluten protein types from wheat, rye, barley and oats for use as reference materials. *PLoS ONE*. (2017) 12:e0172819. doi: 10.1371/journal.pone.0172819
166. Hajas L, Scherf KA, Török K, Bugyi Z, Schall E, Poms RE, et al. Variation in protein composition among wheat (*Triticum aestivum* L.) cultivars to identify cultivars suitable as reference material for wheat gluten analysis. *Food Chem*. (2018) 267:387–94. doi: 10.1016/j.foodchem.2017.05.005
167. Schall E, Scherf KA, Bugyi Z, Hajas L, Török K, Koehler P, et al. Characterisation and comparison of selected wheat (*Triticum aestivum* L.) cultivars and their blends to develop a gluten reference material. *Food Chem*. (2020) 313:126049. doi: 10.1016/j.foodchem.2019.126049
168. Bruins Slot ID, Bremer MGEG, van der Fels-Klerx I, Hamer RJ. Evaluating the performance of gluten ELISA test kits: the numbers do not tell the tale. *Cereal Chem*. (2015) 92:513–21. doi: 10.1094/CCHEM-07-14-0166-R
169. AOAC SMPR 2016.002. *Standard Method Performance Requirements (SMPRs®) for Detection and Quantitation of Selected Food Allergens*. Rockville, MD: AOAC International (2006).
170. Koerner TB, Abbott M, Godefroy SB, Popping B, Yeung JM, Diaz-Amigo C, et al. Validation procedures for quantitative gluten ELISA methods: AOAC allergen community and best practices. *J AOAC Int*. (2013) 96:1033–40. doi: 10.5740/jaoacint.13-043
171. AOAC Appendix D. *Guidelines for Collaborative Study Procedures to Validate Characteristics of a Method of Analysis, Official Methods of Analysis*. 20th ed. Rockville, MD: AOAC International (2016). Available online at: http://www.eoma.aoac.org/app_d.pdf (accessed December 11, 2019).
172. AOAC Appendix F. *Guidelines for Standard Method Performance Requirements, Official Methods of Analysis*. 20th ed. Rockville, MD: AOAC International (2016). Available online at: http://www.eoma.aoac.org/app_f.pdf (accessed December 11, 2019).
173. Koehler P, Chirdo F, Mothes F, Tranquet O, Scherf KA. Considerations concerning methods for gluten quantitation in foods (R5/G12 ELISA). In: Koehler P, editor. *Proceedings of the 30th Meeting, Working Group on Prolamin Analysis and Toxicity*. (Freising) (2017). p. 137–42.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 20 March 2020
doi: 10.3389/fpls.2020.00314

Prolamin Content and Grain Weight in RNAi Silenced Wheat Lines Under Different Conditions of Temperature and Nitrogen Availability

*Miriam Marín-Sanz[1], María J. Giménez[1], Francisco Barro[1] and Roxana Savin[2]**

[1] Department of Plant Breeding, Institute for Sustainable Agriculture (IAS-CSIC), Córdoba, Spain, [2] Department of Crop and Forest Sciences, University of Lleida, Lleida, Spain

Temperature and nitrogen (N) availability are two important environmental factors that may produce important changes in grain composition during grain filling of bread wheat. In this study, four wheat lines with the down-regulation of gliadins by means of RNA interference (RNAi) have been characterized to determine the effect of thermal stress and N availability on grain weight and quality; with focus on gliadin and glutenin protein fractions. Grain weight was reduced with heat stress (HS) in all RNAi lines, whereas gliadin content was increased in the wild-types. With respect to gliadin content, RNAi lines responded to HS and N availability differently from their respective wild-types, except for ω-gliadin content, indicating a very clear stability of silencing under different environmental conditions. In a context of increased temperature and HS events, and in environments with different N availability, the RNAi lines with down-regulated gliadins seem well suited for the production of wheat grain with low gliadin content.

Keywords: gluten proteins, heat stress, transgenic lines, celiac disease, *Triticum aestivum*

OPEN ACCESS

Edited by:
Katharina Anne Scherf, Karlsruhe Institute of Technology (KIT), Germany

Reviewed by:
Ludovit Skultety, Institute of Virology (SAS), Slovakia
Robert Winkler, National Polytechnic Institute, Mexico
Eva Johansson, Swedish University of Agricultural Sciences, Sweden

***Correspondence:**
Roxana Savin
savin@pvcf.udl.cat

Specialty section:
This article was submitted to Plant Proteomics, a section of the journal Frontiers in Plant Science

Received: *20 November 2019*
Accepted: *03 March 2020*
Published: *20 March 2020*

Citation:
Marín-Sanz M, Giménez MJ, Barro F and Savin R (2020) Prolamin Content and Grain Weight in RNAi Silenced Wheat Lines Under Different Conditions of Temperature and Nitrogen Availability. Front. Plant Sci. 11:314. doi: 10.3389/fpls.2020.00314

INTRODUCTION

Grain yield and quality are critical for wheat breeding and management. Both traits are determined during the grain-filling phase. Understanding the processes affecting grain weight and quality during grain filling is important for improving breeding and management strategies. Grain filling is commonly partitioned into three phases: the lag phase, the effective grain filling period, and the maturation drying phase (Egli, 1998). The lag phase is a period of active cell division, characterized by a rapid increase in water content with almost no dry matter accumulation. Grain dry weight then increases rapidly during the effective grain filling period until the maximum dry weight is attained, after which it remains approximately stable while the grain dries. During the effective grain filling period, starch and proteins are deposited in the endosperm (Jenner et al., 1991). It has been shown that the rate of their deposition is controlled by the source-sink balance (Fischer et al., 1977). Approximately 80% of total proteins in wheat grain are gluten (also termed prolamins) whereas the remaining 20% is composed of non-gluten proteins (NGPs) – mainly albumins and globulins (van den Broeck et al., 2009; Wen et al., 2012). Wheat gluten is able to form a network responsible for the viscoelastic properties of wheat flour since it allows the retention of carbon dioxide released during fermentation (Shewry, 2009). Gluten proteins can be further divided into two fractions: glutenins and gliadins (Lafiandra and Kasarda, 1985; Shewry, 2019). The glutenins form polymers linked

by inter-chain disulfide bonds, they are insoluble in alcohol solutions, and can be divided according to their mobility in sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) into high molecular weight (HMW) and low molecular weight (LMW) glutenins. The gliadins are monomeric proteins, soluble in alcohol, and divided into three groups according to their mobility by electrophoresis in polyacrylamide gels at acidic pH (A-PAGE); ω, α and γ-gliadins. The glutenins are responsible for wheat dough elasticity and strength, while gliadins are important for viscosity and extensibility (Shewry and Halford, 2002; Shewry et al., 2003). Gluten proteins, particularly the gliadin fraction, are the primary factors responsible for triggering celiac disease (CD), since they contain the most immunogenic CD epitopes (Arentz-Hansen et al., 2000, 2002; Shan et al., 2002; Molberg et al., 2003; Ludvigsson et al., 2013; **Box 1**).

RNA interference-based (RNAi) techniques are ideal for the down-regulation of specific protein fractions related to CD. Using this technology, γ-gliadins were silenced in two lines of bread wheat, providing reductions of up to 80% in this gliadin fraction (Gil-Humanes et al., 2008). Subsequently, the same workers used chimeric interference RNAs capable of silencing the genes from all the three groups of ω, γ and α-gliadins, to obtain several lines of two wheat genotypes with major reductions (in some cases up to 90%) in total gliadin content (Gil-Humanes et al., 2010). The crossing of the silenced lines with commercial varieties of wheat has allowed the obtention of new lines that effectively express the fragment of silencing in different genetic backgrounds, both to silence γ-gliadins (Gil-Humanes et al., 2012) and total gliadin fractions.

Climate model projections suggest that higher temperatures and heat stress events will become commonplace in most regions where grain crops are produced (Meehl and Tebaldi, 2004). Deleterious effects of high temperature on crop yield and quality are well documented in the literature (e.g. Barnabás et al., 2008). It is also well known that temperate species, such as wheat, maximize their vegetative growth during the period of colder temperatures, and the grain develops as temperature rises. Much works have studied the effects of very high temperature – moderate and short periods – during grain filling in wheat (Wardlaw et al., 2002; Farooq et al., 2011; Nuttall et al., 2018). Typically, high temperature during the grain filling linear phase results in the reduction of grain weight, mainly due to the decrease of soluble starch synthase activity under heat stress (Hawker and Jenner, 1993), reducing starch accumulation (e.g. Bhullar and Jenner, 1986). Simultaneously, heat stress (HS) not only increases protein percentage (Stone, 2001; Wardlaw et al., 2002), but also affects the synthesis of the different prolamin fractions and their ratios, which are responsible of the bread quality (Blumenthal et al., 1993; Stone, 2001).

Nitrogen (N) fertilization is one of the most common management practices used by farmers to improve yields. Consequently, there have been many studies analyzing yield (Foulkes et al., 1998; Barraclough et al., 2014) and protein content (Fischer et al., 1993) in response to soil N availability in wheat. In addition, several studies reported the effects of N on the types of proteins being synthesized during grain filling (Pechanek et al., 1997; Daniel and Triboi, 2000; Johansson et al., 2013), indicating that the synthesis of proteins in cereals is clearly influenced by temperature and N condition under which grain filling proceeds. However, genotypic variability can be found in all these responses (Saint Pierre et al., 2008; Elbashir et al., 2017); and even in the response to interactions between heat and N (Elía et al., 2018; Slafer and Savin, 2018). Therefore, determining the effects of heat and N availability on wheat genotypes with contrasting protein composition are of particular interest in the understanding grain protein distribution and its influence on grain weight and quality.

RNAi lines with low gluten content were subjected to various N and sulfur treatments to study the stability of the gliadin silencing under different fertilization conditions (García-Molina and Barro, 2017). In relation to N, this study showed that the RNAi lines had consistently lower gliadin levels than the wild-type across different N-fertilization regimes, but also that the level of gliadins in RNAi lines was sometimes significantly increased when N availability increased. In that study, N was applied when it would strongly affect grain number and yield and, therefore, may have diluted the availability of N-compounds during grain filling (i.e. more N available for absorption had to be shared between much higher grain numbers). As late N fertilization can be used to maintain green tissues during grain filling and to increase overall N content of the grains (Blandino et al., 2015), it

BOX 1 | Celiac disease (CD) is a chronic enteropathy that results from the ingestion of gluten proteins present in wheat, and other similar proteins in barley and rye (Trier, 1998; Sollid, 2002). After ingestion of gluten, lesions form in the small intestine, characterized by flattening of the microvilli, hyperplasia of crypt cells, and infiltration of leukocytes (Sollid, 2002). As a result, symptoms such as diarrhea and malabsorption of food appear among others, since the spectrum of symptoms can be very broad. The immune response is triggered by the activation of CD4 T cells when they recognize the gluten peptides presented by serotypes HLA-DQ2 and HLA-DQ8. The presence of gluten peptides can be detected by the activity of the tissue transglutaminase 2 enzyme from the intestinal mucosa (Sollid, 2002; Sollid et al., 2012; Gayathri and Rashmi, 2014). CD is present throughout the world and the prevalence in the United States is around 1%, as in Europe, with the highest estimates in Finland and Sweden, and the lowest in Germany (Catassi et al., 2014). Gluten is present in many food products as the main element or as an additive. So far, the only possible treatment for CD is to follow a gluten-free diet for life (Sollid, 2002). The increase in the incidence of the disease was associated with the duration of exposure to gluten (Ventura et al., 1999), which increases the need to obtain wheat with a reduced content of proteins immunogenic for celiac sufferers. In addition to CD, there are other pathologies related to wheat: (i) allergies as wheat-dependent exercise-induced anaphylaxis (WDEIA) – induced by the ω-5 gliadins and the HMW- (Morita et al., 2007; Morita et al., 2009), or baker's asthma associated with non-specific lipid transfer proteins (Brant, 2007; Palacin et al., 2007); and (ii) non-celiac wheat sensitivity (NCWS) (Gibson et al., 2017), with an estimated prevalence ranging from 0.6 to 13% of global population (Aziz et al., 2016). Most of the allergens and proteins related to wheat pathologies have been mapped to the bread wheat Chinese Spring reference genome (RefSeq v1.0, International Wheat Genome Sequencing Consortium) (Appels et al., 2018) contributing to the knowledge of these diseases (Juhász et al., 2018). Moreover, there is a broad study on wheat allergens and CD peptides that allows their identification and composition for diagnostic assays by liquid chromatography-tandem mass spectrometry (Lexhaller et al., 2019).

may be relevant to explore whether the response of the RNAi lines would be even more marked. As mentioned before, grain filling is significantly affected by HS which would also favor the synthesis of proteins compared with that of starch (Barlow et al., 2015). It would thus be of interest to determine whether the synthesis of proteins in general, and gliadins in particular, in these RNAi lines is affected by the combinations of high temperatures and N availability during grain filling.

In this context, the objective of the present work was to determine the effects of contrasting temperature and N availability conditions on the silencing of gliadins. Thus, grain weight, total protein content and gluten protein distribution were studied in a set of RNAi lines and their respective wild-types. The aim was to determine to what degree the silencing of the synthesis of gliadins depends on environmental conditions, which is important for progressing in the development of low-gliadin wheat varieties suitable for new dietary approaches for gluten-related disorders.

MATERIALS AND METHODS

Plant Material, Chamber Experiment and Treatments

Six lines of bread wheat were used: BW208, D770, D793, Gazul, J631, and M959. BW208 is a line from CIMMYT and Gazul is a commercial variety, and both were used as wild-types. D770 and D793 are lines derived from BW208 with RNA interference (RNAi) silencing of all gliadin fractions (Gil-Humanes et al., 2010). J631 and M959 are derived from crossing the Gazul genotype and line D770. Lines J631 and M959 were backcrossed at least four times with Gazul, always selecting the silencing character and the high and low molecular weight glutenins of Gazul, so that both silenced lines maintain the glutenin profile of this parent line.

We carried out a chamber experiment involving six wheat lines (two wild-type cultivars and four RNAi lines), two temperature treatments (control and heat stress, HS) during the linear phase of grain filling period, and two nitrogen (N) availabilities with three replicates, each replicate was composed of 6 plants (all in all there were 18 plants per genotype × N × HS; i.e. 216 plants per chamber). Two seeds were sown in pots (270 cm^3) filled with a mixture of 30% peat and 70% soil. After emergence, one plant was left in each pot.

Plants were grown outdoors until heading when all pots were placed in a growth chamber set at 20/15°C. Different temperature treatments were imposed from 10 days after anthesis (DAA) during 10 consecutive days (**Supplementary Figure 1**). The control was set at 25/18°C in a chamber and the HS treatment to 40/18°C in another chamber. Minimum and maximum temperatures of 18 and 25 or 40°C were maintained for 8 and 6 h, respectively (**Supplementary Figure 1**). After the 10 days of treatment, temperatures were set to 25/18°C until maturity.

Chambers were set under long-day conditions (16 h). Pots inside the chambers were rearranged approximately once a week to minimize the effects of possible differences in microenvironment at different positions within each chamber. Pots were watered regularly to avoid water stress. N (9 mg per pot) was applied as urea diluted in all pots at flag leaf appearance (DC 4.5, Zadoks et al., 1974). At heading (DC 5.9), half of the plants received another dose of N (21 mg per pot).

Grain Weight and Total Protein Determination

At maturity, 18 plants per treatment were sampled. Mature grain weight was determined as the average of all grains from the main spikes harvested. Total grain protein content was determined by Dumas methodology (Dumas, 1831).

Prolamins Quantification by RP-HPLC

For gliadin and glutenin extraction, two grains from three different plants of each line and treatment were weighed and ground using a ball mill, and sequentially extracted following a previous protocol (Pistón et al., 2011) adapted to small samples. Briefly, gliadins were extracted stepwise three times with up to 400 μL of 60% (v/v) ethanol. Samples were centrifuged, and the supernatants collected, mixed together and filtered. The insoluble pellet was re-suspended in 50% (v/v) 1-propanol, 2 M urea, 0.05 M Tris–HCl (pH 7.5), and 2% (w/v) DTT for glutenin extraction, incubated for 30 min at 60°C and centrifuged stepwise three times. For each sample, the three collected supernatants were mixed together and filtered. The protein extracts were used for gliadin and glutenin quantification by Reverse-Phase High-Performance Liquid Chromatography (RP-HPLC, 1200 Series Quaternary LC System liquid chromatography from Agilent Technologies) with a DAD UV-V detector at 210 nm. A 25 cm long column LiChrospher® 100 RP8 (5 μm) (Merck) was used at 50°C and a sample volume equivalent to 2 mg of flour was injected. The flow rate was 0.5 $mL \cdot min^{-1}$. Mobile phase consisted in a mixture of Acetonitrile (ACN 0.1% TFA) and 0.1% aqueous TFA in a linear gradient (0 min 26% ACN, 60 min 54% ACN). The absolute amount of protein was calculated using bovine serum albumin protein as standard (BSA; BSA ≥ 98%, fraction V. Sigma-Aldrich, St Louis, MO, United States cat. no. A3294) (**Supplementary Figure 2**). The intervals of retention time used for the separation of prolamin fractions peaks are indicated in **Supplementary Figure 3** according to Wieser et al. (1998). The integration of the peaks was performed automatically by RP-HPLC software with minor modifications if necessary.

Non-gluten Proteins (NGPs) Determination

The NGPs content was calculated by the difference between the total protein and prolamin content (glutenins and gliadins) for each line. The total protein content (μg/mg) was calculated from the percentage of N obtained by Dumas using the wheat conversion factor (5.83) (Merrill and Watt, 1973).

Data Processing

The retention time (min) and area (mAU) output of the RP-HPLC software was imported into a house developed software made in Python v2.7[1] to obtain the average values from the

[1]https://github.com/MiriamMarinS/prolaminsQuantification

transformed technical repeats using the following formulas, that processes the hundreds of output files in a single run. The output of the software is a file with Microsoft Excel format.

$$\text{Protein}\ (\mu\text{g}\ (\text{mg of flour})^{-1}) =$$

$$0.0005 \cdot \text{Area (mAU)}\ \frac{V_{\text{extraction}}\,(\mu\text{L})}{V_{\text{injection}}\,(\mu\text{L}).\ \text{Grain weight(mg)}}$$

The integration of the profiles, to obtain the area of each peak, and the subsequent transformation using the formula described, allow estimating the amount of protein for the samples. The arithmetic mean of the three technical repetitions was used for the variance analysis.

Statistical Analysis

To determine the effect of HS and N multifactorial univariates ANOVA were performed. Two variants of this model were tested: in the first, genotype, temperature, nitrogen and their interactions were independent variables, while grain weight and protein fractions were dependent variables. It was used to determine the general effect of the treatments on all genotypes. The second, has the same factors and variables, but it was performed for wild-types and RNAi lines separately to determine the effect of the treatments on each of these groups. Principal Components Analysis, PCA, was carried out with grain weight, total gliadin and its fractions, total glutenin and its fractions, and total prolamin as variables to evaluate their contribution to the model variance. The software used for the statistical analysis was R v 3.5.1 (R Core Team, 2018).

RESULTS

Heat Stress and Nitrogen Treatment Effects on Grain Weight and Total Protein

Grain weight was significantly decreased in both wild-types and RNAi lines by heat stress (HS) (**Figure 1A**, **Table 1**, and **Supplementary Table S1**). Additional applications of nitrogen (N) had no significant effect on the RNA interference (RNAi) lines or wild-types (**Figure 1B** and **Table 1**). No significant differences were found in total grain protein content among all genotypes (**Table 1**). HS for a short period did not significantly modify the total protein content for both wild-types and RNAi lines (**Figure 1C**), but the late application of N (N_1) resulted in a significant increase of the total protein for both the wild-types and the RNAi lines (**Figure 1D**).

Heat Stress and Nitrogen Treatment Effects on Gliadins and Glutenins

We confirmed that total gliadin content was significantly higher in the wild-types than in the RNAi lines (**Figure 2A** and **Table 1**).

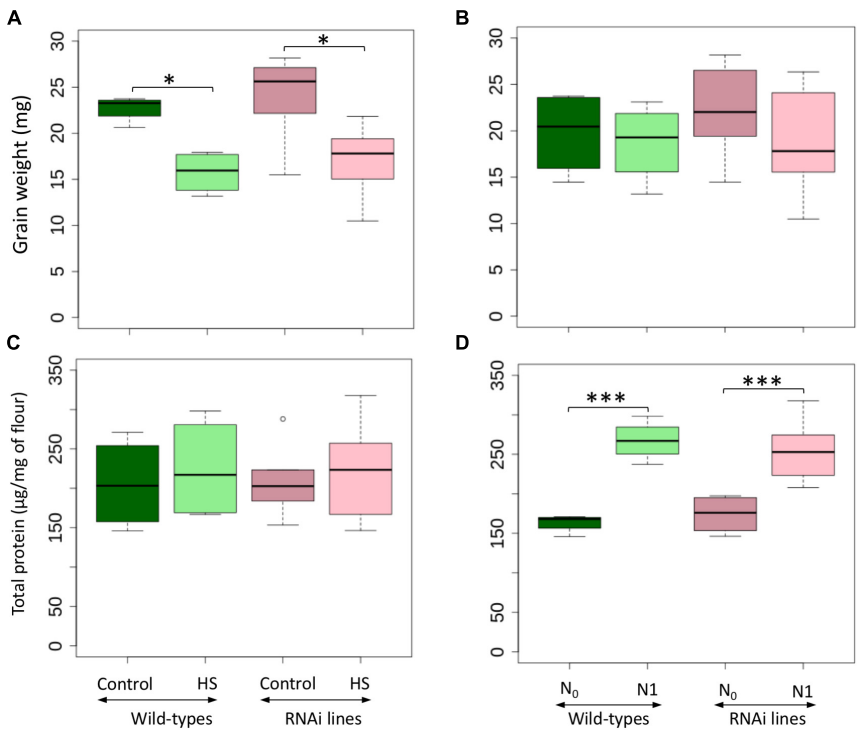


FIGURE 1 | Grain weight and total protein content for wild-types and RNAi lines under control and heat stress temperature treatments **(A,C)** and nitrogen availability **(B,D)**. N_0: no N application after heading, N_1: N application after heading; control: 25/18°C during whole grain filling period, Heat stress (HS): 40/18°C for 10 days during grain filling period. The black line represents the median value. * above the bars indicates significant difference (*$P \leq 0.05$; ***$P \leq 0.001$) between treatments according to the variance analysis.

TABLE 1 | Significance of the variance of effects of genotype (6 lines: BW208, Gazul, D770, D793, J631 and M959), temperature (2 levels: Control temperature and heat stress), nitrogen (2 levels: N_o and N_1) and their interactions for each variable studied (grain weight, protein fractions and protein ratios).

Variables	Factors	*P*-value	Variables	Factors	*P*-value
Grain weight	G	**0.002082****	HMW	G	0.16009
	T	**3.25E-05*****		T	0.09833
	N	**0.003541****		N	0.05425
	GxT	0.604707		GxT	0.55593
	GxN	0.092625		GxN	0.84124
	TxN	0.678413		TxN	0.1672
Total protein	G	0.6512291	LMW	G	**0.005985****
	T	0.2856815		T	0.142497
	N	**0.0008094*****		N	0.104511
	GxT	0.9285325		GxT	0.051446
	GxN	0.5782825		GxN	0.333028
	TxN	0.2003991		TxN	0.929684
ω-gliadins	G	**6.27E-05*****	Total glutenins	G	0.39305
	T	0.07043		T	0.08163
	N	**3.81E-05*****		N	**0.04616***
	GxT	**0.00102****		GxT	0.34164
	GxN	0.06651		GxN	0.65589
	TxN	0.059		TxN	0.24875
α-gliadins	G	**1.30E-05*****	Ratio GLI/GLU	G	**0.0006601*****
	T	0.075602		T	0.0555856
	N	**0.001394****		N	0.3908054
	GxT	**0.014791***		GxT	0.1579138
	GxN	**0.010881***		GxN	0.2167633
	TxN	0.583088		TxN	0.9445682
γ-gliadins	G	**3.01E-07*****	Ratio GLI/TP	G	**2.79E-05*****
	T	0.1145912		T	0.87811
	N	0.1861917		N	**0.009226****
	GxT	**0.0005413*****		GxT	**0.024892***
	GxN	0.180311		GxN	0.072216
	TxN	0.9548613		TxN	0.242135
Total gliadins	G	**1.88E-06*****	Ratio GLU/TP	G	0.23479
	T	0.1784208		T	0.06773
	N	**0.0004081*****		N	0.08808
	GxT	**0.0016167****		GxT	0.30135
	GxN	**0.0127328***		GxN	0.45733
	TxN	0.8348028		TxN	0.17623

*The P-value is presented for significant factors of each variable. *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001. G, genotype; T, temperature; N, nitrogen; HMW, high molecular weight; LMW, low molecular weight; ratio GLI/GLU, ratio total gliadin content/total glutenin content; ratio GLI/TP, ratio total gliadin content/total protein content; ratio GLU/TP, ratio total glutenin content/total protein content; total proteins, total protein content in percent of nitrogen by Dumas. The degree of freedom (d. f.) of the variance analysis (N -1, N: number of observations) for the factors are: G, 5; T, 1; N, 1; GxT, 5; GxN, 5; TxN, 1. P < 0.05 are in bold.*

Among RNAi lines, D793 had lower content of gliadins than that of the rest of RNAi lines (**Supplementary Table S1**). Both HS and the late application of N (N_1) resulted in a significant increase in the total gliadin content for the wild-type lines, whereas no significant variation was observed for the RNAi lines (**Figures 2B,C**).

The degree of silencing of ω-gliadins was lower than that of the rest of the gliadin fractions (**Supplementary Figure 4A** and **Supplementary Table S1**). A significant increase in the content of ω-gliadins due to supplementary N was found on both wild-types and RNAi lines, while HS only had an effect on the wild-types (**Supplementary Figures 4B,C** and **Table 1**). An overall effect of N level on the α-gliadin content (**Supplementary Figure 4C** and **Table 1**), as well as on the total gliadin content (**Figure 2C**), was observed in the wild-types, but the RNAi lines did not show this effect. HS treatment had no effect on the α-gliadin content for wild-types and RNAi lines (**Supplementary Figure 4B**). Conversely, γ-gliadin content was not affected by N availability, but a decrease in the amount of this fraction under HS was seen in wild-types, but not in the RNAi lines (**Supplementary Figure 4B** and **Supplementary Table S1**). Overall, the RNAi line D793 showed the highest reduction in α- and γ-gliadins (**Supplementary Figure 4A** and **Supplementary Table S1**).

The content of glutenins at grain maturity (**Supplementary Table S1**) was not significant higher for most RNAi genotypes

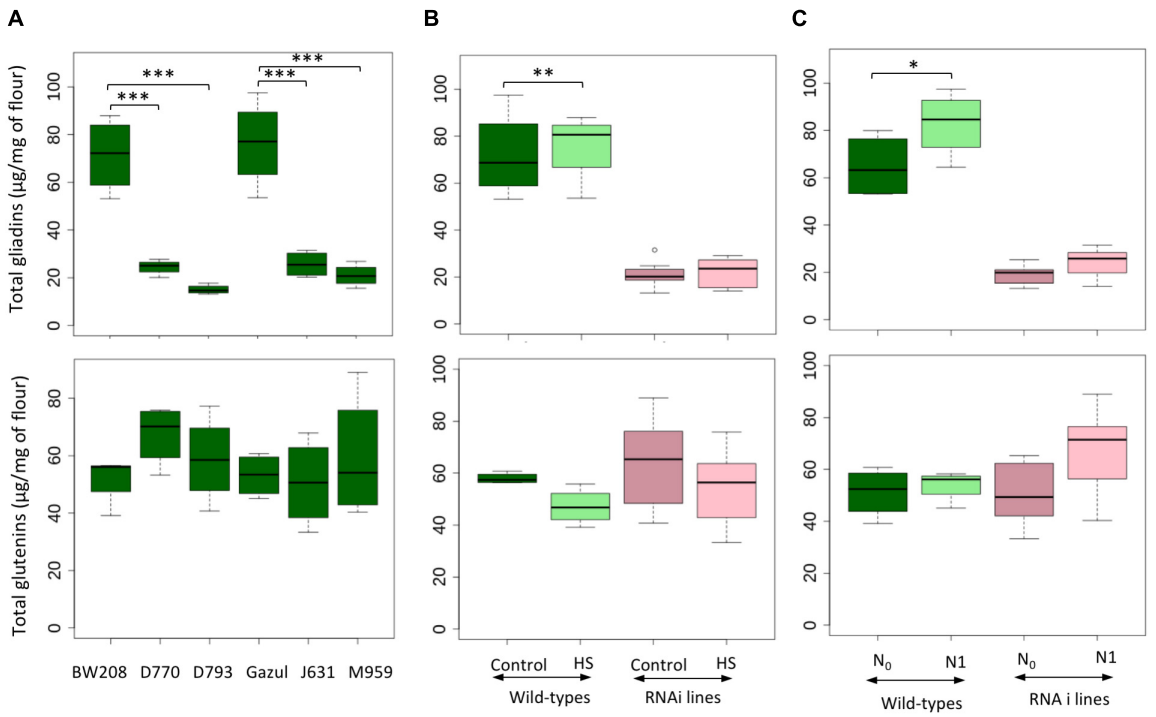


FIGURE 2 | Content of total gliadin and total glutenin content for different genotypes with comparisons between RNAi lines and their wild-type by Dunnett's test **(A)**, wild-types and RNAi lines under different temperature treatments **(B)** and nitrogen availability **(C)**. N_0: no N application after heading, N_1: application after heading; control: 25/18°C during whole grain filling period, Heat stress HS: 40/18°C for 10 days during grain filling period. The black line represents the median value. * above the bars indicates significant difference (*$P \leq 0.05$; **$P \leq 0.01$) between treatments according to the variance analysis.

than that of the wild-type lines (**Figure 2A**). The content of HMW was not statistically different between wild-types and RNAi lines (**Supplementary Figure 5A** and **Table 1**). HS had no effect on the HMW fraction, and for N treatment, RNAi lines tend to have a higher amount of HMW. In contrast, the LMW fraction was decreased in most of the RNAi lines in comparison to that of the wild-types (**Supplementary Figure 5A**, **Table 1**, and **Supplementary Table S1**). In addition, LMW content was affected by HS and N availability; in the wild-types LMW proteins decreased under HS; in RNAi lines LMW fractions increased at higher N availability (**Supplementary Figures 5B,C** and **Supplementary Table S1**).

The total gliadin/total protein ratio (GLI/TP) confirms that RNAi lines have lower gliadin content than the wild-types, particularly line D793 whatever experimental conditions. However, a lower total glutenin/total protein ratio (GLU/TP), was seen in wild-types than in RNAi lines without additional N supply, and in BW208 under HS with N_1 than in RNAi lines (**Figure 3** and **Table 1**).

HS and N availability modified the GLI/TP ratio in the wild-type lines in different ways: it increased with N_1 and HS in BW208 and decreased in Gazul (**Figure 3**). In contrast, minor effects were found in the GLI/TP ratio for the RNAi lines (**Figure 3**). Regarding the GLU/TP ratio, it was also modified by HS and N availability; it was strongly decreased in Gazul wild-type and RNAi derived lines under HS treatment and N_0, while only a minor effect on lines with BW208 background was observed under those conditions; GLI/TP ratio was increased in BW208 RNAi lines under HS and N_1, with only but minor effects in Gazul RNAi lines (**Figure 3**).

A Principal Component Analysis (PCA) was carried out, considering the effect of genotypes, temperature and N availability treatments on the variation of the protein fractions and grain weight (**Figure 4A**). Among the gliadin fractions, ω-gliadin and α-gliadin fractions contribute less and more, respectively, to the variance of the model. HMW proteins were the glutenin fraction that contributed most (**Figure 4A**). Gliadin and glutenin contents varied in perpendicular directions, indicating an independent behavior of both families of proteins (**Figure 4A**). The wild-types are separated from the RNAi lines (**Figure 4B**). The ellipses of 95% confidence level of each genotype indicated that there was a strong association between the variation of the glutenins and the silenced genotypes, and on the other hand, between the variation in the prolamins and the wild-types. D770 and J631 tend toward the direction of variation of the LMW proteins, and D793 and M959 toward that of the total glutenins (**Figure 4**). It is interesting to note that the variation in grain weight was independent of variations in the different protein fractions.

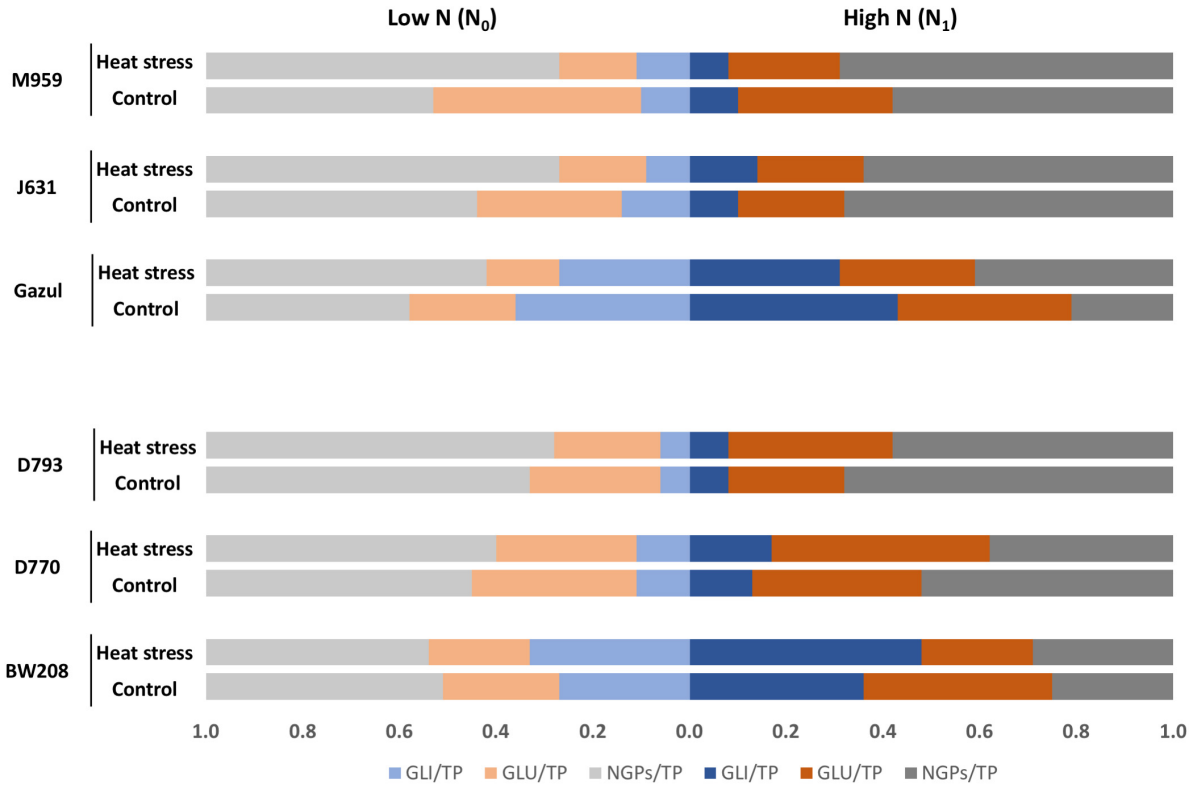


FIGURE 3 | Total gliadin/total grain protein ratio (GLI/TP), total glutenin/total grain protein ratio (GLU/TP) and NGPs/total grain protein ratio (NGPs/TP) between treatments and genotypes. The ratios were obtained with mean values for protein content. Low N: no N application after heading, High N: N application after heading; control: 25/18°C during whole grain filling period, Heat stress HS: 40/18°C for 10 days during grain filling period.

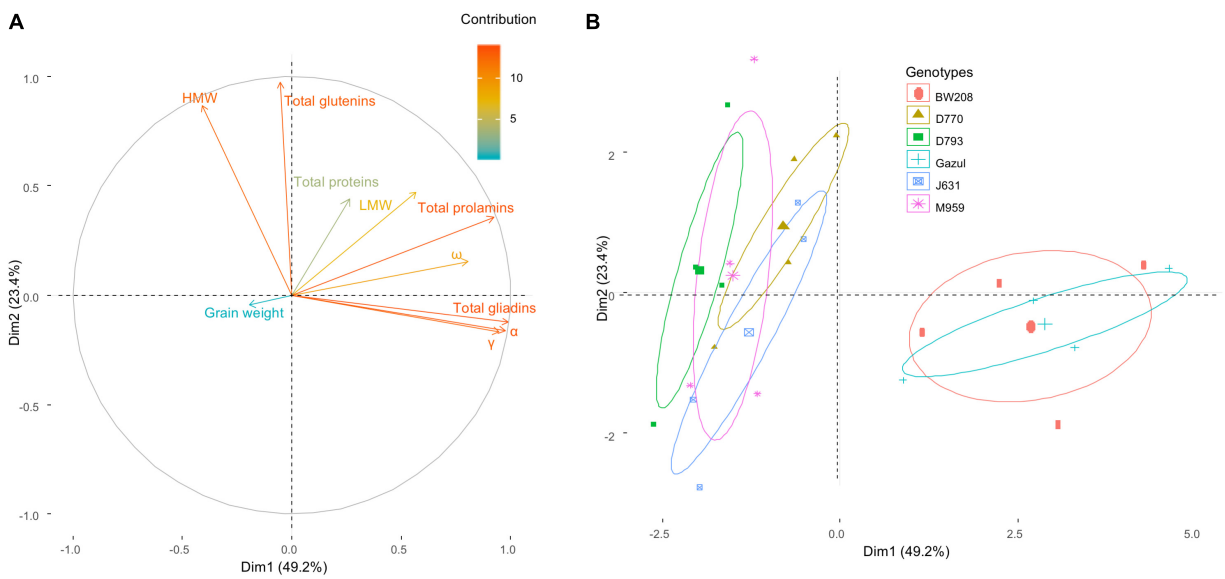


FIGURE 4 | Principal Components Analysis (PCA). Effect of genotypes, temperature and N availability treatments on the variation of the protein fractions and grain weight **(A)**. The high values in the color scale indicates a high contribution to the PCA. The direction and the size of the vectors indicate the relationship between all variables and their contribution to each axis. **(B)** Individuals are represented on the PCA axes with the 95% confidence ellipses showed for each genotype. The largest point for each genotype indicates the intersection of ellipse axes.

DISCUSSION

Grain weight, total grain protein, and prolamin content under control temperature and N_0 (control conditions) resulted in values similar to those previously reported for gliadin down-regulated lines (Gil-Humanes et al., 2010; Pistón et al., 2013). The decrease of the gliadin/total protein (GLI/TP) ratio in the RNA interference (RNAi) lines could be explained by the increase of the NGPs, as the glutenin/total protein (GLU/TP) ratio was higher in RNAi lines. In previous studies with these and other RNAi lines, protein compensation was observed (Altenbach et al., 2014; García-Molina and Barro, 2017) with increments in non-gluten proteins (NGPs) such as serpins, triticins and globulins (Gil-Humanes et al., 2011; Barro et al., 2016).

Brief heat stress (HS) events during the grain filling period generally result in a decrease in grain weight (Bhullar and Jenner, 1986; Savin et al., 1999). The range of variation depends on the genotype, timing and severity of HS (Balla et al., 2019). In the present study, we found a reduction in grain weight of 30% on average under HS and different availability of nitrogen (N). Grain weight was reduced by 35% for lines with BW208 genetic background, and about 24% for Gazul and its RNAi derived lines. This reduction could be mainly due to an extreme dependence on the temperature of starch synthesis, and an irreversible effect of HS on starch production after only a few days or even a few hours per day under control (Daniel and Triboi, 2000; Triboï et al., 2003; Spiertz et al., 2006; DuPont et al., 2006b; Liu et al., 2011; Hurkman et al., 2013) or field conditions (Savin et al., 1996; García et al., 2016; Elía et al., 2018). The percentage of grain protein generally increases under moderately high and very high temperatures (Stone, 2001; Wardlaw et al., 2002), either by a reduction of starch greater than the accumulation of protein, or by a reduction in starch without no change in protein accumulation. However, this response may not always occur for different genetic and environmental backgrounds (Graybosch et al., 1995). Interestingly, in the present study, HS resulted in a higher proportion of gliadin fraction in the wild-types. Other studies reported no effect of HS on total prolamins or even found a decrease with HS (DuPont et al., 2006b; Hurkman et al., 2013). However, Daniel and Triboi (2000) studied each fraction of gliadins and found that the proportion of ω- and α-gliadins increased with HS while γ-gliadins decreased, as found in the wild-types in the present study with the exception of α-gliadins. Also, in this work, grain protein content was increased under HS when post-anthesis N availability was higher (N_1), whereas the total prolamin content of the wild-types differed in the response to HS. The RNAi lines, regardless of their genetic background, did not respond to temperature treatments for the total gliadin content, and for the gliadin fractions. This is an indication that these lines have robust gliadin silencing, independently of the temperature environment. Several authors have described a slight increase in LMW and HMW glutenin fractions with HS (DuPont et al., 2006a,b). Evidence has been also reported that the effect of HS is to cause a reduction in the size of glutenin polymers (e.g. Naeem et al., 2012). Nevertheless, in the present work, total glutenin and their fractions content were not modified significantly with HS in any of the genotypes, except for LMW in wild-types, but there is a non-significant decrease in all fractions in wild-types and RNAi lines.

Under higher N availability (N_1), both wild-type and RNAi lines increase total grain protein, as previously described for other wheat genotypes (Daniel and Triboi, 2000; Triboï et al., 2003). Moreover, when increasing and splitting N doses, grain protein concentration increases and protein composition changes by increasing glutenin fractions (Xue et al., 2016). The response to N_1 in wild-type lines in greater and in RNAi lines in lesser proportion, also confirmed that gliadin to glutenin ratio increase with N_1 (Triboï et al., 2003). The content of ω-gliadins under N_1 was increased, in comparison to N_0, in the wild-types and RNAi lines, whereas the α-gliadin fraction was increased only in the wild-types as previously reported in D793 and other RNAi lines (Gil-Humanes et al., 2010; García-Molina and Barro, 2017). Total gliadin content did not increase in RNAi lines when additional N was supplied (N_1), and this increase of ω-gliadins under N_1 in RNAi lines has to be considered in further designing silencing constructs to improve their effectiveness since the ω-5 gliadins are related to wheat-dependent exercise-induced anaphylaxis (WDEIA) (Inomata, 2009; Morita et al., 2009) and ω-1,2 gliadins to CD (Tye-Din et al., 2010). However, the α-gliadins are reported as the major immunogenic complex in wheat, they contain three major celiac disease (CD) immunogenic peptides (Ozuna et al., 2015), and active peptides from this gliadin fraction were responsible for most of the immune response in patients with CD after eating wheat (Tye-Din et al., 2010). The α-gliadins were strongly reduced in the RNAi lines, and this was not affected either by HS or N application. Although some authors have indicated that LMW proteins decrease with high N availability at moderate temperatures (DuPont et al., 2006b; Hurkman et al., 2013), we found that LMW content increased in RNAi lines when additional N was supplied (N_1). In contrast, the response in the HMW fraction and total glutenin content under N_1, was not statistically significant in any of the genotypes.

CONCLUSION

Wheat grain proteins are important for the breadmaking quality of wheat, but they are also related to human pathologies as celiac disease (CD) and other gluten intolerances. RNA interference (RNAi) technology has provided wheat lines with all the gliadin fractions strongly down-regulated. Heat stress (HS) and nitrogen (N) availability could affect the synthesis and deposition of proteins during grain filling. Wild-types and RNAi lines studied in this work responded similarly for total grain protein and the content of ω-gliadins to additional N supply, as well as for the grain weight under HS. While the wild-types increase their total gliadin content under HS or high N availability the RNAi lines did not. Interestingly, the α-gliadin content, the most CD immunogenic fraction, is unaffected in the RNAi lines under additional N supply, but it was increased in wild-types. Therefore, under the specific scenario of brief events of temperature increase or additional

application of N, studied in this work, the RNAi lines demonstrated a high stability of down-regulation of gliadins. However, further evaluations under field conditions will be necessary to confirm that the silencing of gliadin fractions in RNAi lines can be maintained under different abiotic stress environments.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

FB and RS designed the work. MM-S, MG, and RS carried out the work. All authors wrote the manuscript and approved the manuscript.

FUNDING

The Spanish Ministry of Economy and Competitiveness (Project AGL2016-80566-P) and the European Regional Development Fund (FEDER) supported this research.

ACKNOWLEDGMENTS

The technical assistance of Ana García and the comments on the revised version of the manuscript by Gustavo A. Slafer are gratefully acknowledged. We also thank Dr. Paul A. Lazzeri (Agrasys SL) for his help in English review of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00314/full#supplementary-material

REFERENCES

Altenbach, S. B., Tanaka, C. K., and Seabourn, B. W. (2014). Silencing of omega-5 gliadins in transgenic wheat eliminates a major source of environmental variability and improves dough mixing properties of flour. *BMC Plant Biol.* 14:393. doi: 10.1186/s12870-014-0393-1

Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:eaar7191. doi: 10.1126/science.aar7191

Arentz-Hansen, E., McAdam, S., Molberg, Ø., Kristiansen, C., and Sollid, L. (2000). Production of a panel of recombinant gliadins for the characterisation of T cell reactivity in coeliac disease. *Gut* 46, 46–51. doi: 10.1136/gut.46.1.46

Arentz-Hansen, H., Mcadam, S. N., Molberg, Ø., Fleckenstein, B., Lundin, K. E., Jørgensen, T. J., et al. (2002). Celiac lesion T cells recognize epitopes that cluster in regions of gliadins rich in proline residues. *Gastroenterology* 123, 803–809. doi: 10.1053/gast.2002.35381

Aziz, I., Dwivedi, K., and Sanders, D. S. (2016). From coeliac disease to noncoeliac gluten sensitivity; should everyone be gluten free? *Curr. Opin. Gastroenterol.* 32, 120–127. doi: 10.1097/MOG.0000000000000248

Balla, K., Karsai, I., Bónis, P., Kiss, T., Berki, Z., Horváth, Á., et al. (2019). Heat stress responses in a large set of winter wheat cultivars (*Triticum aestivum* L.) depend on the timing and duration of stress. *PLoS One* 14:e0222639. doi: 10.1371/journal.pone.0222639

Barlow, K., Christy, B., O'leary, G., Riffkin, P., and Nuttall, J. (2015). Simulating the impact of extreme heat and frost events on wheat crop production: a review. *Field Crops Res.* 171, 109–119. doi: 10.1016/j.fcr.2014.11.010

Barnabás, B., Jäger, K., and Fehér, A. (2008). The effect of drought and heat stress on reproductive processes in cereals. *Plant Cell Environ.* 31, 11–38.

