Food Chemistry 234 (2017) 389–397

Contents lists available at ScienceDirect

Food Chemistry

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

Food for thought: Selecting the right enzyme for the digestion of gluten

Michelle L. Colgrave [a,*], Keren Byrne [a], Crispin A. Howitt [b]

[a] CSIRO Agriculture and Food, 306 Carmody Rd, St Lucia, QLD 4067, Australia
[b] CSIRO Agriculture and Food, GPO Box 1700, Canberra, ACT 2601, Australia

ARTICLE INFO

Article history:
Received 8 February 2017
Received in revised form 29 April 2017
Accepted 1 May 2017
Available online 4 May 2017

Keywords:
Gluten
Digestion
Chymotrypsin
Mass spectrometry
Quantification

ABSTRACT

Gluten describes a complex mixture of proteins found in wheat, rye, barley and oats that pose a health risk to people affected by conditions such as coeliac disease and non-coeliac gluten sensitivity. Complete digestion of gluten proteins is of critical importance during quantitative analysis. To this end, chymotrypsin was investigated for its ability to efficiently and reproducibly digest specific classes of gluten in barley. Using proteomics a chymotryptic peptide marker panel was elucidated and subjected to relative quantification using LC-MRM-MS. Thorough investigation of peptide markers revealed robust and reproducible quantification with CVs <15% was possible, however a greater proportion of non-specific cleavage variants were observed relative to trypsin. The selected peptide markers were assessed to ensure their efficient liberation from their parent proteins. While trypsin remains the preferred enzyme for quantification of the avenin-like A proteins, the B-, D- and γ-hordeins, chymotrypsin was the enzyme of choice for the C-hordeins.

1. Introduction

Gluten is the collective term that describes the complex mixture of proteins present in wheat, rye, barley and oats that are rich in proline (Pro) and glutamine (Gln) and are as such called the prolamins. The gluten proteins are responsible for imparting glue-like properties to their food products, for example, the viscoelasticity of dough, noodles and pasta. The high proline content of gluten also renders these proteins resistant to gastrointestinal digestion. Proline is the only amino acid whose side-chain links to the backbone α-amino group thereby hindering protein cleavage by most proteases. As a result relatively long proline and glutamine-rich gluten fragments can reach the small intestine where they elicit an autoimmune response in susceptible individuals.

The ingestion and limited proteolytic processing of gluten proteins in the gastro-intestinal tract are involved in the onset of coeliac disease (CD) (Stamnaes & Sollid, 2015), a condition that affects an estimated 70 million people globally (Fasano et al., 2003). It is of critical importance to the health of those affected by CD or non-coeliac gluten sensitivity (NCGS) (Catassi et al., 2013) that the food industry establish accurate methods for gluten measurement. Enzyme-linked immunosorbent assays (ELISA) remain the method of choice and in recent efforts, harmonised guidelines for gluten measurement have been created (Koerner et al., 2013). Greater adoption of mass spectrometry (MS) for gluten measurement has been observed in recent years owing to its specificity, sensitivity, ability to multiplex and identify hydrolysed gluten (Colgrave, Byrne, Blundell, Heidelberger, et al., 2016; Colgrave, Byrne, Blundell, Howitt, 2016; Colgrave et al., 2015; Fiedler, McGrath, Callahan, & Ross, 2014; Gomaa & Boye, 2015; Sealey-Voyksner, Khosla, Voyksner, & Jorgenson, 2010). The successful application of bottom-up proteomics to the analysis of gluten critically depends upon the efficiency and reproducibility of proteolytic digestion of gluten from the grain or highly processed food product.