Barraclough, P. B., López-Bellido, R., and Hawkesford, M. J. (2014). Genetic variation in the uptake, partitioning and remobilisation of nitrogen during grain-filling in wheat. *Field Crops Res.* 156, 242–248. doi: 10.1016/j.fcr.2013.10.004

Barro, F., Iehisa, J. C., Giménez, M. J., García-Molina, M. D., Ozuna, C. V., Comino, I., et al. (2016). Targeting of prolamins by RNAi in bread wheat: effectiveness of seven silencing-fragment combinations for obtaining lines devoid of coeliac disease epitopes from highly immunogenic gliadins. *Plant Biotechnol. J.* 14, 986–996. doi: 10.1111/pbi.12455

Bhullar, S., and Jenner, C. (1986). Effects of temperature on the conversion of sucrose to starch in the developing wheat endosperm. *Funct. Plant Biol.* 13, 605–615.

Blandino, M., Vaccino, P., and Reyneri, A. (2015). Late-season nitrogen increases improver common and durum wheat quality. *Agron. J.* 107, 680–690. doi: 10.2134/agronj14.0405

Blumenthal, C., Barlow, E., and Wrigley, C. (1993). Growth environment and wheat quality: the effect of heat stress on dough properties and gluten proteins. *J. Cereal Sci.* 18, 3–21. doi: 10.1006/jcrs.1993.1030

Brant, A. (2007). Baker's asthma. *Curr. Opin. Allergy Clin. Immunol.* 7, 152–155.

Catassi, C., Gatti, S., and Fasano, A. (2014). The new epidemiology of celiac disease. *J. Pediatr. Gastroenterol. Nutr.* 59, S7–S9. doi: 10.1097/01.mpg.0000450393.23156.59

Daniel, C., and Triboi, E. (2000). Effects of temperature and nitrogen nutrition on the grain composition of winter wheat: effects on gliadin content and composition. *J. Cereal Sci.* 32, 45–56. doi: 10.1006/jcrs.2000.0313

Dumas, J. B. A. (1831). Procedes de l'analyse organique. *Ann. Chim. Phys.* 2, 198–213.

DuPont, F. M., Hurkman, W. J., Vensel, W. H., Chan, R., Lopez, R., Tanaka, C., et al. (2006a). Differential accumulation of sulfur-rich and sulfur-poor wheat flour proteins is affected by temperature and mineral nutrition during grain development. *J. Cereal Sci.* 44, 101–112. doi: 10.1016/j.jcs.2006.04.003

DuPont, F. M., Hurkman, W. J., Vensel, W. H., Tanaka, C., Kothari, K. M., Chung, O. K., et al. (2006b). Protein accumulation and composition in wheat grains: effects of mineral nutrients and high temperature. *Eur. J. Agron.* 25, 96–107. doi: 10.1016/j.eja.2006.04.003

Egli, D. B. (1998). *Seed Biology and the Yield of Grain Crops*. New York, NY: CAB International, 178.

Elbashir, A. A., Gorafi, Y. S., Tahir, I. S., Elhashimi, A. M., Abdalla, M. G., and Tsujimoto, H. (2017). Genetic variation in heat tolerance-related traits in a population of wheat multiple synthetic derivatives. *Breed. Sci.* 67, 483–492. doi: 10.1270/jsbbs.17048

Elía, M., Slafer, G. A., and Savin, R. (2018). Yield and grain weight responses to post-anthesis increases in maximum temperature under field grown wheat as modified by nitrogen supply. *Field Crops Res.* 221, 228–237. doi: 10.1016/j.fcr.2018.02.030

Farooq, M., Bramley, H., Palta, J. A., and Siddique, K. H. (2011). Heat stress in wheat during reproductive and grain-filling phases. *Crit. Rev. Plant Sci.* 30, 491–507. doi: 10.1080/07352689.2011.615687

Fischer, R., Aguilar, I., and Laing, D. (1977). Post-anthesis sink size in a high-yielding dwarf wheat: yield response to grain number. *Aust. J. Agric. Res.* 28, 165–175.

Fischer, R. A., Howe, G. N., and Ibrahim, Z. (1993). Irrigated spring wheat and timing and amount of nitrogen fertilizer. I. Grain yield and protein content. *Field Crops Res.* 33, 37–56. doi: 10.1016/0378-4290(93)90093-3

Foulkes, M. J., Sylvester-Bradley, R., and Scott, R. K. (1998). Evidence for differences between winter wheat cultivars in acquisition of soil mineral nitrogen and uptake and utilization of applied fertilizer nitrogen. *J. Agric. Sci.* 130, 29–44. doi: 10.1017/s0021859697005029

García, G. A., Serrago, R. A., Dreccer, M. F., and Miralles, D. J. (2016). Post-anthesis warm nights reduce grain weight in field-grown wheat and barley. *Field Crops Res.* 195, 50–59. doi: 10.1016/j.fcr.2016.06.002

García-Molina, M. D., and Barro, F. (2017). Characterization of changes in gluten proteins in low-gliadin transgenic wheat lines in response to application of different nitrogen regimes. *Front. Plant Sci.* 8:257. doi: 10.3389/fpls.2017.00257

Gayathri, D., and Rashmi, B. (2014). Development of Celiac disease; pathogenesis and strategies to control: a molecular approach. *J. Nutr. Food Sci.* 4, 1–10.

Gibson, P. R., Skodje, G. I., and Lundin, K. E. (2017). Non-coeliac gluten sensitivity. *J. Gastroenterol. Hepatol.* 32, 86–89. doi: 10.1111/jgh.13705

Gil-Humanes, J., Pistón, F., Giménez, M. J., Martín, A., and Barro, F. (2012). The introgression of RNAi silencing of γ-gliadins into commercial lines of bread wheat changes the mixing and technological properties of the dough. *PLoS One* 7:e45937. doi: 10.1371/journal.pone.0045937

Gil-Humanes, J., Pistón, F., Hernando, A., Alvarez, J. B., Shewry, P. R., and Barro, F. (2008). Silencing of γ-gliadins by RNA interference (RNAi) in bread wheat. *J. Cereal Sci.* 48, 565–568.

Gil-Humanes, J., Pistón, F., Shewry, P. R., Tosi, P., and Barro, F. (2011). Suppression of gliadins results in altered protein body morphology in wheat. *J. Exp. Bot.* 62, 4203–4213. doi: 10.1093/jxb/err119

Gil-Humanes, J., Pistón, F., Tollefsen, S., Sollid, L. M., and Barro, F. (2010). Effective shutdown in the expression of celiac disease-related wheat gliadin T-cell epitopes by RNA interference. *Proc. Natl. Acad. Sci. U.S.A.* 107, 17023–17028. doi: 10.1073/pnas.1007773107

Graybosch, R. A., Peterson, C., Baenziger, P. S., and Shelton, D. (1995). Environmental modification of hard red winter wheat flour protein composition. *J. Cereal Sci.* 22, 45–51. doi: 10.1016/s0733-5210(05)80006-2

Hawker, J., and Jenner, C. (1993). High temperature affects the activity of enzymes in the committed pathway of starch synthesis in developing wheat endosperm. *Funct. Plant Biol.* 20, 197–209.

Hurkman, W. J., Tanaka, C. K., Vensel, W. H., Thilmony, R., and Altenbach, S. B. (2013). Comparative proteomic analysis of the effect of temperature and fertilizer on gliadin and glutenin accumulation in the developing endosperm and flour from *Triticum aestivum* L. cv. Butte 86. *Proteome Sci.* 11:8. doi: 10.1186/1477-5956-11-8

Inomata, N. (2009). Wheat allergy. *Curr. Opin. Allergy Clin. Immunol.* 9, 238–243. doi: 10.1097/ACI.0b013e32832aa5bc

Jenner, C., Ugalde, T., and Aspinall, D. (1991). The physiology of starch and protein deposition in the endosperm of wheat. *Funct. Plant Biol.* 18, 211–226.

Johansson, E., Malik, A. H., Hussain, A., Rasheed, F., Newson, W. R., Plivelic, T., et al. (2013). Wheat gluten polymer structures: the impact of genotype, environment, and processing on their functionality in various applications. *Cereal Chem.* 90, 367–376. doi: 10.1094/cchem-08-12-0105-fi

Juhász, A., Belova, T., Florides, C. G., Maulis, C., Fischer, I., Gell, G., et al. (2018). Genome mapping of seed-borne allergens and immunoresponsive proteins in wheat. *Sci. Adv.* 4:eaar8602. doi: 10.1126/sciadv.aar8602

Lafiandra, D., and Kasarda, D. D. (1985). One-and two-dimensional (two-pH) polyacrylamide gel electrophoresis in a single gel: separation of wheat proteins. *Cereal Chem.* 62, 314–319.

Lexhaller, B., Colgrave, M. L., and Scherf, K. A. (2019). Characterization and relative quantitation of wheat, rye, and barley gluten protein types by liquid chromatography–tandem mass spectrometry. *Front. Plant Sci.* 10:1530. doi: 10.3389/fpls.2019.01530

Liu, P., Guo, W., Jiang, Z., Pu, H., Feng, C., Zhu, X., et al. (2011). Effects of high temperature after anthesis on starch granules in grains of wheat (*Triticum aestivum* L.). *J. Agric. Sci.* 149, 159–169. doi: 10.1017/s0021859610001024

Ludvigsson, J. F., Leffler, D. A., Bai, J. C., Biagi, F., Fasano, A., Green, P. H., et al. (2013). The Oslo definitions for coeliac disease and related terms. *Gut* 62, 43–52. doi: 10.1136/gutjnl-2011-301346

Meehl, G. A., and Tebaldi, C. (2004). More intense, more frequent, and longer lasting heat waves in the 21st century. *Science* 305, 994–997. doi: 10.1126/science.1098704

Merrill, A., and Watt, B. (1973). *Energy Value of Foods: Basis and Derivation*. Washington, DC: U.S. Department of Agriculture.

Molberg, Ø., Solheim Flaete, N., Jensen, T., Lundin, K. E., Arentz-Hansen, H., Anderson, O. D., et al. (2003). Intestinal T-cell responses to high-molecular-weight glutenins in celiac disease. *Gastroenterology* 125, 337–344. doi: 10.1016/s0016-5085(03)00890-4

Morita, E., Kunie, K., and Matsuo, H. (2007). Food-dependent exercise-induced anaphylaxis. *J. Dermatol. Sci.* 47, 109–117.

Morita, E., Matsuo, H., Chinuki, Y., Takahashi, H., and Dahlström, J. (2009). Food-dependent exercise-induced anaphylaxis—importance of omega-5 gliadin and HMW-glutenin as causative antigens for wheat-dependent exercise-induced anaphylaxis. *Allergol. Int.* 58, 493–498. doi: 10.2332/allergolint.09-RAI-0125

Naeem, H., Paulon, D., Irmak, S., and MacRitchie, F. (2012). Developmental and environmental effects on the assembly of glutenin polymers and the impact on grain quality of wheat. *J. Cereal Sci.* 56, 51–57. doi: 10.1016/j.jcs.2011.10.014

Nuttall, J. G., Barlow, K. M., Delahunty, A. J., Christy, B. P., and O'Leary, G. J. (2018). Acute high temperature response in wheat. *Agron. J.* 110, 1296–1308. doi: 10.2134/agronj2017.07.0392

Ozuna, C. V., Iehisa, J. C., Giménez, M. J., Alvarez, J. B., Sousa, C., and Barro, F. (2015). Diversification of the celiac disease α-gliadin complex in wheat: a 33-mer peptide with six overlapping epitopes, evolved following polyploidization. *Plant J.* 82, 794–805. doi: 10.1111/tpj.12851

Palacin, A., Quirce, S., Armentia, A., Fernández-Nieto, M., Pacios, L. F., Asensio, T., et al. (2007). Wheat lipid transfer protein is a major allergen associated with baker's asthma. *J. Allergy Clin. Immunol.* 120, 1132–1138. doi: 10.1016/j.jaci.2007.07.008

Pechanek, U., Karger, A., Gröger, S., Charvat, B., Schöggl, G., and Lelley, T. (1997). Effect of nitrogen fertilization on quantity of flour protein components, dough properties, and breadmaking quality of wheat. *Cereal Chem.* 74, 800–805. doi: 10.1094/cchem.1997.74.6.800

Pistón, F., Gil-Humanes, J., and Barro, F. (2013). Integration of promoters, inverted repeat sequences and proteomic data into a model for high silencing efficiency of coeliac disease related gliadins in bread wheat. *BMC Plant Biol.* 13:136. doi: 10.1186/1471-2229-13-136

Pistón, F., Gil-Humanes, J., Rodríguez-Quijano, M., and Barro, F. (2011). Down-regulating γ-gliadins in bread wheat leads to non-specific increases in other gluten proteins and has no major effect on dough gluten strength. *PLoS One* 6:e24754. doi: 10.1371/journal.pone.0024754

R Core Team (2018). *A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Saint Pierre, C., Peterson, C. J., Ross, A. S., Ohm, J. B., Verhoeven, M. C., Larson, M., et al. (2008). Winter wheat genotypes under different levels of nitrogen and water stress: changes in grain protein composition. *J. Cereal Sci.* 47, 407–416. doi: 10.1016/j.jcs.2007.05.007

Savin, R., Calderini, D. F., Slafer, G. A., and Abeledo, L. G. (1999). Final grain weight in wheat as affected by short periods of high temperature during pre-and post-anthesis under field conditions. *Funct. Plant Biol.* 26, 453–458.

Savin, R., Stone, P. J., and Nicolas, M. E. (1996). Responses of grain growth and malting quality of barley to short periods of high temperature in field studies using portable chambers. *Aust. J. Agric. Res.* 47, 465–477.

Shan, L., Molberg, Ø., Parrot, I., Hausch, F., Filiz, F., Gray, G. M., et al. (2002). Structural basis for gluten intolerance in celiac sprue. *Science* 297, 2275–2279. doi: 10.1126/science.1074129

Shewry, P. R. (2009). Wheat. *J. Exp. Bot.* 60, 1537–1553. doi: 10.1093/jxb/erp058

Shewry, P. R. (2019). What is gluten-why is it special? *Front. Nutr.* 6:101. doi: 10.3389/fnut.2019.00101

Shewry, P. R., and Halford, N. G. (2002). Cereal seed storage proteins: structures, properties and role in grain utilization. *J. Exp. Bot.* 53, 947–958. doi: 10.1093/jexbot/53.370.947

Shewry, P. R., Halford, N. G., and Lafiandra, D. (2003). Genetics of wheat gluten proteins. *Adv. Genet.* 49, 111–184. doi: 10.1016/s0065-2660(03)01003-4

Slafer, G. A., and Savin, R. (2018). Can N management affect the magnitude of yield loss due to heat waves in wheat and maize? *Curr. Opin. Plant Biol.* 45, 276–283. doi: 10.1016/j.pbi.2018.07.009

Sollid, L. M. (2002). Coeliac disease: dissecting a complex inflammatory disorder. *Nat. Rev. Immunol.* 2, 647–655. doi: 10.1038/nri885

Sollid, L. M., Qiao, S.-W., Anderson, R. P., Gianfrani, C., and Koning, F. (2012). Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenetics* 64, 455–460. doi: 10.1007/s00251-012-0599-z

Spiertz, J., Hamer, R., Xu, H., Primo-Martin, C., Don, C., and Van Der Putten, P. (2006). Heat stress in wheat (*Triticum aestivum* L.): effects on grain growth and quality traits. *Eur. J. Agron.* 25, 89–95. doi: 10.1016/j.eja.2006.04.012

Stone, P. (2001). "The effects of heat stress on cereal yield and quality," in *Crop Responses and Adaptations to Temperature Stress*, ed. A. S. Basra (New York, NY: Food Products Press), 243–291.

Triboï, E., Martre, P., and Triboï-Blondel, A. M. (2003). Environmentally-induced changes in protein composition in developing grains of wheat are related to changes in total protein content. *J. Exp. Bot.* 54, 1731–1742. doi: 10.1093/jxb/erg183

Trier, J. S. (1998). Diagnosis of celiac sprue. *Gastroenterology* 115, 211–216. doi: 10.1016/s0016-5085(98)70383-x

Tye-Din, J. A., Stewart, J. A., Dromey, J. A., Beissbarth, T., van Heel, D. A., Tatham, A., et al. (2010). Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Sci. Transl. Med.* 2:41ra51. doi: 10.1126/scitranslmed.3001012

van den Broeck, H. C., America, A. H., Smulders, M. J., Bosch, D., Hamer, R. J., Gilissen, L. J., et al. (2009). A modified extraction protocol enables detection and quantification of celiac disease-related gluten proteins from wheat. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 877, 975–982. doi: 10.1016/j.jchromb.2009.02.035

Ventura, A., Magazzù, G., and Greco, L. (1999). Duration of exposure to gluten and risk for autoimmune disorders in patients with celiac disease. *Gastroenterology* 117, 297–303. doi: 10.1053/gast.1999.0029900297

Wardlaw, I. F., Blumenthal, C., Larroque, O., and Wrigley, C. W. (2002). Contrasting effects of chronic heat stress and heat shock on kernel weight and flour quality in wheat. *Funct. Plant Biol.* 29, 25–34.

Wen, S., Wen, N., Pang, J., Langen, G., Brew-Appiah, R. A., Mejias, J. H., et al. (2012). Structural genes of wheat and barley 5-methylcytosine DNA glycosylases and their potential applications for human health. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20543–20548. doi: 10.1073/pnas.1217927109

Wieser, H., Antes, S., and Seilmeier, W. (1998). Quantitative determination of gluten protein types in wheat flour by reversed-phase high-performance liquid chromatography. *Cereal Chem.* 75, 644–650. doi: 10.1371/journal.pone.0192804

Xue, C., Rossmann, A., Schuster, R., Koehler, P., and Mühling, K.-H. (2016). Split nitrogen application improves wheat baking quality by influencing protein composition rather than concentration. *Front. Plant Sci.* 7:738. doi: 10.3389/fpls.2016.00738

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 07 July 2020
doi: 10.3389/fpls.2020.00906

Further Steps Toward the Development of Gluten Reference Materials – Wheat Flours or Protein Isolates?

Eszter Schall[1], Katharina A. Scherf[2,3], Zsuzsanna Bugyi[1], Kitti Török[1], Peter Koehler[4], Regine Schoenlechner[5] and Sándor Tömösközi[1]*

[1] Research Group of Cereal Science and Food Quality, Department of Applied Biotechnology and Food Science, Budapest University of Technology and Economics, Budapest, Hungary, [2] Leibniz-Institute for Food Systems Biology at the Technical University of Munich, Freising, Germany, [3] Department of Bioactive and Functional Food Chemistry, Institute of Applied Biosciences, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, [4] Biotask AG, Esslingen am Neckar, Germany, [5] Department of Food Science and Technology, University of Natural Resources and Life Sciences, Vienna, Austria

OPEN ACCESS

Edited by:
Nicolas L. Taylor,
The University of Western Australia, Australia

Reviewed by:
Jose Valero Galvan,
Universidad Autónoma de Ciudad Juárez, Mexico
Maryke T. Labuschagne,
University of the Free State, South Africa
Crispin Alexander Howitt,
Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia

***Correspondence:**
Katharina A. Scherf
katharina.scherf@kit.edu

Specialty section:
This article was submitted to Plant Proteomics, a section of the journal Frontiers in Plant Science

Received: 08 August 2019
Accepted: 03 June 2020
Published: 07 July 2020

Citation:
Schall E, Scherf KA, Bugyi Z, Török K, Koehler P, Schoenlechner R and Tömösközi S (2020) Further Steps Toward the Development of Gluten Reference Materials – Wheat Flours or Protein Isolates? Front. Plant Sci. 11:906. doi: 10.3389/fpls.2020.00906

Celiac disease is a gluten-induced hypersensitivity reaction that requires a lifelong gluten-free diet. Gluten-free foods must not contain more than 20 mg/kg gluten as laid down by Codex Alimentarius. Measuring the presence of gluten with routine immunoanalytical methods in food is a serious challenge as many factors affect accurate determination. Comparability of the results obtained with different methods and method validation are hindered by the lack of a widely accepted reference material (RM). The core questions of RM development from wheat are the number of cultivars to be included and the format of gluten (i.e., flour, gluten, or gliadin isolates) to be applied. Therefore, the aim of our work was to produce an appropriate gluten RM from wheat. For this, five previously selected wheat cultivars and their blend were used to produce flours, gluten and gliadin isolates under laboratory conditions. Protein content, protein composition and responses to different ELISA methods were compared and widely evaluated in our study. The protein contents of the flours were 12.1–18.7%, those of the gluten isolates 93.8–97.4% and those of the gliadin isolates 72.7–101.9%. The gluten and gliadin isolates had similar protein profiles as the source flours. By comparing the different wheat cultivars and their protein isolates, we found that the isolation had a smaller effect on protein composition than genetic variability. The choice of a blend would be more suitable for the production of a RM in case of flours and also isolates. The immunoanalytical results showed that the isolation had an effect on the analytical results, but its extent depended on the ELISA method. The use of flour would be more applicable in this regard, but handling of the material and long-term stability should also be considered in the final decision of gluten RM production.

Keywords: celiac disease, gliadin, gluten, reference material, ELISA, wheat flour, protein isolates

Abbreviations: ANOVA, analysis of variance; ELISA, enzyme-linked immunosorbent assay; HMW-GS, high-molecular-weight glutenin subunits; HPLC, high-performance liquid chromatography; LC-MS, liquid chromatography mass spectrometry; LMW-GS, low-molecular-weight glutenin subunits; LSD, least significant difference; RM, reference material; RP, reversed-phase; SE, size-exclusion; TFA, trifluoroacetic acid; ωb, glutenin-bound ω-gliadin.

INTRODUCTION

Wheat, rye, and barley are widely used cereals in the food industry because of their nutritional quality and beneficial technological properties (Shewry and Tatham, 2016). However, their consumption can cause health problems for some people. One of the most common cereal-induced hypersensitivity reactions is celiac disease, which is a disorder with an autoimmune component associated with serious damage of the small intestinal mucosa. The triggers of celiac disease are the storage proteins of gluten-containing cereals (Leonard et al., 2017). Since patients can only be treated with a gluten-free diet, the availability of gluten-free foods is essential. According to the Codex Alimentarius, foods can be labeled gluten-free if the gluten content does not exceed 20 mg/kg (Codex Stan 118-1979, 2015). Thus, methods for the reliable quantitation of gluten in (gluten-free) foods are needed. One of the analytical problems is that gluten is not a homogeneous, properly defined component, but a mixture of heterogeneous proteins with different physico-chemical properties (Tatham and Shewry, 2012). According to the classical Osborne fractionation, gluten proteins from wheat, rye, and barley can be divided into alcohol-soluble prolamins and glutelins which are not soluble in aqueous alcohol solutions (Osborne, 1907; Koehler and Wieser, 2013). Cereal prolamins have trivial names: gliadins for wheat, secalins for rye and hordeins for barley. However, only wheat glutelin has a trivial name, which is glutenin (Wieser and Koehler, 2009; Koehler and Wieser, 2013). Cereal proteins can be further classified based on their size and electrophoretic mobility. Wheat gluten proteins are subdivided into α-/β-gliadins (QPQPF), γ-gliadins (QQPQQPFP), ω1,2-gliadins (QPQQPFP), and ω5-gliadins (QQQPF), low-molecular-weight glutenin subunits (LMW-GS) (QQPPFS) and high-molecular-weight glutenin subunits (HMW-GS) (QQPGQG, YYPTSP) (Koehler and Wieser, 2013). The typical repetitive amino acid sequences (epitopes) of wheat gluten proteins, examples of which are given in parentheses, are involved in the induction of celiac disease. These sequences have high contents of proline and glutamine, which make them resistant to protein-degrading digestive enzymes (Brouns et al., 2019). Most of the reactive epitopes have been reported in the gliadin fraction and, e.g., α-gliadin contains a peptide with a length of 33 amino acids that was shown to be highly celiac disease-active (Shan et al., 2002). Further studies have shown that other gluten protein types also contain celiac disease-active epitopes (Lexhaller et al., 2019; Sollid et al., 2020).

Several analytical methods based on different mechanisms are available for gluten quantitation (Scherf and Poms, 2016). DNA-based techniques, such as polymerase chain reaction (PCR), sensitively detect DNA segments coding for gluten proteins, but have the disadvantage that gluten proteins are not directly determined (Mujico et al., 2011; Codex Stan 118-1979, 2015). The number of studies using liquid chromatography mass spectrometry (LC-MS) for gluten quantitation has increased, as they are capable of determining all gluten protein types from gluten-containing cereals. However, routine application of LC-MS is limited because of the high level of expertise required and the cost of instrumentation (Schalk et al., 2018). The most common method used in routine analysis is the enzyme-linked immunosorbent assay (ELISA) based on the immunochemical reaction between epitope(s) within gluten proteins and an epitope-specific antibody. The advantages of the method are its relatively easy implementation and the specific and sensitive detection (Diaz-Amigo and Popping, 2013; Bruins Slot et al., 2016). The different ELISA methods available on the market offer various solutions for sample preparation, test format (sandwich or competitive), type of antibody (monoclonal or polyclonal), specificity of the antibody toward different epitopes and calibration material (Lexhaller et al., 2016; Scherf and Poms, 2016). This is why several studies have shown that different ELISA kits give different results when the same samples are analyzed (Bugyi et al., 2013; Rzychon et al., 2017; Scherf, 2017).

The accurate determination of the gluten content of food is a challenge because the identification of factors affecting the analytical results is difficult. One step toward harmonization of analytical methods will be the availability of a universally accepted gluten reference material (RM) (Poms, 2006).

According to the ISO guide 30, a RM is a material that is sufficiently homogenous and stable with respect to one or more specific properties, which has been established to be fit for its intended use in a measurement. Its production must be reproducible, and it should be easy to handle (Diaz-Amigo and Popping, 2013; Scherf and Poms, 2016). Further, a certified RM is a RM that is characterized by a metrologically valid procedure for one or more specified properties, accompanied by a certificate providing the value of the specified property, its associated uncertainty, and a statement of metrological traceability. A certified RM would provide an opportunity to support method validation and to identify the factors influencing gluten analysis. The most widely used standard-like material in gluten analysis is a gliadin isolate called Prolamin Working Group (PWG)-gliadin (van Eckert et al., 2006; van Eckert et al., 2010). The advantage of the material is its high purity, good solubility and detailed characterization. PWG-gliadin was proposed for approval as a certified RM, but it did not meet some of the RM requirements for certification, such as reproducibility of production (Diaz-Amigo and Popping, 2013). Consequently, there is still a need for a gluten RM, but there are a number of questions about its composition (one cultivar or a blend for each species of wheat, rye, and barley) and type (flour or protein isolate) that need to be investigated and answered. In order to provide solid foundations for a comprehensive RM for gluten from wheat, rye, and barley, we started our investigations with wheat, because it is by far the most widely used species of the three. All learnings from our studies on wheat will enable us to easily and efficiently transfer these to the development of RM for rye and barley, because a universal gluten RM should certainly include the relevant proteins of all three species.

Within an international cooperation, the factors affecting gluten analysis (such as genetic and environmental variability) were investigated with the aim to design a gluten RM candidate. For this purpose, 23 wheat cultivars collected from different geographical locations around the world were examined and characterized in detail (Hajas et al., 2018). Based on the results of this study, five cultivars were selected and investigated for the

magnitude of the analytical error of ELISA methods resulting from the use of one cultivar or their blend (Schall et al., 2020). Another major issue of RM production is the decision whether to use flour, a gluten isolate or a gliadin isolate. The flour represents gluten contamination most realistically and its production is relatively simple, but it contains components (e.g., lipids) causing instability during storage (Wang and Flores, 1999). The storage stability and handling of the isolates could be more advantageous, but the protein composition may change during isolation which could affect the analytical results (Diaz-Amigo and Popping, 2012). The gliadin isolate has the advantage of being completely soluble in specific solvents, but it does not contain all protein types that induce celiac disease.

In this work we investigated the effect that the production of protein isolates (gluten or gliadin) from wheat flour has on the amount and composition of proteins compared to the flour. Furthermore, the suitability of the RM material candidates for different analytical methods for gluten quantitation were evaluated to enable the selection of a proper RM. By examining the blend of the five wheat cultivars, the use of individual cultivars and their blend not just as flour but also as protein isolates was possible.

MATERIALS AND METHODS

Wheat Samples

Five wheat (*Triticum aestivum* L.) cultivars were selected in line with a set of selection criteria described in our previous study (Hajas et al., 2018) and collected from the harvest year of 2016 for this work: Akteur (Germany); Carberry (Canada); Mv Magvas (Hungary); Yitpi (Australia), and Yumai-34 (China).

Production of Wheat Flours

The moisture content of the grains was determined by an InfratecTM 1241 Grain Analyzer (Foss Tecator AB, Höganäs, Sweden). The wheat samples were conditioned prior to milling according to Hungarian Standard MSZ 6367-9:1989, 1989. The tempered kernels were milled on a laboratory mill (FQC 109, Metefém, Budapest, Hungary). The whole-meal was sieved on a 250 μm sieve for 20 min (AS 200 basic, Retsch GmbH, Haan, Germany). The blend of the five cultivars was prepared by mixing equal amounts (80 g each) of grains from the single cultivars by shaking in a closed container manually for 10 min before milling. The homogeneity of the blend was confirmed later by chemical composition data in section "Comparison of the Different Gluten and Gliadin Isolates From Individual Cultivars and Their Blend."

Production of Gluten and Gliadin Isolates

Gluten and gliadin isolates were prepared based on the standard for wet gluten production and the study reporting the production of PWG-gliadin (AACC Method 38-12.02, 2000; van Eckert et al., 2006). The Glutomatic System (Perten Instruments, Hägersten, Sweden) was used for the removal of albumins and globulins with 0.4 M NaCl solution from white flours of each cultivar and the blend. The resulting gluten was further washed for another 10 min with tap water to remove residual starch and salt. The gluten was freeze-dried for 24 h (Christ Alpha 1-4 LOC-1M, Martin Christ Gefriertrocknungsanlagen GmbH, Osterode am Harz, Germany). Then the dry gluten was ground with knife mills for 3 × 10 s at 7000 rpm (Grindomix GM200, RETSCH GmbH, Haan, Germany). One-third of the total amount (ca. 10 g) of isolated gluten was collected separately for protein content and composition analysis. The gliadins were then extracted three times with 270 mL 60% (v/v) ethanol from the two-thirds of the total amount of dry gluten. The suspension was stirred for 30 min with a magnetic stirrer followed by centrifugation for 15 min at 4500 *g* (Labofuge 400R, Heraeus, Kendro Laboratory Products, Germany). The supernatants were combined and freeze-dried. In order to verify the reproducibility of the production, two independent batches of samples were prepared from milling to isolation on the arbitrarily selected Akteur cultivar.

Determination of Crude Protein Content

The nitrogen content of the flours and the isolates was determined by a Leco FP 528 nitrogen analyzer (Leco Corporation, St. Joseph, United States) in duplicates following the MSZ En Iso 16634-2:2016, 2016. The nitrogen content was multiplied by 5.7 to obtain the crude protein content.

Calculation of the Relative Amount of Isolates

The following calculations were used to determine the relative amount of materials obtained during the isolation:

Amount of gluten proteins relative to flour proteins (%):

$$\frac{amount\ of\ gluten\ proteins\ extracted\ from\ flour\left(\frac{g}{100g}\right)}{protein\ content\ of\ flour\ \left(\frac{g}{100g}\right)} \times 100$$

Amount of gliadin proteins relative to gluten proteins (%):

$$\frac{amount\ of\ gliadin\ proteins\ extracted\ from\ flour\left(\frac{g}{100g}\right)}{amount\ of\ gluten\ proteins\ extracted\ from\ flour\ \left(\frac{g}{100g}\right)} \times 100$$

Protein Characterization by SE-HPLC

Protein extracts were prepared for size-exclusion high-performance liquid chromatography (SE-HPLC) analyses according to Batey et al. (1991) and Gupta et al. (1993) with minor modifications. Acetonitrile (50%, v/v) containing 0.1% (v/v) trifluoroacetic acid (TFA) was used as the extraction solvent. Wheat flour (15 mg)/gluten isolate (1.5 mg)/gliadin isolate (1.5 mg) was suspended in 1 mL of the extraction solvent and shaken (1,500 rpm, 30 min, 20–22°C) followed by centrifugation (4,500 × g, 20 min, 20°C). The supernatant was collected (extractable protein fraction). The remaining pellet was extracted with 1 mL of the same extraction solution using sonication for 40 s with an amplitude of 90%. Then, samples were shaken (1,500 × rpm, 30 min, 20–22°C) and centrifuged (4,500 *g*, 20 min, 20°C) to obtain a supernatant (unextractable protein

fraction). All supernatants were filtered (Minisart®, 15/0.45 RC, Sartorius AG, Goettingen, Germany) before SE-HPLC analysis. The extractions were done in duplicate for each flour sample. The conditions for the SE-HPLC analyses were the following: instrument: PerkinElmer Series 200 HPLC with TotalChrom Navigator v6.2.1 (PerkinElmer Inc., Shelton, CT, United States); column: BioSep-SEC-s4000 (particle size 5 μm, pore size 50 nm, 300 × 7.8 mm, separation range for proteins 15,000–1,500,000, Phenomenex, Torrance, CA, United States); temperature: 25°C; injection volume: 20 μL; elution solvents: 50% (v/v) acetonitrile containing 0.1% (v/v) TFA; flow rate: 1 mL/min; running time: 20 min, detection: UV absorbance at 214 nm. After each run, the column was equilibrated with the elution solvent for 1 min. The chromatograms of the extractable and unextractable proteins were divided into three sections: the proportion of polymeric, monomeric and albumin/globulin fractions to "total extracted" protein were calculated from the peak areas as percentage of the total peak area.

Protein Characterization by RP-HPLC

Wheat flours (100 mg) were extracted sequentially according to the modified Osborne procedure (Wieser et al., 1998) by magnetic stirring with salt solution (extraction of albumins/globulins), followed by 60% (v/v) ethanol solution (extraction of gliadins), and glutelin extraction solution [containing 1-propanol, tris(hydroxymethyl)aminomethane hydrochloride, dithiothreitol and urea for the extraction of glutenins]. All suspensions were centrifuged (3550 *g*, 25 min, 20°C) and the supernatant filtered (Whatman™ Spartan 13/0.45 RC, GE Healthcare, Freiburg, Germany). The gluten isolates (20 mg) were extracted with 60% (v/v) ethanol solution and glutelin extraction solution in the same way as the flours. The gliadin isolates (5 mg) were extracted with 60% (v/v) ethanol solution in the same way as the flours. The extractions were done in triplicate for each sample. The conditions for reversed-phase high-performance liquid chromatography (RP-HPLC) analyses were the following: the instrument was Jasco XLC with Jasco Chrompass Chromatography Data System (Jasco, Pfungstadt, Germany); column: Acclaim™ 300 C_{18} (particle size 3 mm, pore size 30 nm, 2.1 × 150 mm, Thermo Fisher Scientific, Braunschweig, Germany); temperature: 60°C; elution solvents: TFA (0.1%, v/v) in water (A) and TFA (0.1%, v/v) in acetonitrile (B); linear gradient: 0 min 0% B, 0.5 min 20% B, 7 min 60% B, 7.1–11 min 90% B, 11.1–17 min 0% B for albumins/globulins; 0 min 0% B, 0.5 min 24% B, 20 min 56% B, 20.1–24.1 min 90% B, 24.2–30 min 0% B for gliadins and glutenins; flow rate: 0.2 mL/min; injection volume: 20 mL for albumins/globulins and glutenins, 10 mL for gliadins; detection: UV absorbance at 210 nm. The protein contents of the extracts were calculated from the absorbance areas using 5, 10, 15, and 20 μL of a PWG-gliadin solution (2.5 mg/mL in 60% ethanol) (van Eckert et al., 2006) as calibration reference. The contents of ω5-, ω1,2-, α-, and γ-gliadins were calculated from the absorbance area of each gliadin type relative to the total gliadin content, as were those of glutenin-bound ω-gliadins (ωb-gliadins), HMW-GS and LMW-GS relative to the total glutenin content.

Gliadin/Gluten Quantitation With ELISA Methods

The gliadin/gluten quantitation was performed with two commercially available ELISA test kits: the AgraQuant Gluten G12 Assay (COKAL0200, Romer Labs, Tulln, Austria) and the RIDASCREEN Gliadin Assay (R7001, R-Biopharm, Darmstadt, Germany). They apply different antibodies (monoclonal G12 and monoclonal R5, respectively) and are calibrated differently (vital wheat gluten extract and PWG-gliadin, respectively). ELISA procedures were carried out according to the kit instructions. Three independent extractions were performed for each sample. The absorbances were determined using a microplate reader (iMark™ Microplate Absorbance Reader, Bio-Rad, Hercules, CA, United States). The gliadin/gluten concentration was calculated from the absorbance values by the Bio-Rad Microplate Manager 6 software (Bio-Rad, Hercules, CA, United States) using the curve fit suggested by the manufacturer. The ELISA test kits used for analysis were randomly coded with capital letters (A and B) in section "Results and Discussion."

Statistical Analysis

The analytical results were statistically evaluated with the investigation of means, standard deviations, one-sample *t*-test and analysis of variance (ANOVA) with Fisher's Least Significant Difference (LSD) *post hoc* test at a confidence level of 0.95 using Statistica 13 software (TIBCO Software Inc., Palo Alto, CA, United States).

RESULTS AND DISCUSSION

Investigation of the Reproducibility of Flour, Gluten, and Gliadin Production

An important aspect of choosing a proper RM is the reproducibility of its production. Therefore, we produced two independent batches from the cultivar "Akteur" in parallel with a method based on preliminary experiments and all protein parameters of each type of sample (flour – gluten isolate – gliadin isolate) were investigated. The crude protein content of the laboratory milled batch 1 and 2 flour was 14.6 and 13.9%, respectively (**Table 1**). The difference was significant, but this variation was smaller than the differences between the five cultivars used in this study. Gluten production can be affected by a number of factors that determine the final protein content and yield (Van Der Borght et al., 2005). Depending on the preparation, it may contain starch, lipids and fibers. The amount of starch varies, but with extensive washing a significant reduction of starch embedded in the protein matrix can be observed. However, starch and fiber become entrapped in the cohesive matrix of the protein and become more difficult to remove as the protein content increases (Saulnier et al., 1997; Day et al., 2006). The non-polar lipids of wheat flour interact with the hydrophobic regions of gluten proteins during the washing process, not allowing complete extraction of lipids (Day et al., 2006). A higher protein content could mean greater purity of the gluten isolate, but there may be small amounts of soluble proteins

TABLE 1 | Crude protein content of flours, gluten and gliadin isolates; amount of gluten proteins obtained from flour proteins and amount of gliadin proteins obtained from gluten proteins (all values are expressed on dry matter basis).

Sample	Parameter				
	Crude protein content of flours (%)[abc]	Crude protein content of gluten isolates (%)[abc]	Amount of gluten proteins relative to flour proteins (%)	Crude protein content of gliadin isolates (%)[abc]	Amount of gliadin proteins relative to gluten proteins (%)
Akteur – batch 1	$14.6^{+} \pm 0.0$	96.0 ± 0.4	63.0	$96.4^{+} \pm 0.2$	36.5
Akteur – batch 2	$13.9^{D} \pm 0.0$	$97.1^{A} \pm 0.3$	75.3	$76.0^{E} \pm 1.0$	34.6
Carberry	$18.7^{A} \pm 0.1$	$95.3^{B} \pm 0.0$	76.9	$87.0^{B} \pm 0.8$	34.0
Mv Magvas	$12.1^{E} \pm 0.1$	$97.4^{A} \pm 0.0$	75.5	$72.7^{F} \pm 0.9$	36.9
Yitpi	$16.6^{B} \pm 0.1$	$93.8^{D} \pm 0.2$	73.4	$101.9^{A} \pm 1.5$	58.2
Yumai-34	$16.7^{B} \pm 0.0$	$94.5^{C} \pm 0.3$	76.8	$83.3^{C} \pm 1.6$	52.8
Blend	$15.4^{C} \pm 0.1$	$95.5^{B} \pm 0.1$	78.5	$79.6^{D} \pm 1.5$	28.8
Mean of the five cultivars	15.6	95.6	75.6	84.2	43.3
PWG-gliadin	–	–	–	92.8 ± 0.8	–

[a]Within each column, the mean value of batch 1 Akteur samples marked with a plus sign is significantly different from the mean value of batch 2 Akteur samples ($p < 0.05$; one-way ANOVA). [b]Within each column, the measured value of the blend marked with an asterisk is significantly different from the calculated mean of the five cultivars ($p < 0.05$; one-sample t-test). [c]Within each column, mean values marked with different capital letters are significantly different ($p < 0.05$; factorial ANOVA, Fisher's LSD).

trapped in the gluten matrix as well (Ortolan and Steel, 2017). Therefore, protein content alone is not sufficient to determine gluten quality, which is why the protein profile of isolates should be examined. The crude protein content of the batch 1 gluten isolate was 96.6% while batch 2 had 97.1%, revealing that the isolates had high purity (**Table 1**). The amount of gluten proteins relative to the amount of proteins from flours by weight was 63% for batch 1 and 75.3% for batch 2 (**Table 1**). With our laboratory method, we were able to produce a gluten isolate with a high and constant protein content, but the extra manual washing step in our method could affect the amount of soluble proteins and starch within the gluten matrix and could also mean the loss of gluten proteins. In case of the production of our gliadin isolate, the non-protein components were probably less involved. However, the gluten proteins themselves form a complex system, which makes it difficult to produce a constant quality gliadin isolate. The crude protein content of the batch 1 gliadin isolate was 96.4% while it was only 76% for batch 2 which was a more substantial divergence than between the different batches of gluten isolates (**Table 1**). Interestingly, in both cases similar amounts of gliadin proteins were obtained from gluten isolates: 36.5% for batch 1 and 34.6% for batch 2 (**Table 1**). In the case of PWG-gliadin, the most widely used gliadin standard, a large amount of good quality material was produced, which was tested by several methods, but yield data are not available (van Eckert et al., 2006). Rallabhandi et al. (2015) produced prolamins including gliadin isolates in laboratory conditions. Their gliadin material contained 68% proteins with a yield of 1.44 g/100 g flour (Rallabhandi et al., 2015). The yield and protein content of gliadin isolates may also depend on the methods. Publications for the production of prolamins focus mainly on matching with the source flour, so comparing the protein profile of gliadin isolates with flours and gluten isolates is essential (van Eckert et al., 2006; Huang et al., 2017; Schalk et al., 2017).

The protein composition of the two batches of materials separated by SE-HPLC is shown in **Figure 1**. It can be clearly seen that the protein profiles of the two batches of flours were quite similar both in the soluble and insoluble protein fractions as in the case of gluten isolates. The only conspicuous difference was the higher albumin/globulin peak in the batch 1 gluten isolate. The similarity between the two batches both in flours and gluten isolates was also supported by the distribution of monomeric and polymeric proteins as there were no significant differences between the two batches (**Table 2**). In case of gliadin isolates, the protein profiles between the two batches in the SE-HPLC chromatograms showed differences, as the presence of polymer-like proteins was observed in batch 1 (**Figure 1**). The insoluble fractions of gliadin isolates – with the expectation of a small peak – did not show any higher molecular weight proteins. Glutenin proteins obtained during the isolation were also analyzed in each case (results are not shown), and the two glutenins produced in parallel were similar, which appeared mostly in the insoluble fraction. It is conceivable that the problem with those higher molecular weight proteins appearing in gliadins could be the poor solubility in the solvent used in the SE-HPLC method. The peaks typical for monomeric proteins (between 7.5 and 9.5 min) were similar in the two batches of gliadin isolates. A characteristic value of monomeric proteins could be the ratio of the two peaks appearing in the chromatograms of the soluble fraction, which was 4.98 for batch 1 and 5.92 for batch 2.

The composition of the different protein types within the monomeric and polymeric protein profiles determined by RP-HPLC are shown in **Table 3**. The proportion of different gluten protein types showed very similar values in the two flour batches and there were no significant differences in the proportion of total gliadin, ω1,2- and α-gliadin contents between the two batches. The gliadin/glutenin ratio of batch 1 was 1.7 and 1.6 in batch 2. The two batches of gluten isolates also showed great similarities and the difference was not significant in the ratio of α- and γ-gliadins (**Table 3**). The gliadin/glutenin ratio of batch 1 was 1.2 and 1.1 in batch 2. The distribution of different gliadin types was comparable in the two batches of gliadin isolates (**Table 4**).

TABLE 2 | Ratio of monomeric and polymeric protein fractions of flours, gluten and gliadin isolates determined by SE-HPLC (monomeric and polymeric protein contents are expressed as percentage of total monomeric and polymeric extract; all values are expressed on dry matter basis).

Sample		Parameter			
		Monomeric (%)[abc]	Polymeric (%)[abc]	Monomeric/Polymeric ratio	Monomeric Peak 2/Peak 1 ratio
Flour	Akteur – batch 1	47.8 ± 1.9	52.2 ± 2.1	0.92	4.14
	Akteur – batch 2	46.9^{F} ± 1.9	53.1^{A} ± 2.1	0.88	4.29
	Carberry	51.5^{BCDE} ± 2.1	48.5^{BCD} ± 1.9	1.06	2.68
	Mv Magvas	49.2^{EF} ± 2.0	50.8^{AB} ± 2.0	0.97	5.26
	Yitpi	49.9^{DEF} ± 2.0	50.1^{ABC} ± 2.0	1.00	4.02
	Yumai-34	51.8^{BCDE} ± 2.1	48.2^{BCD} ± 1.9	1.07	4.36
	Blend	50.5^{CDE} ± 2.0	49.5^{BC} ± 2.0	1.02	3.71
Mean of the five cultivars		49.9	50.1	1.00	4.12
Gluten isolate	Akteur – batch 1	48.7 ± 1.9	51.3 ± 2.1	0.95	4.63
	Akteur – batch 2	51.8^{BCDE} ± 2.1	48.2^{BCD} ± 1.9	1.07	4.66
	Carberry	53.8^{ABC} ± 2.2	46.2^{DE} ± 1.8	1.16	3.09
	Mv Magvas	52.8^{BCD} ± 2.1	47.2^{CD} ± 1.9	1.12	6.75
	Yitpi	53.7^{ABC} ± 2.1	46.3^{DE} ± 1.9	1.16	4.76
	Yumai-34	56.8^{A} ± 2.3	43.2^{E} ± 1.7	1.31	5.82
	Blend	54.4^{AB} ± 2.2	45.6^{DE} ± 1.8	1.20	4.58
Mean of the five cultivars		53.8	46.2	1.16	5.02
Gliadin isolate	Akteur – batch 1	–	–	–	4.98
	Akteur – batch 2	–	–	–	5.92
	Carberry	–	–	–	3.13
	Mv Magvas	–	–	–	6.18
	Yitpi	–	–	–	5.08
	Yumai-34	–	–	–	5.62
	Blend	–	–	–	5.96
Mean of the five cultivars		–	–	–	5.19
PWG-gliadin		–	–	–	5.34

[a] Within each column, the mean value of batch 1 Akteur samples marked with a plus sign is significantly different from the mean value of batch 2 Akteur samples ($p < 0.05$; one-way ANOVA). [b] Within each column, the measured value of the blend marked with an asterisk is significantly different from the calculated mean of the five cultivars ($p < 0.05$; one-sample t-test). [c] Within each column, mean values marked with different capital letters are significantly different ($p < 0.05$; factorial ANOVA, Fisher's LSD).

This demonstrates that the production of gliadin isolates yielded similar distributions of alcohol-soluble proteins.

There was a high degree of similarity between the two batches of flour and gluten isolates in protein content and composition and this showed that the reproducibility of the production was similarly satisfactory. This was also confirmed by the ELISA results, as there were no significant differences between gliadin recoveries between the two batches of flours and gluten isolates (**Table 5**). Gliadin isolation appears to be more difficult due to a more complex process. Despite the variations in protein content and protein profile of the two batches of gliadin isolates, the ELISA results showed good similarity (**Table 5**). The amount of proteins obtained from gluten isolates along with the similarity of monomeric protein distribution indicated that the lower protein content and the presence or absence of higher molecular weight proteins on the SE-HPLC chromatograms has no effect on the ELISA results of gliadin fractions.

In addition to reproducible production on laboratory scale, upscaling and reproducible production of larger amounts of material are also important for widespread use. In our previous study we managed to achieve this in the case of the five cultivars and the blend in flour form (Schall et al., 2020). In the case of isolates, the identification of sensitive points in laboratory production may help in upscaling their production.

Comparison of the Different Gluten and Gliadin Isolates From Individual Cultivars and Their Blend

Comparison of the Protein Content of Flours, Gluten, and Gliadin Isolates

After testing the reproducibility of the production of flour and protein isolates from Akteur, the same methods were used for the other individual cultivars and for the blend of five cultivars. The crude protein content of the flours was in the range of 12.1–18.7% and the protein content of the blended flour was 15.4% which represented very well the calculated mean of the five cultivars (15.6%) and demonstrated the homogeneity of our flour blend (**Table 1**).

The crude protein content of gluten isolates from flours varied between 93.8 and 97.4% (**Table 1**). The lowest value belonged to Yitpi while Mv Magvas had the highest one. The high protein content means that comparatively pure gluten isolates could be extracted from each flour. The protein content of our gluten isolates from different flours was in a narrow range, but the

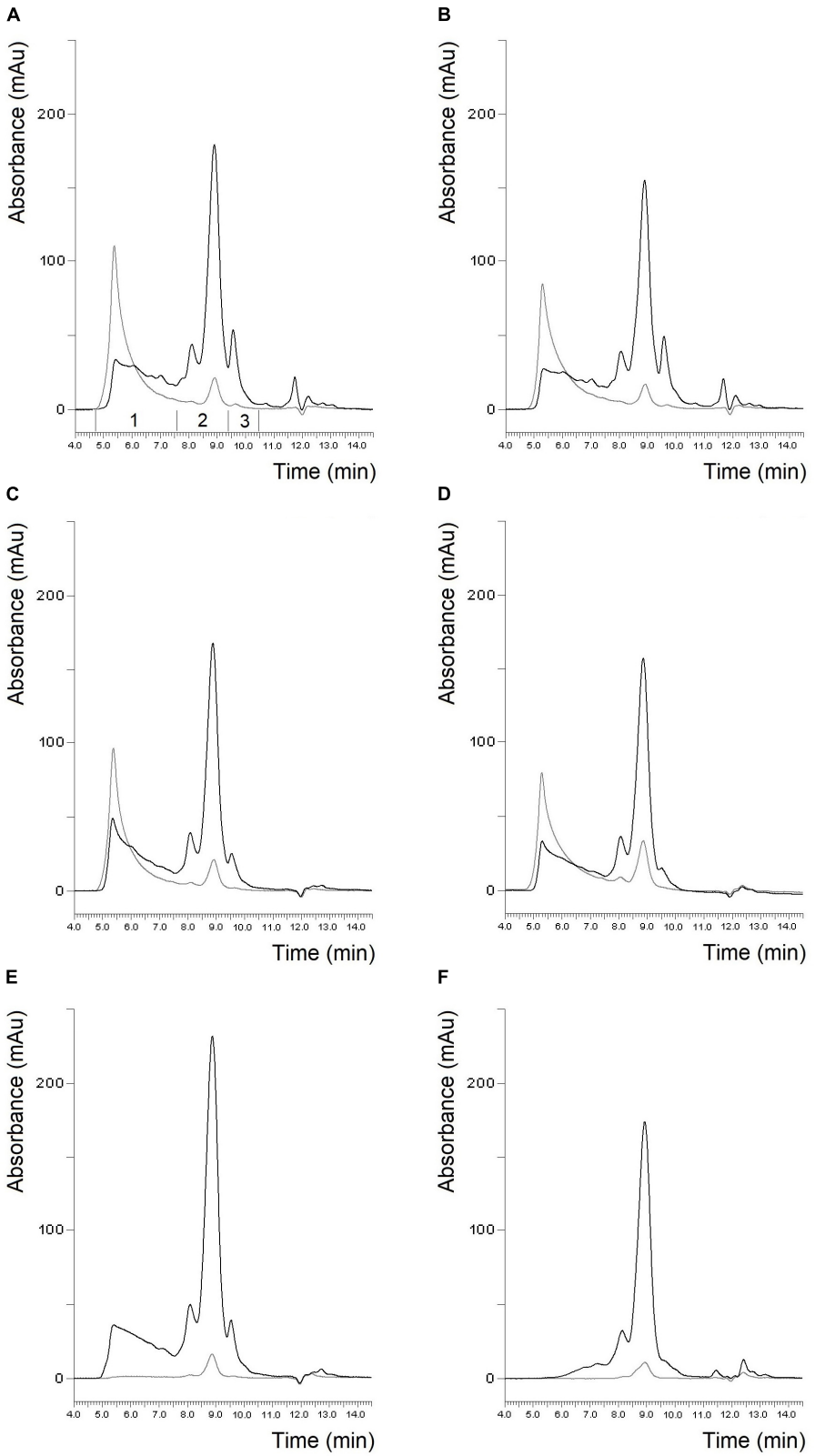


FIGURE 1 | Protein profile of batch 1 flour **(A)**, batch 2 flour **(B)**, batch 1 gluten isolate **(C)**, batch 2 gluten isolate **(D)**, batch 1 gliadin isolate **(E)** and batch 2 gliadin isolate **(F)** from Akteur cultivar determined by SE-HPLC (black line: soluble fraction, gray line: insoluble fraction, 1: polymeric proteins; 2: monomeric proteins; 3: albumins/globulins).

TABLE 3 | Proportion of gliadin, glutenin, and different gluten protein types in wheat flours and gluten isolates determined by RP-HPLC (RP-HPLC results are expressed as percentage of total extractable gluten proteins; all values are expressed on dry matter basis).

Sample		Parameter									
		Gliadins (%)[abc]	**Glutenins (%)[abc]**	**Gliadin/ Glutenin ratio**	**ω5 (%)[abc]**	**ω1,2 (%)[abc]**	**α (%)[abc]**	**γ (%)[abc]**	**ωb (%)[abc]**	**HMW- GS (%)[abc]**	**LMW- GS (%)[abc]**
Flour	Akteur – batch 1	62.6 ± 0.4	37.4^{+} ±	1.7	6.1^{+} ± 0.0	6.4 ± 0.0	30.9 ± 0.2	19.3^{+} ± 0.1	1.7^{+} ± 0.0	11.3^{+} ± 0.1	24.3^{+} ± 0.2
	Akteur – batch 2	61.7^{E} ± 0.4	38.3^{E} ± 0.3	1.6	5.8^{E} ± 0.0	6.5^{B} ± 0.1	30.6^{CD} ± 0.2	18.8^{I} ± 0.1	1.5^{F} ± 0.0	11.8^{B} ± 0.1	25.0^{CD} ± 0.2
	Carberry	63.7^{C} ± 0.5	36.3^{G} ± 0.3	1.8	8.6^{A} ± 0.1	5.9^{E} ± 0.0	26.1^{H} ± 0.2	23.2^{C} ± 0.2	0.9^{H} ± 0.0	10.2^{F} ± 0.1	24.9^{D} ± 0.2
	Mv Magvas	65.9^{AB} ± 0.5	34.1^{IJ} ± 0.2	1.9	3.1^{J} ± 0.0	5.2^{I} ± 0.0	30.9^{C} ± 0.2	26.7^{A} ± 0.2	0.6^{L} ± 0.0	9.3^{G} ± 0.1	24.2^{E} ± 0.2
	Yitpi	65.3^{B} ± 0.5	34.7^{H} ± 0.2	1.9	6.7^{C} ± 0.1	6.2^{D} ± 0.0	29.2^{E} ± 0.2	23.2^{C} ± 0.2	0.7^{K} ± 0.0	10.9^{E} ± 0.1	23.1^{G} ± 0.2
	Yumai-34	65.6^{AB} ± 0.5	34.4^{HI} ± 0.2	1.9	5.4^{G} ± 0.0	6.9^{A} ± 0.1	32.9^{A} ± 0.2	20.3^{G} ± 0.1	0.9^{I} ± 0.0	11.2^{D} ± 0.1	22.4^{H} ± 0.2
	Blend	66.1^{A} ± 0.5	33.9^{J} ± 0.2	2.0	6.4^{D} ± 0.0	6.4^{C} ± 0.0	30.4^{D} ± 0.2	23.0^{CD} ± 0.2	0.8^{J} ± 0.0	10.3^{F} ± 0.1	22.7^{H} ± 0.2
Mean of the five cultivars		64.4*	35.6*	1.8	5.9*	6.1*	29.9	22.4*	0.9*	10.7*	23.9*
Gluten isolate	Akteur – batch 1	54.7^{+} ± 0.4	45.3^{+} ± 0.3	1.2	5.3^{+} ± 0.0	5.5^{+} ± 0.0	26.6 ± 0.2	17.2 ± 0.1	3.2^{+} ± 0.0	12.7^{+} ± 0.1	29.1^{+} ± 0.2
	Akteur – batch 2	53.4^{I} ± 0.4	46.6^{A} ± 0.3	1.1	4.7^{I} ± 0.0	5.3^{H} ± 0.0	26.4^{GH} ± 0.2	16.9^{J} ± 0.1	3.3^{A} ± 0.0	13.5^{A} ± 0.1	29.8^{A} ± 0.2
	Carberry	62.0^{E} ± 0.4	38.0^{E} ± 0.3	1.6	7.9^{B} ± 0.1	5.7^{F} ± 0.0	25.6^{I} ± 0.2	22.9^{D} ± 0.2	2.5^{B} ± 0.0	10.2^{F} ± 0.1	25.2^{C} ± 0.2
	Mv Magvas	57.6^{H} ± 0.4	42.4^{B} ± 0.3	1.4	2.9^{K} ± 0.0	4.6^{J} ± 0.0	26.3^{GH} ± 0.2	23.7^{B} ± 0.2	1.4^{G} ± 0.0	11.0^{E} ± 0.1	30.1^{A} ± 0.2
	Yitpi	59.7^{G} ± 0.4	40.3^{C} ± 0.3	1.5	5.5^{F} ± 0.0	5.4^{G} ± 0.0	26.5^{G} ± 0.2	22.2^{E} ± 0.2	2.2^{C} ± 0.0	12.0^{B} ± 0.1	26.1^{B} ± 0.2
	Yumai-34	62.8^{D} ± 0.4	37.2^{F} ± 0.3	1.7	5.2^{H} ± 0.0	6.3^{D} ± 0.0	31.9^{B} ± 0.2	19.4^{H} ± 0.1	1.7^{E} ± 0.0	11.6^{C} ± 0.1	23.8^{F} ± 0.2
	Blend	60.8^{F} ± 0.4	39.2^{D} ± 0.3	1.6	5.6^{F} ± 0.0	5.7^{F} ± 0.0	28.5^{F} ± 0.2	21.0^{F} ± 0.2	2.1^{D} ± 0.0	11.2^{D} ± 0.1	25.9^{B} ± 0.2
Mean of the five cultivars		59.1*	40.9*	1.5	5.3*	5.5*	27.3*	21.0	2.2*	11.7*	27.0*

[a]Within each column, the mean value of batch 1 Akteur samples marked with a plus sign is significantly different from the mean value of batch 2 Akteur samples ($p < 0.05$; one-way ANOVA). [b]Within each column, the measured value of the blend marked with an asterisk is significantly different from the calculated mean of the five cultivars ($p < 0.05$; one-sample t-test). [c]Within each column, mean values marked with different capital letters are significantly different ($p < 0.05$; factorial ANOVA, Fisher's LSD).

TABLE 4 | Proportion of different gliadin protein types in wheat flours, gluten and gliadin isolates determined by RP-HPLC (RP-HPLC results are expressed as percentage of total extractable gliadin proteins; all values are expressed on dry matter basis).

Sample		Parameter			
		ω5 (%)[abc]	ω1,2 (%)[abc]	α (%)[abc]	γ (%)[abc]
Flour	Akteur – batch 1	9.7 ± 0.1	10.2 ± 0.1	49.3 ± 0.4	30.8 ± 0.2
	Akteur – batch 2	9.5^{F} ± 0.1	10.5^{D} ± 0.1	49.6^{D} ± 0.4	30.5^{M} ± 0.2
	Carberry	13.4^{B} ± 0.1	9.2^{I} ± 0.1	41.0^{L} ± 0.3	36.4^{E} ± 0.3
	Mv Magvas	4.7^{O} ± 0.0	7.9^{J} ± 0.1	46.9^{G} ± 0.3	40.5^{B} ± 0.3
	Yitpi	10.2^{D} ± 0.1	9.6^{G} ± 0.1	44.7^{IJ} ± 0.3	35.5^{G} ± 0.3
	Yumai-34	8.3^{K} ± 0.1	10.5^{D} ± 0.1	50.2^{C} ± 0.4	31.0^{L} ± 0.2
	Blend	9.6^{E} ± 0.1	9.6^{G} ± 0.1	45.9^{H} ± 0.3	34.8^{H} ± 0.2
Mean of the five cultivars		9.2*	9.5	46.5	34.8
Gluten isolate	Akteur – batch 1	9.8 ± 0.1	10.0 ± 0.1	48.5 ± 0.3	31.4 ± 0.2
	Akteur – batch 2	8.9^{I} ± 0.1	10.0^{F} ± 0.1	49.4^{D} ± 0.4	31.7^{K} ± 0.2
	Carberry	12.7^{C} ± 0.1	9.1^{I} ± 0.1	41.2^{L} ± 0.3	36.9^{D} ± 0.3
	Mv Magvas	5.1^{N} ± 0.0	7.9^{J} ± 0.1	45.8^{H} ± 0.3	41.2^{A} ± 0.3
	Yitpi	9.3^{G} ± 0.1	9.1^{I} ± 0.1	44.4^{J} ± 0.3	37.2^{D} ± 0.3
	Yumai-34	8.3^{JK} ± 0.1	10.0^{EF} ± 0.1	50.8^{B} ± 0.4	30.9^{L} ± 0.2
	Blend	9.2^{H} ± 0.1	9.4^{H} ± 0.1	46.9^{FG} ± 0.3	34.5^{H} ± 0.2
Mean of the five cultivars		8.8*	9.2*	46.3*	35.6*
Gliadin isolate	Akteur – batch 1	9.9^{+} ± 0.1	10.1^{+} ± 0.1	49.1^{+} ± 0.3	30.9^{+} ± 0.2
	Akteur – batch 2	8.4^{J} ± 0.1	11.2^{B} ± 0.1	48.3^{E} ± 0.3	32.2^{J} ± 0.2
	Carberry	14.6^{A} ± 0.1	10.1^{E} ± 0.1	42.5^{K} ± 0.3	32.8^{I} ± 0.2
	Mv Magvas	5.3^{M} ± 0.0	9.3^{H} ± 0.1	47.3^{FG} ± 0.3	38.0^{C} ± 0.3
	Yitpi	9.3^{G} ± 0.1	9.6^{G} ± 0.1	45.2^{I} ± 0.3	36.0^{F} ± 0.
	Yumai-34	8.8^{I} ± 0.1	10.9^{C} ± 0.1	52.1^{A} ± 0.4	28.2^{N} ± 0.2
	Blend	7.9^{L} ± 0.1	12.0^{A} ± 0.1	47.4^{F} ± 0.3	32.6^{I} ± 0.2
Mean of the five cultivars		9.3*	10.2*	47.1	33.4*
PWG-gliadin		5.9 ± 0.3	7.2 ± 0.2	50.0 ± 0.4	36.9 ± 0.2

[a]Within each column, the mean value of batch 1 Akteur samples marked with a plus sign is significantly different from the mean value of batch 2 Akteur samples ($p < 0.05$; one-way ANOVA). [b]Within each column, the measured value of the blend marked with an asterisk is significantly different from the calculated mean of the five cultivars ($p < 0.05$; one-sample t-test). [c]Within each column, mean values marked with different capital letters are significantly different ($p < 0.05$; factorial ANOVA, Fisher's LSD).

variations may be due to the different separation behavior of cultivars during gluten processing (Marchetti et al., 2012). But the protein content of the isolates did not depend on the protein content of the flour. The amount of gluten proteins obtained from the proteins from different flours were in the range of 73.4–76.9%. Gluten yield may depend on the protein content of the flour (Van Der Borght et al., 2005), although no correlation was found between yield and protein content in the five samples we examined, meaning that the amount of gluten proteins was reached with similar potency for each sample (**Table 1**). The gluten isolated from the blended flour also had a high crude protein content of 95.5% while the mean of the five cultivars was 95.6%. The amount of gluten proteins relative to the proteins of the blend flour was 78.5% which was close, but a little higher than the mean of the five cultivars (75.6%) (**Table 1**). The crude protein content showed that the cultivars in the blend were affected to the same extent by isolation, however, the amount of proteins that could be obtained from the blend flour was higher compared to the individual cultivars. The difference was probably due to the production, as a difference was also observed in the yield of the two batches of Akteur. However, the protein content of the blend flour and gluten isolate was following the average of the five cultivars well so a detailed analysis of the protein profile is required.

The measured protein content of the gliadin isolates varied in the range of 72.7–101.9%, which was a much wider range than observed for the gluten isolates (**Table 1**). The protein content of PWG-gliadin was 92.8% (van Eckert et al., 2006), and from our samples, the Yitpi gliadin isolate had a higher protein content than PWG-gliadin. Despite the fact that we used exactly the same isolation procedure in each case, the crude protein contents of the gliadin isolates were significantly different. Additionally, it seems that the identified differences did not depend on the protein content of the flours or even the gluten isolates. In case of gluten isolates, it can be assumed that the differences between the samples depended on cultivars. This cannot be clearly stated for the gliadin isolates, because as shown in the investigation of reproducible production, such differences may occur between up to two parallel isolations and the causes of this phenomenon must be revealed. Less pure gliadin isolates thus raise the question if lower protein contents would cause changes in the protein profiles compared to flours and gluten isolates, and consequently in ELISA response. The amount of gliadin proteins obtained from different gluten isolates was between 28.8

TABLE 5 | Gliadin recovery in wheat flours, gluten and gliadin isolates using two different ELISA test kits (all values are expressed on dry matter basis; gliadin recoveries are calculated based on gliadin content measured by RP-HPLC).

Sample		Gliadin recovery (%)[abc]	
		ELISA kit A	ELISA kit B
Flour	Akteur – batch 1	187 ± 2	–
	Akteur – batch 2	163^{BCD} ± 14	156^{F} ± 29
	Carberry	183^{AB} ± 36	179^{EF} ± 24
	Mv Magvas	152^{CDE} ± 13	157^{F} ± 29
	Yitpi	169^{ABCD} ± 21	190^{EF} ± 18
	Yumai-34	165^{BCD} ± 20	185^{EF} ± 16
	Blend	150^{DE} ± 19	183^{EF} ± 25
Mean of the five cultivars		167	173
Gluten isolate	Akteur – batch 1	165 ± 40	–
	Akteur – batch 2	191^{A} ± 17	215^{CDE} ± 12
	Carberry	173^{ABC} ± 26	281^{AB} ± 116
	Mv Magvas	162^{BCD} ± 18	206^{CDEF} ± 10
	Yitpi	170^{ABCD} ± 29	344^{A} ± 38
	Yumai-34	176^{AB} ± 14	215^{CDE} ± 35
	Blend	182^{AB} ± 15	260^{BC} ± 81
Mean of the five cultivars		175	252
Gliadin isolate	Akteur – batch 1	123 ± 15	–
	Akteur – batch 2	132^{E} ± 14	205^{CDEF} ± 4
	Carberry	135^{E} ± 10	187^{EF} ± 8
	Mv Magvas	139^{E} ± 8	223^{BCDE} ± 11
	Yitpi	165^{BCD} ± 27	201^{CDEF} ± 12
	Yumai-34	132^{E} ± 13	202^{CDEF} ± 6
	Blend	180^{AB} ± 11	244^{BCD} ± 24
Mean of the five cultivars		140*	204
PWG-gliadin		125 ± 34	188 ± 55

[a]Within each column, the mean value of batch 1 Akteur samples marked with a plus sign is significantly different from the mean value of batch 2 Akteur samples ($p < 0.05$; one-way ANOVA). [b]Within each column, the measured value of the blend marked with an asterisk is significantly different from the calculated mean of the five cultivars ($p < 0.05$; one-sample t-test). [c]Within each column, mean values marked with different capital letters are significantly different ($p < 0.05$; factorial ANOVA, Fisher's LSD).

and 58.2% (**Table 1**). Gliadin/glutenin ratios based on the weight of the isolated gliadins and glutenins were in the range from 0.5 to 1.4. Such variability in extractable gliadin content and gliadin/glutenin ratios in gluten proteins among cultivars may occur, but this degree of variation did not occur in the production of flour from these cultivars (Hajas et al., 2018; Schall et al., 2020). It is therefore necessary to examine how the differences are affected by the quality of the cultivars and what the effect of isolation is. The protein content of the blend gliadin isolate was 79.6% and the calculated mean of the five cultivars was 84.2%, which was a higher (but still not significant) difference than in case of the blend gluten isolate (**Table 1**). The amount of gliadin obtained from the gluten isolate was 28.8% and the mean of the five cultivars was 43.3% meaning that the theoretical protein content available in the blend gliadin isolate could not be extracted and loss of material could be problematic, as the profile may also change compared to flour. Just as gluten production is affected by different protein interactions, the separation of gliadins and glutenins can also be determined by the different types of proteins in the material. The yield obtained for the blend was lower than in case of any cultivar, and we did not obtain such a difference in the parallel yields of Akteur, so this is difficult to explain with the uncertainty of the isolation method. It is conceivable that glutenins from different cultivars in the blend are able to aggregate strongly, thus negatively affecting the yield of gliadin.

Protein yields in the gluten isolates with a high protein content were well representative of flour values, while the protein contents and the amount of extracted proteins in gliadin isolates were slightly distorted because in some cases gliadin isolates with lower protein content were obtained than for gluten isolates. Protein isolates with high protein content were obtained with our isolation method that are comparable to commercially available materials and PWG-gliadin (van Eckert et al., 2006; Schwalb et al., 2011). Schalk and colleagues dealt with the development of a strategy for the isolation of protein fractions and types from different species, in which the crude protein contents of the different isolates were similar to our isolates (Schalk et al., 2017). However, protein content and yield alone are not sufficiently informative, because a comparative study has already shown that the composition of gliadin isolates is highly dependent on production (Schwalb et al., 2011). In the following paragraphs, the protein composition of the isolates is described to evaluate their identity compared to the gluten proteins of the source flour.

Comparison of the SE-HPLC Protein Profiles of Flours, Gluten, and Gliadin Isolates

The SE-HPLC protein profiles of the six flours and their isolates are shown in **Figures 1**, **2**. Naturally, the sizes of the albumin/globulin peaks were smaller in the gluten isolates compared to the flours, but a small residue was observed in the different gluten samples with varying degrees. The effectiveness of removing water and salt-soluble proteins may depend on cultivars and small variations during isolation. In all six cases, a similar protein composition was seen in the distribution of polymeric and monomeric proteins for gluten isolates compared to the flours which were specific to the cultivars. The greatest change was identified in the higher molecular weight regions of the soluble fractions of all gluten isolates, where the protein profile was slightly modified. There was also a change in monomeric proteins in the insoluble fraction, as their amount increased compared to flours in all cases. Similar findings were made in the blend gluten isolate as in the individual cultivars because the amount of albumins/globulins decreased. A slight change in the size distribution of higher molecular weight proteins could also be observed, as the increase in the amount of monomeric proteins in the insoluble fraction. However, the protein distribution characteristics of the blend flour were also reflected in the gluten isolate. In flour samples, the proportion of monomeric proteins was in the range of 46.9–51.8%, while the values of polymeric proteins were between 48.2 and 53.1% (**Table 2**). There was an increase in the monomeric protein content of the gluten isolates (51.8–56.8%), compared to flour and the extent of change was not the same for all samples, because it was smaller in the case of Carberry than in the other cultivars.

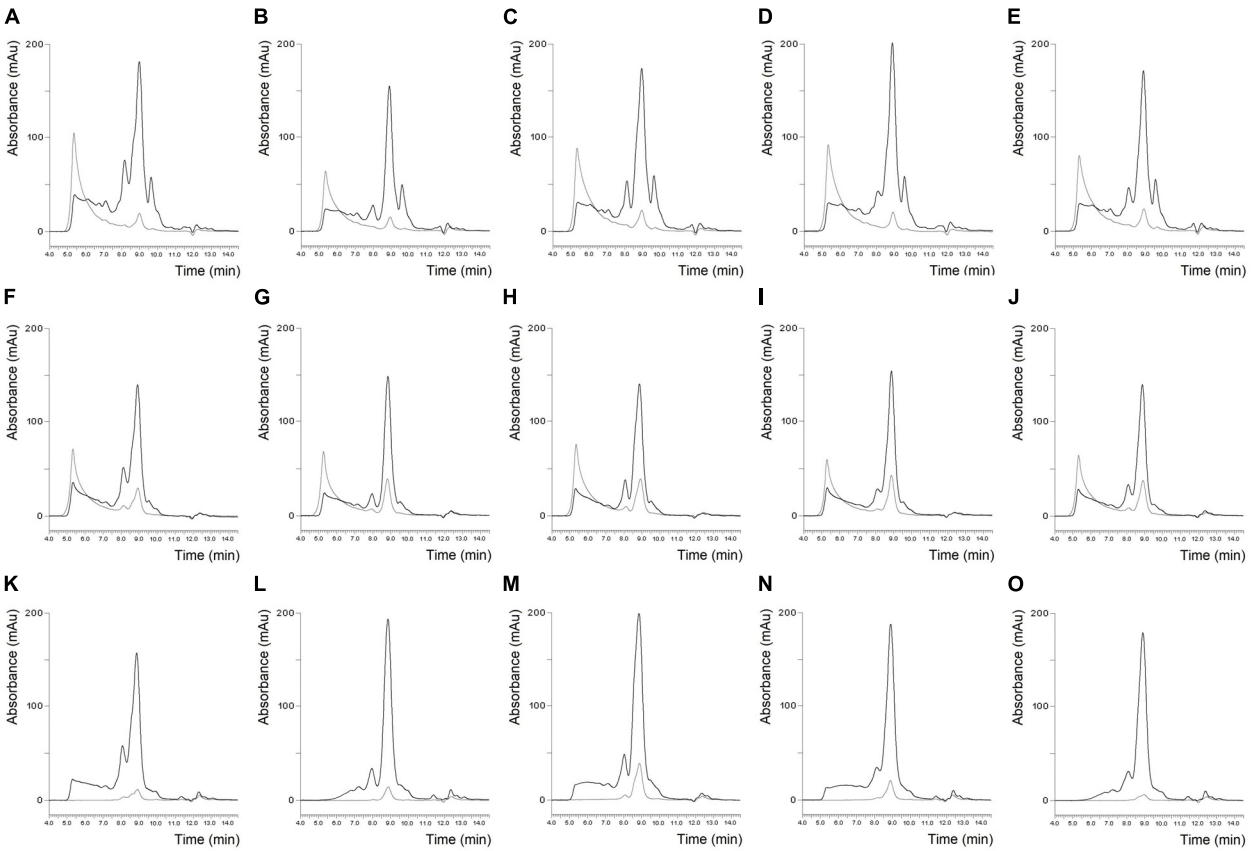


FIGURE 2 | Protein profile of Carberry flour **(A)**, Mv Magvas flour **(B)**, Yitpi flour **(C)**, Yumai-34 flour **(D)**, Blend flour **(E)**, Carberry gluten isolate **(F)**, Mv Magvas gluten isolate **(G)**, Yitpi gluten isolate **(H)**, Yumai-34 gluten isolate **(I)**, Blend gluten isolate **(J)**, Carberry gliadin isolate **(K)**, Mv Magvas gliadin isolate **(L)**, Yitpi gliadin isolate **(M)**, Yumai-34 gliadin isolate **(N)** and Blend gliadin isolate **(O)** determined by SE-HPLC (black line: soluble fraction, gray line: insoluble fraction).

The increase due to isolation in the proportion of monomeric proteins in the gluten isolates was also reflected in changes in the monomeric/polymeric protein ratio (**Table 2**). The increase in the proportion of monomeric proteins also occurred in the blend gluten isolate compared to flour. So the change in the ratio of monomeric and polymeric proteins due to isolation was similar in the blend and the individual cultivars. The monomeric protein content of the blend gluten isolate was 54.4% while the mean of the five cultivars was 53.8%. This similarity was also observed in the amount of polymeric proteins as the blend gluten isolate had 45.6% while the mean was 46.2%. So it represented the average of the five individual gluten isolates. Overall, there were no major differences in the gluten protein composition between gluten isolates and flours, but any change in proportions could be problematic because the flour-specific ratio would be affected, which could cause further uncertainty in methods where only alcohol-soluble proteins are determined.

The size distribution of proteins in the gliadin isolates is shown in **Figures 1**, **2**. As expected, the peak size of albumin/globulin proteins decreased further in gliadin isolates compared to flours and gluten isolates. In most cases albumins/globulins were completely missing, but small residues could be observed, for example, in the case of Yitpi and Mv Magvas. The peaks of monomeric proteins appearing in the soluble fractions (between 7.5 and 9.5 min) followed the pattern observed in the flours and gluten isolates for each sample. The ratio of the two peaks of the monomeric proteins appearing in the gliadin isolates showed higher values than in flours but showed similarities with gluten isolates. It means that the production of gluten isolates had a greater effect on the distribution of these proteins, while the production of the gliadin isolate had a smaller impact. The smallest change of the two peaks in monomeric proteins was observed in the case of Carberry, as the gluten and gliadin isolates showed only a slight increase compared to flour. As mentioned above, the peak of monomeric proteins in the insoluble fraction increased in the gluten isolates compared to flours, while it decreased in most of the gliadin isolates compared to gluten, making it more comparable to flours. As expected, the higher molecular weight proteins were reduced in gliadin isolates, and completely disappeared in the insoluble fraction for each sample. Although we expected the lack of these proteins in the soluble fraction, they did appear in Carberry, Yitpi, and Yumai-34 while they could not be detected in Akteur and Mv Magvas gliadin isolates. However, it was shown during the examination of reproducible production that the presence or absence of these proteins was somehow

influenced by the conditions of production or the sample preparation of the method, not just the cultivars. Similar to most of the individual cultivars, the albumin/globulin peak completely disappeared in the blend gliadin isolate. There was an increase in the ratio of the two peaks of the monomeric proteins in the soluble fractions, but the degree of change was higher compared to the gluten isolate and flour as for the individual cultivars (**Table 2**). The ratio of the blend gliadin isolate was 5.96 while the mean of the five cultivars was 5.19 which showed a slight distortion compared to the theoretical ratio (**Table 2**). This points to uncertainties of production and its reason must be explored in order to standardize the production of RM, and to reduce random errors originating from production. The amount of monomeric proteins in the insoluble fraction decreased compared to gluten isolates but the reduction was higher than in other gliadin isolates (**Figures 1**, **2**). Another important finding is that the higher molecular weight proteins were also missing in the soluble fraction of the blend gliadin isolate.

In case of gliadin isolates, it seems reasonable to compare our results with the well-known PWG-gliadin. As in some of our gliadin isolates, PWG-gliadin also had a small albumin/globulin peak (**Figure 3**; van Eckert et al., 2006). The amount of monomeric proteins in the insoluble fraction was much lower than in our study. As expected and similar to our samples, two peaks appeared in the monomeric proteins of the soluble fraction, with a ratio of 5.34 (**Table 2**). Some of our individual cultivars considerably differed from this value, but the value of our blend gliadin isolate was very close to it. This may mean that a mixture actually represents a number of wheat cultivars, but it may not be necessary to include a large number of cultivars (28 in case of PWG-gliadin), but a smaller number of carefully selected varieties which can facilitate the production of RM. There were no higher molecular weight proteins in the insoluble fraction while they appeared in the soluble fractions of PWG-gliadin (**Figure 3**). In the article on the characterization of PWG-gliadin (van Eckert et al., 2006), these proteins were called oligomeric gliadin and based on our results, the presence of these proteins may be matrix- and/or production-dependent.

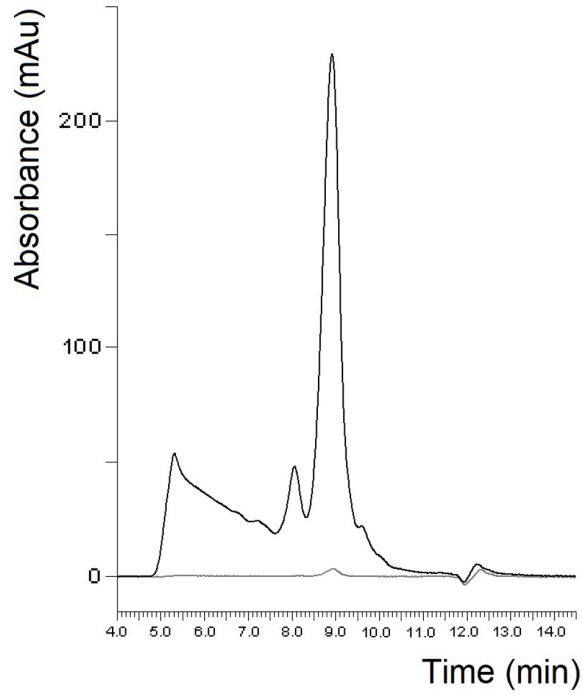


FIGURE 3 | Protein profile of PWG-gliadin determined by SE-HPLC (black line: soluble fraction, gray line: insoluble fraction).

Comparison of the RP-HPLC Protein Profiles of Flours, Glutens, and Gliadin Isolates

The composition of different gluten proteins in flours and gluten isolates separated by RP-HPLC for more detailed examination is shown in **Table 3**. Based on the RP-HPLC chromatograms, the protein composition showed greater similarity between flours and isolates for each sample than in the SE-HPLC chromatograms, so only the RP-HPLC chromatograms for Carberry and Mv Magvas are shown to demonstrate the results in **Figures 4**, **5**. There was a slight decrease in the percentage distribution of ω5- and ω1,2-gliadins in gluten isolates compared to flour, but the extent of change was similar for all samples (**Table 3**). Washing with tap water to remove residual starch in gluten production could cause a slight decrease in the amount of ω-gliadins because these types of proteins are slightly soluble in water. The α- and γ-gliadins also showed a decrease in gluten isolates compared to flours, but the rate of change was not the same. In case of Carberry, the changes were negligible, while Mv Magvas showed the greatest extent both in α- and γ-gliadins. Consequently, the total gliadin content also decreased in the gluten isolates to a different degree relative to the flours, and the slightest change occurred in Carberry. A moderate decrease could be detected in the proportion of different gliadin proteins for blend gluten isolates compared to flours. However, the values of the blend gluten isolate showed great similarity with the average of the five gluten isolates, so it was representative of the five gluten isolates. In accordance with the change in the amount of gliadins, the proportion of glutenins in the gluten isolates was increased compared to the flours. The extent of change was similar for each sample in ωb-gliadins and HMW-GS. While in LMW-GS, the rate of increase was different in the cultivars, the highest change was in Mv Magvas and the lowest was in Carberry, similar to α- and γ-gliadins. In the blend gluten isolate, the proportion of glutenin proteins increased compared to the blend flour to a similar extent as in the individual cultivars. It can also be seen that the values of the blend gluten isolate were very close to the mean of the five isolates from the individual cultivars.

Based on analysis of variance, the variability between cultivars was higher (67% in ω5-, 42% in ω1,2-, 40% in α-, 58% in γ-gliadins and 43% in HMW-GS) than the effect of isolation (19% in ω5-, 38% in ω1,2-, 33% in α-, 24% in γ-gliadins and 31% in HMW-GS) in all types of gliadins and also in HMW-GS. This also means that the potential loss of ω-gliadins during isolation is expected to cause less error than the genetic variability. The

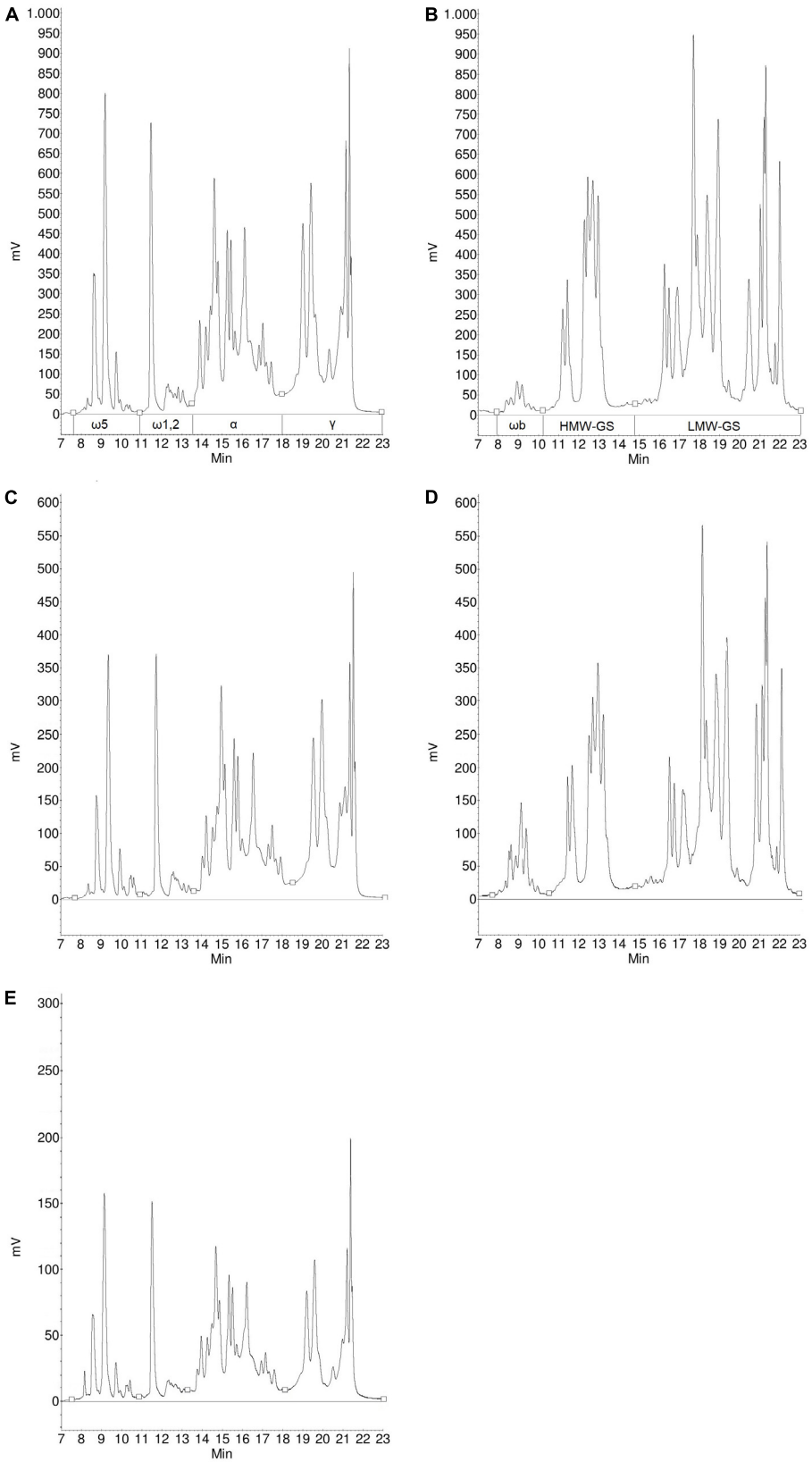


FIGURE 4 | Protein profile of gliadins **(A)** and glutenins **(B)** from Carberry flour, gliadins **(C)** and glutenins **(D)** from Carberry gluten isolate and gliadins **(E)** from Carberry gliadin isolate determined by RP-HPLC.

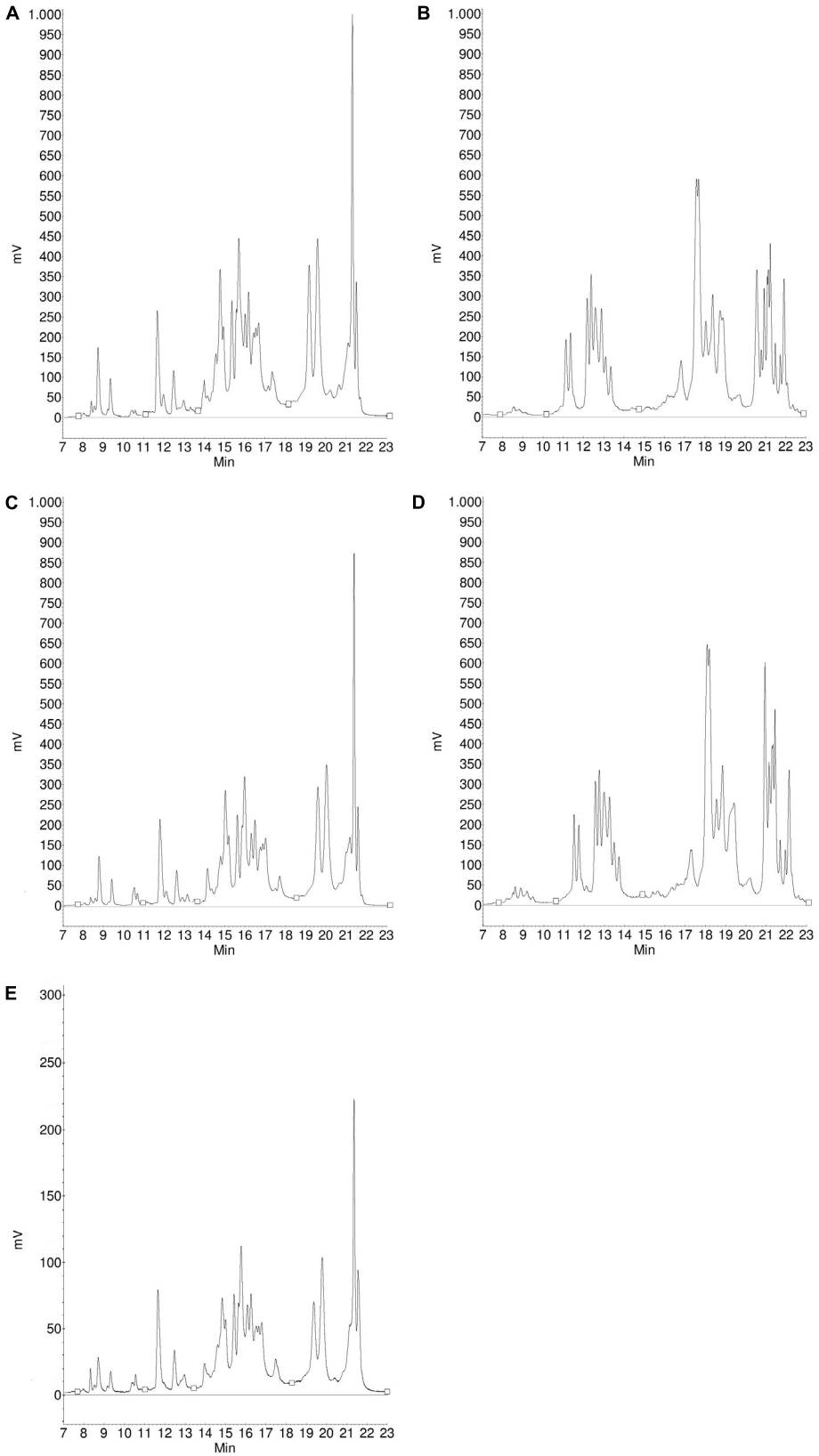


FIGURE 5 | Protein profile of gliadins **(A)** and glutenins **(B)** from Mv Magvas flour, gliadins **(C)** and glutenins **(D)** from Mv Magvas gluten isolate and gliadins **(E)** from Mv Magvas gliadin isolate determined by RP-HPLC.

differences originating from the cultivars should be reduced, which was achieved not only with the blend flour but also with the blend gluten isolate, because the gliadin and HMW-GS contents were close to the mean of the five cultivars. However, the isolation had a higher effect (56% in ωb-gliadins and 42% in LMW-GS) than variability between cultivars (25% in ωb-gliadins and 2% in LMW-GS) in case of ωb-gliadins and LMW-GS.