Trypsin represents the gold standard for proteolytic digestion in proteomics primarily owing to its efficiency and specificity (Tsiatsiani & Heck, 2015), but also due to its low cost. Trypsin cleaves the peptide backbone at lysine or arginine yielding short peptides comprising a C-terminal basic residue, rendering the peptide favourable chromatographic and mass spectrometric (ionisation, fragmentation) properties. However, one of trypsin's benefits is also a detriment. As an example, approximately 50% of yeast tryptic peptides are ≤6 residues in length (Swaney, Wenger, & Coon, 2010), raising challenges in the confident identification of these small fragments using automated database searching algorithms. In contrast, gluten proteins have an inherent lack of trypsin cleavage sites (Ferranti, Mamone, Picariello, & Addeo, 2007). Moreover, the presence of adjacent prolines interferes with the digestion, liberating gluten peptide products that are fewer in number and larger in size. This is particularly

* Corresponding author:
E-mail address: michelle.colgrave@csiro.au (M.L. Colgrave).

http://dx.doi.org/10.1016/j.foodchem.2017.05.008
0308-8146/

problematic with the specific gluten classes of the ω-gliadins (wheat) and C-hordeins (barley). As a result the protein sequence coverage may be compromised. Alternative enzymes such as chymotrypsin, thermolysin and pepsin have been employed to increase the sequence coverage and allow more accurate identification of the closely-related isoforms (Dupont, Vensel, Tanaka, Hurkman, & Altenbach, 2011; Manfredi, Mattarozzi, Giannetto, & Careri, 2015; Prandi et al., 2012; Rombouts, Lagrain, Brunnbauer, Delcour, et al., 2013; Salplachta, Marchetti, Chmelik, & Allmaier, 2005; Vensel, Dupont, Sloane, & Altenbach, 2011; Vensel, Tanaka, & Altenbach, 2014). Chymotrypsin preferentially cleaves at the large hydrophobic residues such as Phe, Trp and Tyr, but also with lower affinity at Leu and Met. From a qualitative proteomics perspective, chymotrypsin yields complementary peptides to that obtained using trypsin, enabling greater sequence coverage.

The analysis of gluten in all of the gluten-containing cereals using bottom-up proteomics presents numerous challenges, with wheat posing a unique subset of challenges due to its hexaploid nature yielding a large and complex genome with at least three copies of every gene. Natural variation between cultivars of gluten-containing cereals further confounds isoform identification. Moreover, the gluten proteins are encoded by large gene families and contain repetitive sequences that make it difficult to distinguish individual proteins. The low proportion of Lys (~1%) and Arg (~2.4%) in hordeins (barley gluten) compared to non-gluten proteins (~5% Lys, ~5% Arg) limits the usefulness of trypsin in qualitative proteomics, but despite this trypsin has been demonstrated to yield useful peptides for hordein quantification (Colgrave, Goswami, Howitt, & Tanner, 2012; Colgrave, Byrne, Blundell, Heidelberger, et al., 2016; Colgrave, Byrne, Blundell, et al, 2016). While many researchers employ chymotrypsin (Fiedler et al., 2014; Manfredi et al., 2015; Rombouts et al., 2013) or chymotrypsin-containing cocktails (Sealey-Voyksner et al., 2010) in quantitative proteomics applications, questions remain as to the digestion efficiency, the reproducible generation of peptides and the optimum conditions for reliable chymotryptic digestion. Many peptides selected for quantitative applications contain missed cleavages in particular at the low affinity sites (Leu and Met). To this end, in this study we have assessed the use of chymotrypsin in the analysis of barley gluten (hordeins) and compare and contrast the results to that obtained for trypsin.

2. Materials and methods

2.1. Plant material

Grains of barley cv Flagship, Hindmarsh, Sloop, Oxford, Baudin, Yagan, Bomi, Fleet, Commander, Gairdner, Scope and Maritime were obtained from the Australian Winter Cereals Collection (Tamworth, Australia). The creation of the ultra-low gluten barley lines used in this study were previously described (Colgrave, Byrne, Blundell, Heidelberger, et al., 2016; Colgrave, Byrne, Blundell, et al, 2016). All grains were milled using a Metefem Hungarian Mill (model FQD2000, Hungary). Fine flour was obtained by sieving the wholemeal with a 300 μm sieve (Endecotts Pty Ltd Sieves, London, England).