The composition of different gliadin proteins in flours, gluten and gliadin isolates separated by RP-HPLC are shown in **Table 4**. The chromatograms of the Carberry and Mv Magvas gliadin isolates are shown in **Figures 4**, **5**. As in the case of gluten isolates, the effect of isolation was less noticeable in the distribution of ω5- and ω1,2-gliadins, because gliadin isolates had similar values to flours and gluten isolates. In ω5-gliadins, the direction of change was also positive and negative depending on the sample compared to glutens and flours while in the case of ω1,2-gliadins, the values of gliadin isolates were higher than in flours and gluten isolates. The α-gliadins also showed very small changes in the distribution compared to flours and gluten isolates. The greatest change due to isolation was found in γ-gliadins. While gluten isolates had a very similar value to flours in every cultivar, the γ-gliadin proportions of gliadin isolates varied and not to the same extent. Interestingly, a high difference occurred in Carberry despite the fact that the gluten isolate showed the best similarity to the flour. The variance analysis with the results of the gliadin isolates also showed that the variability between cultivars had a greater effect (79% in ω5-, 41% in ω1,2-, 71% in α-, and 60% in γ-gliadins) on the distribution of different proteins than the isolation (0% in ω5-, 32% in ω1,2-, 7% in α-, and 17% in γ-gliadins). The change in the blend gliadin isolate was small compared to the gluten isolate and flour and the difference between the values of the blend isolate and the mean of the five gliadin isolates was small. This means that the blend gliadin isolate adequately represented the mean of the five cultivars. Consequently, the isolation did not alter the homogeneity of the proteins and each gliadin type was isolated with the same efficiency. This answered the question above related to protein content and showed that we did not lose information with reduced protein recovery, because the gluten protein composition was very similar in flour and isolates.

In the case of PWG-gliadin, it was found that its composition was very similar to the distribution of gliadin proteins in the source flour (van Eckert et al., 2006). Our gliadin isolates typically had higher ω5- and ω1,2-gliadin contents and lower α- and γ-gliadin contents than PWG-gliadin (**Table 4**). Gluten and gliadin isolates with high protein content are available, but their protein composition is not similar or probably altered compared to the composition of wheat flour (Schwalb et al., 2011). Despite the minor changes observed in the composition during isolation, our isolates – both gluten and gliadins – were similar to the source flour, as well as to PWG-gliadin.

Comparison of the Gliadin Recovery Values of Flours, Gluten, and Gliadin Isolates Measured With ELISA Tests

In the next step we investigated whether the slight and different modifications described above were apparent in the ELISA results. Gliadin recovery values (**Table 5**) of ELISA measurements were calculated relative to the total gliadin content measured by RP-HPLC. For each sample and both ELISAs, the recovery values were above 100%. Lexhaller et al. (2016) noted that both G12 and R5 ELISA kits overestimated the prolamin content relative to RP-HPLC results in wheat. Despite high recovery values, it is informative to compare the results obtained with different samples and methods. The recovery values for gluten isolates obtained by ELISA method A showed good similarity to flours and there were no significant differences between the two. The highest difference between flour and the gluten isolate occurred in cultivar Akteur with an increase of about 30%. The blend had a higher change in the isolate compared to most cultivars, with about the same extent as in Akteur. However, the recovery value of the blend gluten isolate was closer to the mean of the five gluten isolates than in the flour blend. Since the protein profiles of gluten isolates were similar to flours based on both SE- and RP-HPLC results, similarities in ELISA results were expected.

In case of ELISA method B, there was a much greater increase in the recovery values of gluten isolates compared to flour (**Table 5**) than with ELISA method A. The rate of change varied between cultivars; the smallest increase was in Yumai-34, while the highest was in Yitpi. The recovery value of the blend gluten isolate was higher compared to flour, but the blend was comparably close to the mean of the five gluten isolates compared to the blend flour and the mean of the five flours. This is only possible if the blend contains the five cultivars homogeneously in the flour and in the gluten isolate. Considering the results of gluten protein composition, no protein types or even the total gliadin content could be identified that would be associated with the increase in ELISA results in some gluten isolates.

Increases were mainly observed in the recovery values of gluten isolates compared to flour, while the values of gliadin isolates decreased in all cases but not to the same extent in case of method A. The lowest change was observed in Yitpi, while the highest was in Carberry. Interestingly, the value of the blend increased compared to flour in contrast to individual cultivars and its value was very similar to the blend gluten isolate. The recovery value of the blend gliadin isolate differed quite substantially from the mean of the five cultivars. Consequently, it was more affected by isolation than by cultivars. The recovery value of PWG-gliadin was lower than our gliadin samples (**Table 5**), but the mean value of the five cultivars did not show a very high difference from PWG-gliadin.

Interestingly, in case of method B, the recovery values of the gliadin isolates showed better similarity to the flours than the gluten isolates (**Table 5**). In the blend gliadin isolate (similarly to the results of method A) there was an increase compared to flour and its value was more similar to the gluten isolate than the flour. The value of the blend did not differ significantly from the mean of the five cultivars. A lower recovery value was determined for PWG-gliadin than for our samples, except Carberry and there was also a marked difference between the value of our blend gliadin isolate and PWG-gliadin (**Table 5**).

The results obtained with ELISA methods show variation between materials of different formats and compositions (Schwalb et al., 2011; Lexhaller et al., 2016). In addition, the

reactivity of various antibodies can be different, which causes serious uncertainty in the measurement results (Schwalb et al., 2011; Lexhaller et al., 2017). It is difficult to determine which factor is most relevant for the deviation of the results and their relevance for the production of the RM. Protein composition results showed that the isolation had a minimal effect on protein distribution. The results obtained with both ELISA methods were assessed by analysis of variance. The isolation had a significant effect on the results of both ELISA methods because it contributed 28% to the deviation of the results in case of method A and 36% in case of method B. The variation between cultivars also affected the results, but to a lower extent than isolation. The degree of interaction between isolation and genetic variability was 29% for method A, while it was 20% for method B. Genetic variability contributed by an additional 11% to the deviation of the measured values in case of method B. Furthermore, the measurement uncertainty in both methods approximated but did not exceed the effect of isolation and genetic variability. This points to the importance of carefully choosing a RM to minimize these effects. One argument for using flour is that there was no significant difference in the values of the two methods for any of the samples, but the difference between the cultivars could be reduced by the use of the blend, since there was no significant difference between its values and the mean of the five cultivars.

CONCLUSION

We have shown that the isolation of gluten and gliadin proteins from wheat flours has a slight effect on the amount and composition of proteins, which partially depends on the cultivar. However genetic variability still causes higher uncertainty in protein composition than isolation. Obviously, due to the more complex production of gliadins, there is a higher probability of error, and we tried to investigate the causes of it in our study. Immunoanalytical results showed higher effects of isolation on the results and they seem to be method dependent. We demonstrated that the directions and extents of changes are not in direct relation to the other protein properties. Similarly, to our previous results with flours (Schall et al., 2020), here we also confirmed experimentally that the blend is the most appropriate solution to compensate for the effects of genetic variability. As an overall conclusion of our work, we demonstrated first with analytical experiments that similar results can be obtained with isolates than with basic flours. The blends can partly compensate for the effects of genetic and environmental variability and the source and the extent of analytical uncertainty are similar (but not the same) in all investigated materials. In the gliadin isolation process, there is a higher chance of uncertainty that can affect the analytical results. Exploring the sources of these errors, evaluating the results of long-term stability studies, and clarifying the specific analytical goals are necessary to select the suitable form or forms of a widely accepted RM. One limitation of the study is that we have only completed the work for wheat so far, but our experience will now enable us to transfer our findings efficiently to producing RM for rye and barley, as well, and thus provide a comprehensive gluten RM.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

ST, PK, KS, and RS contributed to the conception and design of the study. ES, KS, ZB, and KT contributed to performing the experiments. ES organized the database and performed the statistical analysis. ES and ST evaluated the results and wrote the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was related to the scientific goals of the MoniQA Association, the OTKA ANN 114554 project ("Improving gluten-free dough by a novel hemicellulose network") and the Higher Education Excellence Program of the Ministry of Human Capacities in the frame of Biotechnology research area of Budapest University of Technology and Economics (BME FIKP-BIO). We acknowledge support by the KIT-Publication Fund of the Karlsruhe Institute of Technology.

ACKNOWLEDGMENTS

We gratefully acknowledge the contribution of Prof. Ferenc Békés (FBFD Pty. Ltd.), Dr. Terry Koerner (Head, Allergen and Natural Toxin Section, Food Research Division, Bureau of Chemical Safety, Health Canada), and Dr. Marianna Rakszegi (Agricultural Research Institute of the Hungarian Academy of Sciences) in collecting the wheat samples. We wish to thank Dr. Szilveszter Gergely (Budapest University of Technology and Economics) for his assistance in NIR measurements and Ms. Alexandra Axthelm (Leibniz-LSB@TUM) for excellent technical assistance.

REFERENCES

AACC Method 38-12.02, (2000). *Wet Gluten, Dry Gluten, Water-Binding Capacity, and Gluten index. AACC Approved Methods of Analysis*, 11th Edn. Saint Paul, MN: AACC International.

Batey, I. L., Gupta, R. B., and MacRitchie, F. (1991). Use of size-exclusion high-performance liquid chromatography in the study of wheat flour proteins: an improved chromatographic procedure. *Cereal Chem.* 68, 207–209.

Brouns, F., van Rooy, G., Shewry, P., Rustgi, S., and Jonkers, D. (2019). Adverse reactions to wheat or wheat components. *Compr. Rev. Food Sci. Food Saf.* 18, 1437–1452. doi: 10.1111/1541-4337.12475

Bruins Slot, I. D., van der Fels-Klerx, H. J., Bremer, M. G. E. G., and Hamer, R. J. (2016). Immunochemical Detection Methods for Gluten in Food Products: Where Do We Go from Here? *Crit. Rev. Food Sci. Nutr.* 56, 2455–2466. doi: 10.1080/10408398.2013.847817

Bugyi, Z., Török, K., Hajas, L., Adonyi, Z., Popping, B., and Tömösközi, S. (2013). Comparative study of commercially available gluten ELISA kits using

an incurred reference material. *Qual. Assur. Saf. Crop.* 5, 79–87. doi: 10.3920/QAS2012.0174

Codex Stan 118-1979, (2015). *Codex Standard for Foods for Special Dietary use for Persons Intolerant to Gluten*. Geneva: Codex Alimentarius Commission.

Day, L., Augustin, M. A., Batey, I. L., and Wrigley, C. W. (2006). Wheat-gluten uses and industry needs. *Trends Food Sci. Technol.* 17, 82–90. doi: 10.1016/j.tifs.2005.10.003

Diaz-Amigo, C., and Popping, B. (2012). Gluten and gluten-free: issues and considerations of labeling regulations. Detection methods and assay validation. *J. AOAC Int.* 95, 5–24. doi: 10.5740/jaoacint.SGE

Diaz-Amigo, C., and Popping, B. (2013). Accuracy of ELISA detection methods for gluten and reference materials: a realistic assessment. *J. Agric. Food Chem.* 61, 5681–5688. doi: 10.1021/jf3046736

Gupta, R. B., Khan, K., and MacRitchie, F. (1993). Biochemical basis of flour properties in bread wheats. I. Effects of variation in the quantity and size distribution of polymeric proteins. *J. Cereal Sci.* 18, 23–41. doi: 10.1006/jcrs.1993.1031

Hajas, L., Scherf, K. A., Török, K., Bugyi, Z., Schall, E., Poms, R. E., et al. (2018). Variation in protein composition among wheat (*Triticum aestivum* L.) cultivars to identify cultivars suitable as reference material for wheat gluten analysis. *Food Chem.* 267, 387–394. doi: 10.1016/j.foodchem.2017.05.005

Huang, X., Kanerva, P., Salovaara, H., Stoddard, F. L., and Sontag-Strohm, T. (2017). Proposal for C-hordein as reference material in gluten quantification. *J. Agric. Food Chem.* 65, 2155–2161. doi: 10.1021/acs.jafc.6b05061

Koehler, P., and Wieser, H. (2013). "Chemistry of cereal grains," in *Handbook on Sourdough Biotechnology*, eds M. Gobbetti, and M. Gänzle, (Boston, MA: Springer).

Leonard, M. M., Sapone, A., Catassi, C., and Fasano, A. (2017). Celiac disease and nonceliac gluten sensitivity: a review. *JAMA J. Am. Med. Assoc.* 318, 647–656. doi: 10.1001/jama.2017.9730

Lexhaller, B., Colgrave, M. L., and Scherf, K. A. (2019). Characterization and relative quantitation of wheat, rye, and barley gluten protein types by liquid chromatography-tandem mass spectrometry. *Front. Plant Sci.* 10:1530. doi: 10.3389/fpls.2019.01530

Lexhaller, B., Tompos, C., and Scherf, K. A. (2016). Comparative analysis of prolamin and glutelin fractions from wheat, rye, and barley with five sandwich ELISA test kits. *Anal. Bioanal. Chem.* 408, 6093–6104. doi: 10.1007/s00216-016-9721-7

Lexhaller, B., Tompos, C., and Scherf, K. A. (2017). Fundamental study on reactivities of gluten protein types from wheat, rye and barley with five sandwich ELISA test kits. *Food Chem.* 237, 320–330. doi: 10.1016/j.foodchem.2017.05.121

Marchetti, L., Cardós, M., Campaña, L., and Ferrero, C. (2012). Effect of glutens of different quality on dough characteristics and breadmaking performance. *LWT Food Sci. Technol.* 46, 224–231. doi: 10.1016/j.lwt.2011.10.002

MSZ 6367-9:1989, (1989). *Edible, Fodder and Industrial Seeds and Husked Products. Quality Tests. Production of Wheat Flour in Laboratory*. Budapest: Hungarian Standard Institution.

MSZ En Iso 16634-2:2016, (2016). *Food Products. Determination of the Total Nitrogen Content by Combustion According to the Dumas Principle and Calculation of the Crude protein Content. Part 2: Cereals, Pulses and Milled Cereal Products*. Budapest: Hungarian Standard Institution.

Mujico, J. R., Lombardía, M., Mena, M. C., Méndez, E., and Albar, J. P. (2011). A highly sensitive real-time PCR system for quantification of wheat contamination in gluten-free food for celiac patients. *Food Chem.* 128, 795–801. doi: 10.1016/j.foodchem.2011.03.061

Ortolan, F., and Steel, C. J. (2017). Protein characteristics that affect the quality of vital wheat gluten to be used in baking: a review. *Compr. Rev. Food Sci. Food Saf.* 16, 369–381. doi: 10.1111/1541-4337.12259

Osborne, T. B. (1907). *). The Proteins of Wheat Kernel*. Washington, DC: Carnegie Institute of Washington.

Poms, R. (2006). "Reference materials and method validation in allergen detection," in *Detecting Allergens in Food*, eds J. Koppelman, and S. L. Helfe, (Sawston: Woodhead Publishing).

Rallabhandi, P., Sharma, G. M., Pereira, M., and Williams, K. M. (2015). Immunological characterization of the gluten fractions and their hydrolysates from wheat, rye and barley. *J. Agric. Food Chem.* 63, 1825–1832. doi: 10.1021/jf505716p

Rzychon, M., Brohée, M., Cordeiro, F., Haraszi, R., Ulberth, F., and O'Connor, G. (2017). The feasibility of harmonising gluten ELISA measurements. *Food Chem.* 234, 144–154. doi: 10.1016/j.foodchem.2017.04.092

Saulnier, L., Andersson, R., and Åman, P. (1997). A study of the polysaccharide components in gluten. *J. Cereal Sci.* 25, 121–127. doi: 10.1006/jcrs.1996.0077

Schalk, K., Koehler, P., and Scherf, K. A. (2018). Targeted liquid chromatography tandem mass spectrometry to quantitate wheat gluten using well-defined reference proteins. *PLoS One* 13:e0192804. doi: 10.1371/journal.pone.0192804

Schalk, K., Lexhaller, B., Koehler, P., and Scherf, K. A. (2017). Isolation and characterization of gluten protein types from wheat, rye, barley and oats for use as reference materials. *PLoS One* 24:e0172819. doi: 10.1371/journal.pone.0172819

Schall, E., Scherf, K. A., Bugyi, Z., Hajas, L., Török, K., Koehler, P., et al. (2020). Characterisation and comparison of selected wheat (*Triticum aestivum* L.) cultivars and their blends to develop a gluten reference material. *Food Chem.* 313:126049. doi: 10.1016/j.foodchem.2019.126049

Scherf, K. A. (2017). Gluten Analysis of wheat starches with seven commercial ELISA Test kits-up to six different values. *Food Anal. Methods* 10, 234–246. doi: 10.1007/s12161-016-0573-8

Scherf, K. A., and Poms, R. E. (2016). Recent developments in analytical methods for tracing gluten. *J. Cereal Sci.* 67, 112–122. doi: 10.1016/j.jcs.2015.08.006

Schwalb, T., Wieser, H., and Koehler, P. (2011). "Comparison of different protein references and ELISA kits for the detection of gluten in foods," in *Proceedings of the 24th Workshop on Prolamin Analysis and Toxicity. Zwickau*, ed. M. Stern, Germany.

Shan, L., Molberg, Ø, Parrot, I., Hausch, F., Filiz, F., Gray, G. M., et al. (2002). Structural basis for gluten intolerance in celiac sprue. *Science* 297, 2275–2279. doi: 10.1126/science.1074129

Shewry, P. R., and Tatham, A. S. (2016). Improving wheat to remove coeliac epitopes but retain functionality. *J. Cereal Sci.* 67, 12–21. doi: 10.1016/j.jcs.2015.06.005

Sollid, L. M., Tye-Din, J. A., Qiao, S.-W., Anderson, R. P., Gianfrani, C., and Konig, F. (2020). Update 2020: nomenclature and listing of celiac disease-relevant gluten epitopes recognized by CD4+ T cells. *Immunogenetics* 72, 85–88. doi: 10.1007/s00251-019-01141-w

Tatham, A. S., and Shewry, P. R. (2012). The S-poor prolamins of wheat, barley and rye: Revisited. *J. Cereal Sci.* 55, 79–99. doi: 10.1016/j.jcs.2011.10.013

Van Der Borght, A., Goesaert, H., Veraverbeke, W. S., and Delcour, J. A. (2005). Fractionation of wheat and wheat flour into starch and gluten: overview of the main processes and the factors involved. *J. Cereal Sci.* 41, 221–237. doi: 10.1016/j.jcs.2004.09.008

van Eckert, R., Berghofer, E., Ciclitira, P. J., Chirdo, F., Denery-Papini, S., Ellis, H. J., et al. (2006). Towards a new gliadin reference material-isolation and characterisation. *J. Cereal Sci.* 43, 331–341. doi: 10.1016/j.jcs.2005.12.009

van Eckert, R., Bond, J., Rawson, P., Klein, C. L., Stern, M., and Jordan, T. W. (2010). Reactivity of gluten detecting monoclonal antibodies to a gliadin reference material. *J. Cereal Sci.* 51, 198–204. doi: 10.1016/j.jcs.2009.11.012

Wang, L., and Flores, R. A. (1999). The effects of storage on flour quality and baking performance. *Food Rev. Int.* 15, 215–234. doi: 10.1080/87559129909541187

Wieser, H., Antes, S., and Seilmeier, W. (1998). Quantitative determination of gluten protein types in wheat flour by reversed phase HPLC. *Cereal Chem.* 75, 644–650. doi: 10.1094/cchem.1998.75.5.644

Wieser, H., and Koehler, P. (2009). Is the calculation of the gluten content by multiplying the prolamin content by a factor of 2 valid? *Eur. Food Res. Technol.* 229, 9–13. doi: 10.1007/s00217-009-1020-5

Conflict of Interest: PK was employed by the company Biotask AG (Esslingen am Neckar, Germany).

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 17 July 2020
doi: 10.3389/fnut.2020.00087

Mapping Coeliac Toxic Motifs in the Prolamin Seed Storage Proteins of Barley, Rye, and Oats Using a Curated Sequence Database

Matthew Daly[1], *Sophie N. Bromilow*[1], *Chiara Nitride*[1], *Peter R. Shewry*[2], *Lee A. Gethings*[3] *and E. N. Clare Mills*[1]*

[1] Division of Infection, Immunity and Respiratory Medicine, Faculty of Biology, Medicine and Health, Manchester Institute of Biotechnology, University of Manchester, Manchester, United Kingdom, [2] Centre for Crop Genetic Improvement, Rothamsted Research, Harpenden, United Kingdom, [3] Waters Corporation, Wilmslow, United Kingdom

OPEN ACCESS

Edited by:
Katharina Anne Scherf, Karlsruhe Institute of Technology (KIT), Germany

Reviewed by:
Stefania Masci, University of Tuscia, Italy
Michelle Lisa Colgrave, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia
Vera Muccilli, University of Catania, Italy

***Correspondence:**
E. N. Clare Mills
clare.mills@manchester.ac.uk

Specialty section:
This article was submitted to Nutrition and Food Science Technology, a section of the journal Frontiers in Nutrition

Received: 31 May 2019
Accepted: 12 May 2020
Published: 17 July 2020

Citation:
Daly M, Bromilow SN, Nitride C, Shewry PR, Gethings LA and Mills ENC (2020) Mapping Coeliac Toxic Motifs in the Prolamin Seed Storage Proteins of Barley, Rye, and Oats Using a Curated Sequence Database. Front. Nutr. 7:87. doi: 10.3389/fnut.2020.00087

Wheat gluten, and related prolamin proteins in rye, barley and oats cause the immune-mediated gluten intolerance syndrome, coeliac disease. Foods labelled as gluten-free which can be safely consumed by coeliac patients, must not contain gluten above a level of 20 mg/Kg. Current immunoassay methods for detection of gluten can give conflicting results and may underestimate levels of gluten in foods. Mass spectrometry methods have great potential as an orthogonal method, but require curated protein sequence databases to support method development. The GluPro database has been updated to include avenin-like sequences from bread wheat (n = 685; GluPro v1.1) and genes from the sequenced wheat genome (n = 699; GluPro v 1.2) and *Triticum turgidum* ssp durum (n = 210; GluPro v 2.1). Companion databases have been developed for prolamin sequences from barley (n = 64; GluPro v 3.0), rye (n = 41; GluPro v 4.0), and oats (n = 27; GluPro v 5.0) and combined to provide a complete cereal prolamin database, GluPro v 6.1 comprising 1,041 sequences. Analysis of the coeliac toxic motifs in the curated sequences showed that they were absent from the minor avenin-like proteins in bread and durum wheat and barley, unlike the related avenin proteins from oats. A comparison of prolamin proteins from the different cereal species also showed α- and γ-gliadins in bread and durum wheat, and the sulphur poor prolamins in all cereals had the highest density of coeliac toxic motifs. Analysis of ion-mobility mass spectrometry data for bread wheat (cvs Chinese Spring and Hereward) showed an increased number of identifications when using the GluPro v1.0, 1.1 and 1.2 databases compared to the limited number of verified sequences bread wheat sequences in reviewed UniProt. This family of databases will provide a basis for proteomic profiling of gluten proteins from all the gluten containing cereals and support identification of specific peptide markers for use in development of new methods for gluten quantitation based on coeliac toxic motifs found in all relevant cereal species.

Keywords: gluten, sequence database, barley, rye, oats, coeliac disease, wheat

INTRODUCTION

Wheat is one of the most important crop globally, with the combined production with related cereal species (barley (*Hordeum vulgare*), rye (*Secale cereale*), and oats (A*vena sativa*) exceeding ∼95,026 million tonnes in 2017 (1). The major storage protein fractions in cereal grains are defined as prolamins based on their solubility in mixtures of alcohol and water and their high contents of glutamine and proline. These proteins account for up to 80% of total protein content in wheat, barley and rye (2, 3) but are relatively minor components in oats (4). The gluten proteins of wheat form a visco-elastic network when wheat flour is mixed with water, which enables the production of leavened bread and other products (including pasta and noodles). Although these properties are not shared by the prolamins in related cereals (barley, rye and oats), restricting the use of these cereals in food processing, their sequences are related to those of wheat gluten proteins. Consequently, although the term gluten strictly applies only to wheat prolamins, it is defined in a regulatory context 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" (5).

Cereal seed storage prolamins can be distinguished based on their solubility in aqueous alcohol mixtures as either alcohol-soluble monomeric prolamins or alcohol-insoluble polymeric glutenins (6, 7). The monomeric prolamins can be further classified into α-,γ-, and ω-types based on their electrophoretic mobility whilst the components of the polymeric fractions, can be classified after reduction as belonging to either high molecular weight (HMW) and low molecular weight (LMW) groups (8). The prolamins from different cereal species are termed as either gliadins (wheat), hordeins (barley), or secalins (rye). A further group, originally identified in oats, are called avenins and have previously been classified either into three groups termed α-, β-, and γ-avenins according to electrophoretic mobility at low pH (9) or into eleven groups termed Avn-1-1 to Avn-10 based elution profiles from ion-exchange chromatography followed by RP-HPLC (10). In addition, molecular approaches have been used to classify them into A-, B-, and C-avenins, based on their repetitive domain structure (11). Subsequently sequences encoding proteins related to oat avenins have been identified in bread wheat (12), *T. turgidum* ssp durum (13) and barley (14). Based on sequence homology these have been called "avenin-like" proteins, and have been classified in wheat as being either a or b type avenins, with different subtypes indicated by Arabic numerals (12); it has also been proposed that the avenin-like protiens from wheat be termed farinins (15). They have also been shown to have a positive effect on dough strength in bread wheat (16) as well as pathogen resistance (17). The prolamin seed storage proteins are also important because of their ability to elicit both IgE- and non-IgE immune mediated adverse reactions in some individuals. Coeliac disease is a non-IgE immune-mediated food intolerance, affecting ∼1% of the global population (18) and is triggered by prolamin seed storage proteins present in some cereal grains; wheat, barley, rye and, in some patient populations, oats (18, 19). Ingestion of dietary gluten leads to a variety of symptoms in susceptible individuals such as diarrhoea, abdominal distension, villous atrophy and an increased risk of adenocarcinoma and lymphoma (20). As a consequence of their high contents of proline, these prolamin seed storage proteins are partially resistant to gastric, pancreatic and brush border proteases resulting in longer peptide fragments reaching the small intestinal mucosa. Following the action of tissue transglutaminase (tTG) in the gut epithelium, which deamidates glutamine residues, some of these digestion-resistant fragments contain nine amino acid residue motifs capable of binding to certain variants of the Human Leukocyte Antigen class II receptors, HLA-DQ2 and HLA-DQ8. In addition to stimulating the production of antibodies to both tTG and gluten, the peptides activate gluten-specific naïve $CD4^+$ T cells leading to an inflammatory response that causes the gut mucosa to flatten, reducing its absorptive capacity. These T cell epitopes have been termed coeliac toxic motifs (21, 22). Although the number of coeliac toxic motifs in a protein fragment can be correlated to its immunotoxicity, there are many other factors involved. These include resistance to gastrointestinal digestion, how effective peptides are as substrates for tTG as well as the binding affinity for HLA and capacity to activate T cells. Indeed, there is correlation between the likelihood of a sequence being deamidated by tTG and its ability to activate T cells in individuals with coeliac disease (23, 24). By contrast IgE-mediated food allergies have been associated with sensitisation to particular cereal storage prolamins including wheat-dependent exercise-induced anaphylaxis (WDEIA) a condition associated with sensitisation to ω5-gliadins (also known as Tri a 19). Sensitisation to other seed storage has been described including α- and γ- gliadins, LMW and HMW subunits of glutenin [Tri a 20, 21, 26, and 36; (25, 26)] together with non-gluten proteins, notably the non-specific lipid transfer protein (LTP; Tri a 14).

No cures exist for either coeliac disease, or IgE-mediated food allergies, and the only treatment is strict avoidance of gluten or wheat-containing foods. In order to help patients with coeliac disease avoid gluten the CODEX Alimentarius Commission developed recommendations for gluten-free foods which has been implemented in regulations across the world (27). In the EU, if cereal-derived food ingredients (such as wheat starch or dextrin) contain <20 mg/Kg they can be labelled as gluten-free, although wheat must still be declared on the ingredient label (28, 29). The available validated methods for gluten quantification are immuno-based assays, which suffer from several limitations and can lead to false detection and quantification. The high sequence homology between prolamins in cereal species can cause partial reactivity of the antibodies to wheat, barley, rye and oats, and the potential reactivity with contaminating wild grass species. Moreover, incomplete extraction of proteins and the use of incorrect conversion factors can further compound these issues (30–33).

An alternative to immunoassays is mass spectrometry, which has been used as an orthogonal method of quantifying gluten in complex matrices (34–37). However, accurate identification of proteins using mass spectrometry-based proteomics approaches relies heavily on the quality of the protein database or annotated genome against which the mass spectra are searched. Various

databases are available such as UniProt containing both reviewed (Swiss-Prot) and unreviewed (TrEMBL) protein sequences (38), and the NCBI Protein Database (39). Although curated and draft genomes are available for some plant species, including wheat, barley and rye (40–42). These are inevitably cultivar specific, can be incomplete and often contain partial sequences. Furthermore, the reviewed UniProtKB/SwissProt database contains only 56 prolamin sequences combined from bread wheat, *Triticum turgidum* ssp durum, barley, rye, and oats. Some of these originate from protein sequencing and are not complete protein sequences [e.g., UniProt sequence accession Q09095; (43)]. In order to reduce redundancy in the database UniProtKB/Swiss-Prot the protein produced from a single gene at a species level, is provided as a single entry choosing a canonical sequence based on at set of criteria, one of which is sequence length, with isoforms being provided as alternative sequences under the main entry (44, 45). This curation process means that the number of prolamin sequences in reviewed UniProt has reduced from 61 (accessed 14.5.2019) to 56 accessed 5.12.2019).

An alternative is to create custom databases combining reported protein sequences from other databases such as NCBI and EST sequences in order to facilitate proteomic analysis, although these are not all publicly available (34). One publicly available curated prolamin sequence database is ProPepper, a tool containing ∼2,480 cereal prolamin sequences data (46) although the sequences are not available in a format suitable for direct mining of mass spectrometry data. Other repositories are of curated sequences implicated in IgE-mediated allergies (47) and include the IUIS allergen nomenclature database which seeks to curate well-defined allergen sequences and has 40 sequences from wheat, barley and rye, although they include both inhalant and food allergens (48). Another curated allergen sequence database is AllergenOnline, which contains 2,129 peer-reviewed sequences (49). Such allergen sequence databases are of limited usefulness in searching mass spectrometric data since they are not comprehensive for a given organism and can use conflicting nomenclature. For example, the allergen Tri a 20 is referred to a γ-gliadin in the IUIS database which includes two accessions, but a further six sequence accessions are classified as Tri a 20 in AllergenOnline ver 19.

In order to address the need for a curated sequence database to facilitate analysis of proteomic data, the GluPro database was created containing 630 discrete unique full length bread wheat prolamin protein sequences encompassing both the gliadin and glutenin fraction and applied to characterisation of the bread wheat prolamin proteome (50). However, it does not include the avenin-like sequences from bread wheat and sequences from the wheat genome (cv Chinese Spring) which limits its utility. The sequence database has now been enlarged with avenin sequences to give GluPro v1.1 and further enriched with wheat genome sequences to give GluPro v 1.2. In addition the informatics pipeline developed by Bromilow and co-workers (50) has been applied to develop curated prolamin sequences from other cereal species including pasta wheat (*Triticum turgidum* ssp durum; GluPro v 2.1), barley (GluPro v 3.0), rye (GluPro v 4.0), and oats (GluPro v 5.0). These sequence sets were then compiled into a compendium database of gluten proteins from different cereal species (GluPro v 6.1). The resulting curated sequences were then analysed to determine the distribution of known coeliac toxic motifs using the AllergenOnline Celiac Disease (CD) Novel Protein Risk Assessment Tool (http://www.allergenonline.org/celiachome.shtml) (49). The expanded GluPro v 6.1 database will enable discovery proteomics data to be mined more effectively, in order to identify effective peptide markers. These are required for development of targeted, quantitative mass spectrometry methods for determination of gluten in food, which may originate from bread wheat, *T. turgidum* ssp durum, barley, rye and oats.

MATERIALS AND METHODS

Methods

Database Construction

Sequence sets of seed storage prolamins from *T. turgidum* ssp durum (GluPro v 2.0), *H. vulgare* (GluPro v 3.0), *S. cereale* (GluPro v 4.0) and *A. sativa* (GluPro v 5.0) were created independently and an update of the bread wheat (*T. aestivum*) database was undertaken to enrich it with avenin-like sequences (GluPro v 1.1) (**Figures S1, S2**).

In stage I the entire UniProt (accessed 04.01.2019 for GluPro v 3.0, 4.0 and 5.0, and 29.07.2019 for GluPro v 2.0) and NCBI Protein (accessed 12.02.2019 for GluPro v 3.0, 4.0 and 5.0, and 29.07.2019 for GluPro v 2.0) databases were mined using the search terms; "prolamin," "gluten," "gliadin," "glutenin," "hordein," "secalin," or "avenin" using the origin species set to either "*Triticum turgidum* ssp durum," "*Hordeum vulgare,*" "*Secale cereale,*" or "*Avena sativa.*" When populating the GluPro v 1.1 sequence set, the search term was "avenin" and the origin species was set to "*Triticum aestivum.*" In each case, all sequences were downloaded in FASTA format and combined into origin species-specific sequence sets. Redundant sequences were removed using the DB Toolkit software (51) with UniProt accessions being preferentially retained. Partial, non-seed storage prolamins and sequences containing ambiguous amino acids were then removed from the databases manually (sequence set one) if they lacked homology to reviewed seed storage prolamin sequences (8, 52). This was done, as although the sequence may have some protein level evidence, identifying these proteins experimentally using shotgun proteomics would not be possible. "X" denotes ambiguous amino acids in protein sequences; they arise due to either the presence of multiple sequences showing different amino acids, or poor quality data that is unable to distinguish between amino acids (53, 54).

In Stage II the curated sequence sets for each cereal species were then separately searched against the entire UniProt database (01.03.2019 for GluPro v 3.0, 4.0 and 5.0, and 19.11.2019 for GluPro v 2.0) using protein-protein BLAST (Basic Local Alignment Search Tool). Based on a minimum sequence homology of ∼30% the first 250 sequences were downloaded regardless of origin species. This was below the 40% threshold Addou et al. (55) suggested for inferring homology and was chosen to ensure that all homologous proteins were recovered from searching which were then manually curated (see below). The sequences curated in Stage I and II were combined and

subjected to another round of curation removing duplicates and partial sequences (**Figure S1**) to give databases for bread wheat (GluPro v 1.1), *T. turgidum* ssp durum (GluPro v 2.0), barley (GluPro v 3.0), rye (GluPro v 4.0), and oats (GluPro v 5.0). The species-specific sequence databases were then combined to give a complete seed storage prolamin sequence database (GluPro v 6).

In Stage III the recently published reference genome for *T. aestivum* cv. Chinese Spring (41) and draft genome available for *T. turgidum* ssp durum cv. Svevo (42), were then mined for further prolamin seed storage protein sequences (**Figure S2**). This was not necessary for barley as its draft genome (cv. Morex) is available as a reference proteome on UniProt and sequences from this translated genome were downloaded during creation and curation of GluPro v 3.0. Translated genomes of *T. aestivum* and *T. turgidum* ssp durum were downloaded from Ensembl Plants in FASTA format yielding 133,346 and 196,105 peptide/protein sequences, respectively, for each species. These files were then converted to BLAST databases using the standalone BLAST+ software (56) and the entire GluPro v 6 BLAST searched against them using Genome Workbench v 3.1.0 (57) with an Expect value of 10. After further manual curation (as described for Stages I and II) novel sequences were added to the respective species-specific database to give GluPro v 1.2 and GluPro v 2.1, respectively. These were then added to GluPro v 6 to give GluPro v 6.1. Although a draft genome is available for *S. cereale* cv. Lo7, it is unavailable for download in a translated format (58). However, a BLAST server of the transciptome is available at http://webblast.ipk-gatersleben.de/ryeselect/ (accessed 12.11.2019). Therefore, the GluPro v 6.0 database was BLAST searched against this transcriptome using an Expect value of 10, and homologous sequences were retrieved and manually curated. Where possible, transcript identifiers were replaced with UniProt accession numbers.

Sequence Alignment and Analysis

Sequences were aligned using Clustal Omega (59) and resulting alignments downloaded in Multiple Sequence File (MSF) format and visualised in Jalview (60). A phylogenetic tree was created in Jalview based on average distance (a type of unweighted pair group method with arithmetic mean) and BLOSUM62, viewed and edited in FigTree (v1.4.3). Phylogenetic tree building was undertaken using average distance rather than approaches such as neighbour-joining, as an equal rate of evolution was assumed i.e., a molecular clock. This analysis was only used to cluster proteins into their respective protein groups and not to determine evolutionary origin. Resulting sequence classifications were manually cross-referenced based on available literature regarding N-terminal sequence, mass, repeat sequence and phylogeny (8, 11, 61, 62). Sequences classified as being within the same protein group from the same species were subject to multiple pairwise alignments such that every sequence was compared to every other sequence and average percentage homology calculated (**Tables S1–S4**). Master sequences with protein level evidence were identified where possible for each protein group from each species that represented that protein roup.

Mapping of Coeliac Toxic Motifs

Sequences present in the databases were further analysed with regard to the distribution of coeliac toxic motifs using the online database AllergenOnline (49) that contained 1,013 coeliac toxic peptide sequences at the time of analysis (11.03.2019 and 02.12.2019). It should be noted that some of these peptides are fragments of others and therefore not unique. Using the "Exact Peptide Match" function all 1,013 peptides available were mapped against the full sequences from the curated databases. From this function three measurements were taken: number of unique coeliac toxic motifs per sequence, density of unique coeliac toxic motifs and sequence coverage by coeliac toxic motifs as a percentage of total sequence length. The number of unique coeliac toxic motifs was simply the number of motifs that were present in the sequence, although this excluded instances where unique motifs occurred more than once in the sequence and is irrespective of that fact that some motifs are fragments of others. The density of unique coeliac toxic motifs was calculated by taking the number of unique coeliac toxic motifs present in the sequence and dividing by the sequence length. Sequence coverage by coeliac toxic motifs was calculated using Protein Coverage Summarizer software (v1.3.6794) where all 1,013 sequences in the AllergenOnline CD Tool were mapped against the sequences. This calculation ignores the fact that some sequences present in the AllergenOnline CD Tool are fragments of each other.

Mass Spectrometry Analysis

Seed from *T. aestivum* (cultivars Chinese Spring and Hereward) were obtained from Rothamsted Research (Harpenden, UK), two grains crushed separately and proteins extracted with 50 mM Tris HCl (pH 8.8), 50 mM DTT and 0.02% (w/v) RapiGest^{TM} at 60°C with sonication and vortexing every 5 min (50). Extracts were clarified by centrifugation for 10 min at 10,000 × g, supernatants removed and then further reduced, alkylated with iodoacetamide and digested with chymotrypsin as previously described (50). Resulting peptides were desalted and concentrated using C18 ZipTips (Waters Corporation, Wilmslow, UK). Peptides were subsequently analysed using liquid chromatography ion mobility mass spectrometry (LC-IM-MS-MS). For the chromatography the mobile phases of solvent A consisted [0.1% (v/v) formic acid/99.9% (v/v) water] and solvent B consisted [0.1% (v/v) formic acid/99.9% (v/v) acetonitrile]. Chromatographic separation was undertaken using a linear gradient (flow rate 300 nL/min) from 3 to 40% (v/v) solvent B over 90 min using a M-class ACQUITY UPLC system (Waters Corporation) equipped with a NanoEase 1.8 μm HSS T3 C18 (75 μm × 150 mm) column (Waters Corporation) attached to a SYNAPT G2-Si QTOF mass spectrometer (Waters Corporation). Data were acquired using a data independent approach in positive ion mode over the mass range m/z 50–2,000 with a 0.5 s spectral acquisition time and one cycle of low and elevated energy data was acquired every 1 s (50).

Analysis of Mass Spectrometric Data

IM-MS-MS data were processed using Progenesis QI for Proteomics (v 3) using the Ion Accounting workflow. After

ion detection, low- and high-energy mass events are time-aligned to precursor-product ion tables, and then filtered to remove any precursor ions under 750 Da and all product ions under 350 Da. A searchable database is then selected and a reversed decoy database is appended, and the algorithm completes a pre-search step where, using Bayesian inference, model parameters are adjusted and fine-tuned. The algorithm then completes several passes of database searching to match theoretical peptides to observed mass events. This iterative process of peptide spectrum matching can improve the number of peptides identified from IM-MS-MS compared to other mass spectrometry database search programs such as Mascot and ProteinLynxGlobalSERVER (63). Once imported, sequence sets were searched against the GluPro v 1, 1.1, and 1.2 databases, and reviewed prolamin sequences from *T. aestivum* downloaded from UniProt (downloaded 20.01.2019). Cleavage was set to chymotrypsin with cleavage occurring at Y, W, F or L unless followed by a P with up to two missed cleavages. False discovery rate (FDR) was set to 1% and mass tolerance for peptide and fragment ions were set to 10 and 20 ppm, respectively. The distribution of q-values obtained for all analyses is shown in **Figure S7** with only identifications with q values $\leq$0.01 being considered. Apex 3D parameters were set to 150 counts for low energy intensity threshold and 30 counts for high energy. Carbidomethylation of cysteine was selected as a fixed modification, whereas oxidation of methionine, hydroxylation of proline, deamidation of glutamine or asparagine and N-terminal pyroglutamatic acid were all selected as variable modifications. Protein identifications were only considered valid if at least one unique peptide was identified for that protein in at least 2/3 technical replicates in both biological replicates, and with a peptide score >5.

RESULTS

Database Construction and Sequence Classification

Initially the GluPro v 1.0 database was enriched with avenin-like sequences from *T. aestivum*. A total of 11,917 sequences were downloaded from both UniProt and NCBI Protein databases. Additional prolamin seed storage protein sequence databases were also developed for other cereal species including *T. turgidum* ssp durum, barley, rye and oats (**Figures S1, S2**). The majority of these sequences were duplicates as the different search terms may return the same protein. For example, the protein with UniProt accession P06470 was returned when searched for "gliadin," "glutenin," and "hordein" and was therefore downloaded three times. As such, all databases were reduced to ∼1% of the original size once duplicates had been removed. These included sequences with the same accession number and the same sequence with different accession numbers that have been deposited in the UniProt and NCBI databases more than once. BLAST searching of sequences identified eight avenin-like sequences from bread wheat, 10 sequences from *T. turgidum* ssp durum (two HMW glutenin subunits, four LMW glutenin subunits and four α-gliadin sequences), two C hordeins from barley, and two ω-secalins from rye. Once mis-assigned sequences, partial sequences and sequences with ambiguous amino acids were removed the databases comprised 182 sequences (*T. turgidum* ssp durum; GluPro v 2.0; **Table S1**), 64 (barley; GluPro v 3.0; **Table S2**), 41 (rye; GluPro v 4.0; **Table S3**) and 27 sequences (oats; GluPro v 5.0; **Table S4**), respectively. Fifty-five sequences attributed to avenin-like proteins from *T. aestivum* were added to the original Glu Pro v 1.0 prolamin sequence set (*T. aestivum*; GluPro v 1.1). These were combined to give a more complete "cereals containing gluten" database comprising 998 sequences (Glu Pro v 6).

Mining of the Chinese Spring wheat genome yielded 14 new sequences; nine α-gliadins, three further avenin-like sequences and two δ-gliadin sequences recently reported by Altenbach et al. (64). Interestingly no HMW glutenin subunit sequences were present in the translated genome. Interrogation of the cDNA database indicated five HMW glutenin subunit sequences (Ax, Bx, Dx, By, and Dy), however these were annotated as non-translating CDS and as such did not appear in the translated genome. Four sequences contained "N" in the sequence indicating an unknown nucleotide, and one sequence encoded a protein only 340 amino acids in length. Three HMW glutenin subunit sequences were identified from another sequenced genome using *T. aestivum* cv Chinese Spring (65). Twenty-eight sequences were also added by mining the *T. turgidum* ssp durum translated genome including α-gliadin, avenin-like and low molecular weight glutenin subunit protein sequences. These were added to the bread wheat and *T. turgidum* ssp durum databases GluPro 1.1 and GluPro 2.0 databases to create GluPro 1.2 and 2.1, respectively (**Figure S2**). These were then combined with GluPro v 3.0-5.0 to create GluPro v 6.1 containing 1041 sequences, an increase of 4.2%. It was not necessary to mine the *H. vulgare* cv. Morex translated genome, as it is already available as a reference proteome in UniProt, eighteen sequences in the barley database GluPro 3.0 having originated from the sequenced genome. The *S. cereale* translated genome was also mined but no sequences were identified that were not already present in the rye database GluPro v 4.0. UniProt accession numbers, evidence level and supporting literature for each sequence in the database are available in Supplementary Material (**Tables S1–S4**). In addition all the databases are available in FASTA format from https://figshare.com/search?q=10.6084%2Fm9.figshare.12613154.

Phylogenetic Analysis of Prolamin Sequences From Cereals Containing Gluten

Phylogenetic analysis of all sequences in GluPro v 6.1 revealed clustering into the expected protein groups between and within species similar to that observed previously for the original *T. aestivum* GluPro v 1.0 (50) (**Figure 1**). Briefly, proteins separated into seven groups; the sulphur-rich α-type prolamins, Low Molecular Weight (LMW) glutenin subunits, γ-type prolamins, avenin-like a, b and avenins, δ-type prolamins with γ3-hordeins, the sulphur-poor ω-type prolamins and finally the High Molecular Weight (HMW) glutenin subunits.

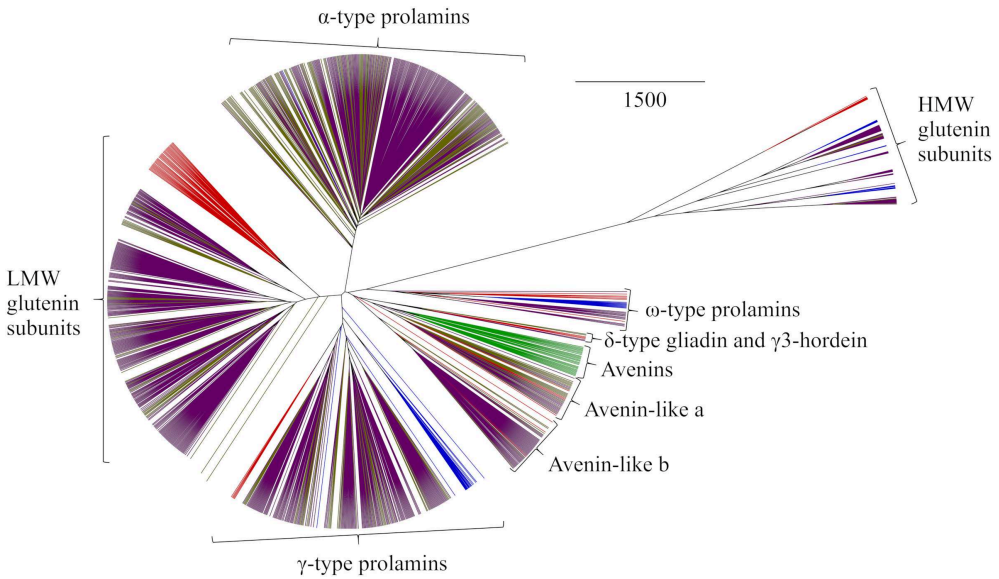


FIGURE 1 | Average distance phylogenetic tree of immature sequences from *T. aestivum* (purple), *T. turgidum* ssp durum (gold), *H. vulgare* (red), *S. cereale* (blue), and *A. sativa* (green). The scale bar indicates the number of amino acid substitutions per site.

The α-type prolamins are only present in bread wheat and *T. turgidum* ssp durum and therefore form a distinct branch on the phylogenetic tree with a single α-type prolamin sequence from rye being identified. This sequence was reported based on a cDNA sequence (66) and may be wrongly assigned or derived from Triticale since rye does not contain α-prolamin genes. Triticale (also called Triticosecale) is derived from hybridization of wheat and rye and therefore contain α-prolamins encoded by the Triticum genome (67). The phylogenetic analysis also revealed the known similarity of the polymeric LMW glutenin subunit types with the monomeric gliadin-like α- and γ-prolamins. LMW glutenin subunits from *T. turgidum* ssp durum clustered largely with sequences from bread wheat, the wheat sequences falling into seven groups which had characteristic N-terminal sequences including into the more phylogenetically distant LMW-i group; the B1 and B3 hordeins from barley also clustered alongside the LMW subunits of glutenin (50).

The sulphur-poor ω-type prolamins were more distantly related, the polymeric HMW subunits of glutenin being the most distantly related type of prolamin sequence. HMW glutenin subunits separate based on length, and in the case of bread wheat, the variation was linked to the chromosomal locations of the encoding genes. The HMW secalins and HMW glutenin subunits from *T. turgidum* ssp durum were less divergent then those from bread wheat. The lower level of variation in HMW subunits sequences in *T. turgidum* ssp durum and rye may relate to the fact that these species are tetraploid and diploid, respectively, whereas, bread wheat is hexaploid. However, the limited variation observed may simply be because fewer sequences were available from rye and *T. turgidum* ssp durum. The D-hordeins were more closely related to the y-type HMW glutenin subunits present in wheat and rye than to the x-type subunits of wheat. Three other types of prolamin were also identified using the phylogenetic analysis which clustered together with the avenins of oats and the avenin-like proteins in bread wheat, *T. turgidum* ssp durum, and barley. The avenins from oats all clustered on one branch with the avenin-like proteins from other cereal species falling into two other clusters corresponding to the avenin-like a and b groups previously identified in wheat (12). Phylogenetic analysis also allowed identification of the recently discovered δ-type prolamin present in bread wheat (64, 68, 69), and now also identified in *T. turgidum* ssp durum. Interestingly, these sequences clustered with the three γ3-hordein sequences from barley, and appear related to the avenins and avenin-like proteins, demonstrating the homologous nature of these proteins but further complicating nomenclature regarding the prolamins.

In order to interrogate the sequence relationships between the different types of prolamin, master sequences were identified for which protein level evidence existed (**Table 1**) and aligned C-terminal segments shown in **Figure S3**. Protein-level evidence was lacking for δ-gliadin and avenin-like proteins from *T. turgidum* ssp durum, certain avenin-like proteins from barley, α-type prolamin from rye and A-type avenin from oats. For these classes of prolamins candidate master sequences were selected with a proline plus glutamine content >30% to confirm they were prolamins and a high sequence homology to every other sequence in the protein group (**Table 1**). Within-protein group sequence homology between species was also high (>50%), further demonstrating the correct classification of these sequences. Extremely high homology (of 92.25 and 91.73%, respectively), was observed between the avenin-like a and b proteins from bread wheat, *T. turgidum* ssp durum and barley. This analysis also confirmed that, although avenins

TABLE 1 | Sequence similarity within protein groups between species.

Protein group	Origin species	UniProt accession number	Evidence level	Supporting literature reference	% Sequence similarity	Proline + glutamine (%)
α-type prolamins	*T. aestivum*	X2KVH9	Protein	(70)	84.15	48.67
	T. turgidum ssp durum	D2X6C9	Protein	(71)		49.46
	S. cereale	H8Y0F9	Genome	(66)		50.00
δ-type prolamins	*T. aestivum*	A0A2U8JD37	Protein	(64)	89.64	45.54
	T. turgidum ssp durum	A0A446IHB0	Genome	Manual submission L. Milanese Sep 2017		37.31
γ-type prolamins	*T. aestivum*	K7X1R6	Protein	(70)	65.29	50.55
	T. turgidum ssp durum	Q6EEW5	Protein	(71)		43.10
	H. vulgare	M0XYT2	Protein	(72)		47.22
		P17990	Protein	(72)		44.76
		P80198	Protein	(70)		44.98
	S. cereale	E5KZQ5	Protein	(70)		42.73
		E5KZP9	Protein	(70)		61.09
Group I avenins	*T. aestivum*	P0CZ07	Protein	(73)	92.25	35.00
	T. turgidum ssp durum	182970*	Genome	(42)		32.60
	H. vulgare	F2EGD5	Protein	(74)		31.82
Group II avenins	*T. aestivum*	P0CZ05	Protein	(73)	91.73	34.83
	T. turgidum ssp durum	A0A446WXS7	Genome	Manual submission L. Milanese Sep 2017		35.71
	H. vulgare	A7XUQ7	Genome	(75)		34.59
Group III avenins	*A. sativa*	L0L8A4	cDNA	(11)	62.98	35.29
		P80356	Protein	(70)		41.79
		Q09114	Protein	(70)		41.76
LMW glutenin subunits	*T. aestivum*	B2Y2S3	Protein	(70)	75.98	50.86
	T. turgidum ssp durum	A0A2P1BXV0	Protein	(71)		50.15
	H. vulgare	P06470	Protein	(70)		49.64
		I6TEV5	Protein	(70)		51.20
ω-type prolamins	*T. aestivum*	Q402I5	Protein	(70)	53.87	72.86
	H. vulgare	A0A287EIM7	Protein	(70)		69.98
	S. cereale	C4NFN9	Protein	(70)		68.64
HMW glutenin subunits	*T. aestivum*	G3FLC7	Protein	(70)	69.31	49.68
		Q94IJ6	Protein	(70)		43.78
	T. turgidum ssp durum	Q8RVX0	Protein	(71)		45.99
		A0A0E4G9A4	Protein	1(76)		44.40
	H. vulgare	Q84LE9	Protein	(70)		36.68
	S. cereale	Q94IK8	Protein	(70)		45.38
		Q94IL2	Protein	(70)		48.16

*Sequence similarity was calculated using pairwise alignment of master sequences, alongside UniProt accession number, protein group, origin species, evidence level with supporting literature reference and proline and glutamine percentage. Accession number indicated by *was retrieved from mining of the translated T. turgidum ssp durum genome (42).*

from oats were distinctly separated from the gliadins (**Figure 1**), that they are indeed prolamins, although their proline plus glutamine content is lower (32–42%) than other prolamin sequences (**Table 1**) (4). This is because the avenin proteins lack the long repetitive domains present in other prolamins, indicating that the coding genes could either be related to ancestral forms of seed storage genes that have since evolved a repetitive domain, or the result of a more recent evolution

that have removed the repetitive domain (77). The avenin-like sequences of *T. turgidum* ssp durum typically had a content of proline plus glutamine ranging from 22 to 34%. This lower level is due to a subset of avenin-like a proteins having shorter sequences, together with point mutations and deletions in the short polyglutamine region. The characteristic conserved skeleton of eight cysteine residues of the prolamin superfamily is demonstrated in all sequences apart from the ω-type prolamins and the HMW glutenin subunits (**Figure S3**) (78). The ω-type prolamins contain no cysteine residues and consist mostly of repeat motifs, and HMW glutenin subunits contain a longer central domain of repeat motifs that disrupts the characteristic cysteine residue backbone.

The relationships between the avenin-like proteins from the different cereal species were then analysed separately (**Figure 2A**). The avenin-like a proteins comprised sequences annotated as "avenin-like a" and "avenin-like" sequences from bread wheat and barley, respectively, and included "uncharacterised" protein sequences from *T. turgidum* ssp durum. The avenin-like b proteins, comprised protein sequences from bread wheat, *T. turgidum* ssp durum and barley that were annotated in Uniprot as being "avenin-like b" proteins.

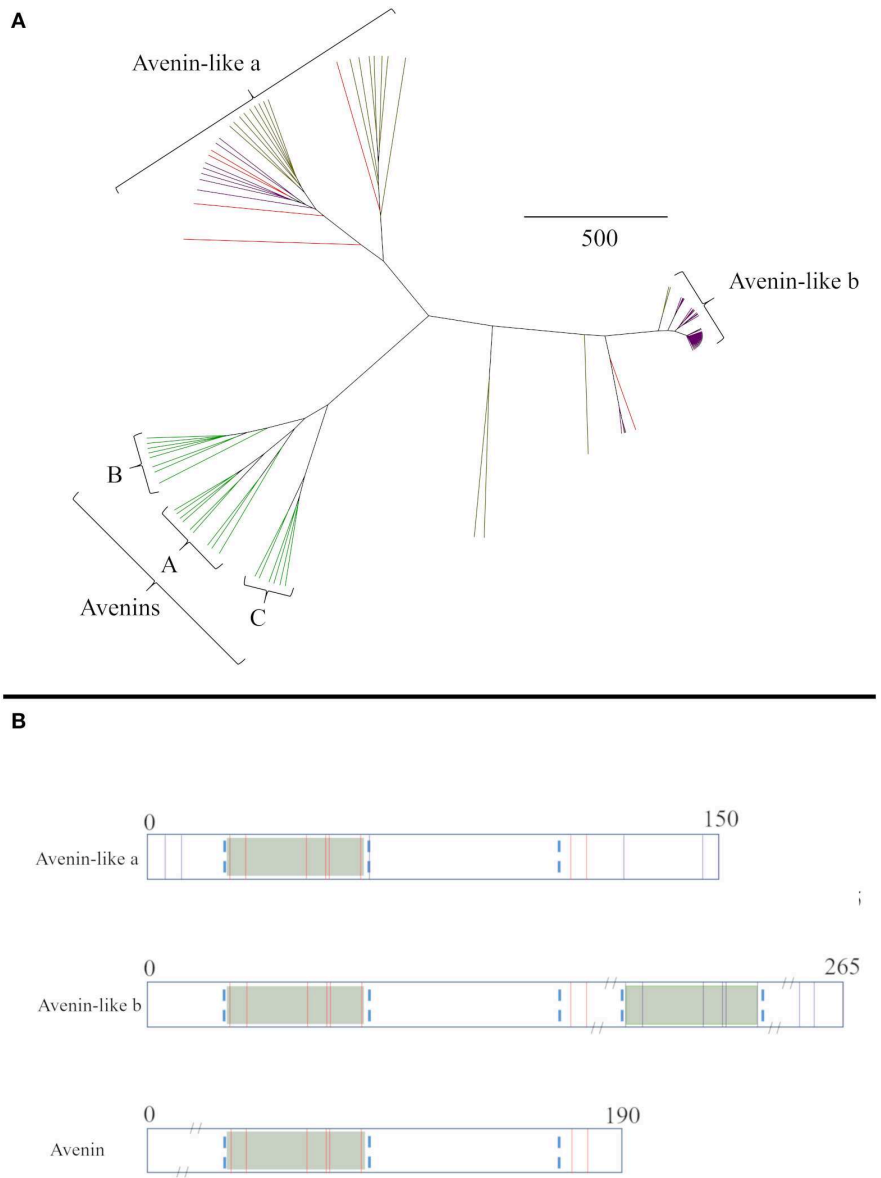


FIGURE 2 | (A) Average distance phylogenetic tree of mature avenin and avenin-like sequences from *T. aestivum* (purple), *H. vulgare* (red) and *A. sativa* (green). The scale bar indicates the number of amino acid substitutions per site. **(B)** Schematic depiction including sequence length and position of cysteine residues present in Group I, Group II and Group III avenins. Conserved cysteine residues between all three groups are coloured red and non-conserved are shown in purple. Conserved domains that contain the characteristic prolamin cysteine residue skeleton are distinguished by green boxes and regions are outlined by blue dashed lines.

This separation results from differences in amino acid sequence, with "avenin-like b" sequences containing four to five more cysteine residues than "avenin-like a" due to duplication of a polyglutamine region containing the conserved prolamin cysteine residue skeleton (**Figure 2B**) (12, 77, 79). Avenins from oats clustered separately from sequences from the other cereal species (**Figure 2A**). Three avenin-like proteins from bread wheat, one from *T. turgidum* ssp durum and one from barley did not cluster with the other avenin-like proteins or avenins but formed separate branches, closer to the avenin-like b proteins. Interrogation of aligned amino acid sequences indicated high sequence homology between these sequences and avenin-like b sequences. However, a deletion of seven amino acids at residue number 209 and a further two amino acid insertion at position 242 was observed which explains the distance seen between the sequences on the phylogenetic tree.

Individual species-specific phylogenetic trees provide further insights into the variations between the gluten components (**Figure 3**). Sequence homology was also determined within protein groups of the same species using all sequences available and is shown in Supplementary Material (**Tables S1–S4**). Similar to bread wheat, *T. turgidum* ssp durum prolamin sequences clustered into α-, δ- and γ-gliadins, low-molecular weight subunits of glutenin and avenin-like proteins together with the more distantly related HMW subunits of glutenin (**Figure 3A**). Interestingly, no ω-gliadins were identified in this organism despite the encoding regions being present on the short arm of all group 1 homoeologous chromosomes in bread wheat (80). Polypeptides with molecular weights consistent with ω-gliadins have been observed in electrophoretically separated extracts of *T. turgidum* ssp durum after immunoblotting with anti-ω5 gliadin antibodies (81). Peptide fragments of ω-gliadins have also been identified in simulated gastro-duodenal digests of pasta (82), although no sequences are available in either the UniProt or NCBI databases at present.

Analysis of the barley prolamin sequences allowed hordeins to be classified into avenin-like sequences, B1-, B3-, γ1-, γ2-, γ3- and C-hordeins together with the more distantly related D-hordeins (**Figure 3B**). Examination of the aligned γ-hordein sequences demonstrated that γ2-hordeins have extremely high homology to γ1-hordeins. Tanner et al. (83) suggested this is probably because γ2-hordeins are encoded by the γ1-hordein gene but have a post-translational deletion in the sequence. This results in proteins being expressed that are shorter by ∼30 amino acids although evidence to support this suggestion is currently lacking.

Rye secalins could be classified into α-prolamins (the single sequence referred to above), ω-, 40 k γ-, 75 k γ-secalins and two types of HMW secalin subunit (**Figure 3C**). Differences in mass

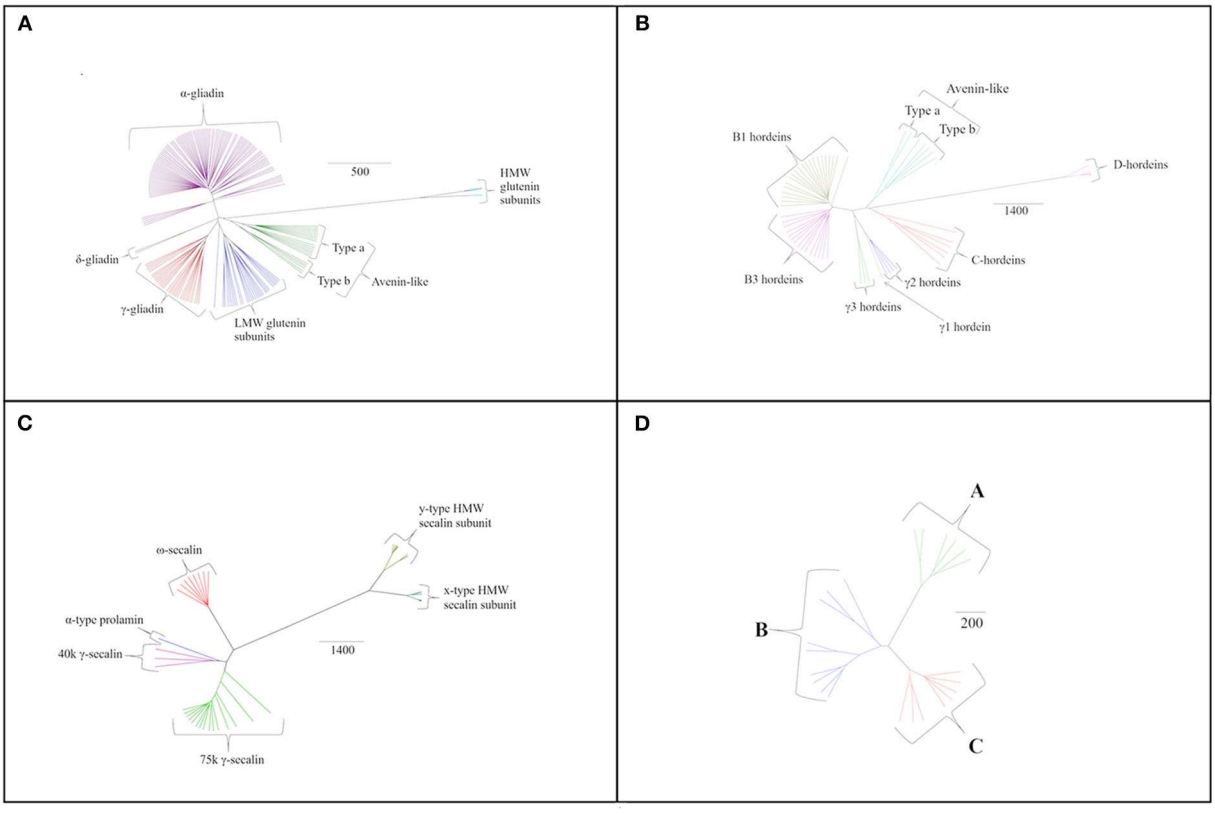


FIGURE 3 | Species specific phylogenetic trees based on average distance and BLOSUM62 with different protein grouping for **(A)** *T. turgidum* ssp durum, **(B)** *H. vulgare* sequences, **(C)** *S. cereale* sequences, and **(D)** *A. sativa* sequences. The scale bar indicates the number of amino acid substitutions per site.

TABLE 2 | Classification of avenin sequences from *Avena sativa* (Glu Pro 5.0 database).

Phylogenetic classification	UniProt accession number	Shewry (62) classification	Repeat motif region I	Repeat motif region II	No. of cysteine residues
Group A avenins	L0L6J7	Avenin-1-1, 1-2,−2,−4	PFM[$Q_{(1-5)}$]	No repeat	9
	L0L5H3	pAv10 genomic clone			
	Q09071	pAV10 genomic clone			
	L0L8A0	Avenin-1-1, 1-2,−2,−4			
	L0L833	Avenin-1-1, 1-2,−2,−4			
	L0L837	Avenin-1-1, 1-2,−2,−4			
	L0L8A4	Avenin-1-1, 1-2,−2,−4			
	I4EP78	Avenin-1-1, 1-2,−2,−4			
	I4EP85	Avenin-1-1, 1-2,−2,−4			
	I4EP86	Avenin-1-1, 1-2,−2,−4			
Group B avenins	I4EP88	Avenin-1-1, 1-2,−2,−4	No repeat	VFQPQLQQ	8
	Q38794	AV45-X1 genomic clone	MLL[$Q_{(3-6)}$]	FFQPQMQQ + VTQG	
	L0L4J1	Avenin-3			
	P80356	Avenin-3			
	L0L5I0	Avenin-3			
	L0L6J5	Avenin-3			
	Q2EPY2	Avenin-3			
	L0L4I8	Avenin-3		VFQPQLQQ	
	L0L6J0	Avenin-7/-8			
Group C avenins	L0L5H5	Avenin-5,−6.−7,−8,−9,−10	PFV[$Q_{(2-4)}$]	FFQPQMQQ + VTQG	
	L0L5G8	Avenin-5,−6.−7,−8,−9,−10		VFQPQLQQ	
	Q09072	Avenin-5,−6.−7,−8,−9,−10			
	Q09114	Avenin-9			
	L0L8B6	Avenin-5,−6.−7,−8,−9,−10			
	L0L841	Avenin-5,−6.−7,−8,−9,−10			
	L0L6K5	Avenin-5,−6.−7,−8,−9,−10			
	L0L6K1	Avenin-5,−6.−7,−8,−9,−10			

Avenin sequences from A. sativa UniProt accession number alongside previous classification according to Shewry (62) and the criteria used to classify them into either group A, B or C.

were used to separate the 40 k and 75 k γ-secalins, and the y-type and x-type HMW secalin subunits (x-type subunits being larger).

The avenins from oats could be further classified into A-, B-, or C-avenins based on their amino acid sequences (11) (**Figure 3D** and **Table 2**). Two B–avenins and one C-avenin classified by phylogenetic analysis contained repeat motifs that could place them either in group B or C (**Figure 3** and **Figure S4**). An additional distinction can be made based on the number of cysteine residues: A-avenins contain nine cysteine residues and B- and C-avenins contain eight. A-avenins could therefore form intermolecular disulphide bonds due to the odd number of cysteine residues therefore making A-avenins polymeric (11). However, it should be noted that protein level evidence for the existence of A-avenins is currently lacking.

Analysis of Coeliac Toxic Motifs and IgE-Reactive Allergens

The average number of unique coeliac toxic motifs per sequence, coeliac toxic motif density and sequence coverage by coeliac toxic motifs was evaluated using the "exact peptide match" function from AllergenOnline and the repository of 1,013 coeliac toxic motifs contained within the database (**Figure 4** and **Table S5**). There were large similarities between all metrics of coeliac toxic motif analysis within homologous protein groups across species although there were some differences compared to bread wheat (50). Thus, the S-poor prolamins in barley (C hordeins) and rye (ω-secalins) together with the rye 75 k γ-secalins and the α- and γ-gliadins from *T. turgidum* ssp durum generally carried the greatest number of coeliac toxic motifs across all the measures applied. Only the density of coeliac toxic motifs

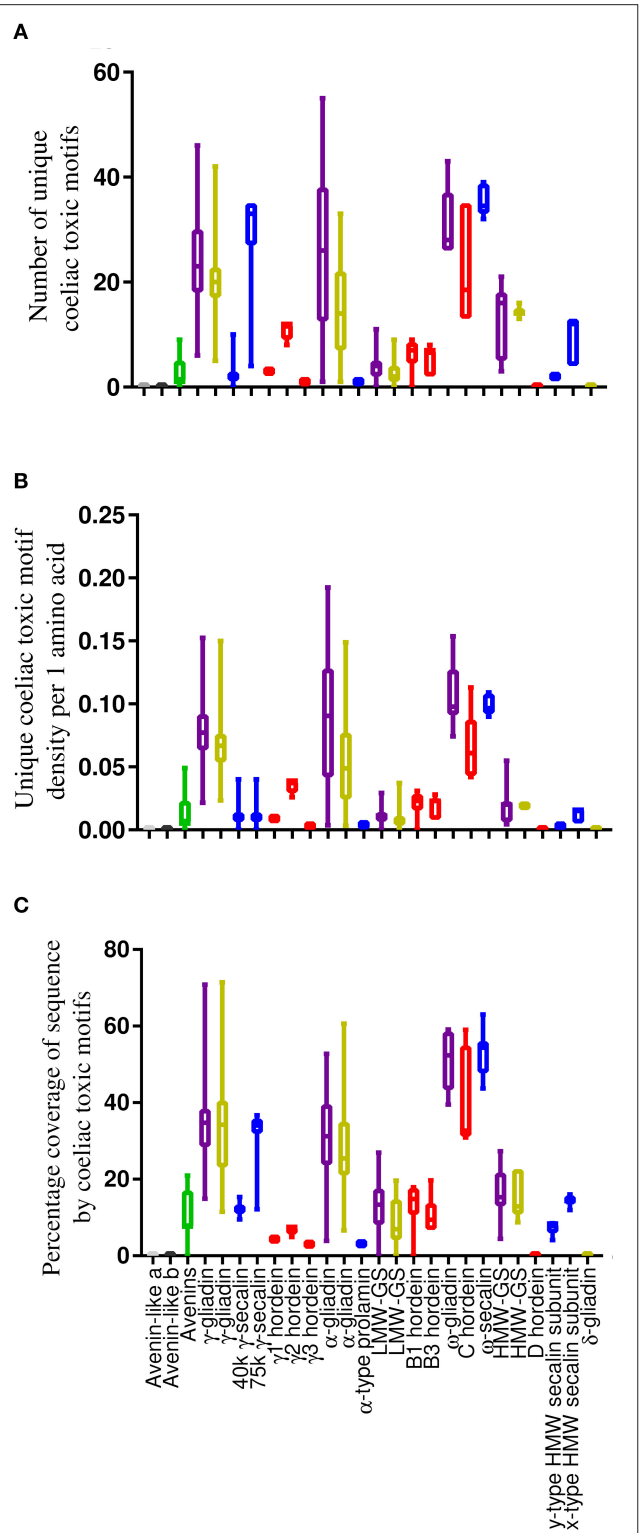


FIGURE 4 | Analysis of sequences from the *T. turgidum* ssp durum, *H. vulgare, Secale cereale,* and *A. sativa* databases and avenin-like sequences from *T. aestivum* in the context of coeliac toxic motifs from the AllergenOnline database shown using box and whisper plots with sequences arranged according to structural classification. **(A)** shows the number of unique coeliac *(Continued)*

FIGURE 4 | toxic motifs per sequence, **(B)** showing the density of motifs and **(C)** showing the sequence coverage by motifs calculated using the Protein Coverage Summarizer software. Bars coloured grey indicate sequences from *T. aestivum*, *T. turgidum* ssp durum, and *H. vulgare,* gold are sequences from *T. turgidum* ssp durum, red are sequences from *H. vulgare,* blue are sequences from *S. cereale* and green corresponds to sequences from *A. sativa*.

per residue varied, which was much lower for the rye 75 k γ-secalins. This is unlike bread wheat where α-gliadins contained the largest number and the highest density of coeliac toxic motifs (50) although it should be noted that this protein fraction is absent from barley and oats with only one unconfirmed sequence reported for rye. The avenins from oats contained a moderate load of coeliac toxic motifs and although 10.5% of A-avenins had no coeliac toxic motifs, the remainder carried at least one, as did the B- and C-type avenins. In contrast, no coeliac toxic motifs were identified in the avenin-like a and b proteins and δ-gliadins in any of the cereal species and were either low (e.g., the y-type HMW secalin subunit sequences) or absent (e.g., barley D hordeins) from the HMW subunits of glutenin.

Analysis of IgE-reactive proteins, using the allergen sequences defined in the IUIS Allergen Nomenclature database (www.allergen.org) identified seven seed storage prolamin food allergens in bread wheat as follows: ω5-gliadin (Tri a 19; UniProt accession Q402I5), γ-gliadin (Tri a 20; UniProt accession A0A060N479 and Q9SYX8), α-gliadin (Tri a 21; UniProt accession D2T2K3), HMW GS Dx5 and Bx7 (Tri a 26; UniProt accession P10388 and Q45R38) and LMW GS GluB3-23 (Tri a 36; UniProt accession B2Y2Q7). *T. turgidum* ssp durum only contains one known allergenic protein, the non-specific lipid transfer protein (Tri tu 14; GenBank accession JF799976.1) Barley and rye only contain allergenic seed storage prolamin proteins located in the γ-type protein group; γ3-hordein (Hor v 20; UniProt accession P80198) and 75 k γ-secalin (Sec c 20; UniProt accession Q9S8B0 and Q9S8A7). The database contained no known allergenic proteins that mapped to oats.

Application of the GluPro Bread Wheat Databases for Searching of Mass Spectrometry Data

The curated sequences from bread wheat (GluPro v 1.0, 1.1 and 1.2) were then used to analyse IM-MS-MS spectral libraries for bread wheat from cultivars Chinese Spring and previously published data from cv Hereward (50) and compared with searching against the reviewed UniProt sequences. Searching was undertaken using variable modifications for deamidation of glutamine and hydroxylation of proline, as these have previously been observed in plant proteomic data sets (84, 85). The distribution of q-values (adjusted *p*-values found using an optimised FDR approach) is shown in **Figure S7** when mining the spectral libraries using the different databases. These density histograms show the distribution was as expected where the null features represent the flat portion whilst the "true"

features all lie very close to zero. Since the FDR was set at 1% only proteins with a q value $\leq$ 0.01 were accepted as identifications. Example extracted ion chromatograms for selected peptides are shown in **Figure S5**. Using the UniProt reviewed prolamin sequences allowed a total of 16 and 19 proteins, respectively, in cvs Chinese Spring and Hereward to be identified (**Table 3**; **Supplementary Datasheets 1**, **2**). In comparison searching using the curated gluten protein sequence databases yielded a much larger number of identifications, which were greater (40–42) for cv Hereward, compared to cv Chinese Spring (19–20). Modifying the searching databases to include the avenin sequences (GluPro v1.1) and the additional bread wheat accessions from the Chinese Spring Genome (GluPro v 1.2) had little impact on the total numbers of proteins identified but it did affect, in some cases, the numbers of proteins belonging to a specific class or the specific protein accessions identified. Thus, as expected, avenin-like proteins were identified using GluPro v1.1 and 1.2 although the numbers varied. Similarly the δ-gliadins were only identified using GluPro v 1.2, the database which actually contained these protein sequences as has previously been reported (64). Thus, using the curated sequence databases did offer an advantage over using a simple UniProt download.

Comparison of the number of identifications made with the number of genes present using only genes encoding full length proteins showed that the number of identifications made varied between protein group, being only 7.69% of total α-gliadins compared with 90% of LMW glutenin subunits whilst none of the HMW subunits of glutenin were identified (**Table 3**). The number of δ-, γ-, ω-gliadins and LMW glutenin subunits matched to gene sequence data was in line with the identifications made by Altenbach et al. (64). The low number of α-gliadins and HMW glutenin subunits identified in Chinese Spring is most likely due to incomplete extraction of the prolamin protein fraction due to lack of aqueous alcohol in the extraction buffer. However, there are some anomalies in the reference proteome since currently it includes sequences for 1Dx5 and 1Dy10 rather than the actual HMW subunits.

There were a number of anomalies regarding the identifications particularly with regards annotation of the HMW glutenin subunits (**Table 4**, **Table S6**) and **Figure S6**). The cvs Chinese Spring and Hereward have well described HMW subunit compositions of 6+8, 2+12, and 7+9, 3+12, respectively. Using the UniProt download five HMW subunits were identified in each cv including an Ax subunit (P02861), despite both cultivars being Glu 1A Null. When the same MS libraries were analysed using the curated sequence databases many of these peptides were no longer identified as being "unique" to one accession, altering the pattern of identifications. For example one unique 15 residue peptide (YPTSPQQSGQGQQGY), which was reproducibly identified with a score of between 5.041 and 5.231 probably arises from the 1Bx subunits in both cvs, as it appears as a tandem repeat in 1Bx sequences including G4Y3Y2 (1Bx7.3), Q6UKZ5 (1Bx14) sequences which share 95.7% sequence identity. Since neither of these sequences are in the reviewed UniProt database, the peptide was mis-identified as being unique to the Ax subunit (P02861). Similar reasons may explain other misidentifications, such as subunits 1Dx5 (P10388) and 1Dy10 (P10387). For example, the unique peptide, QQPGQGQQGHY, was found in the Chinese Spring data set with a score of 6.4 may have originated from a 1By sequence, such as Q52JL2, and was miss assigned to the 1Dy10 subunit again due to the restricted nature of the reviewed UniProt download.

A second factor that affected the sequence accessions identified was that the peptide scores changed with each database. This meant that peptides with scores close to the cut-off of 5.0 were falling in and out of significance. Such a phenomenon probably results from the way in which the decoy database is developed that underpins the reduction of false positive identifications which requires that predicted peptides in the decoy sequence lists are absent from the target sequence list (86). Short motifs, such as those found in the repetitive domain

TABLE 3 | Summary of proteins identified by analysis of IM-MS-MS data for bread wheat cultivars Chinese Spring (CS) and Hereward using different bread wheat gluten protein sequence databases.

Protein group	No of CS genes/ proteins	UniProt reviewed prolamins		GluPro v 1		GluPro v 1.1		GluPro v 1.2	
		CS	Hereward	CS	Hereward	CS	Hereward	CS	Hereward
Avenin-like	19	0	0	0	0	0	1	0	4
α-gliadins	26	3	6	1	6	2	8	2	7
δ-gliadins	2	0	0	0	0	0	0	1	1
γ-gliadins	11	6	5	6 (5)	9	6	8	7 (5)	8
ω-gliadins	5	0	0	3	8	2	8	2	8
LMW-GS	10	2	3	8 (6)	13	9 (6)	12 (9)	6	11 (7)
HMW-GS	4	5	5	2 (1)	4 (3)	0	4	1 (0)	4
Total	77	16	19	20 (16)	40 (39)	19 (16)	41 (38)	19 (17)	43 (39)

The number of full length gluten protein sequences in Chinese Spring was recovered from the annotated genome sequence (64) supplemented with the total number of avenin-like sequences from the CS reference proteome available on UniProt (accessed16.01.2020). Databases used in searching were as follows; GluPro v 1 (n = 630), GluPro v 1.1 (n = 685) and GluPro v 1.2 (n = 699). Identifications were made using unique peptides of any length; those with unique peptides ≥5 amino acids in length are given in parentheses.

TABLE 4 | High molecular weight glutenin subunits identified by IM-MS-MS analysis of bread wheat cvs Chinese Spring (CS) and Hereward.

Database	Cultivar	Accession number	Subunit type	No of peptides (unique peptides)	% Sequence coverage	Protein score
UniProt reviewed prolamins	CS	P02861	Ax	2 (1)	26.07	11.08
		P08489	Dx2	44 (7)	50.82	424.10
		P10388	Dx5	40 (1)	29.70	304.93
		P10387	Dy10	18 (3)	24.25	150.34
		P08488	Dy12	20 (5)	38.46	200.36
	Hereward	P02861	Ax	2 (1)	52.48	5.42
		P08489	Dx2	82 (14)	78.76	499.03
		P10388	Dx5	82 (11)	53.24	442.85
		P10387	Dy10	48 (11)	53.09	304.67
		P08488	Dy12	44 (8)	58.01	290.11
GluPro v 1	CS	Q41553	Ax2	11 (1) [0]	19.08	69.83
		G4Y3Y2	Bx7.3	19 (1)	39.49	139.83
	Hereward	A0MZ38	Ax	10 (1) [0]	22.21	65.73
		Q6UKZ5	Bx14	12 (3)	35.58	80.36
		Q52JL2	By	31 (2) [1]	45.63	197.64
		G3FLC7	Dx2/3	48 (1)	57.36	315.01
GluPro v 1.1	CS	None identified	None identified	None identified	None identified	None identified
	Hereward	Q6UKZ5	Bx14	13 (3)	31.90	87.28
		Q52JL2	By	31 (3) [2]	47.83	196.86
		G3FLC7	Dx2/3	47 (1)	56.16	303.04
		Q52JL3	Dy12	22 (1)	46.81	175.38
GluPro v 1.2	CS	A0MZ38	Ax	6 (1) [0]	10.43	50.83
	Hereward	Q6UKZ5	Bx14	15 (5) [3]	46.18	97.57
		Q52JL2	By	28 (2)	46.72	182.53
		G3FLC7	Dx2/3	46 (1)	55.99	293.40
		Q52JL3	Dy12	21 (1)	46.18	170.81

Identifications of HMW glutenin subunits arising from interrogation of mass spectrometry against different curated databases, the UniProt accession number, subunit type, number of peptides identified, sequence coverage and protein score for that identification. Identifications were made using unique peptides of any length; those with unique peptides ≤5 amino acids in length are given in square backets.

of prolamins, could give rise to ambiguous identifications by appearing in both the decoy and target databases. To take account of this the mass spectra for the unique peptides were visually inspected and included some very short peptides ≤5 residues in length, which could map to different proteins. Excluding these peptides reduced the total numbers of gluten proteins identified but did not generally change the nature of the identifications made (**Tables 3**, **4**).

DISCUSSION

Creation of an expanded gluten protein sequence database has highlighted the large number of partial or fragment sequences and the high degree of redundancy present in UniProt and the NCBI Protein database as well as genome sequences. We also found, as others have observed, that these databases contain sequences that are not always fully annotated, curated or complete, limiting their usefulness for searching MS data (87) including gluten protein proteomics (34). BLAST searching to recover homologous sequences proved important and necessary as this recovered more sequences, especially for *T. turgidum* ssp durum where an additional 45 sequences were identified. Mining of genomes also proved useful for identifying sequences from cereal species, such as *T. turgidum* ssp durum. However, no new sequences were added through mining the rye genome although it only covers the low copy portion representing 2.8Gbp of the total 7.9Gbp, as highly repetitive sequences are difficult to assemble (58). Development of the manually curated databases presented here has addressed these issues and allowed an increased the number of identifications to be made when mining MS data, compared to searching against prolamins in reviewed UniProt.

The number of sequences in the respective cereal species databases correlates well to the number of sequences suggested by genomic and proteomic data (88, 89). Therefore, although the numbers of sequences for barley, rye and oats are relatively low, they should represent almost all of the prolamin sequences that would be observed experimentally. In comparison to ProPepper, GluPro v 6.1 database contains a larger number of sequences attributed to wheat, barley and rye, but fewer for oats. The AllergenOnline database contains fewer sequences from all species because it only includes allergen sequences, which are either IgE-reactive or carry coeliac toxic motifs. Although ω-gliadins have been identified in durum

wheat using bread wheat prolamin sequences (71) no ω-gliadins sequences have been attributed to durum wheat in UniProt at present.

Evaluation of coeliac toxicity of prolamins in the GluPro database family using sequences representing T-cell epitopes present in the AllergenOnline database demonstrated that the C hordeins of barley and ω-secalins of rye (both homologues of wheat ω-gliadin) contained, on average, the highest number, density and percentage coverage by coeliac toxic motifs. Interestingly, the 75 k γ-secalins, like the γ-gliadins in both bread wheat and *T. turgidum* ssp durum (50, 71), also contained a relatively high number of coeliac toxic motifs, in contrast to both the 40 k γ-secalins and γ-hordeins. As the 75 k γ-secalins comprise ∼50% of the total seed proteins of rye, this could pose a high risk to individuals with coeliac disease (52). Unlike barley and rye, there was little variation in the potential coeliac toxicity of the gliadins between bread wheat and *T. turgidum* ssp durum (50). Gliadins from bread wheat are often considered the most coeliac toxic group, with a 33 mer peptide derived from α2-gliadin described at the most important coeliac toxic fragment (90, 91). In addition to the content of coeliac toxic motifs, the total prolamin content and proportions of each prolamin type within a given cereal species needs to be taken into account in assessing potential coeliac toxicity. For example, although ω-type prolamins (ω-gliadins, ω-secalins and C hordeins) contained a relatively large number of coeliac toxic motifs, these proteins only constitute a minor fraction of total expressed prolamins in these grains. In contrast, the LMW glutenin fraction present in bread wheat, *T. turgidum* ssp durum and barley could pose a greater risk to coeliac sufferers, as although they contain fewer coeliac toxic motifs, they account for ∼30% of total seed storage prolamins (92).

The avenin-like proteins from bread wheat *T. turgidum* ssp durum and barley did not contain any known coeliac toxic motifs, and consequently may pose little or no risk to those with coeliac disease. However, this will require confirmation through, for example, assessing the capacity of these proteins to stimulate T-cells. In contrast, avenin proteins from oats contained many coeliac toxic motifs. In addition, since avenins comprise the minor fraction of seed storage proteins in oats further reducing the total content of coeliac toxic motifs in oats compared to wheat, barley and rye. This supports observations that oats cannot be tolerated by some of those individuals with coeliac disease (19) and calls in to question claim that oats should be included in a gluten-free diet. With regards IgE-mediated food allergy, only seven of the eleven sequence accessions corresponding to seed storage prolamin food allergens mapped to full length protein sequences in the GluPro databases. Several prolamins contained IgE epitopes identified by Juhasz et al. (93), particularly in the ω-type prolamins from all the cereal species except oats, with one epitope (QQFPQQQ) only being present in bread wheat and *T. turgidum* ssp durum.

The development of a suite of curated prolamin sequences from bread wheat, *T. turgidum* ssp durum, barley, rye and oats into a family of databases will support mining of mass spectrometric data in future. It will also potentially provide the protein level evidence currently lacking for protein sequences contained in the databases, such as the α-prolamins in rye, avenin-like proteins, and δ-gliadin in *T. turgidum* ssp durum, avenin-like b proteins in barley and A-avenins in oats. The mapping of coeliac toxic motifs within the database will allow peptide markers for coeliac toxic motifs to be identified using mass spectrometry. This could thereby support the development of new analytical methods, which can quantify the burden of toxic motifs in gluten-containing and gluten-free food.

DATA AVAILABILITY STATEMENT

The curated sequence sets are available for download in FASTA format through the Figshare data repository (doi: 10.6084/m9.figshare.12613154).

AUTHOR CONTRIBUTIONS

MD generated the databases and completed the phylogenetic analysis of the sequences along with coeliac toxic motif evaluation. SB undertook proteomic analysis of wheat grain samples. MD and EM conceived and wrote the manuscript. CN, PS, and LG contributed to interpretation and discussion of data generated and EM wrote the manuscript and revision of the paper.

FUNDING

This work was funded through a BBSRC CASE grant (Grant No: BB/M011208/1) awarded to MD and BBSRC CADE grant (Grant No: BB/K01305X/1) awarded to SB. Rothamsted Research receives grant-aided support from the Biotechnology and Biological Sciences Research Council (BBSRC) of the UK and the work forms part of the Designing Future Wheat Strategic Programme (BB/P016855/1).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2020.00087/full#supplementary-material

REFERENCES

1. Un Food and Agriculture Organization Corporate Statistical Database (Faostat). *Crops*. Food and Agriculture Organization (FAO) of the UN) (2017). Available online at: http://www.fao.org/faostat/en/#data/QC
2. Tosi P, Gritsch CS, He J, Shewry PR. Distribution of gluten proteins in bread wheat (*Triticum Aestivum*) grain. *Ann Bot*. (2011) 108:23–35. doi: 10.1093/aob/mcr098
3. Shewry PR, Tatham AS. Improving wheat to remove coeliac epitopes but retain functionality. *J Cereal Sci*. (2016) 67:12–21. doi: 10.1016/j.jcs.2015.06.005
4. Anderson OD. The spectrum of major seed storage genes and proteins in oats (*Avena Sativa*). *PLoS ONE*. (2014) 9:e83569. doi: 10.1371/journal.pone.0083569
5. Codex Alimentarius Commission. *Foods for Special Dietary Use for Persons Intolerant to Gluten* Rome: FAO-WHO (2008).