2.2. Gluten extraction and protein digestion

The samples (milled flour) for proteomic analysis were prepared as previously described (Colgrave, Byrne, Blundell, Heidelberger, et al., 2016; Colgrave, Byrne, Blundell, et al, 2016) with minor modifications. Briefly, flour was extracted using 55% IPA/2% DTT solution (200 μL, 10:1 w/v) for 30 min at 50 °C. The supernatant (100 μL) after centrifugation (20,800*g*, 15 min) was applied to a 10 kDa molecular weight cut-off filter and washed twice with a buffer consisting of 8 M urea in 0.1 M Tris-HCl (pH 8.5). The proteins were alkylated by incubation with 50 mM iodoacetamide for 20 min at RT in the dark. Buffer exchange using 100 mM ammonium bicarbonate (pH 8.5) was undertaken by two consecutive wash/centrifugation steps (20,800*g*, 10 min) before application of the digestion enzyme (trypsin or chymotrypsin, 20:1 protein:enzyme w/w) in 100 ammonium bicarbonate, 1 mM $CaCl_2$ (pH 8.5) and 16 h incubation at 37 °C. For pepsin, buffer exchange was accomplished with acidified ammonium bicarbonate (pH 1.2) prior to addition of pepsin. Filtrates containing the proteolytically digested peptides were collected by centrifugation (20,800*g*, 10 min) and the filters were washed with 100 mM ammonium bicarbonate.

A collection of 60 commercial beers was also subjected to chymotryptic digestion as described previously (Colgrave et al., 2012) with minor modifications. Briefly, aliquots (100 μL) of degassed beer were taken and diluted 1:1 with 100 mM ammonium bicarbonate (pH 8.5). The beers were reduced by addition of 20 μL of 50 mM DTT under N_2 for 30 min at 60 °C. To these solutions, 20 μL of 100 mM iodoacetamide (IAM) was added and the samples were incubated for 20 min at RT. To each solution 5 μL of 1 mg/mL chymotrypsin was added and the samples incubated at 37 °C for 16 h. The digested peptide solution was acidified by addition of 10 μL of 5% formic acid, then lyophilised.

The lyophilized samples were reconstituted in 100 μL of 1% formic acid and stored at 4 °C until analysis.

2.3. Global proteomic profiling

Gluten-enriched fractions (5 μL) were analysed precisely as described previously (Colgrave, Goswami, Blundell, Howitt, & Tanner, 2014) with chromatographic separation using a nano HPLC system (Shimadzu Scientific, Rydalmere, Australia) directly coupled to a 5600 TripleTOF MS (SCIEX, Foster City, USA). ProteinPilot™ 4.0 software (SCIEX) with the Paragon Algorithm (Shilov et al., 2007) was used for protein identification. Tandem mass spectrometry data was searched against *in silico* proteolytic digests of Poaceae proteins of the Uniprot database (version 2017/02; 2,891,190 sequences) appended with a custom database of gluten sequences as described previously (Colgrave et al., 2012). The search parameters were defined as iodoacetamide modified for cysteine alkylation and trypsin, chymotrypsin or pepsin as the digestion enzyme. When the data generated from all enzymatic digests were combined in a single database search, no enzyme specificity was employed. ProteinPilot generates a score for each protein based on the confidence, wherein a confidence of 99% is assigned a score of 2.00 and a confidence of 95% is assigned a score of 1.30. The database search results from the combined barley cultivar analyses were manually curated to yield the protein identifications (Supp. Table 1) using a 1% global false discovery rate (FDR) determined by the in-built FDR tool within ProteinPilot software (Tang, Shilov, & Seymour, 2008).

2.4. Identification of peptides for monitoring extraction efficiency

Peptide summaries generated by ProteinPilot were used to select peptides that yielded intense peaks and were fully chymotryptic, *i.e.* no variable or missed cleavages. 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. Using chymotrypsin (Supplementary Table 2), 30 peptides (spanning five gluten families) were assessed. Only peptides with ≥95% confidence were used for sequence coverage analysis and targeted peptide quantification.