6. Shewry PR, Halford NG, Belton PS, Tatham AS. The structure and properties of gluten: an elastic protein from wheat grain. *Philos Trans R Soc Lond, B Biol Sci*. (2002) 357:133–42. doi: 10.1098/rstb.2001.1024
7. Tatham AS, Shewry PR. The S-poor prolamins of wheat, barley and rye: revisited. *J Cereal Sci*. (2012) 55:79–99. doi: 10.1016/j.jcs.2011.10.013
8. Shewry PR, D'ovidio R, Lafiandra D, Jenkins JA, Mills ENC, Békés F. Wheat grain proteins. In: Khan K, Shewry PR, editors. Fourth edition. *Wheat: Chemistry and Technology*. St. Paul, MN: AACC International (2009). p. 223–98. doi: 10.1094/9781891127557.008
9. Kim SI, Charbonnier L, Mosse J. Heterogeneity of avenin, the oat prolamin. Fractionation, molecular weight and amino acid composition. *Biochim Biophys Acta*. (1978) 537:22–30. doi: 10.1016/0005-2795(78)90599-8
10. Egorov TA, Musolyamov AK, Kochergin AA, Andersen JS, Roepstorff P. Isolation, characterization by mass-spectrometry and partial amino-acid sequencing of avenins. *J Cereal Sci*. (1994) 20:107–17. doi: 10.1006/jcrs.1994.1051
11. Real A, Comino I, De Lorenzo L, Merchan F, Gil-Humanes J, Gimenez MJ, et al. Molecular and immunological characterization of gluten proteins isolated from oat cultivars that differ in toxicity for celiac disease. *PLoS ONE*. (2012) 7:e48365. doi: 10.1371/journal.pone.0048365
12. 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. doi: 10.1016/j.jcs.2006.04.004
13. De Caro S, Ferranti P, Addeo F, Mamone G. Isolation and characterization of avenin-like protein type-B from durum wheat. *J Cereal Sci*. (2010) 52:426–31. doi: 10.1016/j.jcs.2010.07.005
14. Picariello G, Bonomi F, Iametti S, Rasmussen P, Pepe C, Lilla S, et al. Proteomic and peptidomic characterisation of beer: immunological and technological implications. *Food Chem*. (2011) 124:1718–26. doi: 10.1016/j.foodchem.2010.07.111
15. Kasarda DD, Adalsteins E, Lew EJ, Lazo GR, Altenbach SB. Farinin: characterization of a novel wheat endosperm protein belonging to the prolamin superfamily. *J Agric Food Chem*. (2013) 61:2407–17. doi: 10.1021/jf3053466
16. Ma F, Li M, Li T, Liu W, Liu Y, Li Y, et al. Overexpression of avenin-like b proteins in bread wheat (*Triticum Aestivum L.*) improves dough mixing properties by their incorporation into glutenin polymers. *PLoS ONE*. (2013) 8:e66758. doi: 10.1371/journal.pone.0066758
17. Zhang Y, Cao X, Juhasz A, Islam S, Qi P, She M, et al. Wheat avenin-like protein and its significant Fusarium Head Blight resistant functions. *bioRxiv*. (2018) 406694. doi: 10.1101/406694
18. Mustalahti K, Catassi C, Reunanen A, Fabiani E, Heier M, Mcmillan S, et al. The prevalence of celiac disease in Europe: results of a centralized, international mass screening project. *Ann Med*. (2010) 42:587–95. doi: 10.3109/07853890.2010.505931
19. Arentz-Hansen H, Fleckenstein B, Molberg O, Scott H, Koning F, Jung G, et al. The molecular basis for oat intolerance in patients with celiac disease. *PLoS Med*. (2004) 1:e1. doi: 10.1371/journal.pmed.0010001
20. Marsh MN. Gluten, major histocompatibility complex, and the small intestine. A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). *Gastroenterology*. (1992) 102:330–54. doi: 10.1016/0016-5085(92)91819-P
21. Abadie V, Sollid LM, Barreiro LB, Jabri B. Integration of genetic and immunological insights into a model of celiac disease pathogenesis. *Annu Rev Immunol*. (2011) 29:493–525. doi: 10.1146/annurev-immunol-040210-092915
22. Sollid LM, Qiao SW, Anderson RP, Gianfrani C, Koning F. Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenetics*. (2012) 64:455–60. doi: 10.1007/s00251-012-0599-z
23. Dorum S, Qiao SW, Sollid LM, Fleckenstein B. A quantitative analysis of transglutaminase 2-mediated deamidation of gluten peptides: implications for the T-cell response in celiac disease. *J Proteome Res*. (2009) 8:1748–55. doi: 10.1021/pr800960n
24. Dorum S, Arntzen MO, Qiao SW, Holm A, Koehler CJ, Thiede B, et al. The preferred substrates for transglutaminase 2 in a complex wheat gluten digest are peptide fragments harboring celiac disease T-cell epitopes. *PLoS ONE*. (2010) 5:e14056. doi: 10.1371/journal.pone.0014056
25. Tatham AS, Shewry PR. Allergens to wheat and related cereals. *Clin Exp Allergy*. (2008) 38:1712–26. doi: 10.1111/j.1365-2222.2008.03101.x
26. Scherf KA, Brockow K, Biedermann T, Koehler P, Wieser H. Wheat-dependent exercise-induced anaphylaxis. *Clin Exp Allergy*. (2016) 46:10–20. doi: 10.1111/cea.12640
27. Bucchini L, Daly M, Mills ENC. Food allergen labelling regulation. In: Astley SB, ed. *Health Claims and Food Labelling*. London: Royal Society of Chemistry (2020). doi: 10.1039/9781788013031-00107
28. European Union. Commission implementing regulation (EU) No 828/2014 on the requirements for the provision of information to consumers on the absence or reduced presence of gluten in food. *Official J EU*. (2014) 5–8. Available online at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2014:228:TOC
29. European Commission. (2017). *Commission Notice 2017/C 428/01 of 13.7.2017 Relating to the Provision of Information on Substances or Products Causing Allergies or Intolerances as Listed in Annex II of Regulation (EU) No 1169/2011 on the Provision of Food Information to Consumers*. (ed.) E. Commission. Available online at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:C:2017:428:TOC
30. Wieser H, Koehler P. Is the calculation of the gluten content by multiplying the prolamin content by a factor of 2 valid? *Eur Food Res Technol*. (2009) 229:9–13. doi: 10.1007/s00217-009-1020-5
31. Diaz-Amigo C, Popping B. Accuracy of ELISA detection methods for gluten and reference materials: a realistic assessment. *J Agric Food Chem*. (2013) 61:5681–8. doi: 10.1021/jf3046736
32. Martínez-Esteso MJ, Brohée M, Nørgaard J, O'connor G. Label-free proteomic analysis of wheat gluten proteins and their immunoreactivity to ELISA antibodies. *Cereal Chem J*. (2017) 94:820–6. doi: 10.1094/CCHEM-11-16-0266-R
33. Hajas L, Scherf KA, Torok K, Bugyi Z, Schall E, Poms RE, et al. Variation in protein composition among wheat (*Triticum Aestivum L.*) cultivars to identify cultivars suitable as reference material for wheat gluten analysis. *Food Chem*. (2018) 267:387–94. doi: 10.1016/j.foodchem.2017.05.005
34. Colgrave ML, Goswami H, Howitt CA, Tanner GJ. What is in a beer? Proteomic characterization and relative quantification of hordein (gluten) in beer. *J Proteome Res*. (2012) 11:386–96. doi: 10.1021/pr2008434
35. Fiedler KL, Mcgrath SC, Callahan JH, Ross MM. Characterization of grain-specific peptide markers for the detection of gluten by mass spectrometry. *J Agric Food Chem*. (2014) 62:5835–44. doi: 10.1021/jf500997j
36. Schalk K, Koehler P, Scherf KA. Targeted liquid chromatography tandem mass spectrometry to quantitate wheat gluten using well-defined reference proteins. *PLoS ONE*. (2018) 13:e0192804. doi: 10.1371/journal.pone.0192804
37. Schalk K, Koehler P, Scherf KA. Quantitation of specific barley, rye, and oat marker peptides by targeted liquid chromatography-mass spectrometry to determine gluten concentrations. *J Agric Food Chem*. (2018) 66:3581–92. doi: 10.1021/acs.jafc.7b05286
38. Uniprot Consortium T. UniProt: the universal protein knowledgebase. *Nucleic Acids Res*. (2018) 46:2699. doi: 10.1093/nar/gky092
39. Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, et al. The NCBI biosystems database. *Nucleic Acids Res*. (2010) 38:D492–6. doi: 10.1093/nar/gkp858
40. International Barley Genome Sequencing Consortium. A physical, genetic and functional sequence assembly of the barley genome. *Nature*. (2012) 491:711–6. doi: 10.1038/nature11543
41. International Wheat Genome Sequencing Consortium. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*. (2018) 361:eaar7191. doi: 10.1126/science.aar7191
42. Maccaferri M, Harris NS, Twardziok SO, Pasam RK, Gundlach H, Spannagl M, et al. Durum wheat genome highlights past domestication signatures and future improvement targets. *Nat Genet*. (2019) 51:885–95. doi: 10.1038/s41588-019-0381-3
43. Pernollet JC, Huet JC, Galle AM, Sallantin M. N-terminal sequences of oat avenins compared to other cereal prolamins. *Biochimie*. (1987) 69:683–9. doi: 10.1016/0300-9084(87)90189-1
44. The Uniprot Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res*. (2017) 45:D158–69. doi: 10.1093/nar/gkw1099

45. The Uniprot Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* (2019) 47:D506–15. doi: 10.1093/nar/gky1049
46. Juhasz A, Haraszi R, Maulis C. ProPepper: a curated database for identification and analysis of peptide and immune-responsive epitope composition of cereal grain protein families. *Database.* (2015) 2015:bav100. doi: 10.1093/database/bav100
47. Sircar G, Sarkar D, Bhattacharya SG, Saha S. Allergen databases. *Methods Mol Biol.* (2014) 1184:165–81. doi: 10.1007/978-1-4939-1115-8_9
48. Pomes A, Davies JM, Gadermaier G, Hilger C, Holzhauser T, Lidholm J, et al. WHO/IUIS allergen nomenclature: providing a common language. *Mol Immunol.* (2018) 100:3–13. doi: 10.1016/j.molimm.2018.03.003
49. Goodman RE, Ebisawa M, Ferreira F, Sampson HA, Van Ree R, Vieths S, et al. AllergenOnline: a peer-reviewed, curated allergen database to assess novel food proteins for potential cross-reactivity. *Mol Nutr Food Res.* (2016) 60:1183–98. doi: 10.1002/mnfr.201500769
50. Bromilow SN, Gethings LA, Buckley M, Bromley M, Shewry PR, Langridge JI, et al. A curated gluten protein sequence database to support development of proteomics methods for determination of gluten in gluten-free foods. *J Proteomics.* (2017) 163:67–75. doi: 10.1016/j.jprot.2017.03.026
51. Martens L, Vandekerckhove J, Gevaert K. DBToolkit: processing protein databases for peptide-centric proteomics. *Bioinformatics.* (2005) 21:3584–5. doi: 10.1093/bioinformatics/bti588
52. Balakireva AV, Zamyatnin AA. Properties of gluten intolerance: gluten structure, evolution, pathogenicity and detoxification capabilities. *Nutrients.* (2016) 8:644. doi: 10.3390/nu8100644
53. Bachor R, Kluczyk A, Stefanowicz P, Szewczuk Z. New method of peptide cleavage based on edman degradation. *Mol Divers.* (2013) 17:605–11. doi: 10.1007/s11030-013-9453-y
54. Stadnik D, Bierczynska-Krzysik A, Zielinska J, Antosik J, Borowicz P, Bednarek E, et al. Identification of lysine misincorporation at asparagine position in recombinant insulin analogs produced in *E. Coli. Pharm Res.* (2019) 36:79. doi: 10.1007/s11095-019-2601-z
55. Addou S, Rentzsch R, Lee D, Orengo CA. Domain-based and family-specific sequence identity thresholds increase the levels of reliable protein function transfer. *J Mol Biol.* (2009) 387:416–30. doi: 10.1016/j.jmb.2008.12.045
56. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics.* (2009) 10:421. doi: 10.1186/1471-2105-10-421
57. Sayers EW, Beck J, Brister JR, Bolton EE, Canese K, Comeau DC, et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* (2019). doi: 10.1093/nar/gky1069
58. Bauer E, Schmutzer T, Barilar I, Mascher M, Gundlach H, Martis MM, et al. Towards a whole-genome sequence for rye (*Secale Cereale L.*). *Plant J.* (2017) 89:853–69. doi: 10.1111/tpj.13436
59. Sievers F, Higgins DG. Clustal omega for making accurate alignments of many protein sequences. *Protein Sci.* (2018) 27:135–45. doi: 10.1002/pro.3290
60. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics.* (2009) 25:1189–91. doi: 10.1093/bioinformatics/btp033
61. Kreis M, Forde BG, Rahman S, Miflin BJ, Shewry PR. Molecular evolution of the seed storage proteins of barley, rye and wheat. *J Mol Biol.* (1985) 183:499–502. doi: 10.1016/0022-2836(85)90017-8
62. Shewry PR. Avenins: the prolamins of oats. In: Shewry PR, R Casey, editors. *Seed Proteins*. Dordrecht: Springer (1999) 79–92. doi: 10.1007/978-94-011-4431-5_4
63. 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. doi: 10.1002/pmic.200800564
64. Altenbach SB, Chang HC, Simon-Buss A, Mohr T, Huo N, Gu YQ. Exploiting the reference genome sequence of hexaploid wheat: a proteomic study of flour proteins from the cultivar Chinese Spring. *Funct Integr Genomics.* (2019) 20:1–6. doi: 10.1007/s10142-019-00694-z
65. Clavijo BJ, Venturini L, Schudoma C, Accinelli GG, Kaithakottil G, Wright J, et al. An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations. *Genome Res.* (2017) 27:885–96. doi: 10.1101/gr.217117.116
66. Qi PF, Chen Q, Ouellet T, Wang Z, Le CX, Wei YM, et al. The molecular diversity of alpha-gliadin genes in the tribe *Triticeae*. *Genetica.* (2013) 141:303–10. doi: 10.1007/s10709-013-9729-2
67. Pruska-Kedzior A, Makowska A, Kedzior Z, Salmanowicz BP. Rheological characterisation of gluten from triticale (*x Triticosecale* Wittmack). *J Sci Food Agric.* (2017) 97:5043–52. doi: 10.1002/jsfa.8555
68. Anderson OD, Dong L, Huo N, Gu YQ. A new class of wheat gliadin genes and proteins. *PLoS ONE.* (2012) 7:e52139. doi: 10.1371/journal.pone.0052139
69. Huo N, Zhang S, Zhu T, Dong L, Wang Y, Mohr T, et al. Gene duplication and evolution dynamics in the homeologous regions harboring multiple prolamin and resistance gene families in hexaploid wheat. *Front Plant Sci.* (2018) 9:673. doi: 10.3389/fpls.2018.00673
70. Schalk K, Lexhaller B, Koehler P, Scherf KA. Isolation and characterization of gluten protein types from wheat, rye, barley and oats for use as reference materials. *PLoS ONE.* (2017) 12:e0172819. doi: 10.1371/journal.pone.0172819
71. Pilolli R, Gadaleta A, Di Stasio L, Lamonaca A, De Angelis E, Nigro D, et al. A comprehensive peptidomic approach to characterize the protein profile of selected durum wheat genotypes: implication for coeliac disease and wheat allergy. *Nutrients.* (2019) 11:2321. doi: 10.3390/nu11102321
72. Colgrave ML, Goswami H, Byrne K, Blundell M, Howitt CA, Tanner GJ. Proteomic profiling of 16 cereal grains and the application of targeted proteomics to detect wheat contamination. *J Proteome Res.* (2015) 14:2659–68. doi: 10.1021/acs.jproteome.5b00187
73. Martinez-Esteso MJ, Norgaard J, Brohee M, Haraszi R, Maquet A, O'connor G. Defining the wheat gluten peptide fingerprint via a discovery and targeted proteomics approach. *J Proteomics.* (2016) 147:156–68. doi: 10.1016/j.jprot.2016.03.015
74. Iimure T, Kihara M, Sato K, Ogushi K. Purification of barley dimeric alpha-amylase inhibitor-1 (BDAI-1) and avenin-like protein-a (ALP) from beer and their impact on beer foam stability. *Food Chem.* (2015) 172:257–64. doi: 10.1016/j.foodchem.2014.09.012
75. Chen P, Wang C, Li K, Chang J, Wang Y, Yang G, et al. Cloning, expression and characterization of novel avenin-like genes in wheat and related species. *J Cereal Sci.* (2008) 48:734–40. doi: 10.1016/j.jcs.2008.04.002
76. Santagati VD, Sestili F, Lafiandra D, D'ovidio R, Rogniaux H, Masci S. Characterization of durum wheat high molecular weight glutenin subunits Bx20 and By20 sequences by a molecular and proteomic approach. *J Mass Spectrom.* (2016) 51:512–7. doi: 10.1002/jms.3776
77. Chen XY, Cao XY, Zhang YJ, Islam S, Zhang JJ, Yang RC, et al. Genetic characterization of cysteine-rich type-b avenin-like protein coding genes in common wheat. *Sci Rep.* (2016) 6:30692. doi: 10.1038/srep30692
78. Shewry PR, Halford NG. Cereal seed storage proteins: structures, properties and role in grain utilization. *J Exp Bot.* (2002) 53:947–58. doi: 10.1093/jexbot/53.370.947
79. Dupont FM, Vensel WH, Tanaka CK, Hurkman WJ, Altenbach SB. Deciphering the complexities of the wheat flour proteome using quantitative two-dimensional electrophoresis, three proteases and tandem mass spectrometry. *Proteome Sci.* (2011) 9:10. doi: 10.1186/1477-5956-9-10
80. Wang DW, Li D, Wang J, Zhao Y, Wang Z, Yue G, et al. Genome-wide analysis of complex wheat gliadins, the dominant carriers of celiac disease epitopes. *Sci Rep.* (2017) 7:44609. doi: 10.1038/srep44609
81. De Santis MA, Giuliani MM, Giuzio L, De Vita P, Lovegrove A, Shewry PR, et al. Differences in gluten protein composition between old and modern durum wheat genotypes in relation to 20th century breeding in Italy. *Eur J Agron.* (2017) 87:19–29. doi: 10.1016/j.eja.2017.04.003
82. Pilolli R, Gadaleta A, Mamone G, Nigro D, De Angelis E, Montemurro N, et al. Scouting for naturally low-toxicity wheat genotypes by a multidisciplinary approach. *Sci Rep.* (2019) 9:1646. doi: 10.1038/s41598-018-36845-8
83. Tanner GJ, Blundell MJ, Colgrave ML, Howitt CA. Creation of the first ultra-low gluten barley (Hordeum vulgare L.) for coeliac and gluten-intolerant populations. *Plant Biotechnol J.* (2016) 14:1139–50. doi: 10.1111/pbi.12482
84. Johnson PE, Sayers RL, Gethings LA, Balasundaram A, Marsh JT, Langridge JI, et al. Quantitative proteomic profiling of peanut allergens in food ingredients used for oral food challenges. *Anal Chem.* (2016) 88:5689–95. doi: 10.1021/acs.analchem.5b04466
85. Bromilow SN, Gethings LA, Langridge JI, Shewry PR, Buckley M, Bromley MJ, et al. Comprehensive proteomic profiling of wheat gluten using a combination

of data-independent and data-dependent acquisition. *Front Plant Sci*. (2017) 7:2020. doi: 10.3389/fpls.2016.02020
86. Elias JE, Gygi SP. Target-decoy search strategy for mass spectrometry-based proteomics. *Methods Mol Biol*. (2010) 604:55–71. doi: 10.1007/978-1-60761-444-9_5
87. Huo N, Zhu T, Zhang S, Mohr T, Luo MC, Lee JY, et al. Rapid evolution of alpha-gliadin gene family revealed by analyzing Gli-2 locus regions of wild emmer wheat. *Funct Integr Genomics*. (2019) 19:993–1005. doi: 10.1007/s10142-019-00686-z
88. Dubois B, Bertin P, Mingeot D. Molecular diversity of alpha-gliadin expressed genes in genetically contrasted spelt (*Triticum Aestivum ssp*. spelta) accessions and comparison with bread wheat (*T aestivum ssp Aestivum*) and related diploid *Triticum* and *Aegilops* species. *Mol Breed*. (2016) 36:152. doi: 10.1007/s11032-016-0569-5
89. Dawson C, Mendoza-Porras O, Byrne K, Hooper T, Howitt C, Colgrave M. Oat of this world: defining peptide markers for detection of oats in processed food. *Peptide Sci*. (2018) 110:e24045. doi: 10.1002/pep2.24045
90. Morón B, Cebolla A, Manyani H, Álvarez-Maqueda M, Megías M, Del Carmen Thomas M, et al. Sensitive detection of cereal fractions that are toxic to celiac disease patients by using monoclonal antibodies to a main immunogenic wheat peptide. *Am J Clin Nutr*. (2008) 87:405–14. doi: 10.1093/ajcn/87.2.405
91. Schalk K, Lang C, Wieser H, Koehler P, Scherf KA. Quantitation of the immunodominant 33-mer peptide from alpha-gliadin in wheat flours by liquid chromatography tandem mass spectrometry. *Sci Rep*. (2017) 7:45092. doi: 10.1038/srep45092
92. Liu L, Ikeda TM, Branlard G, Pena RJ, Rogers WJ, Lerner SE, et al. Comparison of low molecular weight glutenin subunits identified by SDS-PAGE, 2-DE, MALDI-TOF-MS and PCR in common wheat. *BMC Plant Biol*. (2010) 10:124. doi: 10.1186/1471-2229-10-124
93. Juhasz A, Belova T, Florides CG, Maulis C, Fischer I, Gell G, et al. Genome mapping of seed-borne allergens and immunoresponsive proteins in wheat. *Sci Adv*. (2018) 4:eaar8602. doi: 10.1126/sciadv.aar8602

Conflict of Interest: LG is employed by Waters Corporation, a manufacturer and vendor of mass spectrometers used for proteomics.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Nutrition

REVIEW
published: 21 October 2020
doi: 10.3389/fnut.2020.517313

The Two Faces of Wheat

Herbert Wieser[1], *Peter Koehler*[2] *and Katharina A. Scherf*[3]*

[1] Retired, Freising, Germany, [2] Biotask AG, Esslingen am Neckar, Germany, [3] Department of Bioactive and Functional Food Chemistry, Institute of Applied Biosciences, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

Wheat-based foods have been staple foods since about 10,000 years and constitute a major source of energy, dietary fiber, and micronutrients for the world population. The role of wheat in our diet, however, has recently been scrutinized by pseudoscientific books and media reports promoting the overall impression that wheat consumption makes people sick, stupid, fat, and addicted. Consequently, numerous consumers in Western countries have started to question their dietary habits related to wheat consumption and voluntarily decided to adopt a wheat-free diet without a medical diagnosis of any wheat-related disorder (WRD), such as celiac disease, wheat allergy, or non-celiac gluten sensitivity. The aim of this review is to achieve an objective judgment of the positive aspects of wheat consumption as well as adverse effects for individuals suffering from WRDs. The first part presents wheat constituents and their positive nutritional value, in particular, the consumption of products from whole-grain flours. The second part is focused on WRDs that affect predisposed individuals and can be treated with a gluten-free or -reduced diet. Based on all available scientific knowledge, wheat consumption is safe and healthy for the vast majority of people. There is no scientific evidence to support that the general population would benefit from a wheat-free diet.

Keywords: allergy, baking, breeding, celiac disease, gluten, non-celiac gluten sensitivity (NCGS), nutritional value, wheat

OPEN ACCESS

Edited by:
José M. Alvarez-Suarez, University of the Americas, Ecuador

Reviewed by:
Peter Shewry, Rothamsted Research, United Kingdom
Francisco Cabrera-Chavez, Autonomous University of Sinaloa, Mexico

***Correspondence:**
Katharina A. Scherf
katharina.scherf@kit.edu

Specialty section:
This article was submitted to Nutrition and Food Science Technology, a section of the journal Frontiers in Nutrition

Received: *03 December 2019*
Accepted: *16 September 2020*
Published: *21 October 2020*

Citation:
Wieser H, Koehler P and Scherf KA (2020) The Two Faces of Wheat. Front. Nutr. 7:517313. doi: 10.3389/fnut.2020.517313

INTRODUCTION

Wheat is one of the major crops grown worldwide with a production of 7.34×10^8 tons on an area of 2.14×10^6 km^2, which is about the size of Greenland (1). Wheat-based foods have been staple foods since wheat was domesticated about 10,000 years ago, and they constitute a major source of macro- and micronutrients and energy (15–20% of the required intake) for the world population, especially in developing countries (2, 3). Many health benefits such as favorable weight management and reductions in the risks for cardiovascular diseases and type 2 diabetes have been shown to be associated with the consumption of wheat-based foods, especially made of whole grains (4–6). Moreover, many non-food products for daily use contain wheat constituents as valuable ingredients. As a result, the positive aspects of wheat were commonly unquestioned. On the other hand, wheat-based foods are known to cause wheat-related disorders (WRDs), such as celiac disease (CD), wheat allergy (WA), and non-celiac gluten sensitivity (NCGS) in predisposed individuals (7). In the last decade, wheat received an increasingly negative image due to a number of pseudoscientific books and press reports, which recommended the avoidance of wheat consumption for the general population, not only for those suffering from WRDs. As a consequence, increasing numbers of individuals in Western countries decided to adopt a gluten-/wheat-free diet even without clear indications of WRDs or medical advice. The percentages of individuals self-reporting a WRD among the general population were 15% in Australia (8), 13% in the United Kingdom (9), 10%

in Brazil (10), 8% in Mexico (11), 6% in the Netherlands (12), and 3% in El Salvador (13). In light of this controversial debate, the aims of this review are to provide an objective judgment of the positive aspects of wheat consumption as well as adverse effects for individuals suffering from WRDs.

ORIGINS AND IMPORTANCE OF WHEAT

Wheat plants are grasses belonging to the monocot family Poaceae. Cultivated wheat (*Triticum* spp.) consists of three species: diploid (genome A^mA^m) einkorn (*T. monococcum* ssp. *monococcum*), tetraploid (AABB) emmer (*T. turgidum* ssp. *dicoccum*) and durum wheat (*T. turgidum* ssp. *durum*), and hexaploid (AABBDD) common wheat (*T. aestivum* ssp. *aestivum*) and spelt (*T. aestivum* ssp. *spelta*) (14). Using recent advances in sequencing techniques, the International Wheat Genome Sequencing Consortium (IWGSC) recently published a detailed description of the total genome of common wheat (cultivar Chinese Spring) and enabled access to 107,891 gene sequences (15). Einkorn developed from wild einkorn 1 (*T. monococcum* ssp. *boeticum*) and the cultivation of einkorn started around 10,000 years ago in the Fertile Crescent. The hybridization of a different wild einkorn (*T. monococcum* ssp. *urartu*) with an *Aegilops speltoides*-related species (BB) resulted in wild emmer (*T. turgidum* ssp. *dicoccoides*), the ancestor for domesticated emmer. Some subspecies of wild emmer developed free-threshing (naked) grains, known as durum wheat. When durum wheat crossed with *Aegilops tauschii* (DD), naked hexaploid common wheat evolved. Spelt most likely emerged from hybridization between *T. aestivum* and *T. dicoccum* (16).

Common (bread) wheat makes up about 95% of all wheat cultivated globally, and most of the remaining 5% is durum (pasta) wheat. Despite its ability to grow in variable environmental conditions, the selection of qualified sites and soils for wheat cultivation as well as suitable varieties and optimal crop management are important factors to ensure high yields (17, 18). With an estimated additional two billion people on the planet by 2050, food production needs to be increased despite challenges arising from climate change (19, 20). Improved wheat plants, resistant to frost, heat, drought, and/or salty soils, may increase the grain yield and the area suitable for wheat production and thus help ensure food security.

Wheat grains are dry one-seeded fruits (caryopses), in which fruit and seed coats are tightly linked. The husk is fused to the fruit coat in the hulled species einkorn, emmer, and spelt, which means that the husk cannot be separated from the grain by threshing. The grains consist of five main compartments with different constituents and biological functions: Fruit coat (pericarp) (4–5% of grain weight) and seed coat (testa) (≈1%) are the outer layers and surround the whole grain. The inner tissues (endosperm) comprise the aleurone layer (6–9%) and the starchy endosperm (80–85%). The germ (3%) located at the dorsal side of the caryopsis is the embryo, which includes a storage cotyledon and the embryonic axis. Dry (moisture content <12.5%), cool (<10°C), and pest-free storage of wheat grains protects against crop failure by providing a buffer to ensure nutrition security worldwide.

THE SMILING FACE: WHEAT CONSTITUENTS AND THEIR HEALTH BENEFITS

The chemical composition of mature grains (water content ≈13%) varies in a relatively small range, although it is influenced by species, variety, and growing conditions. Carbohydrates, mainly present as starch (≈58%) and non-starch polysaccharides (NSP) (≈13%), are predominant, followed by proteins (≈11%), lipids (≈2%), and minerals (≈2%) (**Table 1**). Vitamins and phytochemicals occur in very small amounts (<0.1%) but are important due to their contribution to human health.

Mono-, Di-, and Oligosaccharides

Wheat grains contain only minor amounts (<0.1%) of the monosaccharides D-glucose and D-fructose, but about 0.5–1.6% of sucrose and 0.1–0.2% of maltose as disaccharides and 0.2–0.7% of the trisaccharide raffinose. The predominant oligosaccharides of wheat are 0.8–1.9% of fructans (22). Wheat fructans are of the graminan-type and comprise three or more fructose monomers linked via β-(2→1) and β-(2→6) glycosidic bonds and may also contain one glucose monomer. The degree of polymerization is 5–7 on average, but may also be below 5 or up to 17–19. Fructans are enriched in the bran and adhering endosperm (3–4%), but also present in the germ (≈2%) and endosperm (1–2%) (23). As part of fermentable oligo-, di-, and monosaccharides and polyols (FODMAPs), fructans and raffinose are metabolized by gut microbiota in the colon and may exert positive health effects similar to NSPs (see the Non-starch Polysaccharides section) by acting as prebiotics (24). In addition to their stimulatory effect on the composition and/or activity of beneficial gut bacteria, inulin-type fructans were shown to have direct immunomodulating and antioxidant protective effects on the gut mucosa (25, 26), but it is still unclear whether graminan-type fructans have the same effects as inulin-type fructans because studies with isolated cereal fructans are still needed (27). However, FODMAPs are also associated with intestinal complaints, as discussed below.

Starch

Starch is restricted to the endosperm, where it is present as large lenticular granules (A-starch) and small spherical granules (B-starch). The major carbohydrates of both granule types are the polysaccharides amylose and amylopectin in a mass ratio of about 25%/75%. Amylose consists of mainly linear α-(1→4)-linked D-glucopyranose units that form a helical structure and has molecular masses between 80,000 and 1 million. Apart from α-(1→4) glycosidic bonds, amylopectin has a branched structure with α-(1→6) bonds occurring every 24–30 glucose units and reaches molecular masses as high as 100 million (28). Starch is important for end-product quality and important for human nutrition. It constitutes the main source of energy (11.4 kJ/g of wholegrain wheat flour and 13.1 kJ/g of white wheat flour)

TABLE 1 | Average contents of grain constituents of common wheat (21).

Main constituents	(g/100 g)	Minerals	(mg/100 g)	Vitamins	(μg/100 g)
Carbohydrates	73.2	Potassium	380	B_3	5,100
Starch	58.2	Phosphorus	342	E	1,400
Non-starch polysaccharides	13.3	Magnesium	97	B_5	1,200
Mono-, di-, oligosaccharides	1.7	Calcium	33	B_1	455
Water	12.7	Sodium	7.7	B_6	269
Protein	10.6	Iron	3.2	B_2	94
Lipids	1.8	Manganese	3.1	B_9	87
Minerals	1.7	Zinc	2.6	B_7	6

because starch is readily degraded by amylases during human gastrointestinal digestion to its glucose monomers.

Resistant starch is the proportion of starch that escapes digestion in the upper gastrointestinal tract and small intestine (29), resulting in a caloric value of 8 kJ/g compared to 15 kJ/g for rapidly digestible starch. It is inaccessible to amylases depending on the size, shape, and crystallinity of the starch granule, the complexation of amylose with lipids, proteins, and phosphate as well as food processing (30, 31). Similar to NSP, resistant starch forms part of dietary fiber (DF) because it is fermented by microbiota in the colon to short-chain fatty acids. Especially butyrate promotes normal colon function by serving as a source of energy for the epithelial cells and by lowering luminal pH, which, in turn, facilitates the growth, and proliferation of beneficial gut microbiota (32, 33). Further, health benefits of resistant starch include improved insulin sensitivity (34), reduced oxidative stress in the colon (35), and lower body fat levels (36). Despite the different effects discussed before, the EU register on Nutrition and Health Claims contains only one authorized entry asserting the reduction of post-prandial glycemic responses for foods where resistant starch is at least 14% of total starch. Wheat whole grain flour contains about 3% of resistant starch and thus contributes to the estimated intake of 3–6 g/day in Europe (37). However, this is still far from the recommended intake of 20 g of resistant starch/day that is needed to achieve the positive health effects (38).

Non-starch Polysaccharides

NSPs include the cell wall polysaccharides arabinoxylans (AX, 5.5–7.4%), cellulose (1.7–3.0%), and β-glucans (0.5–1.0%) (22), but also minor contents of glucomannan, callose, xyloglucan, and pectins. Compared to amylose and amylopectin, NSPs are non-granular and belong to DF because human gastrointestinal enzymes are not able to cleave the predominant β-glycosidic linkages of the monosaccharide units. Depending on the wheat grain tissue, the composition of NSP varies between the endosperm, the bran, the aleurone, and the outer pericarp (39).

AX consist of a chain of β-(1→4)-linked D-xylopyranose residues carrying substitutions via (2→1)- and/or (3→1)-glycosidic bonds to α-L-arabinofuranose residues. Some arabinofuranose residues may additionally be linked to ferulic acid at C_5, so that two adjacent AX chains may become crosslinked via diferulate following oxidation. The extent of diferulate crosslinking affects the solubility and viscosity of AX that can be subdivided into water-extractable (WE)-AX and water-unextractable (WU)-AX. Cellulose consists of linear β-(1→4)-linked D-glucopyranose units, whereas β-glucans are β-D-glucopyranose units linked via (1→3)- and (1→4)-glycosidic bonds. Compared to oats and barley, wheat β-glucans are poorly soluble. Wheat bran composition is significantly influenced by genotype × environment interactions, and the content of neutral detergent fiber ranged from 19 to 31%, that of acid detergent fiber from 5 to 10%, that of cellulose from 3 to 9% and that of hemicellulose from 14 to 21% (40).

All NSPs as well as lignin and fructans are summarized as DF. Wholegrain wheat flour contains 10.3–15.5% of total DF, whereas white flour only has 1.9–6.3% (41). Countless studies support the beneficial effects of wheat NSP on human health, including the comprehensive review by the UK Scientific Advisory Committee on Nutrition (SACN) (42). The most important effects include the regulation of colonic functions, protection against colonic cancer, normalization of serum lipid levels, and attenuation of post-prandial glucose response. A number of studies have reported the protective effects of wheat DF against colon, small intestinal, pancreatic, prostate, and breast cancer, with the effects on colorectal cancer being most evident (43–45). The influence of DF on the incidence of cardiovascular diseases has been the subject of many studies, and the relation between whole-grain intake and improved cardiovascular functions were clearly demonstrated with reduced risks for coronary events, stroke, elevated blood pressure, and hypertension (46). Many studies demonstrated that DF, enriched in whole grain products, has the ability to reduce insulin and glucose response significantly and, thus, reduce the occurrence of type 2 diabetes (47–49). Whether the consumption of DF-rich foods reduces appetite and contributes to weight control needs further clinical investigations (50). Currently, the EU Register on Nutrition and Health Claims contains two authorized entries for wheat bran fiber, one claiming reduced intestinal transit time and the other increased fecal bulk. Wheat AX is listed as suitable for the reduction of post-prandial glycemic responses. About 40 more entries related to DF were submitted, but not authorized, mostly due to the fact that the food constituent was not sufficiently characterized in relation to the claimed effects and a missing cause and effect relationship.

Proteins

Together with yield, grain protein content is of primary importance for wheat breeding because the content and

composition of wheat proteins largely determine the bread making quality of wheat (51). Both protein content and composition are influenced by genetic and environmental factors, but the interaction between the two complicates the identification of molecular markers linked to these traits (15). Among environmental factors, nitrogen fertilization is the most prominent determinant linked to protein content and composition, but other factors such as soil fertility, precipitation, temperature, and altitude also play a role (52). The protein content of wheat may range from 7 to 22%, but mostly lies between 10 and 15% (53). The highest percentages of proteins within the grain are found in the germ (34%), followed by the aleurone (23%) and 5–6% in the outer layers. Consequently, the protein content of whole-grain flour is usually about 2% higher compared to white flour.

The protein fraction of wheat consists of over 100 individual proteins, which can be classified according to their functions: storage proteins, metabolic proteins, protective proteins, and miscellaneous proteins with further specific functions. Gluten proteins are storage proteins located in the endosperm (54), representing around 80% of total grain protein and can be grouped into monomeric gliadins soluble in aqueous alcohols and insoluble polymeric glutenins (55, 56). Gliadins have molecular masses between 30,000 and 55,000 and are structurally differentiated into four types: ω5-, ω1,2-, and α- and γ-gliadins. Glutenins are linked by interchain disulfide bonds and have molecular masses between 600,000 and more than 10 million. The respective monomers are classified into low-molecular-weight (LMW) and high-molecular-weight (HMW) glutenin subunits (GS) with molecular masses around 30,000 and 75,000, respectively. Metabolic proteins include enzymes like hydrolases, which cleave starch (amylases), proteins (peptidases), and lipids (lipases), as well as other oxidoreductases, transferases, and further enzymes (57). The majority of protective proteins are enzyme inhibitors that inhibit external amylase and peptidase activities, some of which are bifunctional like α-amylase/trypsin inhibitors (ATIs). The group of miscellaneous proteins includes, e.g., puroindolines, purothionins, and agglutinins (58, 59).

The nutritional value of wheat proteins is determined by their relative contents of the essential amino acids valine, leucine, isoleucine, phenylalanine/tyrosine, tryptophan, threonine, methionine/cysteine, lysine, and the semi-essential arginine and histidine (60). Lysine is the first limiting amino acid in wheat grains, whereas the other essential amino acids are present in adequate amounts (41). The biological value of white wheat flour is estimated to be 52 and that of whole-grain wheat flour is 17–26% higher (61). This difference is due to the fact that white flour contains higher proportions of gluten proteins compared to whole-grain flour, and the amino acid composition of gluten is characterized by exceptionally high contents of non-essential glutamine (26–53%) and proline (10–29%) (62).

Lipids

Wheat lipids constitute about 2–2.5% of the flour and can be classified into non-polar lipids (acylglycerols and free fatty acids) and polar lipids (phospholipids and glycolipids). The major components of non-polar lipids are triacylglycerols (≈40% of lipids) and free fatty acids (≈15%), while the percentages of mono- and diacylglycerols (≈1 and 4%, respectively) and sterol lipids (<1%) are low. Wheat phospholipids are composed of lysophosphatidylcholine (≈2%), phosphatidylcholine (≈1%), and <1% each of phosphatidylethanolamine, *N*-acylphosphatidylethanolamine, phosphatidylglycerol, and phosphatidyl inositol. Wheat glycolipids are comprised of digalactosyldiacylglycerol (DGDG) (≈15%), monogalactosyldiacylglycerol (MGDG) (≈4%), as well as DGMG and MGMG (about 1% each) (63, 64). Regarding nutritional benefits, wheat has high amounts of oleic acid (≈14%) as well as linoleic acid (≈60%) and linolenic acid (≈4%) as unsaturated fatty acids, and therefore a favorable ratio of unsaturated to saturated fatty acids of about 78%/22% (21). Wheat also contains phenolic lipids, known as alkylresorcinols (1,3-dihydroxybenzene derivatives with an odd-numbered alkyl chain at position 5 of the benzene ring) that may serve as markers of whole-grain cereals in food (65) and as biomarkers of whole-grain wheat intake (66). They have been reported to prevent colon cancer in mouse and *in vitro* models based on their antimutagenic and apoptotic activity (44, 45).

Vitamins and Minerals

Wheat grains are important sources of vitamin E (mainly α-tocopherol) and B vitamins, especially thiamine (B_1), riboflavin (B_2), niacin (B_3), pantothenic acid (B_5), pyridoxine (B_6), and folates (B_9) (**Table 1**). Whole-grain flours have considerably higher vitamin contents than white flours because vitamins are predominantly found in the bran and germ. One important point to consider related to nutrition is the bioavailability of vitamins. For example, most of the niacin present in wheat bran is bound, and only about 10–20% was found to be bioavailable (67).

The major minerals are potassium, phosphorus, magnesium, and calcium, followed by zinc, manganese, and iron in lower amounts (**Table 1**). Copper and selenium are trace minerals. All vitamins and minerals present in wheat have well-known functions in supporting growth and in maintaining the health and well-being of humans.

Phytochemicals

Wheat grains contain small amounts of phytochemicals that are defined as non-nutritive biologically active molecules that function in the human body to achieve health benefits, promote well-being, and prevent certain disease processes. The two major classes of phytochemicals found in wheat are phenolic compounds and terpenoids, derived from the shikimate and mevalonate or methylerythritol phosphate biosynthetic pathways, respectively (22, 68). Phenolic compounds are a structurally diverse group comprising phenolic acids (e.g., cinnamic acid and benzoic acid derivatives), flavonoids (e.g., flavanols and anthocyanins/anthocyanidins), and lignans. Terpenoids include sterols and stanols, with β-sitosterol as the primary compound in wheat, that may occur either as free form, esterified, glycosylated, or acylated and glycosylated. Other terpenoids are tocopherols and tocotrienols (α-tocopherol is commonly known as vitamin E) as well as carotenoids that can be subdivided into oxygen-containing xanthophylls (e.g.,

lutein and zeaxanthin) and oxygen-free carotenes (e.g., α- and β-carotene). Depending on their molecular structures, some carotenes may be converted to vitamin A in humans. With phytochemicals being mainly located in the aleurone and bran, they occur in the mg/kg range in whole-grain flours, but with a wide range of concentrations determined by natural genetic and environmental variations (69, 70).

Phenolic acids are known as strong antioxidants (71, 72), and there is evidence that phenolic compounds improve vascular functions in humans (73) and may have antitumor activity (44, 74). Considering the low overall concentrations of phytochemicals in whole-grain flours and absorbance rates of 5–10% in the human small intestine, direct antioxidant effects appear to be unlikely. However, the remaining 90–95% of phytochemicals are transferred to the colon where they are metabolized by microbiota and may exert positive effects on colon health through this route.

Wheat Constituents Related to Consumption

In summary, wheat grains contain low amounts of sugars, sodium, fat, saturated fatty acids, and are free of cholesterol, all of them considered to restrict health. As most components related to health are concentrated in the outer layers (bran) and germ of the grain, their contents are reduced in flours with low extraction rates (e.g., white flour), which are used to make the majority of wheat products (e.g., bread, pasta, and noodles) consumed in Western countries (**Figure 1**). Products made from white flour are usually preferred by consumers because the high bran content of wholegrain flour may give a dark color, bitterness, gritty texture, and short shelf life. Changing consumer preferences is difficult, and thus, whole-grain intake is still below daily recommendations in most countries. To improve the situation, research and development should be focused on combining innovative processing with better product quality to increase the utilization of whole-grain flours. Moreover, physicians, nutritionists, and the media should advocate the consumption of products made from whole grains (75).

Altogether, the consumption of wheat-based foods, especially whole-grain products, is highly recommended owing to excellent nutritional profiles and their importance as sources of energy, proteins, DF, B, and E vitamins, minerals, and different micronutrients, all contributors to a healthy diet. Depending on the average annual per capita consumption, wheat contributes different amounts of nutrients in relation to the overall diet. The average per capita consumption of wheat worldwide was 67.0 kg (2016–2018), but with considerable regional variations ranging from 50.4 kg in Africa to 109 kg in Europe, considering continents, but even ranging from 16.2 kg in Thailand to 209.7 kg in Turkey, considering countries (76). In Germany, wheat consumption contributes 23% of energy, 34% of digestible carbohydrates, 34% of protein, 24% of DF, and 20–30% of vitamins and 10–20% of minerals compared to the recommended average intake values (77). However, considering the wide span of per capita wheat consumption, these values will vary in the same wide range and the contribution of wheat to dietary nutrient intake needs to be considered individually for each country.

THE SAD FACE: WHEAT-RELATED DISORDERS

Immune-mediated adverse reactions to wheat may occur in predisposed individuals. These hypersensitivities commonly referred to as wheat-related disorders (WRDs) can be classified into CD, gluten ataxia, and dermatitis herpetiformis characterized by an autoimmunogenic response (IgA and IgG antibodies), into IgE- and non-IgE-mediated WA and into NCGS characterized by an innate immune response (78) (**Figure 2**). A subgroup of patients with diarrhea-predominant irritable bowel syndrome may also be affected by wheat consumption.

Celiac Disease

Definition and Prevalence

CD belongs to the most common food-related lifelong disorders worldwide. It is defined as a chronic immune-mediated small intestinal enteropathy precipitated by exposure to dietary gluten in genetically predisposed individuals (7). The term "gluten" comprises the closely related storage proteins of wheat (gliadins and glutenins), rye (secalins), and barley (hordeins). No species or variety of these cereals is currently safe for patients with CD. Recent epidemiological data suggest a mean worldwide prevalence of 1.4% [1.1–1.7%] based on positive serology and of 0.7% [0.5–0.9%] based on biopsy-confirmed diagnosis (79). The serology-confirmed prevalence values were 1.1% in Africa, 1.3% in Europe and South America, 1.4% in North America and Oceania, and 1.8% in Asia. Only people from sub-Saharan Africa appear to be hardly affected. There are considerable regional differences in seroprevalence ranging as high as 2.1–8.5% in Algeria, Czech Republic, India, Israel, Mexico, Malaysia, Saudi Arabia, Sweden, Portugal, and Turkey, or as low as 0.2–0.8% in Estonia, Germany, Iceland, Libya, Poland, Spain, and Switzerland. CD can occur at any age, but the biopsy-confirmed prevalence was significantly greater in children (0.9%) than in adults (0.5%) and also higher in female (0.6%) compared to male (0.4%) individuals. The worldwide prevalence of CD increased from 0.6% [0.5–0.7%] (years 1991–2000) to 0.8% [0.5–1.0%] (years 2001–2016), but the reasons are still unknown, although environmental factors are most likely (80).

Causes

A combination of environmental and genetic factors is necessary to trigger CD in susceptible individuals. Consumption of gluten from wheat, rye, or barley is the decisive environmental factor necessary for CD onset. Although comparative data on the CD activity of single gluten proteins is unavailable, most findings suggest that all gluten proteins are relevant. The unique structural features of gluten proteins are long repetitive amino acid sequences rich in glutamine and proline (62). These sequences are resistant to human gastrointestinal digestive enzymes, so that high amounts of long gluten peptides reach the upper small intestinal mucosa, pass the epithelium, and arrive at the lamina propria, where CD-specific immune reactions are induced (81).

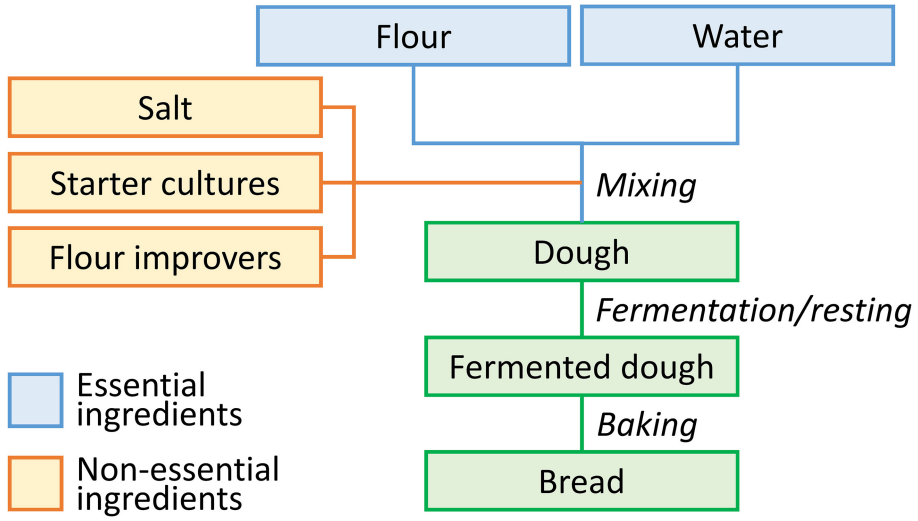
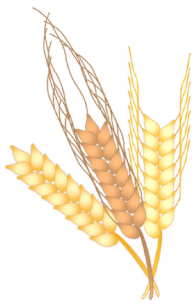


FIGURE 1 | Schematic representation of the bread-making process.



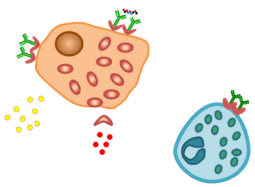
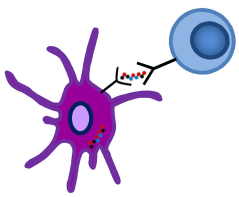
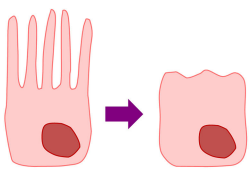
	Wheat allergies	NCGS	Celiac disease
			
Immunity	allergic	innate	adaptive & innate
Serology	IgE Abs	IgG Abs?	IgA, IgG Abs
Prevalence	≈ 0.5 %	0.6 - 6 %	≈ 1.4 %
Onset of symptoms	min-hours	hours	days-weeks
Therapy	Wheat-free diet	Low gluten diet	Gluten-free diet
Causes	-	-	HLA-DQ2/8 genes
	Gluten	Gluten?	**Gluten**
	ATIs	**ATIs**	
	Non-gluten proteins	Non-gluten proteins?	

FIGURE 2 | Overview of wheat allergy, non-celiac gluten sensitivity, and celiac disease. Abs, antibodies; ATIs, α-amylase/trypsin-inhibitors; HLA, human leukocyte antigen; Ig, immunoglobulin; NCGS, non-celiac gluten sensitivity.

Recent evidence suggests that non-gluten proteins, such as ATIs, may also be involved in fueling the CD-specific immune response (82, 83).

The genetic susceptibility to develop CD is associated with the major histocompatibility class II genes on chromosome 6 coding for human leukocyte antigens (HLA)-DQ2 and -DQ8. HLA-DQ proteins are heterodimeric receptors expressed on the surface of antigen-presenting cells (APCs) that are responsible for binding gluten peptides and presenting them to gluten-specific $CD4^+$ T cells. Depending on genetic expression patterns, either the HLA-DQ2.5, -DQ8, or -DQ2.2 heterodimers are present and associated with a very high, high, or low predisposition for CD, respectively (84, 85). These different levels of CD risk are due to HLA-DQ2.5 being capable of binding a large repertoire of gluten peptides that are resistant to gastrointestinal degradation, whereas HLA-DQ8 and -DQ2.2 bind a small to very small selection of gluten peptides that are also less resistant to degradation (86). About 95–97% of CD patients are HLA-DQ2/8 positive, but this genetic predisposition is also present in about 30% of the healthy population. Thus, the absence of HLA-DQ2/8 is a reliable criterion to exclude CD, but its presence is not sufficient to cause CD. A variety of non-HLA genes, mostly encoding for T cells or APCs, has been associated with CD development, but each of these genes most likely only contributes a small percentage to increase the risk of CD (87). It is interesting to note that there is a significant correlation between the level of wheat consumption, the frequency of HLA-DQ2/8 and the prevalence of CD, but with several outlier populations in regions such as northwestern India, northern Africa, Mexico, Finland, and Russia. For example, the prevalence of CD in Algeria is among the highest worldwide (5.6%), whereas that of Tunisia is very low (0.3%), although both countries share similar levels of wheat and barley consumption and frequencies of HLA-DQ2/8 (84). This discrepancy can only be explained by further environmental factors that cause a loss of tolerance to dietary gluten and initiate CD. The most likely factors are infections (rotavirus, adenovirus 12), changes of intestinal microbiota, increased small intestinal permeability, and the so-called hygiene hypothesis that proposed a lower incidence of infections in early childhood as an explanation for the rise in immune-mediated hypersensitivities (88). Although childbirth (natural vs. cesarean section), duration of breastfeeding, and the time of gluten introduction into the child's diet have been discussed, the associations are far from clear, and the cumulative incidence of CD in later childhood was similar independent of breastfeeding or timing of gluten introduction (89, 90). Currently, there is no possibility to prevent CD (91).

Symptoms

The clinical appearance of CD is highly variable and can range from asymptomatic to full-blown symptoms due to the multisystemic nature of CD (92). Classical gastrointestinal symptoms are chronic diarrhea, abdominal pain, vomiting, and steatorrhea. Extraintestinal manifestations include chronic fatigue, night blindness, anemia, osteoporosis, thyroid dysfunction, and reproductive disease (93) and are usually caused by generalized malabsorption of essential nutrients, e.g., vitamins and minerals. Slow growth rates and delayed sexual maturation are known complications of CD in children and adolescents. The characteristic feature of CD is damage to the upper small intestine (duodenal bulb, duodenum, proximal jejunum) characterized by increased infiltration of intraepithelial lymphocytes (IELs), crypt hyperplasia, and partial to total villous atrophy. The degree of mucosal damage is classified according to Marsh–Oberhuber (into types 0, 1, 2, 3a–c, and 4) (94) or Corazza (into grades normal, A, B1, and B2) considering the numbers of IELs per 100 enterocytes, ratio of villous height to crypt depth, and degree of villous atrophy [overview in Ludvigsson et al. (7)]. Asymptomatic CD patients have minimal or no symptoms, but they have increased CD-specific serum antibody levels and villous atrophy. Potential CD patients have no symptoms and a normal mucosa, but increased levels of antibodies indicating an increased risk of developing CD. Most individuals with asymptomatic or potential CD have not yet been diagnosed for CD. They may be diagnosed through case-finding approaches in at-risk populations such as first-degree relatives of CD patients or patients with autoimmune diseases (e.g., diabetes mellitus type I, autoimmune hepatitis, autoimmune thyroid disease) or genetic disorders (e.g., Down syndrome) with known associations to CD (95–97).

Refractory CD, classified into type I (normal IEL phenotype) and II (aberrant IEL phenotype), is a rare but very serious complication, which is characterized by persistence of CD-specific symptoms and mucosal lesions despite a permanent, strict gluten-free diet (GFD) (98). Type I can usually be treated with corticosteroids, but type II imposes a serious risk of progression to enteropathy-associated T-cell lymphoma, small intestinal adenocarcinoma, and ulcerative jejunitis.

Diagnosis

The diagnosis of CD requires a high level of clinical suspicion and a stepwise approach. The diagnostic scheme may consist of five steps: (a) clinical history and symptoms, (b) serology, (c) small intestinal histology, (d) response to a GFD, and (e) HLA status. The most recent diagnostic algorithm developed for children by the European Society for Pediatric Gastroenterology, Hepatology, and Nutrition (ESPGHAN) recommends testing symptomatic patients or risk groups for IgA anti-tissue transglutaminase antibodies (TGA) and total serum IgA to exclude IgA deficiency. If IgA TGA is negative and total IgA is normal, CD is unlikely. If IgA TGA levels are $<10\times$ the upper limit of normal (ULN), the patient has to undergo upper endoscopy with multiple biopsies including at least four from the descending part of the duodenum and at least one from the duodenal bulb. If IgA TGA levels are $>10\times$ ULN, the patient is additionally tested for IgA anti-endomysium antibodies (EMAs). If EMAs are positive, the diagnosis of CD is established, and a GFD is initiated with subsequent follow-up for improvement of symptoms and decline of antibodies to substantiate the correct diagnosis. If EMA is negative, additional testing including biopsies is needed. Gluten challenge and repetitive biopsies are only necessary in ambiguous cases. Especially in children,

biopsies are avoided, if possible, but are still recommended, if there are uncertainties related to the performance of the kit used to measure IgA TGA (99). The same algorithm essentially applies to adults, with the only difference that a duodenal biopsy is advisable in almost all cases because adults are more likely to have an alternative diagnosis or to be non-responsive to a GFD (100). There are many clinical situations that require special considerations and additional tests, such as patients under the age of two and patients with selective IgA deficiency, other immunodeficiencies, or on immunosuppressive medication (101). In case of potential CD, regular follow-up on a normal diet is recommended, whereas patients with suspected, but undocumented CD already adhering to a GFD need to undergo gluten challenge (102). The final confirmation of CD is established by a clinical, serological, and histological response to a strict GFD. Genetic tests based on HLA-DQ2 and -DQ8 alleles can be used to rule out CD in ambiguous cases because of their high negative predictive value. The need for mass screening for CD in the general population is discussed controversially (103). At present, screening is recommended for close relatives of CD patients and for persons with diseases known to be associated with CD such as autoimmune diseases.

Pathomechanism

The pathomechanism of CD is complex and involves both adaptive and innate immune responses (104). Due to their high contents of proline and glutamine, gluten proteins are resistant to complete digestion by human gastrointestinal enzymes. As a consequence, peptides with a length of nine amino acid residues and more stay intact and pass the small intestinal epithelium, either by the trans- or the paracellular routes into the lamina propria (105). More than 1,000 CD-active (CD-toxic and/or -immunogenic) gluten peptides derived from gliadins, glutenins, secalins, hordeins, and avenins have been identified, seven of which are classified as CD-toxic (tested *in vivo* or on organ cultures), five of which are classified as CD-toxic and CD-immunogenic (tested by T-cell proliferation assays), and the vast majority of which are classified as CD-immunogenic (106). Typical features of CD-active peptides are high contents of proline (P) residues and a left-handed polyproline II helical conformation that protects from enzymatic degradation as well as high contents of glutamine (Q) residues that serve as substrates for deamidation or transamidation by human tissue transglutaminase (TG2) (107). Having reached the lamina propria, gluten peptides with QXP or QXXJ motifs (X, any amino acid, J, hydrophobic amino acid) are specifically deamidated (Q $\rightarrow$ E, introduction of a negatively charged glutamic acid residue) or transamidated by TG2 (either to itself or to other lysine donors) (108, 109), whereas QP and QXXP motifs are left unmodified (110). Then, gluten peptides are bound to the heterodimeric HLA-DQ2 or -DQ8 receptors on the surface of APCs. Deamidation at positions 4, 6, and 7 of the gluten peptides strongly favors binding to HLA-DQ2, while HLA-DQ8 prefers deamidation at positions 1 and 9 (respective locations of positively charged amino acid residues in the binding pockets of HLA-DQ2 or -DQ8) (86). APCs subsequently present gluten peptides to the receptor of naive $CD4^+$ T cells and promote T-cell activation and differentiation into gluten-specific inflammatory effector T cells. On the one hand, these gluten-specific $CD4^+$ T cells secrete proinflammatory cytokines, such as interferon-γ and tumor necrosis factor (TNF)-α, which stimulate the release and activation of matrix metallopeptidases. These break down extracellular matrix proteins and thus lead to the destruction of the small intestinal epithelium (proinflammatory Th1-pathway, adaptive immune response). On the other hand, gluten-specific $CD4^+$ T cells help B cells that carry internalized TG2-gluten peptide complexes. This results in B-cell activation and differentiation into plasma cells that produce IgA and IgG antibodies against gliadin and deamidated gluten peptides as external antigens, and against endomysium and TG2 as autoantigens (anti-inflammatory Th2-pathway, adaptive immune response) (84, 111).

Additionally, gluten peptides stimulate the innate immune response and trigger the secretion of interleukin (IL)-15 by activating enterocytes, macrophages, and dendritic cells. As a result, lymphocytes are stimulated to express the receptor NKG2D and epithelial cells to express MICA (major histocompatibility complex class I chain-related molecule A), the ligand for NKG2D. Once MICA has bound to NKG2D, IELs start to destroy epithelial cells. The key factor explaining why tolerance to gluten is lost involving the switch from a tolerogenic $Foxp3^+$ regulatory T-cell response to a proinflammatory Th1 response still remains elusive (112).

Treatment

A strict GFD with a daily gluten intake below 20 mg is currently the only safe and efficient therapy for CD to fully restore patient health. Dietetic gluten-free products are made from safe cereals (e.g., corn, rice, sorghum, or millet), pseudocereals (e.g., amaranth, buckwheat, or quinoa), other sources of flour or starch (potatoes or chestnut), and "gluten replacers" (xanthan or guar gum) (113). Owing to the restricted availability and high costs of gluten-free alternatives and their poorer quality (114), many CD patients regard the GFD as a substantial burden that decreases quality of life, especially when eating out or traveling. Therefore, their most pronounced desire is the development of a pill or a vaccine that will allow them to eat gluten-containing foods, at least sometimes. Therefore, alternative therapies targeting different steps in the pathomechanism of CD are in various stages of development. Therapies that have already reached clinical trial stage include glucocorticoids (budesonide), oral administration of gluten-degrading enzymes (ALV003, a mixture of cysteine endopeptidase B, isoform 2, and a prolyl endopeptidase from *Sphingomonas capsulata*), oral intake of gluten-sequestering polymeric resins [poly(2-hydroxyethylmethacrylate-co-styrene-4-sulfonic acid, sodium salt)], zonulin antagonists (larazotide acetate), vaccination (Nexvax2), probiotics, and hookworm infection (115). All alternative therapies still need to demonstrate that they are tolerable and safe and have no adverse long-term side-effects. Especially for gluten-degrading or -removing agents, the amount of gluten that can be safely ingested needs to be determined considering all other ingredients of the meal that may hinder the efficacy of the agent. Finally, the benefits and risks of alternative therapies have to be carefully weighed against

the GFD to ensure that CD patients receive the best treatment option available.

Dermatitis Herpetiformis and Gluten Ataxia

With a prevalence of 0.03–0.07%, dermatitis herpetiformis (DH) is often referred to as the skin manifestation of CD. Both diseases are caused by gluten, respond to treatment with a GFD, and share the genetic predisposition caused by HLA-DQ2 or -DQ8. Typical symptoms are intense itching and burning papules, macules, and blisters especially on the elbows, knees, and buttocks. DH is ideally diagnosed by direct immunofluorescence biopsy of unaffected skin close to an active lesion that reveals granular IgA deposits in the papillary dermis. The autoantigen in DH is epidermal transglutaminase (TG3). The most likely pathogenic route starts from potential or asymptomatic CD in the small intestine with secretion of IgA against TG2 and TG3 into the blood circulation and results in the deposition of TG3 and IgA against TG3 in the papillary dermis. Due to the presence of active TG3, IgA–TG3 complexes are formed and crosslinked to fibrinogen in the skin (116).

Gluten ataxia (GA) can be regarded as rare neurological manifestation of CD and is defined as idiopathic sporadic ataxia characterized by the presence of IgA or IgG against gliadin in the blood. GA presents with gait and lower limb ataxia, nystagmus, and other visual disorders. Up to 40% of patients show small intestinal damage as in CD and up to 60% of patients have evidence of cerebellar atrophy. Much like in DH, the autoantigen in GA is the primarily brain-expressed transglutaminase (TG6) and IgA against TG6 are serological markers for GA. TG6 and IgA against TG6 accumulate in the brain stem and cerebellum causing infiltration of white matter with lymphocytes and irreversible loss of Purkinje cells in the cerebellar cortex. Therefore, a fast diagnosis of GA and treatment with a GFD are essential to prevent progression of cerebellar dysfunction (117).

IgE-Mediated Wheat Allergies

WAs are defined as adverse immune responses to wheat proteins that reproducibly occur in affected individuals within minutes to hours after exposure (118, 119). A wide variety of wheat proteins including gluten and non-gluten proteins have been identified as allergens [overviews in Brouns et al. (120), Juhasz et al. (121), and Tatham and Shewry (122)]. Depending on the route of allergen exposure and the underlying pathomechanism, WAs can be classified into food allergy, wheat-dependent exercise-induced anaphylaxis (WDEIA), respiratory allergy, and skin allergy. The prevalence estimates for WAs depend on the assessment method used, but range from 0.1% (positive food challenge) to 3.6% (lifetime self-reported prevalence) (123).

Common diagnostic procedures include patient history reporting reproducible symptoms after allergen exposure, skin prick tests, analysis of specific IgE antibodies, or functional assays such as *in vitro* basophil activation tests and oral food challenge. While oral food challenge is generally regarded as a gold standard for the diagnosis of wheat allergy and can offer clarity in ambiguous cases, it is difficult to undertake in routine clinical practice and puts patients at risk of experiencing a severe allergic reaction. The treatment for wheat allergy mostly involves avoidance of exposure to allergens, either in the form of flours and flour dust or elimination of wheat products from the diet. Antihistamines or corticosteroids can be used to treat acute cases (124, 125).

The mechanism of IgE-mediated allergies includes two phases. The first step is sensitization to the allergen upon initial contact, followed by an allergic reaction upon reexposure to the allergen (126). When an allergen is encountered for the first time, it is internalized by APCs (e.g., dendritic cells) and presented to naive $CD4^+$ T cells. In the presence of cytokines, the naive $CD4^+$ T cells become activated and differentiate into Th2 cells that subsequently produce different ILs. IL-4 turns on IgE-producing B cells and sustains the development of Th2 cells, IL-5 activates eosinophils, IL-9 enhances IgE production, mast cell growth and expression of the high-affinity IgE receptor (FcεRI), and IL-13 acts on epithelial cells to stimulate mucus secretion. The secreted IgE antibodies bind to FcεRI on the surface of mast cells and basophils. When the mast cell carrying IgE antibodies is reexposed to the allergen, the multivalent allergen crosslinks two adjacent IgE antibodies and the underlying FcεRI. This bridging leads to mast cell degranulation with discharge of primary (preformed) mediators (histamine, neutral peptidases, acid hydrolases, and proteoglycans such as heparin and chondroitin sulfate) and synthesis and release of secondary mediators (leukotrienes, prostaglandin D2, platelet-activating factor, cytokines, and chemokines). The rapid release of histamine and leukotrienes is responsible for the intense early allergic response characterized by wheezing, sneezing, urticaria, and mucus secretion. The survival of mast cells and enhanced expression of the FcεRI receptor is sustained by signals from these receptors, thus providing a mechanism of amplification. In the following phase, eosinophils are activated by IL-5 and attracted to the site of the immediate reaction by chemokines (e.g., eotaxin). By producing cytokines, leukotrienes, and proteins (major basic protein and eosinophil cationic protein) that are toxic to epithelial cells, the inflammatory response is amplified and sustained without additional exposure to the allergen. These events lead to the late allergic response, which involves further wheezing, nasal blockage, and eczema. Basophils play a similar role to that of mast cells, but they circulate in the blood rather than being present in the affected tissues.

Respiratory and Skin Allergy to Wheat

Respiratory WA comprises baker's asthma and allergic rhinitis, which are allergic responses to the inhalation of flours and dust from wheat and other cereals (rye, barley) known since Roman times. Both allergies rank among the most prevalent occupational diseases and affect 1–10% (baker's asthma) and 18–29% (allergic rhinitis) of bakers, millers, and confectioners. Depending on the severity of the reaction, vocational retraining may be necessary. More than 100 IgE-binding proteins were identified, of which chloroform–methanol-soluble (CM-) ATIs, lipid transfer (LTP) and non-specific lipid transfer proteins (nsLTP) are the major allergens (127).

Contact urticaria is an allergic reaction on the skin following contact with an eliciting allergen (128). Typical symptoms of urticaria are localized wheal-and-flare reactions such as hives and a raised, burning, and/or pruritic swelling of the skin,

often accompanied by angioedema that appear within 10–30 min after allergen exposure and fade away within hours. In contrast, contact dermatitis is accompanied by the appearance of large, burning, and itchy rashes, blisters, and wheals that take several days to weeks to heal (129). Similar to respiratory WA, millers, bakers, and flour handlers are most frequently affected.

Food Allergy to Wheat

Wheat is the third most common cause of food allergy (only surpassed by milk and egg) and has to be labeled on prepacked foods according to Codex Alimentarius Standard 1–1985 (130) and also non-prepacked foods according to EU regulation 1169/2011. Wheat food allergy occurs within a few hours of wheat ingestion and may present with symptoms on the skin (e.g., atopic dermatitis, urticaria, angioedema), in the respiratory tract (e.g., wheezing, bronchial obstruction), in the gastrointestinal tract (e.g., abdominal pain, bloating, diarrhea), and even anaphylaxis. In comparison to other allergies such as peanut allergy, the dose of wheat proteins needed to trigger allergic reactions is usually quite high (about 1 g), but may also be lower (10 mg) depending on individual sensitivities (131). The causative factors are non-gluten proteins as well as gluten proteins, with ATIs 0.19, CM1, CM2, CM3, and CM16, LTP, and nsLTP as well as α-gliadins, γ-gliadins, and HMW-GS as major allergens (132). The IgE-binding epitopes derived from the repetitive sequences of gluten proteins such as QFPQQQFPQQ (ω5-gliadins), QQSFPLQPQQ (ω1,2-gliadins), VQQQQFPGQQ (α-gliadins), QQLPQPQQP (γ-gliadins), and SQQQPPF (LMW-GS) with the consensus motif QQX_1PX_2QQ (with X_1 being L, F, S, or I and X_2 being Q, E, or G) are different to those reported for CD (133).

Wheat-Dependent Exercise-Induced Anaphylaxis

WDEIA is a special form of WA because wheat intake alone does not trigger the allergic reaction, but in combination with augmenting cofactors, such as physical exercise. Further cofactors are alcohol, acetylsalicylic acid (aspirin®), and other non-steroidal anti-inflammatory drugs and stress (134, 135). The estimated prevalence of WDEIA is <0.1%. Clinical features range from urticaria and angioedema to dyspnea, hypotension, collapse, and anaphylactic shock. The major triggers are ω5-gliadins and HMW-GS, but other gluten protein types may also be involved. A special form of WDEIA is an allergic reaction caused by epicutaneous sensitization with hydrolyzed wheat proteins in cosmetics (136). Although WDEIA is regarded as the best-studied model of cofactor-induced anaphylaxis, it is still not clear how exactly the cofactors act as adjuvants causing the reaction. A decreased activation threshold of mast cells and basophils has been discussed, but it appears to be more likely that cofactors increase the bioavailability of allergens by promoting small intestinal permeability (137).

Non-IgE-Mediated Wheat Allergies

Non-IgE-mediated food allergies are well-known in children under the age of three, and they are divided into three main clinical conditions: food protein-induced (FPI) enterocolitis syndrome, FPI proctocolitis, and FPI enteropathies. The pathomechanism is currently not well-known, but these conditions are characterized by high levels of IL-13 and TNF-α as drivers of intestinal epithelial damage and eosinophil infiltration. The most common trigger for all three conditions is cow's milk, but soy, rice, and wheat have also been reported (138). Non-IgE-mediated food allergies are less well-recognized in adults (139), but there is an increasing body of evidence that they may be under-recognized. NCGS might in fact be a non-IgE-mediated food allergy because of patient history (food allergy during childhood or presence of atopic diseases) and serological and histological findings such as positive serum anti-gliadin antibodies, *in vitro* basophil activation, and presence of eosinophils in the intestinal mucosa (140).

Non-Celiac Gluten Sensitivity

NCGS, frequently termed gluten or (non-celiac) wheat sensitivity, may be defined as a gluten (wheat)-dependent disorder with symptoms similar to CD, but usually normal small intestinal mucosa (141). Moreover, NCGS is characterized by the lack of serum TG2 antibodies and the missing association to HLA-DQ2/8 alleles. The exact prevalence is still unknown and may be similar to that of CD with reports ranging from about 1% in El Salvador (13) and Mexico (11) to 1.7% in Brazil (10), but may also be higher with up to 6% (142). The clinical differentiation of NCGS from other WRDs is difficult due to several common features: similar symptoms, wheat proteins as the triggering factor, and a GFD as recommended treatment (78). Typical gastrointestinal manifestations of NCGS are abdominal pain, bloating, and chronic diarrhea. Frequent extraintestinal complaints include headache, "foggy mind," fatigue, anxiety, depression, numbness in the legs, arms and fingers, and joint pain (143). In contrast to CD, NCGS is not associated with malabsorption, nutritional deficiencies, or increased risk for autoimmune diseases or malignancy.

Because symptoms disappear on a GFD, gluten proteins have been considered as a precipitating factor (144), but it is still not clear whether gluten is responsible. Other wheat constituents such as non-gluten proteins and FODMAPs might be additionally responsible for NCGS (145). Recent studies proposed that ATIs contribute to the development of NCGS by an innate immune response mediated by the toll-like receptor pathway (146). Due to the absence of NCGS-specific biomarkers, the diagnosis is currently made by exclusion of CD, WA, other food intolerances, and irritable bowel syndrome (147). The diagnosis is definitely proven by oral wheat challenge after at least 3 weeks on a GFD and the subsequent occurrence of typical symptoms (148). A GFD is recommended as treatment, whereby symptoms usually improve rapidly. In contrast to CD, where strict GFD has to be maintained, patients with NCGS could adapt a more liberal diet reducing the gluten intake by ≈90% (146).

Irritable Bowel Syndrome

Irritable bowel syndrome (IBS) is the most common and extensively evaluated functional bowel disorder (149). Its prevalence in adult individuals has been estimated in a range from 5 to 20%. IBS is not a single disease but rather a symptom cluster resulting from diverse pathologies (150).

Typical symptoms, including abdominal discomfort and pain, gas, bloating, and diarrhea with and without constipation, are similar to those of CD, NCGS, intestinal bacterial overgrowth, and lactose intolerance (151). The Rome IV criteria categorize IBS by the most predominant presenting symptoms: diarrhea, constipation, mixed, or unspecified (152). The majority of patients perceive their symptoms as being related to specific meals, in particular, foods rich in carbohydrates. Wheat is regarded as one of the most relevant IBS triggers, although which component of this cereal is involved remains unclear (153). Gluten, FODMAPs, other wheat proteins, for example, ATIs, have been suggested as possible factors for symptom generation. The experimental and clinical evidence on the role of gluten/wheat in IBS has been presented by Volta et al. (154). Some types of IBS, especially diarrhea-predominant cases, show symptomatic improvement on a GFD (155). Today, the diagnosis is not only based on the exclusion of other food hypersensitivities, but on positive diagnosis using symptom-related criteria ("Rome III diagnostic criteria") (156). The pathomechanism of IBS is not completely understood; factors important to the development of IBS include alterations in the gut microbiome, intestinal permeability, gut immune function, motility, visceral sensation, brain–gut interactions, and psychosocial status (157). Strategies for treatment are based on general recommendations such as avoidance of foods rich in fat and carbohydrates, dairy products, caffeine, and alcohol, on the one hand, and increased intake of DF, on the other hand. If patients can associate the ingestion of certain foods with the complaints, an improvement may be achieved after restriction of these foods. Diarrhea-predominant patients are advised to test a GFD for several months (155).

THE SOURCES OF CONFUSION AROUND THE SAD FACE OF WHEAT

Unsubstantiated Statements Blaming Wheat

Over the last decade, wheat has been the center of a vigorous debate related to health and nutrition, and it has gained an increasingly negative reputation among the Western population. The first and main source of confusion arose from several pseudoscientific books such as "Wheat Belly" (158) and "Grain Brain" (159), numerous media reports and statements by celebrities promoting the overall impression that wheat consumption has adverse health effects for the general population. The main statements claimed that wheat consumption was the cause of overweight and obesity, that wheat caused a whole number of other disorders such as diabetes type 2, asthma, reflux disease, sleep disorders, neuronal complaints, etc., that wheat bread had an overly high glycemic index, that wheat contained opioids that caused addiction, and that modern wheat had been genetically modified and contained unique toxic proteins that caused the higher prevalence of WRDs. In conclusion, the general population was advised to avoid the consumption of wheat products.

However, these books intermingle sound scientific evidence with theoretical, controversial, and wrong statements and make it virtually impossible for consumers to separate the truth from myths. The essential summary that "foods made from wheat make people sick, stupid, fat, and addictive" (158, 159) led to great uncertainties among consumers. As a consequence, the popularity of a GFD has increased, with up to 5% of the population in New Zealand reporting gluten avoidance (160) and up to 13% of the UK population self-reporting a WRD (9). The consumption of gluten-free foods has significantly increased over the last years, up from a global retail sales value of 1.95×10^9 USD in 2012 to 3.84×10^9 USD in 2017, with further increases projected in the coming years (161). The reasons why some people voluntarily adopt a GFD include that they think it might help them reduce weight, that they perceive a GFD as healthier and better for their overall well-being and that they self-diagnosed a WRD.

To counteract the increasing uncertainty among consumers, numerous counterstatements of the scientific community [e.g., (120, 162–165)] have emphasized that wheat consumption is safe for the vast majority of the population and that wheat avoidance is only necessary after medical diagnosis of a true WRD. These reviews have compiled convincing evidence to refute the abovementioned statements and assert that the regular consumption of whole-grain products is associated with reduced risks of type 2 diabetes and of colorectal cancer, likely reduced risks of colon cancer and cardiovascular diseases, and more favorable weight management (166).

The Increasing Prevalence of Wheat-Related Disorders

The second source of confusion related to wheat consumption is based on the increasing amount of evidence from well-founded epidemiological studies that show a rise in the prevalence of WRDs over the past 50 years. While this rise can be partially explained by better diagnostics and improved awareness, recent reports, e.g., from Denmark (167), Italy (168), the United States (169), and worldwide (79) show that the prevalence of CD has indeed increased over time. The same is reported for NCGS (170) and also for allergies and autoimmune diseases, in general (171). Despite ongoing research, the underlying causative factors have not been unambiguously identified so far. WRDs are initiated through a loss of immunotolerance to wheat proteins at a certain point in time, but the factors causing this initial loss are the subject of ongoing investigations and may also be different depending on the genetic predisposition, dietary habits, and overall lifestyle of each individual.

The most likely factors include the hygiene hypothesis that was originally based on the observation that the decreased frequency of overall infectious and parasitic diseases was inversely correlated to the increased frequency of allergic and autoimmune diseases seen in industrialized countries since the 1950s (88). The hygiene hypothesis is supported by studies on migrants, who are as likely to develop an autoimmune disease as individuals in the host country with a high incidence of autoimmune diseases, even if they originally came from a country with a low incidence of autoimmune diseases, but moved at a young age (172). Studies in mice support the protective effects

of pathogenic bacteria, viruses, and parasites on autoimmunity and the additional positive effects of commensals that stimulate innate and adaptive immune regulatory pathways (88). In contrast, a lack of physical activity, a lot of time spent indoors, and a diet rich in saturated fats and digestible carbohydrates, but poor in DF, are associated with a loss of microbial diversity of the gut, skin, and other tissues and a subsequent loss of symbiotic relationships with parasites and bacteria that used to exist during human evolution (173–175). Additional factors contributing to changes in the microbiome are antibiotics and vaccinations and decreased exposure to airborne bacteria, all of which may contribute to alterations in intestinal permeability. Intestinal barrier dysfunction has been associated with a variety of intestinal and systemic diseases, including CD (176).

Seen from the side of wheat and related cereals, changes in protein composition due to breeding, heat and cold stress, or agricultural practices have been postulated as potential contributors to a higher immunostimulatory potential of modern wheat species compared to landraces and heritage wheats. While protein expression patterns do differ between wheat species, i.e., diploid einkorn, tetraploid emmer, and durum wheat as well as hexaploid spelt and common wheat (177–179), there are currently too few comparative *in vitro* or *in vivo* studies available to allow a precise assessment as to whether these differences might be related to the prevalence of WRDs or not. For example, the 33-mer peptide from α2-gliadin that is frequently described as the immunodominant peptide in CD, was only detected in common wheat and spelt, but it was not present in emmer, durum wheat, or einkorn (180). Ancient wheats like einkorn, emmer and spelt were suggested to provide health benefits compared to common wheat, but recent reviews collected evidence demonstrating that they differ little in their composition. Thus, ancient wheats do not appear to be "healthier" than modern wheats (41), with some exceptions, e.g., high lutein and steryl ferulate contents in einkorn (181).

Several comparative studies on old and modern cultivars within the species *Triticum aestivum* set out to study the influence of breeding during the last century (182). So far, the results are somewhat inconclusive because one study from Canada reported an increase (183), two studies from the United States essentially reported no change (184, 185), and three others from the United Kingdom, Germany, and the United States reported a decrease in protein contents over time (186–188). All studies report a substantial influence of the growing conditions on the content and composition of wheat proteins. Regarding protein composition, an increase in glutenins and a decrease in gliadins and gliadin/glutenin ratios, but essentially no changes for albumins/globulins and gluten, were observed in German winter wheat cultivars from 1891 to 2010, all grown at the same location in three consecutive years (189). Similar results were reported by Ozuna and Barro (190). Principal component analysis of the chromatographic fingerprints of albumins/globulins, gliadins, and glutenins showed a cluster formation of the most modern cultivars (first registered from 1981 to 2010) and of the oldest ones (1891–1920), but with exceptions and samples from 1951 to 1980 in between. The largest variability in protein profiles was observed for the samples from 1921 to 1950 [**Figure 3**, (191)]. Altogether, the evidence, so far, points to the conclusion that old and modern wheat cultivars do show changes in protein composition due to breeding, but so far, none of these changes seems to be linked to the prevalence of WRDs.

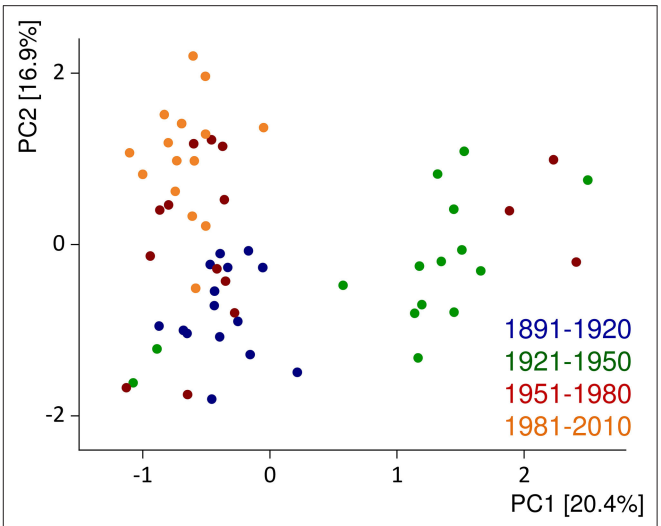


FIGURE 3 | Principal component (PC) analysis biplot of protein fingerprints of albumins/globulins, gliadins, and glutenins relative to the sum of extractable proteins. The data are displayed for 60 German winter wheat cultivars first registered from 1891 to 2010 and show the average of three harvest years (2015–2017). Figure modified from Pronin et al. (191).

As wheat is part of a huge variety of products, the amounts of immunoreactive proteins in the end-product that is eventually consumed is more important than in the original flour. Wheat processing techniques as well as strategies to reduce exposure have been extensively reviewed recently (165). The use of ungerminated grains, of refined white flour instead of wholegrain flour, of fast straight-dough yeast fermentation instead of diverse and long sourdough fermentations, as well as the use of wheat gluten as a technofunctional additive in a number of food products (192) have been discussed as additional factors that may contribute to causing WRDs (165, 193). Consumption of these wheat products may have increased the total amount of gluten in the diet, thus surpassing a certain threshold level necessary to trigger WRDs. However, credible data about adverse effects of modern wheat processing are not available, and no epidemiological studies have evaluated the contribution of modern processing on the increasing prevalence of WRDs. For example, there is no proof that modern bread making, including short and non-acidic fermentation of doughs or addition of vital gluten, resulted in higher gluten immunoreactivity. Epidemiological data demonstrate that countries, in which long-fermented sourdough breads are common (e.g., Finland and Sweden) even have higher CD prevalences than countries, where short-fermented yeast-leavened breads are consumed almost exclusively (e.g., Italy and Spain) (194).

The Unknowns

The third source of confusion mostly arises from the fact that the triggers for NCGS and wheat-sensitive IBS have not been clearly identified, so far (140), partly due to short-comings in the design of the nutritional intervention studies, placebo and nocebo effects (195), and/or insufficient characterization of the wheat product or wheat extract administered. The constituents that are discussed as causes of NCGS are gluten, non-gluten proteins (e.g., ATIs), and FODMAPs. In many cases, products containing wheat flour were used for oral food challenge, and in this case, gluten, non-gluten proteins, and FODMAPs were present (196). In other cases, wheat gluten was used, but these isolates also contain non-gluten proteins, so that the effects of gluten and ATIs cannot be clearly distinguished (195, 197). Gluten, as a cause of NCGS, has mostly been inferred from the fact that NCGS patients' symptoms are alleviated if they follow a GFD, but several studies have reported that gluten may not be the causative factor in NCGS, but rather FODMAPs (195, 198). ATIs were identified as triggers of innate immunity via the toll-like receptor 4, and they have been implicated in causing NCGS (146) and acting as adjuvants of other inflammatory diseases (78).

Furthermore, there is a significant overlap of symptoms of CD, NCGS, and IBS. It is clear that CD patients need to follow a strict GFD, and no positive effects of a GFD or a low FODMAP diet have been proven for healthy individuals (199). In between, both NCGS and IBS patients benefit from a low FODMAP diet, but even more so of a GFD, probably because of a multifactorial etiology of NCGS that combine a function effect caused by FODMAPs with a mild immune reaction combined by a dysbalance of microbiota (200). In most cases, gluten-free raw materials are naturally low in FODMAPs and also low in ATIs, so that a GFD is low in all potential causes for NCGS. In this context, sourdough fermented breads may also be better tolerated by NCGS and IBS patients, because *Lactobacillaceae* and *Bifidobacteriaceae*, as well as fungi and yeasts, possess enzymes capable of degrading gluten and FODMAPs (201). Due to the overlap between a low FODMAP diet and a GFD, it is likely that patients with NCGS, especially self-diagnosed ones, are more likely to suffer from IBS. *Vice versa*, there is a subgroup of NCGS patients among the IBS patients (140). A FODMAP-restricted diet is recommended to treat IBS, but not in the long term, because FODMAPs are also part of DF, and a complete elimination will have a negative effect on gut microbial diversity. Further well-designed dietary intervention studies with appropriate controls and sufficient characterization of the challenge materials are still needed to identify the causes of NCGS and also differentiate which patients will benefit from a GFD or rather a low FODMAP diet.

THE SMILING FACE OF WHEAT PREVAILS

In the last 10 years, wheat has received much negative attention because several pseudoscientific books and numerous media reports fueled the overall assumption that wheat consumption makes people sick. Despite the common consumer perception of a GFD as being healthy, gluten-free foods had higher contents of fat, saturated fat, sugar, and salt compared to gluten-containing foods (202). The most common deficiencies on a GFD are insufficient amounts of DF, vitamins, calcium, iron, magnesium, and zinc. Moreover, a strict GFD decisively reduces the quality of life, as CD patients can confirm, and may lead to low compliance (203).

Wheat is among the oldest and most extensively grown crops and one foundation to ensure food security for the increasing world population. Technological advances in breeding, farming, and processing have paved the way for wheat to become one of the most widespread and cheapest raw materials for food and non-food applications. No other food crop supplies humans with such a huge diversity of products from bread to other baked goods and pasta products that serve as staple foods all over the world. Wheat-based foods provide valuable nutrients such as proteins, DF, vitamins, minerals, and bioactive phytochemicals and supply up to 20% of the energy intake of the global population. Additionally, wheat is important for many non-food applications, and a number of wheat constituents like starch and gluten are present in items of daily use. Considering all available evidence, so far, there is no reason to eliminate wheat from the diet, except for individuals suffering from WRDs.

AUTHOR CONTRIBUTIONS

HW wrote the original draft. PK and KS wrote, reviewed, and edited the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We acknowledge support by the KIT-Publication Fund of the Karlsruhe Institute of Technology.

REFERENCES

1. FAOSTAT. *Food and Agriculture Organization of the United Nations* (2020). Available online at: http://www.fao.org/faostat/en/#data
2. Balfourier F, Bouchet S, Robert S, De Oliveira R, Rimbert H, Kitt J, et al. Worldwide phylogeography and history of wheat genetic diversity. *Sci. Adv.* (2019) 5:eaav0536. doi: 10.1126/sciadv.aav0536
3. Shiferaw B, Smale M, Braun HJ, Duveiller E, Reynolds M, Muricho G. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Sec.* (2013) 5:291–317. doi: 10.1007/s12571-013-0263-y
4. Chen GC, Tong X, Xu JY, Han SF, Wan ZX, Qin JB, et al. Whole-grain intake and total, cardiovascular, and cancer mortality: a systematic review and meta-analysis of prospective studies. *Am J Clin Nutr.* (2016) 104:164–72. doi: 10.3945/ajcn.115.122432
5. Jonnalagadda SS, Harnack L, Hai Liu R, McKeown N, Seal C, Liu S, et al. Putting the whole grain puzzle together: health benefits associated with whole grains—summary of american society for nutrition 2010 satellite symposium. *J Nutr.* (2011) 141:1011S−22S. doi: 10.3945/jn.110.132944
6. Reynolds AN, Akerman AP, Mann J. Dietary fibre and whole grains in diabetes management: systematic review and meta-analyses. *PLoS Med.* (2020) 17:e1003053. doi: 10.1371/journal.pmed.1003053

7. Ludvigsson JF, Leffler DA, Bai JA, Biagi F, Fasano A, Green PHR, et al. The oslo definitions for coeliac disease and related terms. *Gut.* (2013) 62:43–52. doi: 10.1136/gutjnl-2011-301346
8. Potter MD, Walker MM, Jones MP, Koloski NA, Keely S, Talley NJ. Wheat intolerance and chronic gastrointestinal symptoms in an Australian population-based study: association between wheat sensitivity, celiac disease and functional gastrointestinal disorders. *Am J Gastroenterol.* (2018) 113:1036–44. doi: 10.1038/s41395-018-0095-7
9. Aziz I, Lewis NR, Hadjivassiliou M, Winfield SN, Rugg N, Kelsall A, et al. A UK study assessing the population prevalence of self-reported gluten sensitivity and referral characteristics to secondary care. *Eur J Gastroenterol Hepatol.* (2014) 26:33–9. doi: 10.1097/01.meg.0000435546.87251.f7
10. Arámburo-Gálvez JG, Beltrán-Cárdenas CE, Geralda André T, Carvalho Gomes I, Macêdo-Callou MA, Braga-Rocha ÉM, et al. Prevalence of adverse reactions to gluten and people going on a gluten-free diet: a survey study conducted in Brazil. *Medicina.* (2020) 56:163. doi: 10.3390/medicina56040163
11. Ontiveros N, López-Gallardo JA, Vergara-Jiménez MJ, Cabrera-Chávez F. Self-reported prevalence of symptomatic adverse reactions to gluten and adherence to gluten-free diet in an adult Mexican population. *Nutrients.* (2015) 7:6000–15. doi: 10.3390/nu7075267
12. Van Gils T, Nijeboer P, IJssennagger CE, Sanders DS, Mulder CJ, Bouma G. Prevalence and characterization of self-reported gluten sensitivity in The Netherlands. *Nutrients.* (2016) 8:714. doi: 10.3390/nu8110714
13. Ontiveros N, Rodríguez-Bellegarrigue CI, Galicia-Rodríguez G, Vergara-Jiménez MDJ, Zepeda-Gómez EM, Arámburo-Galvez JG, et al. Prevalence of self-reported gluten-related disorders and adherence to a gluten-free diet in Salvadoran adult population. *Int J Environ Res Public Health.* (2018) 15:786. doi: 10.3390/ijerph15040786
14. Feuillet C, Langridge P, Waugh R. Cereal breeding takes a walk on the wild side. *Trends Genet.* (2008) 24:24–32. doi: 10.1016/j.tig.2007.11.001
15. International Wheat Genome Sequencing Consortium (IWGSC). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science.* (2018) 361:eaar7191. doi: 10.1126/science.aar7191
16. Pont C, Leroy T, Seidel M, Tondelli A, Duchemin W, Armisen D, et al. Tracing the ancestry of modern bread wheats. *Nat Genet.* (2019) 51:905–11. doi: 10.1038/s41588-019-0393-z
17. Mondal S, Rutkoski JE, Velu G, Singh PK, Crespo-Herrera LA, Guzmán C, et al. Harnessing diversity in wheat to enhance grain yield, climate resilience, disease and insect pest resistance and nutrition through conventional and modern breeding approaches. *Front Plant Sci.* (2016) 7:991. doi: 10.3389/fpls.2016.00991
18. Reganold JP, Wachter JM. Organic agriculture in the twenty-first century. *Nat Plants.* (2016) 2:15221. doi: 10.1038/nplants.2015.221
19. Asseng S, Ewert F, Martre P, Rötter RP, Lobell DB, Cammarano D, et al. Rising temperatures reduce global wheat production. *Nat Clim Change.* (2015) 5:143–7. doi: 10.1038/nclimate2470
20. Trnka M, Hlavinka P, Semenov MA. Adaptation options for wheat in Europe will be limited by increased adverse weather events under climate change. *J R Soc Interface.* (2015) 12:20150721. doi: 10.1098/rsif.2015.0721
21. Souci SW, Fachmann W, Kraut H. *Lebensmitteltabelle für die Praxis (5. Aufl).* Germany: Wissenschaftliche Verlagsgesellschaft Stuttgart (2011). p. 241.
22. Andersson AA, Andersson R, Piironen V, Lampi AM, Nyström L, Boros D, et al. Contents of dietary fibre components and their relation to associated bioactive components in whole grain wheat samples from the HEALTHGRAIN diversity screen. *Food Chem.* (2013) 15:1243–8. doi: 10.1016/j.foodchem.2012.09.074
23. Haska L, Nyman M, Andersson R. Distribution and characterisation of fructan in wheat milling fractions. *J Cereal Sci.* (2008) 48:768–74. doi: 10.1016/j.jcs.2008.05.002
24. Roberfroid M, Gibson GR, Hoyles L, McCartney AL, Rastall R, Rowland I, et al. Prebiotic effects: metabolic and health benefits. *Brit J Nutr.* (2010) 104:S1–63. doi: 10.1017/S0007114510003363
25. Pasqualetti V, Altomare A, Guarino MPL, Locato V, Cocca S, Cimini S, et al. Antioxidant activity of inulin and its role in the prevention of human colonic muscle cell impairment induced by lipopolysaccharide mucosal exposure. *PLoS ONE.* (2014) 9:e98031. doi: 10.1371/journal.pone.0098031
26. Vogt LM, Meyer D, Pullens G, Faas MM, Venema K, Ramasamy U, et al. Toll-Like receptor 2 activation by β2→ 1-fructans protects barrier function of T84 human intestinal epithelial cells in a chain length-dependent manner. *J Nutr.* (2014) 144:1002–8. doi: 10.3945/jn.114.191643
27. Verspreet J, Dornez E, Van den Ende W, Delcour JA, Courtin CM. Cereal grain fructans: structure, variability and potential health effects. *Trends Food Sci Technol.* (2015) 43:32–42. doi: 10.1016/j.tifs.2015.01.006
28. Delcour JA, Hoseney RC. Starch. In: *Principles of Cereal Science and Technology*. 3rd Ed. St. Paul: AACCI (2010). p. 23–51. doi: 10.1094/9781891127632
29. Englyst HN, Cummings JH. Digestion of the polysaccharides of some cereal foods in the human small intestine. *Am J Clin Nutr.* (1985) 42:778–87. doi: 10.1093/ajcn/42.5.778
30. Birt DF, Boylston T, Hendrich S, Jane JL, Hollis J, Li L, et al. Resistant starch: promise for improving human health. *Adv Nutr.* (2013) 4:587–601. doi: 10.3945/an.113.004325
31. Hasjim J, Lee SO, Hendrich S, Setiawan S, Ai Y, Jane JL. Characterization of novel resistant-starch and its effects on postprandial plasma-glucose and insulin responses. *Cereal Chem.* (2010) 87:257–62. doi: 10.1094/CCHEM-87-4-0257
32. Ashwar BA, Gani A, Shah A, Wani IA, Masoodi FA. Preparation, health benefits and applications of resistant starch – a review. *Starch Stärke.* (2015) 68:287–301. doi: 10.1002/star.201500064
33. Topping DL, Bajka BH, Bajka AR, Bird JM, Clarke L, Cobiac MA, et al. Resistant starches as a vehicle for delivering health benefits to the human large bowel. *Microb. Ecol. Health Dis.* (2008) 20:103–8. doi: 10.1080/08910600802106541
34. Robertson MD, Bickerton AS, Dennis AL, Vidal H, Frayn KN. Insulin-sensitizing effects of dietary resistant starch and effects on skeletal muscle and adipose tissue metabolism. *Am J Clin Nutr.* (2005) 82:559–67. doi: 10.1093/ajcn/82.3.559
35. Hamer HM, Jonkers DM, Bast A, Vanhoutvin SA, Fischer MA, Kodde A, et al. Butyrate modulates oxidative stress in the colonic mucosa of healthy humans. *Clin Nutr.* (2009) 28:88–93. doi: 10.1016/j.clnu.2008.11.002
36. Pawlak DB, Kushner JA, Ludwig DS. Effects of dietary glycemic index on adiposity, glucose homeostasis, and plasma lipids in animals. *Lancet.* (2004) 364:778–85. doi: 10.1016/S0140-6736(04)16937-7
37. Bednar GE, Patil AR, Murray SM, Grieshop CM, Merchen NR, Fahey GC Jr. Starch fiber fractions in selected food feed ingredients affect their small intestinal digestibility fermentability their large bowel fermentability *in vitro* in a canine mode. *J Nutr.* (2001) 131:276–86. doi: 10.1093/jn/131.2.276
38. Baghurst K, Baghurst PA, Record SJ. Dietary fiber, non-starch polysaccharide, and resistand starch intakes in Australia. In: Spiller GA, editor. *CRC Handbook of Dietary Fiber in Human Health.* Boca Raton, FL: CRC Press (2001). p. 583–91.
39. Shewry PR, Piironen V, Lampi AM, Edelmann M, Kariluoto S, Nurmi T, et al. The HEALTHGRAIN wheat diversity screen: effects of genotype and environment on phytochemicals and dietary fiber components. *J Agric Food Chem.* (2010) 58:9291–8. doi: 10.1021/jf100039b
40. Hossain K, Ulven C, Glover K, Ghavami F, Simsek S, Alamri MS, et al. Interdependence of cultivar and environment on fiber composition in wheat bran. *Aust J Crop Sci.* (2013) 7:525–31.
41. Shewry PR, Hey S. Do "ancient" wheat species differ from modern bread wheat in their contents of bioactive components? *J Cereal Sci.* (2015) 65:236–43. doi: 10.1016/j.jcs.2015.07.014
42. Scientific Advisory Committee on Nutrition Draft Carbohydrate and Health Report 2014 (2015). Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/445503/SACN_Carbohydrates_and_Health.pdf
43. Wcislo G, Szarlej-Wcislo K. Colorectal cancer prevention by wheat consumption: a three-valued logic – true, false, or otherwise? In: Watson RR, Preedy VR, Zibadi S, editors. *Wheat Rice in Disease Prevention Health.* London; Waltham, MA; San Diego: Academic Press, Elsevier (2014). p. 91–111. doi: 10.1016/B978-0-12-401716-0.00008-8
44. Sang S, Zhu Y. Bioactive phytochemicals in wheat bran for colon cancer prevention. In: Wheat Rice In: Watson RR, Preedy VR, Zibadi S, editors.*Disease Prevention Health.* London;

Waltham, MA; San Diego: Academic Press; Elsevier (2014). p. 121–9. doi: 10.1016/B978-0-12-401716-0.00010-6
45. Zhu Y, Conklin DR, Chen H, Wang L, Sang S. 5-alk(en)ylresorcinols as the major active components in wheat bran inhibit human colon cancer cell growth. *Bioorg Med Chem.* (2011) 19:3973–82. doi: 10.1016/j.bmc.2011.05.025
46. Aune D, Keum N, Giovannucci E, Fadnes LT, Boffetta P, Greenwood DC, et al. Whole grain consumption and risk of cardiovascular disease, cancer, and all cause and cause specific mortality: systematic review and dose-response meta- analysis of prospective studies. *BMJ.* (2016) 353:i2716. doi: 10.1136/bmj.i2716
47. Nieman KM, Rains TM, Dicklin MR, Maki KC. Wheat fiber in postprandial metabolic profile health. In: Watson RR, Preedy VR, Zibadi S, editors. *Wheat Rice in Disease Prevention Health*. London; Waltham, MA; San Diego: Academic Press; Elsevier (2014) p. 59–66. doi: 10.1016/B978-0-12-401716-0.00005-2
48. Chanson-Rolle A, Meynier A, Aubin F, Lappi J, Poutanen K, Braesco V. Systematic review and meta-analysis of human studies to support a quantitative recommendation for whole grain intake in relation to type 2 diabetes. *PLoS ONE.* (2015) 10:e0131377. doi: 10.1371/journal.pone.0131377
49. Aune D, Norat T, Romundstad P, Vatten L. Whole grain and refined grain consumption and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis of cohort studies. *Eur J Epidemiol.* (2013) 28:845–58. doi: 10.1007/s10654-013-9852-5
50. Svihus B. Nutritive digestive effects of starch fiber in whole wheat. In: Watson RR, Preedy VR, Zibadi S,editors. *Wheat Rice in Disease Prevention Health*. London; Waltham, MA; San Diego: Academic Press; Elsevier (2014). p. 81–87. doi: 10.1016/B978-0-12-401716-0.00007-6
51. Shewry P. Wheat. *J Exp Botany.* (2009) 60:1537–53. doi: 10.1093/jxb/erp058
52. Triboi E, Martre P, Triboi-Blondel AM. Environmentally-induced changes in protein composition in developing grains of wheat are related to changes in total protein content. *J Exp Bot*. (2003) 54:1731–42. doi: 10.1093/jxb/erg183
53. Groos C, Robert N, Bervas E, Charmet G. Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat. *Theor Appl Genet.* (2003) 106:1032–40. doi: 10.1007/s00122-002-1111-1
54. Shewry P. What is gluten – why is it special? *Front Nutr.* (2019) 6:101. doi: 10.3389/fnut.2019.00101
55. Wieser H, Koehler P, Konitzer K. Gluten – the precipitating factor. In: *Celiac disease and gluten – Multidisciplinary challenges and opportunities.* 1st ed. London; Waltham, MA; San Diego: Academic Press; Elsevier (2014). p. 97–148. doi: 10.1016/B978-0-12-420220-7.00002-X
56. Scherf KA. Immunoreactive cereal proteins in wheat allergy, non-celiac gluten/wheat sensitivity (NCGS) and celiac disease. *Curr Opin Food Sci.* (2019) 25:35–41. doi: 10.1016/j.cofs.2019.02.003
57. Brijs K, Courtin CM, Goesaert H, Gebruers K, Delcour JA, Shewry PR, et al. Enzymes and enzyme inhibitors endogenous to wheat. In: Khan K, Shewry PR, et al. *Wheat – Chemistry and Technology*. St Paul, MN: AACC (2009). p. 401–35. doi: 10.1094/9781891127557.011
58. Altenbach SB, Vensel WH, Dupont FM. The spectrum of low molecular weight α-amylase/protease inhibitor genes expressed in the US bread wheat cultivar Butte 86. *BMC Res Notes.* (2011) 4:242–54. doi: 10.1186/1756-0500-4-242
59. Dupont FM, Vensel WH, Tanaka CK, Hurkman WJ, Altenbach SB. Deciphering the complexities of the wheat flour proteome using quantitative two-dimensional electrophoresis, three proteases and tandem mass spectrometry. *Proteome Sci.* (2011) 9:10. doi: 10.1186/1477-5956-9-10
60. FAO/WHO/UNO. *Protein and amino acid requirements in human nutrition*. WHO Press (2007). p. 265. Available online at: https://www.who.int/nutrition/publications/nutrientrequirements/WHO_TRS_935/en/
61. Chick H. Biological value of the proteins contained in wheat flours. *Lancet*. (1942) 242:405–8. doi: 10.1016/S0140-6736(00)79573-0
62. Scherf KA, Koehler P, Wieser H. Gluten and wheat sensitivities – an overview. *J Cereal Sci.* (2016) 67:2–11. doi: 10.1016/j.jcs.2015.07.008
63. González-Thuillier I, Salt L, Chope G, Penson S, Skeggs P, Tosi P, et al. Distribution of lipids in the grain of wheat (cv. Hereward) determined by lipidomic analysis of milling and pearling fractions. *J Agric Food Chem.* (2015) 63:10705–16. doi: 10.1021/acs.jafc.5b05289
64. Schaffarczyk M, Østdal H, Koehler P. Lipases in wheat breadmaking: analysis functional effects of lipid reaction products. *J Agric Food Chem*. (2014) 62:8229–37. doi: 10.1021/jf5022434
65. Chen Y, Ross AB, Aman P, Kamal-Eldin A. Alkylresorcinols as markers of whole grain wheat and rye in cereal products. *J Agric Food Chem.* (2004) 52:8242–6. doi: 10.1021/jf049726v
66. Koskela A, Samaletdin A, Aubertin-Leheudre M, Adlercreutz H. Quantification of alkylresorcinol metabolites in plasma by high-performance liquid chromatography with coulometric electrode array detection. *J Agric Food Chem.* (2008) 56:7678–81. doi: 10.1021/jf801252s
67. Shewry PR, Van Schaik F, Ravel C, Charmet G, Rakszegi M, Bedo Z, et al. Genotype and environment effects on the contents of vitamins B1, B2, B3, and B6 in wheat grain. *J Agric Food Chem.* (2011) 59:10564–71. doi: 10.1021/jf202762b
68. Liu RH. Whole grain phytochemicals and health. *J Cereal Sci.* (2007) 46:207–19. doi: 10.1016/j.jcs.2007.06.010
69. Chung OK, Ohm JB, Ram MS, Park SH, Howitt CA. Wheat Lipids. In Khan K, Shewry PR, editors. *Wheat: chemistry technology,* 4th ed. St Paul, MN: AACC (2009). p. 363–99. doi: 10.1094/9781891127557.010
70. Piironen V, Lampi AM, Ekholm P, Salmenkallio-Marttila M, Liukkonen H, et al. Micronutrients phytochemicals in wheat grain. In: Khan K, Shewry PR, editors. *Wheat: chemistry and technology*, 4th ed. St Paul, MN: AACC (2009). p. 179–222. doi: 10.1094/9781891127557.007
71. Beta T, Nam S, Dexter JE, Sapirstein HD. Phenolic content and antioxidant activity of pearled wheat and roller-milled fractions. *Cereal Chem.* (2005) 82:390–3. doi: 10.1094/CC-82-0390
72. Zhu Y, Sang S. Phytochemicals in whole grain wheat and their health-promoting effects. *Mol Nutr Food Res.* (2017) 61:1600852. doi: 10.1002/mnfr.201600852
73. Badawy D, El-Bassossy HM, Fahmy A, Azhar A. Aldolase reductase inhibitors zopolrestat and ferulic acid alleviate hypertension associated with diabetes: effect on vascular reactivity. *Can J Physiol Pharmacol.* (2013) 91:101–7. doi: 10.1139/cjpp-2012-0232
74. Drankhan K, Carter J, Madl R, Klopfenstein C, Padula F, Lu Y. Antitumor activity of wheat with high orthophenolic content. *Nutr Cancer.* (2003) 47:188–94. doi: 10.1207/s15327914nc4702_12
75. Hemdane S, Jacobs PJ, Dornez E, Verspreet J, Delcour JA, Courtin CM. Wheat (*Triticum aestivum* L.) bran in bread making: a critical review. *Compr Rev Food Sci Food Safety*. (2016) 15:28–42. doi: 10.1111/1541-4337.12176
76. OECD-FAO. *Agricultural Outlook OECD Agriculture Statistics (Database)* (2019). Available online at: https://stats.oecd.org/
77. Zentgraf H. Unser Brotgetreide Nr. 1. *Mehlreport*. (2013) 22:1–4.
78. Schuppan D, Zevallos V. Wheat amylase trypsin inhibitors as nutritional activators of innate immunity. *Dig Dis.* (2015) 33:260–3. doi: 10.1159/000371476
79. Singh P, Arora A, Strand TA, Leffler DA, Catassi C, Green PH, et al. Global prevalence of celiac disease: systematic review and meta-analysis. *Clin Gastroenterol Hepatol.* (2018) 16:823–36. doi: 10.1016/j.cgh.2017.06.037
80. Ludvigsson JF, Murray JA. Epidemiology of celiac disease. *Gastroenterol. Clin. North Am.* (2019) 11:54–60. doi: 10.1016/j.gtc.2018.09.004
81. Plugis NM, Khosla C. Therapeutic approaches for celiac disease. *Best Pract Res Clin Gastroenterol.* (2015) 29:503–21. doi: 10.1016/j.bpg.2015.04.005
82. Huebener S, Tanaka CK, Uhde M, Zone JJ, Vensel WH, Kasarda DD, et al. Specific nongluten proteins of wheat are novel target antigens in celiac disease humoral response. *J Proteome Res.* (2015) 14:503–11. doi: 10.1021/pr500809b
83. Junker Y, Zeissig S, Kim SJ, Barisani D, Wieser H, Leffler DA, et al. Wheat amylase trypsin inhibitors drive intestinal inflammation via activation of toll-like receptor 4. *J Exp Med.* (2012) 209:2395–408. doi: 10.1084/jem.20102660
84. Abadie V, Sollid LM, Barreiro LB, Jabri B. Integration of genetic and immunological insights into a model of celiac disease pathogenesis. *Annu Rev Immunol.* (2011) 29:493–525. doi: 10.1146/annurev-immunol-040210-092915
85. Ricano-Ponce I, Wijmenga C, Gutierrez-Achury J. Genetics of celiac disease. *Best Pract Res Clin Gastroenterol.* (2015) 29:399–412. doi: 10.1016/j.bpg.2015.04.004
86. Koning F. Celiac disease: quantity matters. *Semin Immunopathol*. (2012) 34:541–9. doi: 10.1007/s00281-012-0321-0

87. Pes GM, Bibbo S, Dore MP. Celiac disease: beyond genetic susceptibility and gluten. *Narrative Rev Ann Med.* (2019) 9:1–47. doi: 10.1080/07853890.2019.1569254
88. Bach JF. The hygiene hypothesis in autoimmunity: the role of pathogens and commensals. *Nat Rev Immunol.* (2018) 18:105–20. doi: 10.1038/nri.2017.111
89. Lebwohl B, Murray JA, Verdu EF, Crowe SE, Dennis M, Fasano A, et al. Gluten introduction, breastfeeding, and celiac disease: back to the drawing board. *Am J Gastroenterol.* (2016) 111:12–4. doi: 10.1038/ajg.2015.219
90. Szajewska H, Shamir R, Mearin L, Ribes-Koninckx C, Catassi C, Domellöf M, et al. Gluten introduction and the risk of coeliac disease: a position paper by the european society for pediatric gastroenterology, hepatology, and nutrition. *J Pediatr Gastroenterol Nutr.* (2016) 62:507–13. doi: 10.1097/MPG.0000000000001105
91. Meijer C, Shamir R, Szajewska H, Mearin L. Celiac disease prevention. *Front Pediatr.* (2018) 6:368. doi: 10.3389/fped.2018.00368
92. Green PH, Krishnareddy S, Lebwohl B. Clinical manifestations of celiac disease. *Dig Dis.* (2015) 33:137–40. doi: 10.1159/000370204
93. Leffler DA, Green PH, Fasano A. Extraintestinal manifestations of coeliac disease. *Nat Rev Gastroenterol Hepatol.* (2015) 12:561–71. doi: 10.1038/nrgastro.2015.131
94. Oberhuber G, Granditsch G, Vogelsang, H. The histopathology of coeliac disease: time for a standardized report scheme for pathologists. *Eur J Gastroenterol. Hepatol.* (1999) 11:1185–94. doi: 10.1097/00042737-199910000-00019
95. Conti L, Lahner E, Galli G, Esposito G, Carabotti M, Annibale B. Risk factors associated with the occurrence of autoimmune diseases in adult coeliac patients. *Gastroenterol Res Pract.* (2018) 2018:3049286. doi: 10.1155/2018/3049286
96. Kaukinen K, Lindfors K, Collin P, Koskinen O, Mäki M. Coeliac disease – a diagnostic and therapeutic challenge. *Clin Chem Lab Med.* (2010) 48:1205–16. doi: 10.1515/cclm.2010.241
97. Kim HS, Unalp-Arida A, Ruhl CE, Choung RS, Murray JA. Autoimmune and allergic disorders are more common in people with celiac disease or on a gluten-free diet in the United States. *J Clin Gastroenterol.* (2019) 53:e416–23. doi: 10.1097/MCG.0000000000001100
98. Malamut G, Cellier C. Refractory celiac disease. *Gastroenterol Clin North Am.* (2019) 48:137–44. doi: 10.1016/j.gtc.2018.09.010
99. Husby S, Koletzko S, Korponay-Szabó I, Kurppa K, Mearin ML, Ribes-Koninckx C, et al. European society paediatric gastroenterology, hepatology nutrition guidelines for diagnosing coeliac disease 2020. *J Pediatr Gastroenterol Nutr*. (2020) 70:141–56. doi: 10.1097/MPG.0000000000002497
100. Reilly N, Husby S, Sanders DS, Green PH. Coeliac disease: to biopsy or not? *Nat Rev Gastroenterol Hepatol.* (2018) 15:60–6. doi: 10.1038/nrgastro.2017.121
101. Hujoel IA, Reilly NR, Rubio-Tapia A. Celiac disease: clinical features and diagnosis. *Gastroenterol Clin North Am.* (2019) 48:19–37. doi: 10.1016/j.gtc.2018.09.001
102. Korponay-Szabó IR, Troncone R, Discepolo V. Adaptive diagnosis of coeliac disease. *Best Pract Res Clin Gastroenterol.* (2015) 29:381–98. doi: 10.1016/j.bpg.2015.05.003
103. Lundin KE, Wijmenga C. Coeliac disease and autoimmune disease – genetic overlap and screening. *Nat Rev Gastroenterol Hepatol.* (2015) 12:507–15. doi: 10.1038/nrgastro.2015.136
104. Schuppan D, Junker Y, Barisani D. Celiac disease: from pathogenesis to novel therapies. *Gastroenterology.* (2009) 137:1912–33. doi: 10.1053/j.gastro.2009.09.008
105. Valitutti F, Fasano A. *Breaking down* barriers: how understanding celiac disease pathogenesis informed the development of novel treatments. *Dig Dis Sci*. (2019) 64:1748–58. doi: 10.1007/s10620-019-05646-y
106. Goodman RE, Ebisawa M, Ferreira F, Sampson HA, van Ree R, Vieths S, et al. AllergenOnline: A peer-reviewed, curated allergen database to assess novel food proteins for potential cross-reactivity. *Mol Nutr Food Res*. (2016) 60:1183–98. doi: 10.1002/mnfr.201500769
107. Dieterich W, Ehnis T, Bauer M, Donner P, Volta U, Riecken EO, et al. Identification of tissue transglutaminase as the autoantigen of celiac disease. *Nat Med.* (1997) 3:797–801. doi: 10.1038/nm0797-797
108. Fleckenstein B, Qiao SW, Larsen MR, Jung G, Roepstorff P, Sollid LM. Molecular characterization of covalent complexes between tissue transglutaminase and gliadin peptides. *J Biol Chem.* (2004) 279:17607–16. doi: 10.1074/jbc.M310198200
109. Lexhaller B, Ludwig C, Scherf KA. Identification of isopeptides between human tissue transglutaminase and wheat, rye, and barley gluten peptides. *Sci Rep.* (2020) 10:7426. doi: 10.1038/s41598-020-64143-9
110. Vader LW, de Ru A, van der Wal Y, Kooy YMC, Benckhuijsen W, Mearin ML, et al. Specificity of tissue transglutaminase explains cereal toxicity in celiac disease. *J Exp Med.* (2002) 195:643–9. doi: 10.1084/jem.20012028
111. Sollid LM, Iversen R, Steinsbo O, Qiao SW, Bergsen E, Dorum S, et al. Small bowel celiac disease and adaptive immunity. *Dig Dis*. (2015) 33:115–21. doi: 10.1159/000369512
112. Kim SM, Mayassi T, Jabri B. Innate immunity: actuating the gears of celiac disease pathogenesis. *Best Pract Res Clin Gastroenterol.* (2015) 29:425–35. doi: 10.1016/j.bpg.2015.05.001
113. Rai S, Kaur A, Chopra CS. Gluten-free products for celiac susceptible people. *Front Nutr.* (2018) 5:116. doi: 10.3389/fnut.2018.00116
114. White LE, Bannerman E, Gillett PM. Coeliac disease and the gluten-free diet: a review of the burdens; factors associated with adherence and impact on health-related quality of life, with specific focus on adolescence. *J Hum Nutr Diet.* (2016) 29:593–606. doi: 10.1111/jhn.12375
115. Serena G, Kelly CP, Fasano A. Nondietary therapies for celiac disease. *Gastroenterol Clin North Am.* (2019) 48:145–63. doi: 10.1016/j.gtc.2018.09.011
116. Reunala T, Salmi TT, Hervonen K, Kaukinen K, Collin. P. Dermatitis herpetiformis: a common extraintestinal manifestation of coeliac disease. *Nutrients.* (2018) 10:e602. doi: 10.3390/nu10050602
117. Hadjivassiliou M, Sanders DD, Aeschlimann DP. Gluten-related disorders: gluten ataxia. *Dig Dis.* (2015) 33:264–8. doi: 10.1159/000369509
118. 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. doi: 10.1186/1741-7015-10-13
119. Pasha I, Saeed F, Sultan MT, Batool R, Aziz M, Ahmed W. Wheat allergy and intolerance: recent updates and perspectives. *Crit Rev Food Sci Nutr*. (2016) 5:13–24. doi: 10.1080/10408398.2012.659818
120. Brouns F, van Rooy G, Shewry P, Rustgi S, Jonkers D. Adverse reactions to wheat or wheat components. *Compr Rev Food Sci Food Saf.* (2019) 18:1437–52. doi: 10.1111/1541-4337.12475
121. Juhasz A, Belova T, Florides CG, Maulis C, Fischer I, Gell G, et al. Genome mapping of seed-borne allergens and immunoresponsive proteins in wheat. *Sci Adv*. (2018) 4:eaar8602. doi: 10.1126/sciadv.aar8602
122. Tatham AS, Shewry PR. Allergens to wheat and related cereals. *Clin Exp Allergy.* (2008) 38:1712–26. doi: 10.1111/j.1365-2222.2008.03101.x
123. Nwaru BI, Hickstein L, Panesar SS, Roberts G, Muraro A, Sheikh A. Prevalence of common food allergies in Europe: a systematic review and meta- analysis. *Allergy.* (2014) 69:992–1007. doi: 10.1111/all.12423
124. Muraro A, Werfel T, Hoffmann-Sommergruber K, Roberts G, Beyer K, Bindslev-Jensen C, et al. EAACI food allergy and anaphylaxis guidelines: diagnosis and management of food allergy. *Allergy.* (2014) 69:1008–25. doi: 10.1111/all.12429
125. Sicherer SH, Sampson HA. Food allergy: epidemiology, pathogenesis, diagnosis, and treatment. *J Allergy Clin Immunol.* (2014) 133:291–307. doi: 10.1016/j.jaci.2013.11.020
126. Hawrylowicz CM, O'Garra A. Potential role of interleukin−10-secreting regulatory T cells in allergy and asthma. *Nat Rev Immunol*. (2005) 5:271–83. doi: 10.1038/nri1589
127. 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. doi: 10.1067/mai.2001.113761
128. Bhatia R, Alikhan A, Maibach HI. Contact urticaria: present scenario. *Indian J Dermatol.* (2009) 54:264–8. doi: 10.4103/0019-5154.55639
129. Matsuo H, Uemura M, Yorozuya M, Adachi A, Morita E. Identification of IgE-reactive proteins in patients with wheat protein contact dermatitis. *Contact Dermatitis.* (2010) 63:23–30. doi: 10.1111/j.1600-0536.2010.01741.x
130. Codex Standard 1-1985. *Codex Alimentarius Commission. Amendment* 7 (2010). Available online at: http://www.fao.org/3/Y2770e/y2770e02.htm#bm02

131. Hischenhuber C, Crevel R, Jarry B, Maeki M, Moneret-Vautrin DA, Romano A, et al. Review article: safe amounts of gluten for patients with wheat allergy or coeliac disease. *Aliment Pharmacol Therapeut*. (2006) 23:559–75. doi: 10.1111/j.1365-2036.2006.02768.x
132. Mäkelä MJ, Eriksson C, Kotaniemi-Syrjaenen A, Palosuo K, Marsh J, Borres M, et al. Wheat allergy in children – new tools for diagnostics. *Clin Exp Allergy*. (2014) 44:1420–30. doi: 10.1111/cea.12393
133. Battais F, Pineau F, Popineau Y, Aparicio C, Kanny G, Moneret-Vautrin DA, et al. Food allergy to wheat: identification of immunoglobulin E and immunoglobulin G-binding proteins with sequential extracts and purified proteins from wheat flour. *Clin Exp Allergy*. (2003) 33:962–70. doi: 10.1046/j.1365-2222.2003.01592.x
134. Scherf KA, Brockow K, Biedermann T, Koehler P, Wieser H. Wheat-dependent exercise-induced anaphylaxis. *Clin Exp Allergy*. (2016) 46:10–20. doi: 10.1111/cea.12640
135. Wölbing F, Fischer J, Köberle M, Kaesler S, Biedermann T. About the role and underlying mechanismens of cofactors in anaphylaxis. *Allergy*. (2013) 68:1085–92. doi: 10.1111/all.12193
136. Yokooji T, Kurihara S, Murakami T, Chinuki Y, Takahashi H, Morita E, et al. Characterization of causative allergens for wheat-dependent exercise-induced anaphylaxis sensitized with hydrolyzed wheat proteins in facial soap. *Allergol Int*. (2013) 62:435–45. doi: 10.2332/allergolint.13-OA-0561
137. Suzuki T. Regulation of intestinal epithelial permeability by tight junctions. *Cell Mol Life Sci*. (2013) 70:631–59. doi: 10.1007/s00018-012-1070-x
138. Turnbull JL, Adams HN, Gorard DA. Review article: the diagnosis and management of food allergy and food intolerances. *Aliment Pharmacol Therapeut*. (2015) 41:3–25. doi: 10.1111/apt.12984
139. Yu Y, Blokhuis BR, Garssen J, Redegeld FA. Non-IgE mediated mast cell activation. Eur. J. *Pharmacol*. (2016) 778:33–43. doi: 10.1016/j.ejphar.2015.07.017
140. Catassi C, Alaedini A, Bojarski C, Bonaz B, Bouma G, Carroccio A, et al. The overlapping area of non-celiac gluten sensitivity (NCGS) and wheat-sensitive irritable bowel syndrome (IBS): an update. *Nutrients*. (2017) 9:1268. doi: 10.3390/nu9111268
141. Volta U, de Giorgio R, Caio G, Uhde M, Manfredini R, Alaedini A. Nonceliac wheat sensitivity: an immune-mediated condition with systemic manifestations. *Gastroenterol Clin North Am*. (2019) 48:165–82. doi: 10.1016/j.gtc.2018.09.012
142. Mansueto P, Seidita A, D'Alcamo A, Carroccio A. Non-celiac gluten sensitivity: literature review. *J Am Coll Nutr*. (2014) 33:39–54. doi: 10.1080/07315724.2014.869996
143. Losurdo G, Principi M, Iannone A, Amoruso A, Ierardi E, di Leo A, et al. Extra-intestinal manifestations of non-celiac gluten sensitivity: an expanding paradigm. *World J Gastroenterol*. (2018) 24:1521–30. doi: 10.3748/wjg.v24.i14.1521
144. Di Sabatino A, Volta U, Salvatore C, Biancheri P, Caio G, de Giorgio R, et al. Small amounts of gluten in subjects with suspected nonceliac gluten sensitivity: a randomized, double-blind, placebo-controlled, cross-over trial. *Clin Gastroenterol Hepatol*. (2015) 13:1604–12.e3. doi: 10.1016/j.cgh.2015.01.029
145. Gibson PR, Muir JG, Newnham ED. Other dietary confounders: FODMAPS et al. *Dig Dis*. (2015) 33:269–76. doi: 10.1159/000371401
146. Zevallos VF, Raker V, Tenzer S, Jimenez-Calvente C, Ashfaq-Khan M, Rüssel N, et al. Nutritional wheat amylase-trypsin inhibitors promote intestinal inflammation via activation on myeloid cells. *Gastroenterology*. (2017) 152:1100–13.e12. doi: 10.1053/j.gastro.2016.12.006
147. Branchi F, Aziz I, Conte D, Sanders DS. Noncoeliac gluten sensitivity: a diagnostic dilemma. *Curr Opin Clin Nutr Metab Care*. (2015) 18:508–14. doi: 10.1097/MCO.0000000000000207
148. Catassi C, Elli L, Bonaz B, Bouma G, Carroccio A, Castillejo G, et al. Diagnosis of non-celiac gluten sensitivity (NCGS): the Salerno experts' criteria. *Nutrients*. (2015) 7:4966–77. doi: 10.3390/nu7064966
149. Rej A, Sanders DS. Gluten-free diet and its 'cousins' in irritable bowel syndrome. *Nutrients*. (2018) 10:e1727. doi: 10.3390/nu10111727
150. Chey W, Kurlander J, Eswaran S. Irritable bowel syndrome: a clinical review. *JAMA*. (2015) 313:949–58. doi: 10.1001/jama.2015.0954
151. Borghini R, Donato G, Alvaro D, Picarelli A. New insights in IBS-like disorders. Pandora's box has been opened; a review. *Gastroenterol Hepatol Bed Bench*. (2017) 10:79–89.
152. Drossmann DA. Functional gastrointestinal disorders: history, pathophysiology, clinical features and Rome IV criteria. *Gastroenterology*. (2016) 150:1262–79. doi: 10.1053/j.gastro.2016.02.032
153. De Giorgio R, Volta U, Gibson PR. Sensitivity to wheat, gluten and FODMAPs in IBS: facts or fiction? *Gut*. (2016) 65:169–78. doi: 10.1136/gutjnl-2015-309757
154. Volta U, Pinto-Sanchez MI, Boschetti E, Caio G, de Giorgio R, Verdu EF. Dietary triggers in irritable bowel syndrome: is there a role for gluten? *J Neurogastroenterol Motil*. (2016) 22:547–57. doi: 10.5056/jnm16069
155. Zanwar VG, Pawar SV, Gambhire PA, Jain SS, Surude RG, Shah VB, et al. Symptomatic improvement with gluten restriction in irritable bowel syndrome: a prospective, randomized, double-blinded placebo controlled trial. *Intest Res*. (2016) 14:343–50. doi: 10.5217/ir.2016.14.4.343
156. Longstreth GF, Thompson WG, Chey WD, Houghton LA, Mearin F, Spiller RC. Functional bowel disorders. *Gastroenterology*. (2006) 130:1480–91. doi: 10.1053/j.gastro.2005.11.061
157. Bellini M, Gambaccini D, Stasi C, Urbano, MT, Marchi S, Usai-Satta P. Irritable bowel syndrome: a disease still searching for pathogenesis, diagnosis and therapy. *World J Gastroenterol*. (2014) 20:8807–20. doi: 10.3748/wjg.v20.i27.8807
158. Davis W. *Wheat Belly*. New York, NY: Rodale Inc. (2011).
159. Perlmutter D. *Grain Brain*. New York, NY; Boston; London: Little, Brown and Company (2013).
160. Tanpowpong P, Ingham TR, Lampshire PK, Kirchberg FF, Epton MJ, Crane J, et al. Coeliac disease and gluten avoidance in New Zealand children. *Arch Dis Child*. (2012) 97:12–6. doi: 10.1136/archdischild-2011-300248
161. Statista Database (2018). Available online at: https://www.statista.com/statistics/819377/global-gluten-free-retail-sales/
162. Jones J. Wheat belly – an analysis of selected statements and basic theses from the book. *Cereal Foods World*. (2012) 57:177–89. doi: 10.1094/CFW-57-4-0177
163. Brouns FJ, van Buul VJ, Shewry PR. Does wheat make us fat and sick? *J Cereal Sci*. (2013) 58:209–15. doi: 10.1016/j.jcs.2013.06.002
164. Gaesser GA, Angadi SS. Navigating the gluten-free boom. *JAAPA*. (2015) 28:1–7. doi: 10.1097/01.JAA.0000469434.67572.a4
165. Rustgi S, Shewry P, Brouns F, Deleu LJ, Delcour JA. Wheat seed proteins: factors influencing their content, composition, and technological properties, and strategies to reduce adverse reactions. *Compr Rev Food Sci Food Saf*. (2019) 18:1751–69. doi: 10.1111/1541-4337.12493
166. Tieri M, Ghelfi F, Vitale M, Vetrani C, Marventano S, Lafranconi A, et al. Whole grain consumption human health: an umbrella review of observational studies. *Int J Food Sci Nutr*. (2020) 71:668–77. doi: 10.1080/09637486.2020.1715354
167. Grode L, Bech BH, Jensen TM, Humaidan P, Agerholm IE, Plana-Ripoll O, et al. Prevalence, incidence, and autoimmune comorbidities of celiac disease: a nation-wide, population-based study in Denmark from 1977 to 2016. *Eur J Gastroenterol Hepatol*. (2018) 30:83–91. doi: 10.1097/MEG.0000000000000992
168. Gatti S, Lionetti E, Balanzoni L, Verma AK, Galeazzi T, Gesuita R, et al. Increased prevalence of celiac disease in school-age children in Italy. *Clin Gastroenterol Hepatol*. (2020) 18:596–603. doi: 10.1016/j.cgh.2019.06.013
169. Almallouhi E, King KS, Patel B, Wi C, Juhn YJ, Murray JA, et al. Increasing incidence and altered presentation in a population-based study of pediatric celiac disease in North America. *J Pediatr Gastroenterol Nutr*. (2017) 65:432–7. doi: 10.1097/MPG.0000000000001532
170. Fasano A, Sapone A, Zevallos V, Schuppan D. Nonceliac gluten sensitivity. *Gastroenterology*. (2015) 148:1195–204. doi: 10.1053/j.gastro.2014.12.049
171. Verdu EF, Danska JS. Common ground: shared risk factors for type 1 diabetes and celiac disease. *Nat Immunol*. (2018) 19:685–95. doi: 10.1038/s41590-018-0130-2
172. Kuehni CE, Strippoli MP, Low N, Silverman M. Asthma in young south Asian women living in the United Kingdom: the importance of early life. *Clin Exp Allergy*. (2007) 37:47–53. doi: 10.1111/j.1365-2222.2006.02627.x
173. Hanski I, von Hertzen L, Fyhrquist N, Koskinen K, Torppa K, Laatikainen T, et al. Environmental biodiversity, human microbiota,

and allergy are interrelated. *Proc Nat Acad Sci USA*. (2012) 109:8334–9. doi: 10.1073/pnas.1205624109
174. Lambrecht B, Hammad H. The immunology of the allergy epidemic and the hygiene hypothesis. *Nat Immunol.* (2017) 18:1076–83. doi: 10.1038/ni.3829
175. Rook GA. Hygiene hypothesis and autoimmune diseases. *Clin Rev Allergy Immunol.* (2012) 42:5–15. doi: 10.1007/s12016-011-8285-8
176. Odenwald MA, Turner JR. The intestinal epithelial barrier: a therapeutic target? *Nat Rev Gastroenterol Hepatol.* (2017) 14:9–21. doi: 10.1038/nrgastro.2016.169
177. Prandi B, Tedeschi T, Folloni S, Galaverna G, Sforza S. Peptides from gluten digestion: a comparison between old and modern wheat varieties. *Food Res Int.* (2017) 91:92–102. doi: 10.1016/j.foodres.2016.11.034
178. Geisslitz S, Ludwig C, Scherf KA, Koehler P. Targeted LC-MS/MS reveals similar contents of α-amylase/trypsin-inhibitors as triggers of non-celiac gluten sensitivity in all wheat species except einkorn. *J Agric Food Chem.* (2018) 66:12395–403. doi: 10.1021/acs.jafc.8b04411
179. Geisslitz S, Longin CFH, Scherf KA, Koehler P. Comparative study on gluten protein composition of ancient (einkorn, emmer and spelt) and modern wheat species (durum and common wheat). *Foods.* (2019) 8:409. doi: 10.3390/foods8090409
180. Schalk K, Lang C, Wieser H, Koehler P, Scherf KA. Quantitation of the immunodominant 33-mer peptide from α-gliadin in wheat flours by liquid chromatography tandem mass spectrometry. *Sci Rep.* (2017) 7:45092. doi: 10.1038/srep45092
181. Ziegler JU, Schweiggert RM, Würschum T, Longin CFH, Carle R. Lipophilic antioxidants in wheat (*Triticum* spp.): a target for breeding new varieties for future functional cereal products. *J Func Foods.* (2016) 20:594–605. doi: 10.1016/j.jff.2015.11.022
182. Ribeiro M, Rodriguez-Quijano M, Nunes FM, Carrillo JM, Branlard G, Igrejas G. New insights into wheat toxicity: breeding did not seem to contribute to a prevalence of potential celiac disease's immunostimulatory epitopes. *Food Chem.* (2016) 213:8–18. doi: 10.1016/j.foodchem.2016.06.043
183. Hucl P, Briggs C, Graf RJ, Chibbar RN. Genetic gains in agronomic and selected end-use quality traits over a century of plant breeding of canada western red spring wheat. *Cereal Chem.* (2015) 92:537–43. doi: 10.1094/CCHEM-02-15-0029-R
184. Kasarda DD. Can an increase in celiac disease be attributed to an increase in the gluten content of wheat as a consequence of wheat breeding? *J Agric Food Chem.* (2013) 61:1155–9. doi: 10.1021/jf305122s
185. Malalgoda M, Meinhardt SW, Simsek S. Detection and quantitation of immunogenic epitopes related to celiac disease in historical and modern hard red spring wheat cultivars. *Food Chem.* (2018) 264:101–7. doi: 10.1016/j.foodchem.2018.04.131
186. Shewry PR, Pellny TK, Lovegrove A. Is modern wheat bad for health? *Nat Plants.* (2016) 2:16097. doi: 10.1038/nplants.2016.97
187. Gulati P, Brahma S, Graybosch RA, Chen Y, Rose DJ. *In vitro* digestibility of proteins from historical and modern wheat cultivars. *J Sci Food Agric.* (2020) 100:2579–84. doi: 10.1002/jsfa.10283
188. Laidig F, Piepho H-P, Rentel D, Drobek T, Meyer U, Huesken A. Breeding progress, environmental variation and correlation of winter wheat yield and quality traits in German official variety trials and on-farm during 1983-2014. *Theor Appl Genet.* (2017) 130:223–45. doi: 10.1007/s00122-016-2810-3
189. Pronin D, Börner A, Weber H, Scherf KA. Wheat (*Triticum aestivum* L.) breeding from 1891 to 2010 contributed to increasing yield glutenin contents but decreasing protein gliadin contents. *J Agric Food Chem.* (2020). doi: 10.1021/acs.jafc.0c02815. [Epub ahead of print].
190. Ozuna CV, Barro F. Characterization of gluten proteins and celiac disease-related immunogenic epitopes in the Triticeae: cereal domestication and breeding contributed to decrease the content of gliadins and gluten. *Mol Breeding.* (2018) 38:22. doi: 10.1007/s11032-018-0779-0
191. Pronin D, Geisslitz S, Börner A, Scherf KA. Fingerprinting of wheat protein profiles for improved distinction between wheat cultivars and species. *Cereal Chem.* (2020). 97:999–1009. doi: 10.1002/cche.10323
192. Ortolan F, Steel CJ. Protein characteristics that affect the quality of vital wheat gluten to be used in baking: a review. *Compr Rev Food Sci Food Safety.* (2017) 16:369–81. doi: 10.1111/1541-4337.12259
193. Kucek LK, Veenstra LD, Amnuaycheewa P, Sorrells ME. A grounded guide to gluten: how modern genotypes and processing impact wheat sensitivity. *Compr Rev Food Sci Food Saf.* (2015) 14:285–302. doi: 10.1111/1541-4337.12129
194. Lionetti E, Catassi C. Co-localization of gluten consumption and HLA-DQ2 and-DQ8 genotypes, a clue to the history of celiac disease. *Dig Liv Dis.* (2014) 46:1057–63. doi: 10.1016/j.dld.2014.08.002
195. Biesiekierski JR, Peters SL, Newnham ED, Rosella O, Muir JG, Gibson PR. No effects of gluten in patients with self-reported non-celiac gluten sensitivity after dietary reduction of fermentable, poorly absorbed, short-chain carbohydrates. *Gastroenterology.* (2013) 145:320–8. doi: 10.1053/j.gastro.2013.04.051
196. Carroccio A, Mansueto P, Iacono G, Soresi M, D'Alcamo A, Cavataio F, et al. Non-celiac wheat sensitivity diagnosed by double-blind placebo-controlled challenge: exploring a new clinical entity. *Am J Gastroenterol.* (2012) 107:1898–906. doi: 10.1038/ajg.2012.236
197. Zanini B, Baschè R, Ferraresi A, Ricci C, Lanzarotto F, Marullo M, et al. Randomised clinical study: gluten challenge induces symptom recurrence in only a minority of patients who meet clinical criteria for non-coeliac gluten sensitivity. *Aliment Pharmacol Therapeut.* (2015) 42:968–76. doi: 10.1111/apt.13372
198. Skodje GI, Sarna VK, Minelle IH, Rolfsen KL, Muir JG, Gibson PR, et al. Fructan, rather than gluten, induces symptoms in patients with self-reported non-celiac gluten sensitivity. *Gastroenterology.* (2018) 154:529–39. doi: 10.1053/j.gastro.2017.10.040
199. Dieterich W, Zopf Y. Gluten and FODMAPS - Sense of a restriction/When is restriction necessary? *Nutrients.* (2019) 11:1957. doi: 10.3390/nu11081957
200. Dieterich W, Schuppan D, Schink M, Schwappacher R, Wirtz S, Agaimy A, et al. Influence of low FODMAP and gluten-free diets on disease activity and intestinal microbiota in patients with non-celiac gluten sensitivity. *Clin Nutr.* (2019) 38:697–707. doi: 10.1016/j.clnu.2018.03.017
201. Muir JG, Varney JE, Ajamian M, Gibson PR. Gluten-free and low-FODMAP sourdoughs for patients with coeliac disease and irritable bowel syndrome: a clinical perspective. *Int J Food Microbiol.* (2019) 290:237–46. doi: 10.1016/j.ijfoodmicro.2018.10.016
202. Fry L, Madden AM, Fallaize R. An investigation into the nutritional composition and cost of gluten-free versus regular food products in the UK. An investigation into the nutritional composition and cost of gluten-free versus regular food products in the UK. *J Hum Nutr Diet.* (2018) 31:108–20. doi: 10.1111/jhn.12502
203. Jadresin O, Misak Z, Sanja K, Sonicki Z, Zizic V. Compliance with gluten-free diet in children with coeliac disease. *J Pediatr Gastroenterol Nutr.* (2008) 47:344. doi: 10.1097/MPG.0b013e31816f856b

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership