Patron: Her Majesty The Queen

ROTHAMSTED RESEARCH

Rothamsted Research
Harpenden, Herts, AL5 2JQ

Telephone: +44 (0)1582 763133
Web: http://www.rothamsted.ac.uk/

Rothamsted Repository Download

A - Papers appearing in refereed journals

Lafiandra, D. and Shewry, P. R. 2022. Wheat glutenin polymers 1. structure, assembly and properties. *Journal of Cereal Science.* 106, p. 103486. https://doi.org/10.1016/j.jcs.2022.103486

The publisher's version can be accessed at:

- https://doi.org/10.1016/j.jcs.2022.103486

The output can be accessed at: https://repository.rothamsted.ac.uk/item/98897/wheat-glutenin-polymers-1-structure-assembly-and-properties.

© 28 April 2022, Please contact library@rothamsted.ac.uk for copyright queries.

05/05/2022 16:16 repository.rothamsted.ac.uk library@rothamsted.ac.uk

Rothamsted Research is a Company Limited by Guarantee
Registered Office: as above. Registered in England No. 2393175.
Registered Charity No. 802038. VAT No. 197 4201 51.
Founded in 1843 by John Bennet Lawes.

Review

Wheat Glutenin Polymers 2, The role of wheat glutenin subunits in polymer formation and dough quality

Domenico Lafiandra[a] and Peter R. Shewry[b]

Dedication: This paper is dedicated to our mentor, friend and colleague Dr. Donald D. Kasarda (1933-2021).

[a]Department of Agriculture and Forest Sciences (DAFNE), University of Tuscia, Viterbo 01100, Italy

[b]Rothamsted Research, West Common, Herts AL5 2JQ, UK

Abstract

The high- and low-molecular-weight glutenin subunits are major contributors to dough strength and elasticity, through their influence on glutenin polymer structure. The study of their genetics, diversity, and role in quality characteristics has been facilitated by the development of genetic materials and analytical tools. Induced and natural mutants have been used to confirm the importance of the number and location of cysteine residues in modulating the size distribution of the glutenin polymers with the presence of an additional cysteine in the central repetitive domain of the HMW-GS having a negative role on polymer formation. The application of genomics and proteomics the integration of datasets from these technologies have facilitated the identification, cloning and functional analysis of genes encoding novel HMW-

GS and LMW-GS and have also allowed the identification of post-translational modifications (PTMs) and elucidated their roles in determining dough quality.

1. Introduction

It is generally accepted that the gluten proteins of wheat are the major determinants of the unique biophysical properties of doughs that allow them to be used for making bread, other baked goods, pasta and noodles. These proteins form a continuous network when wheat flour is kneaded with water to make dough, conferring cohesion, elasticity and extensibility. Gluten can be isolated as a cohesive mass by washing dough in water to remove soluble components, starch granules and particulate matter. It comprises mainly proteins with the major components being classified into two types called gliadins and glutenins. The gliadins and glutenins each consist of multiple individual proteins (usually stated as 50-100 in total) which have related amino acid sequences and properties. However, the gliadins and glutenins differ in one important respect: the gliadins are monomeric whereas the glutenins form polymers stabilised by inter-chain disulphide bonds. In addition, the gliadin monomers and glutenin polymers interact by non-covalent forces particularly hydrogen bonds. The difference in subunit interactions result in different biophysical properties, with the glutenin polymers conferring elasticity to dough and the gliadins viscosity and extensibility. Breadmaking requires highly elastic (often called strong) doughs and hence there has been extensive research on the glutenin polymers as a basis for improving breadmaking quality.

Glutenin polymers range in mass from oligomers with masses below 100,000 which are soluble in aqueous alcohols (60-70% aq. ethanol) to large polymers with masses of up to 20 million (or more) which are insoluble in their native state. Hence, most studies have been carried out after reducing the polymer size, after shearing by sonication or reducing the interchain disulphide bonds to release the individual subunits.

Glutenin polymers are formed by high molecular weight- (HMW-GS) and low molecular weight- glutenin subunits (LMW-GS), which are cross-linked by intermolecular disulphide bonds. The HMW-GS and LMW-GS make up about 30 and 60% of the glutenin polymer, respectively, whereas the remaining 10% is contributed by chain terminator proteins and thiols (Wieser et al., 2006).

The existence of strong positive correlation between the amounts of large sized polymers, as measured by the relative amount of unextractable polymeric proteins (%UPP) or Glutenin Macropolymer (GMP) (Gupta et al., 1995; Don et al., 2006), and dough processing properties has been investigated by many authors. This has demonstrated that the HMW-GS play a major role in determining dough strength, accounting for between 45-70 % of the variation in bread making characteristics (Payne et al., 1988), despite representing only about 10-12% of the total grain proteins (corresponding to about 1-1.7% of the flour dry weight).

2. High Molecular Weight Glutenin Subunits (HMW-GS)

2.1. Genetics

The HMW-GS are encoded by three orthologous pairs of genes, located at the *Glu-A1*, *Glu-B1* and *Glu-D1* loci present on the long arms of the

homoeologous (homologous but non-pairing) group 1 chromosomes of the three genomes (A, B and D) of hexaploid bread wheat. Each locus contains two tightly linked paralogous genes encoding different types of HMW-GS. These are designated as x- and y-type and have higher and lower molecular weights, respectively. It has been suggested that a tandem duplication of an ancestral region of ~7.2-kb, containing genes encoding a globulin protein and a HMW-GS, resulted in the presence of the two paralogous x- and y-type genes. The x-type HMW-GS genes arose from duplication and divergence of the y-type HMW-GS genes, after the divergence of the ancestors of wheat and barley (Kong et al., 2004). Recombination between the x- and y-type genes is very rare, Payne et al., (1984) reported of a recombination at the *Glu-B1* locus, at a rate of about one recombinant in 1000 progeny. Hence, all x-type and y-type subunits encoded by *Glu-D1*, and some subunits encoded by *Glu-B1*, are inherited as “allelic pairs”.

2.2. Polymorphism

Genotypes of bread wheat show extensive polymorphism in the number of expressed HMW-GS (between three and five) and in their mobilities when separated by SDS-PAGE (Figure 1). In particular, whereas the x-type gene at the *Glu-A1* locus is expressed in some genotypes the y-type *Glu-A1* gene is almost always silent in cultivated tetraploid and hexaploid wheats, as result of a premature termination codon (PTC) or WIS 2-1A retrotransposon insertion (Gu et al., 2004). However, a single hexaploid wheat line from Sweden expressing both x- and y-type genes at the *Glu-A1* locus was reported by Margiotta et al., (1996). Either one or two HMW-GS encoded by the *Glu-B1*

locus and two subunits encoded by the *Glu-D1* locus are present in all modern bread wheat genotypes. However, lines expressing only one (x- or y-type) or no *Glu-D1* subunits have been identified in the landraces grown historically (Payne et al., 1984).

In contrast to cultivated polyploid wheats, the y-type genes present at the *Glu-A1* locus are frequently expressed in related species and subspecies with the A genome: in the cultivated *(T. monococcum* ssp *monococcum*) and wild (*T. monococcum* ssp*. boeoticum,* and *T. urartu*) diploid wheat, in wild tetraploid wheat *T. turgidum* ssp. *dicoccoides* and in wild and cultivated tetraploid *T. araraticum* and *T. timopheevii* (AAGG) (Figure. 2), (Jang et al., 2009; Shewry et al.,2006).

A large number of allelic variants have been identified at the *Glu-1* loci of durum and bread wheat, initially using SDS-PAGE (Payne et al., 1984) and subsequently using 2D electrophoresis, reversed-phase high performance liquid chromatography (RP-HPLC), capillary electrophoresis and mass spectrometry (Margiotta et al., 1993; Uthayakumaran et al., 2005; Jang et al., 2021).

The availability of extensive libraries of genomic sequences (Bromilow et al., 2017), including sets of gluten genes from the reference genotype Chinese Spring (Altenbach et al., 2020), provides information on the molecular basis for this polymorphism and the mechanisms responsible for associations with processing properties. Genomic and functional genomics studies have also provided understanding of the structure of gluten chromosomal loci and genes and the mechanisms regulating their protein expression, allowing changes at

whole genome level in response to environmental and changing climatic conditions to be monitored (Wang et al., 2020).

2.3. Sequences and structures

The complete sequences of numerous HMW-GS genes present in cultivated wheat and of related wild relatives in the *Triticeae* tribe have been reported (Bromilow et al., 2017). These show that the subunits have a highly conserved structure comprising three domains. These are a short N- terminal domain, ranging from about 70 to 100 amino acid residues, and a short C-terminal domain, comprising 42 amino acid residues, separated by a central domain which is much more variable in length (between 300 and 1100 amino acid residues (Shewry et al., 2009). The N- and C-terminal domains are non-repetitive and contain most or all of the cysteine residues whereas the central domain comprises tandem and interspersed repeats of short peptide motifs: hexa-, nona-, and tripeptide (consensus PGQGQQ, GYYPTSPQQ, GQQ) in the x-type subunits and by hexa- and nona-peptides (consensus PGQGQQ and GYYPTSLQQ) in the y-type subunits. The variation in length of the domain arises from deletions and insertions within the central repetitive domain, which may result from replication slippage, transposition, gene conversion and unequal crossing over (Shewry et al., 2009). The variation in the length of the repetitive domain is largely responsible for the variation in molecular masses of the whole HMW-GS (Figure 1), from 45 to 106 kDa (Tahir et al., 1996; Shewry et al., 2009). Variation in size of the y-type subunit is also evident moving from the wild diploid species (AA) to tetraploid wheats (AABB and AAGG), as shown in Figure 2.

Whereas the N- and C-terminal domains may have globular structures, the repeated sequences in the central domain may form a loose spiral based on β-turns which is stabilised by hydrogen bonds formed between glutamine residues and may be intrinsically elastic (Belton, 1999).

The number of cysteine residues differs between x- and y-type subunits, which usually contain four (three in the N-terminal domain and one in the C-terminal domain) and seven (five in the N-terminal domain, one at the end of the repetitive domain and one in the C-terminal domain) cysteine residues, respectively (Table 1). However, subunits with different number of cysteine residues have been reported (Table 1). In particular, subunit 1Dx5, which is associated with superior dough properties, has an extra cysteine residue within the repetitive domain, close to the N-terminal region. This extra cysteine residue appears to increase the formation of large glutenin polymers and has been shown to be linked by an intermolecular bond with the cysteine present in the C-terminal region of all x-type HMW-GS (Lutz et al., 2012). Subunits 1Bx20, 1Bx26 and 1Bx14 have similar sequences and all lack the second and third cysteine residues which are usually present in the N-terminal domain (Shewry et al.; 2003; Margiotta et al., 2000; Li et al., 2004).

Most of 1Ay-type subunits have only six conserved cysteine residues, with the residue that is normally present at the end of the repetitive domain being missing (Jang et al., 2009; Yu et al., 2019), though subunits with the complete set of seven cysteine residues have also been identified in 1Ay-type subunits in wild diploid relatives of the A genome (Jang et al., 2009).

3. Relationship between HMW-GS type and number and the size distribution of glutenin polymers

Many studies have indicated that number of HMW-GS influences the breadmaking characteristics of wheat. The identification of natural *null* mutants at the *Glu-1* loci and the generation of mutations using radiation or chemical mutagens, have made it possible to produce lines varying in their number of subunits (Fig. 3) and better define their functional role (Payne et al., 1984).

3.1. Comparison of genotypes

Halford et al., (1992) compared 22 cultivars of bread wheat showing that the presence of either subunit 1Ax1 or 1Ax2*, when compared with a *null* allele, resulted in an increase in the proportion of HMW-GS from about 8 to 10% of the total grain protein with a consequent increase in the amount of high M_r glutenin polymers. Differences in the proportions of individual HMW-GS were reported by Wieser and Zimmermann (2000) who analysed a set of 29 different bread wheat varieties grown in different years and locations. In general, HMW-GS 1Dx2, 1Dx5, 1Bx7, 1Dy10 and 1Dy12 were present at higher concentrations while HMW-GS 1Ax1, 1Ax2*, 1By6, 1By8 and 1By9 were minor components. For example, the proportion of subunit 1Bx7 was more than twice that of subunit 1By6.

Correlation analyses showed that dough development time, maximum resistance of dough, amount of gluten and bread volume showed significant positive correlations with the amounts of HMW-GS subunits with x-type having stronger effects than the y-type subunits. Furthermore, subunits 1Dx5 and

1Bx7 showed stronger effects on dough characteristics than other x-type subunits, probably as a consequence of the extra cysteine present in 1Dx5 and the greater amount of 1Bx7, demonstrating both qualitative and quantitative effects of HMW-GS on breadmaking.

Wieser and Kieffer (2001) compared the relative effects of LMW-GS and HMW-GS on breadmaking quality by determining the amounts of different gluten protein type in a set of 14 bread wheat varieties. They showed that both HMW-GS and LMW-GS were correlated with quality, as reported by Gupta et al (1995), but regression analysis showed that twice the amount of LMW GS was required to have the same effect on dough strength as with HMW-GS. In addition, they showed that x-type HMW-GS had greater effects on rheological properties than the y-type HMW-GS.

3.2. Comparison of mutant series

Lawrence et al., (1988) produced a set of lines which varied in number of HMW-GS from zero to five by crossing a mutant line of the cultivar Olympic, which was *null* at the *Glu-B1* locus, and an isogenic line of the cultivar Gabo, which was *null* at the *Glu-A1* and *Glu-D1* loci. Analysis of the mixing properties by Mixograph showed a strong relationship between the number of expressed HMW-GS and dough strength. Similarly, Gupta et al. (1995) compared two sets of genetic lines with varying numbers (from three to zero) of HMW-GS or LMW-GS, showing that the absence of all HMW-GS (*Glu*-1) or LMW-GS (*Glu*-3) had significant effects on the amounts of total, extractable and unextractable polymeric proteins (UPP) and on the dough and gluten properties. In particular, the dough and gluten properties were significantly

correlated with the proportions of both total polymers and unextractable polymers (a measure of the relative molecular size distribution of polymeric proteins), although more strongly with the proportions of unextractable polymers in the case of *Glu-1 null* lines. The proportion of total polymeric proteins decreased more markedly when all the *Glu-3* subunits were deleted than when all the *Glu-1* subunits were absent, in accordance with the relative quantities of these two types of subunits in the grains. By contrast, loss of all the *Glu-1* subunits, on an equal weight basis, reduced the amounts of the larger polymers to a much greater extent than the loss of all the *Glu-3* subunits. Yang et al (2014) used ion beam radiation to generate deletion lines of HMW-GS in the bread wheat Xiaoyan 81 which has the subunit composition 1Ax1, 1Bx14+1By15 and 1Dx2+1Dy12 at the *Glu-A1, Glu-B1* and *Glu-D1* loci. This gave single mutant lines at the three *Glu-1* loci which were crossed to give double mutants. Wang et al. (2017) used this material together with two mutant lines lacking the 1Dx2 or 1Dy12 subunits (obtained by ethyl methane sulfonate (EMS) treatment of the wheat variety Xiaoyan 54) to study the effects of the absence of the different HMW-GS on the quality characteristics of the doughs. Comparisons of the single and double deletion mutants confirmed that the subunit pair 1Dx2+1Dy12 had the greatest effect on functional properties of gluten and dough and on breadmaking quality, confirming the report of Rogers et al. (1991) who showed that the 1Dx2+1Dy12 allelic pair had a stronger positive effect on baking quality compared to the allelic pair 1Bx7+1By8. In addition, double deletion mutants lacking *Glu-D1* in combination with *Glu-A1* or *Glu-B1* showed strong decreases in dough functional characteristics. Knockout mutants of 1Dx2 or

1Dy12, showed that 1Dx2 subunit was more effective than 1Dy12 in promoting larger amount of glutenin macropolymers.

3.3. Effects of subunit over-expression

Several studies have shown that subunit 1Bx7 is overexpressed in bread wheat varieties with the *Glu-B1al* allele, due to a functional gene duplication (Lukow et al., 1992,). Ragupathy et al (2008) reported that the overexpressed subunit (called $1Bx7^{OE}$) probably resulted from a gene duplication mediated by the insertion of a retroelement and showed that it was present in 40 out of 316 land races and varieties of bread wheat originating from different countries. More recently, Geng et al (2014) used fusions comprising promoter sequences of x-type genes at the *Glu-B1* locus with the β-glucuronidase (GUS) reporter to show that the $1Bx7^{OE}$ promoter present in cv. Yumai 33 was markedly stronger than other promoters. Geng et al., (2014) therefore suggested that the relationship between the strong activity of the $1Bx7^{OE}$ promoter and high level of protein produced demonstrates the importance of transciptional regulation in determining the overexpression of the 1Bx7 subunit. The presence of $1Bx7^{OE}$ has been shown to be associated with very strong dough characteristics (Butow et al., 2003).

Juhász et al., (2003) determined the gluten protein composition of different lines isolated from the old landrace Bankuti 1201, showing wide allelic variation in gliadins and glutenin subunits, with the exception of the *Glu-D1* locus with 1Dx2+1Dy12 being present in 95% of the population. Despite the predominance of subunits 1Dx2+1Dy12 which are associated with lower quality than the allelic pair 1Dx5+1Dy10, the lines contained large amounts of

UPP (average of 48.40%). Unusually high contents of UPP were present in three lines (58.18, 58.89 and 60.07%), higher than that determined for the strong Canadian wheat variety Glenlea (57.8%) was associated with the presence of the *Glu-B1al* allele.

3.4. Introduction of a 1Ay subunit into bread wheat

The introduction of an active 1Ay gene into bread wheat results in improvement of dough characteristics. Two alleles with active x- and y-type genes at the *Glu-A1* locus were identified in accessions of *T. boeoticum.* These alleles (designated *Glu-A1r and Glu-A1s*) and the encoded HMW-GS pairs (1Ax39+1Ay40 and 1Dx41+1Dy42) were introgressed into bread wheat (*Triticum aestivum* L.) cv. Sicco (Rogers et al., 1997). The introduction of *Glu-A1r* resulted in reduced dough stickiness and improved stability during mixing compared with the *Glu-A1a* allele encoding subunit 1Ax1, with a small improvement in gluten strength being shown by the SDS-sedimentation test.

The importance of the number of subunits in determining dough characteristics was indicated by Roy et al., (2018) who introgressed x- and y-type subunits at the *Glu-A1* locus (designated 1Ax21* and 1Ay21y*), identified in a Swedish bread wheat line (Margiotta et al., 1996), into the Australian cultivars Livingston and Bonnie Rock. These both have subunit 1Ax2* at the *Glu-A1* locus and subunits 1Bx17+1By 18 at the *Glu-B1* locus but differ at *Glu-D1* with Livingstone having 1Dx5+1Dy10 and Bonnie Rock 1Dx2+1Dy12. Several lines derived from each variety had subunits 1Ax21* and 1Ay21y* instead of 1Ax2* subunit, resulting in an increase in subunit number from five to six. It was assumed that the differences observed between the lines and

the two parent varieties were mainly associated with the additional subunit 1Ay21y*, as subunit 1Ax21* replaced subunit 1Ax2*. Furthermore, these subunits were assumed to have similar effects on dough characteristics since analysis by RP-HPLC indicated that they both contain four cysteine residues (Roy et al., 2018).

Analysis of three sister lines derived from the two crosses showed that the presence of the active 1Ay subunit was associated with increases in the %UPP and the glutenin/gliadin ratio. The expressed 1Ay subunit also resulted in an up to 10% increase of gluten content and a 5% increase in glutenin content, resulting in an increase in the HMW-GS to LMW-GS ratio without affecting the relative amounts of other subunits.

Recently, Cao et al., (2021) determined the differences between the grain proteomes of three bread wheat cultivars and corresponding near-isogenic lines carrying the 1Ax21*+1Ay21* HMW-GS. In addition to increased abundance of the 1Ay HMW-GS, 115 differentially expressed proteins were also identified showing that introgression of the two HMW-GS at the *Glu-A1* locus increase wheat grain protein content and breadmaking quality in association with a wider reshaping of the grain proteome network.

Dumur et al., (2010) adopted a different approach to increase the number of HMW-GS, by manipulating the mechanism controlling chromosome pairing in wheat. Chromosome pairing in wheat is under control of a major gene present at the *Ph1* locus on the long arm of chromosome 5B (Sears, 1976). This gene restricts pairing to homologous chromosomes within the polyploid wheat genome. In its absence, homoeologous pairing and crossing over can occur between chromosomes of the different wheat genomes and with those present

in more distant related species thus providing an opportunity to transfer genetic variation. This methodology, which allows segments of chromosomes carrying genes of interest to be transferred between genomes and species, was given the name of chromosome engineering by Sears (1976).

Chromosome engineering has been used to transfer segments of the D genome containing the genes encoding the pair of HMW-GS 1Dx5+1Dy10 or 1Dx2+1Dy12 from bread wheat to chromosome 1R of triticale and then to different durum wheat varieties, replacing the *null* allele at the *Glu-A1* locus (Lukaszevsky, 2003). Dumur et al. (2010) used the same methodology to transfer a segment of wheat chromosome 1D containing the genes encoding subunits 1Dx2+1Dy12 into the bread wheat variety Courtout (HMW-GS composition 1Ax2*, 1Bx7+1By8, 1Dx2+1Dy12) where it was translocated to replace the long arm of chromosome 1A, thus replacing the *Glu-A1* locus encoding HMW-GS 1Ax2* with a duplicated *Glu-D1* locus. This duplication of *Glu-D1* had significant effects on dough strength, mixing resistance and the Zeleny sedimentation volume. However, no significant effects were observed on breadmaking and loaf volume. The authors suggested that the French bread-making test which was used may have been less sensitive to small differences in breadmaking quality than other types of test baking. Alternatively, that changes in the ratio of x-type to y-type HMW-GS, which was lower in the modified lines, may also have affected quality. Additional work is therefore required with different varieties and quality testing methods to provide conclusive evidence on the effect of the duplicated 1Dx2+1Dy12 allele on bread making quality.

4. Relationships between HMW subunit structure, disulphide bond formation and functional properties of bread wheat

Cysteine residues account for about 2.5g per 100g of proteins (Wieser et al., 2006) and largely contribute to the structure and functional properties of the gluten complex through the formation of intra- and inter-molecular disulphide bonds (Shewry and Tatham 1997; Wieser et al., 2006). Considerable efforts have been directed to determine role of number, position and environment of cysteine residues and to identify the pattern of formation of disulphide bonds. The incorporation *in vitro* of synthetic peptides and model proteins related to gluten proteins has shown that both the number and position of cysteine residues are important factors in determining gluten structure and functionality as they affect the size-distribution of the glutenin polymer and the rheological properties of the dough (Tamás et al., 2002). A few HMW-GS with variation in number of cysteine residues have been identified, allowing their influences on the rheological characteristics of doughs and size distribution of polymeric glutenins to be determined(Table 1).

Pirozi et al. (2008) studied the relationships between the number of cysteine residues in HMW-GS, the size of glutenin polymers, dough properties and bread making characteristics, using biotypes and lines of the varieties Avocet and Halberd differing in HMW-GS present at the *Glu-1* loci and the Italian bread wheat cultivar Fiorello and a derived line in which the subunit pair 1Bx7+1By8 was replaced by subunits 1Bx26+1By27 present in the older cultivar Cologna. Subunit 1Bx26 was previously demonstrated to possess only two cysteine residues, one in the N-terminal and one in the C-terminal region of the subunit (Margiotta et al. 2000).

As discussed above, subunit 1Bx20 lacks two cysteine residues in the N-terminal domain and is known to negatively affect dough characteristics, probably by affecting polymer formation (Shewry et al., 2003). The same effect therefore is expected for the 1Bx26 which also has two cysteine residues .

Substitution of HMW-GS 1Bx7+1By8 and 1Bx7+1By9, with the allelic pairs 1Bx20+1By20 or 1Bx26+1By27 resulted in a marked decrease in the % UPP and weaker dough properties with the presence of the extra subunit 1Ax1 in the Avocet line C failing to compensate for the negative effect of the 1Bx20. The presence of 1Bx20 and 1Bx26 is also associated with an increase in extensibility, which is also associated with the subunit pair 1Dx2+1Dy12 present in Avocet line A compared with Avocet line B which possess the subunit pair 1Dx5+1Dy10. The breadmaking characteristics of the lines are consistent with %UPP, rheological data and loaf volume (Pirozi et al., 2008).

Gupta and Macritchie (1994) used biotypes and recombinant inbred lines differing in composition of allelic HMW-GS present at the *Glu-B1* (1Dx17+1Dy18 vs 1Dx20x+1Dy20y) and at the *Glu-D1* loci (1Dx5+1Dy10 vs 1Dx2+1Dy12) to show that the 1Bx17+1By18 or 1Dx5+1Dy10 pairs of subunits resulted in a larger amount of large glutenin polymers and consequent higher dough strength than the allelic pairs of subunits 1Dx2+1Dy12 and 1Dx20x+1Dy20y. Similarly, Popineau et al (1994) reported that the HMW-GS 1Dx5+1Dy10, which are normally present in the cultivar Sicco, were associated with higher proportions of unextractable large glutenin than HMW-GS 1Dx2+1Dy12 when compared is near-isogenic lines. No direct comparisons of glutenin polymer structure in the lines used in these studies have been reported but the differences in numbers of cysteine residues in the

399 subunits compared are consistent with the effects resulting from differences
400 in cross-linking of the polymers.
401 As discussed above, a subunit containing an additional cysteine residue
402 encoded by the *Glu-A1* locus was identified in the bread wheat landrace
403 Bankuti 1201 (Juhász et al., 2003) (Table 1). Analysis of processing quality
404 showed that the effect of this allele (designated 1Ax2*^{B}) on polymer size and
405 dough properties was less than expected based on the assumption that the
406 additional cysteine residue would lead to higher cross-linking. This may relate
407 to the position of the additional cysteine, which is in the middle of the repetitive
408 domain compared to the additional cysteine in HMW-GS 1Dx5 which is close
409 to the N-terminal end of the repetitive domain (Juhász et al. 2003). This
410 different location may account for the fact that the extra cysteine residue in
411 HMW-GS 1Dx5 has a significant effect on dough strength compared with the
412 limited effect of HMW-GS 1Ax2*^{B}. This could result from the residues forming
413 different types of disulphide bonds (inter-chain or intra-chain) or having
414 different effects on the properties of the polymers formed.
415 Further evidence that the position of the cysteine residues is important in
416 determining the amount and properties of the glutenin polymers was reported
417 by Gao et al., (2012). They showed that the Australian cultivar Galaxy H45
418 which has the 1Bx7^{OE} subunit (which as discussed above results from the
419 presence of two x-type of gene) contained a lower amount of UPP and poorer
420 rheological characteristics than were expected based on its HMW-GS
421 composition.
422 Gao et al., (2012) determined and aligned the sequences of the two linked
423 1Bx7 genes (which they called *Bx7.1* and *Bx7.2*) present in Galaxy H45, the

breeding line VQ0437 and the cultivar Glenlea, revealing the presence of single nucleotide polymorphisms (SNPs). In particular, one of the SNPs resulted in a change of a tyrosine present in the middle of the repetitive domain in Glenlea into a cysteine in Galaxy H45, resulting in five cysteines compared to the four conserved cysteine residues typical of x-type subunits (Table 1). Comparison of the sequences of the two 1Bx genes of VQ0437 with those of Glenlea revealed two SNPs in the first gene and one SNP in the second one, consequently the alleles of H45 and VQ0437 were designated *Glu-B1br* and *Glu-B1bs*, respectively.

Previous work of Tao et al., (1992) followed by Luz et al., (2012) has led to the identification of the disulphide bonds formed by subunits 1Bx17 and 1Dy10. In particular, the first two cysteine residues present in the N-terminal domain of the 1Bx7 form an intrachain bond, whereas the third cysteine present in the same region and the single cysteine in the C-terminal region, of the 1Dy subunit, form an interchain bond leading to a head to tail cross-link. (Lutz et al., 2012). Gao et al., (2012) therefore hypothesized that the extra cysteine present in HMW-GS 1Bx7.1 could interfere with the formation of this head -to-tail cross-link, by forming an intra-chain bond with the C-terminal cysteine.

More recently, Li et al (2015) used a mutant population produced by treating the bread wheat variety Xiaoyan 54 with EMS. This variety has five HMW-GS, 1Ax1, 1Bx14, 1By15, 1Dx2, and 1Dy12 and SDS-PAGE analysis of the population identified knock out mutants for all the five subunits and three missense mutants for subunits 1Ax1, 1Dx2 and 1Dy12. The 1Ax1 subunit had a single amino acid substitution (G330E) which influenced its mobility, being slower, compared to the corresponding subunit of Xiaoyan 54.

Six rounds of backcrosses with Xiaoyan 54, were performed In order to remove background mutations and the mixing characteristics and rheological properties of doughs were determined for Xiaoyan 54 and six mutants: two mutants without subunits 1Ax1 and 1Bx14, one double mutant lacking both 1Ax1 and 1Bx14, and the 1Ax1 mutant with a single amino acid substitution. Dough development time and dough stability were significantly lower in the five knockout mutants compared to Xiaoyan 54 with the effects being greater for the double mutants, whereas both parameters were both greatly increased in the 1Ax1 missense mutants of the 1Ax1 subunit. Similarly, maximum resistance was reduced in the five knockout mutants with the effect being greatest for the double mutants, intermediate for the two 1Ax mutants and lower for the 1Bx14 mutants. In addition, the two 1Bx14 mutants and the double 1Ax1/1Bx14 mutant showed significantly reduced dough extensibility compared with the control, whereas no change in the extensibility was associated with the three 1Ax mutations, indicating that subunit 1Bx14 was responsible for the increased extensibility.

Subunit 1Bx14 is very similar to Bx20 in sequence but lacks the second and third cysteine residues in the N-terminal region (Li et al., 2004) (Table1). It is therefore likely that the effects of 1Bx14 on dough rheology are associated with the reduced numbers of cysteine residues, consistent with the results of Pirozi et al., (2008).

The 1Ax1 missense mutant with a single amino acid substitution showed different characteristics from the knock out mutants, exhibiting dough development and stability times and a maximum resistance value higher than those of the control variety. The loaf volumes of the mutants mirrored the

rheological analyses, with the five knock-out mutants showing significant reductions in loaf volume and the 1Ax1 missense mutant an increase of 16% compared to the control variety.

More extreme variation in HMW-GS structure, such as length of the repetitive domain and numbers of cysteine residues, has also been identified in more distant wheat relatives within the tribe of the *Triticeae.* For example, Jang et al (2013), identified and characterised two novel HMW-GS (designated 1Fx3.7 and 1Fy1.5) in *Eremopyrum bonaepartis* (Spreng.) Nevski, an annual diploid species of the *Triticeae* tribe. The deduced amino acid sequence showed that subunit 1Fx3.7 had a very high molecular weight, comprising 1223 amino acid residues compared to 817 for HMW-GS 1Dx2 which has the highest mass of commonly occurring HMG-GS in bread wheat, and an N-terminal region similar to those of typical y-type subunits. It also contains eight cysteine residues, with two additional cysteine residues in the N-terminal region and two in the repetitive domain when compared with x-type HMW-GS in bread wheat. The last cysteine residue normally present in the C-terminal region of x-type HMW-GS is absent due to a mutation but an additional cysteine residue is present at the beginning of the C-terminal region. This cysteine is also present in the y-type subunit of rye and in D hordein of barley (Jang et al., 2013). The 1Fy1.5 subunit has six cysteine residues, rather than seven as normally present in y-type subunits, due to the deletion of the GHCPTSPQQ nonapeptide at the end of the repetitive domain. The effects of these novel HMW-GS on glutenin polymer formation have not been reported but could be easily established using in-vitro incorporation, as described by Tamas et al., (2002).

5. Low Molecular Weight Glutenin Subunits (LMW-GS)

5.1. Genetics and polymorphism

The LMW-GS comprise a large and heterogenous group of subunits, represent 60% of the total glutenins (Masci et al 1988; Wieser et al., 2006). It is also difficult to distinguish between the LMW-GS released by reduction of polymers and gliadins using SDS-PAGE. It has therefore proved a challenge to analyse the multigene loci which encode the LMW-GS, and in particular to identify recombination within loci (Dong et al., 2010; Ibba et al., 2017a).

Jackson et al., (1983) initially separated the LMW-GS using two different two-dimensional electrophoretic systems that combined isoelectric focusing or non-equilibrium pH gradient electrophoresis in the first dimension, followed by SDS-PAGE in the second dimension. They classified the LMW-GS into three groups (B, C and D types) based on size and isoelectric point (pI).

The C group of subunits have mobilities similar to those of α-, β- and γ-gliadins when separated by SDS-PAGE and have a wide range of isoelectric points, whereas the D group are more acidic in pI than the other subunits and comprise a small number of components. Comparisons of amino acid sequences clearly show that the C and D LMW-GS correspond to mutant forms of gliadins. However, their presence in glutenin polymers is consistent with their classification as types of LMW-GS (D'Ovidio et al., 1995).

The B-type represent the major group of LMW-GS. They have slower mobilities on SDS-PAGE compared to the gliadins and are more basic in pI than other wheat endosperm proteins (D'Ovidio and Masci., 2004). They are encoded by genes at the orthologous *Glu-3* loci, present on the short arms of

524 the homoeologous group 1 chromosomes in bread and durum wheats, closely
525 linked with the *Gli-1* loci that encode gliadins (Gupta and Shepherd, 1993;
526 Payne, 1987).
527 Comparison of the two-dimensional separations of the C-type subunits
528 present in Chinese Spring and aneuploid (nullisomic-tetrasomic and
529 ditelosomic) lines has shown that they are encoded by genes present on the
530 chromosomes of the homoeologous groups 1 and 6 (Masci et al., 2002). This
531 is consistent with the suggestion that the C subunits may be encoded by
532 genes either present in or tightly linked to the *Gli-1* and *Gli-2* loci whereas the
533 typical B-type LMW-GS are likely to be encoded by the *Glu-3* loci.
534 In bread wheat a different number of alleles at the *Glu-3* loci have been
535 detected (McIntosh et al., 2013),, eight at the *Glu-A3* locus, fourteen at the
536 *Glu-B3* locus, and thirteen at the *Glu-D3* locus.Zhang et al., (2013)
537 determined the LMW-GS genes present in a micro-core collections of Chinese
538 common wheat covering >70% of the genetic diversity of Chinese wheat
539 germplasm and found that a variable number from nine to thirteen active
540 LMW-GS genes were present in each accession. In general, the LMW-GS
541 are associated with dough resistance and extensibility and some allelic forms
542 of LMW-GS show greater effects on these properties than HMW-GS (Gupta
543 et al., 1995; Ibba et al., 2017b).
544 The correct identification of LMW-GS genes and their association with
545 chromosomes has been limited by their high complexity and polymorphism.
546 For example, Liu et al., (2010) used SDS-PAGE, 2-DE, MALDI-TOF-MS and
547 PCR to determine the LMW-GS compositions of wheat cultivars from 12

different countries showing that the identification of certain alleles was possible only with a combination of different methods.

More recently, Cho et al., (2021), analysed a set of 12 near isogenic lines (NILs), developed in the bread wheat variety Aroona, containing five different alleles at the *Glu-A3* and *Glu-D3* loci and eight alleles at the *Glu-B3* locus, using reversed-phase (RP)-HPLC and a proteomic approach combining two-dimensional electrophoretic separation with liquid chromatography–tandem mass spectrometry (MS/MS). This allowed them to identify individual LMW-GSs corresponding to the different alleles at the *Glu-A3, Glu-B3*, and *Glu-D3* loci. Comparison of the patterns separated by RP-HPLC and 2-D gels showed that the LMW-GS alleles were consistent with a series of standard wheat cultivars with known allelic composition of LMW-GS, with 12 previously uncharacterised components.

5.2. Sequences and structures of LMW-GS proteins and genes

The importance of the LMW-GS in determining grain processing quality has led to detailed studies to determine their structure, genomic organization, gene expression and polymorphism in cultivated wheat and wild relatives (Dong et al., 2010; Wang et al., 2020; Shen et al., 2018).

Tao and Kasarda (1989) prepared a fraction enriched in glutenin polymers by ion exchange chromatography and after reduction separated the subunits by two-dimensional electrophoresis. Spots corresponding to the B and C type LMW-GS were characterised by N-terminal amino acid sequencing showing that the C group comprised proteins with sequences similar to the monomeric α- and γ-gliadins. Similarly, the D-type subunits correspond to ω-gliadins

573 which differ from “typical” ω-gliadins in containing a single cysteine residue
574 (Masci et al., 1999).
575 The B group of LMW-GS have been further divided into LMW-m, LMW-s and
576 LMW-i types (D’Ovidio and Masci, 2004) based on the first amino acid residue
577 of the mature protein: methionine, serine, isoleucine (Figure 3). The LMW-s
578 are the most abundant type whereas the presence of the LMW-i type has been
579 reported more recently and is associated with the A genome (Cloutier et al.,
580 2001; Ikeda et al., 2002, D’Ovidio and Masci, 2004). Gao et al. (2007)
581 demonstrated that the i-type LMW-glutenin originated from a deletion of 33-
582 bps in the 5’ coding region of the m-type gene in the A genome before wheat
583 polyploidization.
584 Masci et al., (2002) used precipitation with increasing concentrations of
585 propan-1-ol to separate B- and C- type LMW-GS present in the bread wheat
586 variety Chinese Spring and determined their N-terminal sequences after
587 separation by Reversed Phase-High Performance Liquid Chromatography (RP-
588 HPLC). The fraction enriched in B-type subunits consisted mostly of typical
589 LMW-GS which were classified as LMW-s (52%) or LMW-m (24%), whereas
590 95% of the C-type subunits had gliadin-like sequences with a slight prevalence
591 of γ-type (55%) over α-type (40%) gliadin sequences (Table 2).
592 The structure of the B-type LMW-GS was described by Cassidy et al (1998)
593 who identified four major structural domains: a short N-terminal region, not
594 present in the LMW-i type, a repetitive domain and a C-terminal domain which
595 included a cysteine-rich region (I), a glutamine-rich region (II), and a C-
596 terminal conserved region (III) (Figure 3) . As above for the HMW-GS,
597 variation in the size of the repetitive domain of the LMW-GS is responsible for

598 the variation in molecular masses that range from 31 to 43 kDa (Shewry et al.,
599 2009). The sequences of the C-type and D-type LMW-GS are similar to those
600 of the related α-, γ- and ω-gliadins but with the presence of additional cysteine
601 residues. Anderson and Greene (1997) sequenced 25 α-gliadin genes and
602 found that five had five cysteine residues instead of six while Qi et al. (2009)
603 reported sequences encoding proteins with seven, nine and ten cysteine
604 residues as well as “typical” sequences with eight cysteine residues among
605 170 γ-gliadin genes from bread wheat and related species. However, it is not
606 known whether all of these variant genes are expressed.
607 As with the HMW-GS, wider variation in LMW-GS has been identified in wild
608 relatives of wheat. For example, Tang et al., (2018) identified an i-type LMW-
609 GS with a very high molecular weight (48.4 kDa) and an additional cysteine
610 residue in the first insertion of the repetitive domain in *Aegilops uniaristata*
611 (NN, 2*n* = 2*x* = 14). Transgenic lines expressing this large LMW-GS were
612 produced using the bread wheat variety Bobwhite and Mixograph analyses
613 indicated that the transgenic lines incorporating the glutenin subunit had
614 improved flour mixing characteristics (Du et al., 2020).
615
616 **5.3. Role of the relative position of cysteine residues**
617 Most B-type LMW-GS contain eight cysteine residues, one of which is present
618 in the N-terminal domain or at the beginning the repetitive domain of the LMW-
619 m and LMW-s forms. By contrast, all the eight cysteine residues in the LMW-i
620 proteins are located in the large C-terminal region (Figure 3). It has been
621 shown that six conserved cysteine residues form inter-chain disuphide bonds

and with the remaining two cysteines available to form interchain disulphide bonds (Shewry and Tatham, 1997).

Masci et al., (1998) used the flexibility modelling to predict that the repetitive domain of a 42 kDa B-type LMW-GS had a high degree of flexibility particularly where stretches of glutamine residues were present. The modelling also showed a high degree of flexibility in the repetitive regions surrounding the first and the penultimate cysteine residues which are involved in the formation of inter-chain disulphide bonds and suggested that this flexibility favours inter-chain bond formation.

Juhàsz and Gianibelli (2006) noted that the number of residues separating the two cysteine residues involved in the formation of inter-chain disulphide bonds in B-type LMW-GS were different, with 236 residues separating the cysteine in the s-type subunits, 200 on average in the m-type and only 77 in the i-type, and suggested that also these differences could be relevant to their different contribution to polymer formation.

Recently, Markgren et al (2020) used Monte Carlo simulations to predict the formation of disulphide bonds in gliadins and low molecular weight glutenin subunits (LMW-GS). Their results suggest that the importance of the hydrophobicity of neighbouring peptide sections, synthesis chronology and chain flexibility are factors in determining the specificity for the formation of intra-chain disulphide bonds formed directly after synthesis, as suggested by Masci et al., (1998).

Lutz et., (2012) introduced a new method to determine of disulphide cross-links using liquid chromatography (LC-MS) with alternating electron transfer dissociation (ETD) and collision-induced dissociation (CID). CID produces

fragments with intact disulphide bonds, whereas ETD promotes the cleavage of disulphide bonds. This allowed the simultaneous detection of disulphide-linked and disulphide-cleaved peptide ions in a fraction partially hydrolysed with thermolysin. In particular, a new peptide was identified with a head-to-tail cross-link between the extra cysteine residue present in the HMW-GS 1Dx5 and cysteine residue present in the C-terminal region of an x-type HMW-GS. In addition, six new peptides were identified with inter-chain disulphide bonds between α-gliadins, γ-gliadins, and B-type LMW-GS.

5.4. C-type subunits, high molecular weight gliadins and the role of LMW-GS as chain terminators

The existence of mutant forms of gliadins that could have acquired or lost cysteine residues and become incorporated in the glutenin polymers is well established at the protein level (Lew et al., 1992; Wieser et al., 2006) and their existence confirmed by sequencing of corresponding genes (Anderson and Greene 1997). Hence, their incorporation into the glutenin polymers through intermolecular disulphide bonds could affect the functional properties of doughs.

Schmid et al. (2017a) reported that 128 sequences of gluten proteins with odd number of cysteine residues were present in the UniProtKB/SwissProt, European NucleotideDatabase: 42 sequences of α-gliadins,19 of γ -gliadins, 3 of ω-gliadins, 22 of LMW-GS and 42 of HMW-GS. However, it is not known whether these are all expressed in the grain or their levels of expression.

Gliadin-related C- and D-type LMW-GS subunits with odd numbers of cysteine residues will behave as chain terminators, being incorporated into glutenin

672 polymers but preventing further growth. By contrast, most B-type subunits
673 have two cysteine residues available to form inter-chain disulphide bonds and
674 will act as chain extenders of the growing polymers (Kasarda, 1989). At the
675 molecular level the major consequence of the presence of chain terminators
676 will be to reduce the average molecular weight of the glutenin polymers, with
677 a reduction in dough quality, whereas chain extenders will have an opposite
678 effect on polymer size and improve the dough characteristics.
679 More details of the compositions of the large polymeric fraction associated
680 with flour quality (Unextractable Polymeric Proteins, UPP) and the fraction
681 containing smaller polymers (Extractable Polymeric Proteins) were reported
682 by Vensel et al., (2014). They used 0.5% (w/v) SDS with and without
683 sonication to separate wheat flour proteins into SDS-extractable polymeric
684 proteins (EPP) and SDS-unextractable polymeric proteins (UPP). Both
685 fractions were then separated by size exclusion chromatography, into
686 monomeric and polymeric fractions and analysed by quantitative two-
687 dimensional gel electrophoresis coupled with tandem mass spectrometry
688 (MS/MS). This showed the presence of HMW-GS and LMW-GS in both
689 fractions. Tandem mass spectroscopy (MS/MS) was used to distinguish
690 monomeric gliadins from gliadins containing an odd number of cysteine
691 residues. These latter were detected in all the fractions but formed a much
692 higher proportion of the SDS-EPP (14.7%), compared to the SDS-UPP
693 (5.4%).
694 Further information on the relationship between chain terminators and gluten
695 polymer structure has been provided by analysis of polymer fractions which
696 differ in their solubility and in particular of an alcohol-soluble polymeric fraction

which has been called aggregated gliadins, HMW gliadins and alcohol-soluble reduced glutenin. This fraction, which is called “alcohol-soluble glutenins” here, is known to comprise oligomers and small polymers consisting largely of LMW-GS, including C- and D-type LMW GS (which are now known to be chain-terminators) and well as the major B-type LMW GS (Shewry et al., 1983).

Schmid et al., (2016) reported detailed studies of this fraction, separating a fraction extracted from the cv Akteur with 60% ethanol and separating it by gel-permeation HPLC. The molecular mass distribution ranged from 66 kDa to 680 kDa with an average degree of polymerization of 13. Three sub-fractions were separated, GP1 (21.5%) which consisted of oligomeric HMW gliadins, GP2 (15.2%) of ω5-gliadins and GP3 (63.3%) of ω1,2-gliadins, α-gliadins and γ -gliadins. The fraction was also reduced and the subunits separated by reversed-phase HPLC into four subfractions which were characterized by SDS-PAGE and semiquantitative *N*-terminal sequencing. This showed that the alcohol-soluble glutenins comprised all types of gluten proteins: 48% LMW-GS, 18% γ-gliadins, 13% α-gliadins, 9% ω1,2-gliadins, 8% HMW-GS, and 4% ω5-gliadins. They also showed that 16.5% of the terminator positions were occupied by LMW thiols, glutathione and cysteine and that proteins with an odd number of cysteine were present at 83.5% of the terminator positions.

Subsequently, Schmid et al (2017b) compared a modified Osborne fractionation procedure in the absence or presence of an alkylating agent N-ethylmaleinimide (NEMI) to alkylate free thiols groups. Alkylation led to an increase in alcohol-soluble glutenins with a higher content of LMW-GS. The

alcohol-soluble glutenins contained gluthathione, cysteine and gliadins with odd numbers of cysteine residues (ie. C- and D-type LMW-GS) which is consistent with their role as chain terminators.

These studies indicate that the polymerisation of glutenin subunits may be blocked by the incorporation of “chain terminators”, such C- and D-type LMW-GS and LMW thiols. They also suggested that the alcohol-soluble glutenins form a well-defined protein fraction together with the monomeric gliadins and polymeric glutenins. According to Schmid et al. (2017a) this fraction can account for between 10-15% of total gluten proteins, is soluble in aqueous alcohols and its components have molecular masses ranging between 70 kDa and 700 kDa. Schmid et al (2017b) also determined the disulphide bonds present in alcohol-soluble glutenins using the fractions prepared with and without NEMI. The fractions were treated with thermolysin and the peptides released separated by gel permeation chromatography and analysed by liquid chromatography/mass spectrometry using alternating ETD/CID. Most of the disulphide bonds identified had been reported before but eleven new bonds were identified. Fifteen peptides containing disulphide bonds were unique to the alcohol-soluble glutenins and contained cysteine residues from ω5-, α- and γ-gliadins with an odd number of cysteines (ie. C- and D-type LMW-GS). This therefore supports the suggestion that gliadins with an odd number of cysteines, glutathione and cysteine acted as terminators of glutenin polymerisation.

No major differences were observed between the disulphide bonds in fractions prepared without and with NEMI treatment showing that polymerisation was complete in the flour. However, the non-treated sample did contain higher

amounts of most of the ten peptides with disulphide bonds between gluten proteins and low-molecular-weight thiols such as glutathione.

Taken together, the studies of Schmid et al., (2017a,b) show that the alcohol-soluble glutenins are a discrete fraction which may further polymerise during extraction (and presumably also processing) by rearrangements of disulphide bonds with low molecular weight thiols to form protein-protein disulphide bonds.

6. Role of post translational modification of HMW-GS and LMW-GS

Proteins may undergo a range of post-translational modifications and processing events, including phosphorylation, glycosylation and proteolytic processing. Only the latter appears to occur to a significant extent in wheat gluten proteins and this has only been characterised in detail in the last twenty years.

Dupont et al., (2004) showed that the ω-gliadins encoded by chromosomes 1A and 1D of bread wheat have two different N-terminal sequences, ARELNPQNKEL- and KELQSPQQSF, with the latter the result of a partial cleavage at the NK site. This suggests that the cleavage is catalysed by a legumain-like asparaginyl endoprotease, which have been shown to be involved in post-translational proteolysis of seed storage protein in the vacuole (Muntz and Shutov, 2002).

The complete nucleotide sequences show that the protein sequences of LMW-s and LMW-m types are very similar, with the LMW-s lacking three N-terminal amino acids (MET- MEN- or IEN) corresponding to the N-termini of the LMW-m GS. Masci et al., (1998) identified a major LMW-GS present in the bread

772 wheat cultivar Yecora Rojo (designated 42K) and compared the sequences of
773 the protein and corresponding putative gene. The deduced amino acid
774 sequence of the gene revealed the presence of the MEN tripeptide adjacent
775 to the protein N-terminus. This differs from the MET sequence present in the
776 LMW-m type GS and Masci et al. (1998) suggested that the asparagine (N)
777 residue provides a site for proteolysis.
778 In order to explore this suggestion, Egidi et al., (2014) used site-directed
779 mutagenesis to replace the asparagine present at position 23 of an LMW-s
780 subunit with threonine, and similarly mutated a LMW-m type gene to replace
781 threonine at position 23 to with asparagine. The mutated versions of the LMW-
782 m and LMW-s genes, together with the wild-type counterpart of the LMW-m
783 gene, were then introduced into the durum wheat variety Svevo. Proteomic
784 comparisons followed by MS/MS analyses and Edman N-terminal sequencing
785 showed that the processing is dependent on the presence of a threonine or
786 asparagine residue in position 23 of the coding sequences: when an
787 asparagine residue is present, proteins are processed as LMW-s types;
788 conversely, when a threonine residue is present, the mature proteins are
789 LMW-m types. Egidi et al., (2014) therefore suggested that the processing of
790 the LMW-s proteins is the result of a post-translational processing, probably
791 by a legumain-like asparaginyl endoprotease, as discussed above.
792 Bacala et al. (2020a) analysed two parents and 28 F5 progeny of a cross
793 between them. They identified 24 pairs of LMW-GS which differed in mass by
794 163 Da. Subsets of these protein pairs were present in the individual lines with
795 the proteins with lower masses always being less abundant and eluting faster.
796 They then purified the most abundant LMW-GS pair (40,344/40,181Da) from

797 the Canadian bread wheat variety Cardale and used peptide mapping to show
798 that the 163 Da difference in mass resulted from cleavage of a conserved
799 tyrosine which is present at the C-terminus of all B-type LMW-GS (Bacala et
800 al., 2020b).
801 Using the same set of lines Bacala et al (2020a) demonstrated that
802 proteolytical processing also occurs for HMW-GS. They showed the presence
803 of co-eluting lower mass variants for HMW-GS subunits 1Bx7 and 1Bx7* and
804 1Dy. In particular, the masses of the two 1Bx subunits were reduced from
805 83,544 Da to 82,888 Da (a difference of 656 Da), whereas that of the 1Dy
806 subunit was reduced from 68,212 Da to 67,654 Da (a difference of 558 Da).
807 The same mass reduction was also detected for the 1Bx7 and 1Dy HMW-GS
808 present in the bread wheat Cardale (Bacala et al., 2020b). The authors stated
809 that the truncations occurred at the C-terminal end but did not determine
810 precise sites. This agrees with the studies of Cunsolo et al. (2003, 2004) who
811 showed that a proteolytic truncation removed amino acid residues from the C-
812 terminus of HMW-GS 1Bx7, 1Bx20, Dy10 and Dy12.
813 More recently, Nunes-Miranda et al. (2017), used proteomics to characterise
814 the faster subunit bands associated with 1By HMW-GS when separated by
815 SDS-PAGE. They showed that the two fast migrating bands accompanying
816 the major 1By HMW-GS subunits arose by proteolytic cleavage at two
817 asparagine residues present in the beginning of their C-terminal domain. This
818 processing generated two isoforms that had lost 36 or 41 amino acid residues
819 including a unique cysteine residue present at the C-terminal domain of the
820 subunits. This cysteine residues is believed to be involved in the formation of
821 an inter-chain disulphide bond with a cysteine residue of a x-type HMW-GS

thus contributing to the extension of the glutenin polymers. The importance of this bond in the formation of glutenin polymers was demonstrated by Wang et al (2021) who studied the bread wheat variety Shumai 482 (Ax1, Bx7 + By9, and Dx5 +Dy10) after treatment with EMS. SDS-PAGE of the HMW-GS showed the presence of an additional band, with a faster electrophoretic mobility than 1Dy10 subunit on SDS-PAGE, which resulted from a missense mutation (G to A) which led to a serine-to-asparagine substitution at residue 619 residue within the C-terminal domain of HMW-GS 1Dy10. This change resulted in the loss of a sequence 29 residues with the mutant form accounting for 52.6% of the total HMW-GS Dy10 and was associated with drastic changes in quality traits, rheological properties, and baking quality.

7. **Conclusions**

Early studies demonstrated the importance of gluten proteins in determining dough rheological characteristics, with gliadins imparting extensibility and glutenin strength. The development in the past 20 years of mass-based proteomics, with its technological evolution and versatility, has greatly expanded our knowledge of glutenin subunits including discovering new diversity in the number of cysteine subunits, identifying intra- inter-chain disulphide bonds demonstrating the existence of post-translational modifications. MS-based proteomics will also allow high throughput analysis of the large gene pool of the *Triticeae* in order to identify novel variation which can be exploited in improving durum and bread wheat. However, this progress has mainly related to the HMW-GS and the identification of LMW-GS proteins and corresponding genes that confer improved quality remains a challenge.

The future landscape will also be strongly influenced by the development of genome editing, with the CRISPR (clustered regularly interspaced short palindromic repeats) Cas9 technology being most used due to its simplicity and versatility. Genome edited allows the precise modification of gene sequences including single base changes, insertions, and deletions. Genome editing has already been applied to many crop species, including wheat, but is having the greatest impact in species such as maize in which the relationships between gene sequences and phenotypic traits are well understood. We therefore require further information on the relationships between the sequences of HMW and LMW GS and the structures and functional properties of glutenin polymers to its full potential.

References

Altenbach, S.B., Chang, H.-C., Simon-Buss, A., Mohr, T., Huo, N., Gu, Y.Q., 2020. Exploiting the Reference Genome Sequence of Hexaploid Wheat: A Proteomic Study of Flour Proteins from the Cultivar Chinese Spring. Funct. Integr. Genom. 20, 1–16.

Anderson, O.D., Greene, F.C., 1997. The α-gliadin gene family. II. DNA and protein sequence variation, subfamily structure, and origins of pseudogenes. Theor. Appl. Genet. 95, 59–65.

Bacala, R., Fu, B.X., Perreault, H., Hatcher, D.W., 2020a. Quantitative LC-MS proteoform profiling of intact wheat glutenin subunits. J. Cereal Sci. 94, 102963.

Bacala, R., Fu, B.X., Perreault, H., Hatcher, D.W., 2020b C-terminal tyrosine removal from wheat low-molecular weight glutenin subunits (LMW-GS); biologically relevant or mistaken substrate?. J. Cereal Sci. 95, 103060.

873 Belton, P.,1999. On the elasticity of wheat gluten. J. Cereal Sci. 29, 103-107.
874 Bromilow, S., Gethings, L.A., Buckley, M., Bromley, M., Shewry, P.R., Langridge, J.I.,
875 Clare Mills, E.N., 2017. A curated gluten protein sequence database to support
876 development of proteomics methods for determination of gluten in gluten-free
877 foods. J. Proteomics 163, 67-75.
878 Butow, B.J., Ma, W., Gale, K.R., et al., Butow, B.J., Ma, W., Gale, K.R., Cornish,G.B.,
879 Rampling, L., Larroque, O., Morell, M.K., Békés, F., 2003. Molecular discrimination
880 of Bx7 alleles demonstrates that a highly expressed high-molecular weight
881 glutenin allele has a major impact on wheat flour dough strength. Theor. Appl.
882 Genet. 107, 1524–1532.
883 Cao, H., Duncan, O., Islam, S., Zhang, J., Ma, W., Millar, A.H., 2021.Increased Wheat
884 Protein Content via Introgression of an HMW Glutenin Selectively Reshapes
885 the Grain Proteome Mol. Cell. Proteom. 20, 100097.
886 Cassidy, B.G., Dvorak, J., Anderson, O.D., 1998. The wheat low-molecular-weight
887 glutenin genes: characterization of six new genes and progress in
888 understanding gene family structure Theor. Appl. Genet. 96, 743-750.
889 Cho, K., Jang, Y.R., Lim, S.H., Altenbach, S.B.,. Gu, Y.Q., Simon-Buss, A., Lee, J-Y.,
890 2021. Proteomic determination of low molecular weight glutenin subunit
891 composition in Aroona near-isogenic lines and standard wheat cultivars. Int.
892 J. Mol. Sci. 22, 7709.
893 Cloutier, S., Rampitsch, C., Penner, G.A., Lukow, O.M., 2001. Cloning and expression
894 of a LMW-i glutenin gene. J. Cereal Sci. 33, 143–154.
895 Cunsolo, V., Foti, S., Saletti, R., Gilbert, S., Tatham, A.S., Shewry, P.R., 2003.
896 Structural studies of glutenin subunits 1Dy10 and 1Dy12 by matrix-assisted
897 laser desorption/ ionisation mass spectrometry and high-performance liquid

chromatography/ electrospray ionisation mass spectrometry. Rapid Commun. Mass Spectrom. 17, 442–454.

Cunsolo, V., Foti, S., Saletti, R., Gilbert, S., Tatham, A.S., Shewry, P.R., 2004. Structural studies of the allelic wheat glutenin subunits 1BX7 and 1Bx20 by matrix-assisted laser desorption/ionization mass spectrometry and high-performance liquid chromatography/electrospray ionization mass spectrometry. J. Mass Spectrom. 39, 66–78. https://doi.org/10.1002/jms.558.

D'Ovidio R., Masci., 2004. The low-molecular-weight glutenin subunits of wheat gluten. J. Cereal Sci. 39, 321-339.

D'Ovidio, R., Simeone, M., Masci, S., Porceddu, E., Kasarda, D.D., 1995. Nucleotide sequence of a gamma-type glutenin gene from a durum wheat: Correlation with a gamma-type glutenin subunit from the same biotype. Cereal Chem. 72,. 443-449.

Don, C., Mann G., Bekes, F., Hamer, R.J., 2006. HMW-GS affect the properties of glutenin.

Dong, L., Zhang. X., Liu, D, Fan, H., Sun, J., Zhang, Z., Bin, H., Li, B., Hao, S., Li, Z., Wang, D., Zhang, A. Ling, H-Q., 2010. New Insights into the Organization, Recombination, Expression and Functional Mechanism of Low Molecular Weight Glutenin Subunit Genes in Bread Wheat. PLoS ONE 5, (10).

Du, X., Wei, J., Luo, X., , Liu, Z., Qian, Y., Zhu, B., Weng, Q,, Tang, H., 2020. Low-molecular-weight glutenin subunit LMW-N13 improves dough quality of transgenic wheat. Food Chem. 327, 127048.

Dumur, J., Jahier, J., Dardevet, M.,Chiron, H., Tanguy, A.M., Branlard, G., 2010. Effects of the replacement of *Glu-A1* by *Glu-D1* locus on agronomic

922 performance and bread-making quality of the hexaploid wheat cv. Courtot. J.
923 Cereal Sci. 51, 175-18.
924 Dupont, F.M., Vensel, W., Encarnacao,T., Chan, R., Kasarda, D.D., 2004. Similarities
925 of omega gliadins from *Triticum urartu* to those encoded on chromosome 1A
926 of hexaploid wheat and evidence for their post-translational processing,
927 Theor. Appl. Genet. 108, 1299–1308.
928 Egidi, E.,Sestili, F., Janni, M., D'Ovidio, R., Lafiandra, D., Ceriotti, A., Vensel, W.H.,
929 Kasarda, D.D., Masci, S., 2014. An asparagine residue at the N-terminus
930 affects the maturation process of low molecular weight glutenin subunits of
931 wheat endosperm, BMC Plant Biol. 14, 64.
932 Gao, S., Gu, Y.Q., Wu, J., Coleman-Derr, D., Huo, N., Crossman, C., Jia, J., Zuo, Q.,
933 Ren, Z., Anderson, O.D., Kong, X., 2007. Rapid evolution and complex
934 structural organization in genomic regions harboring multiple prolamin genes
935 in the polyploid wheat genome Plant Mol. Biol., 65, 189-203.
936 Gao, X., Appelbee, M.J., Mekuria, G.T., Chalmers, K.J., Mather, D.E., 2012. A second
937 'overexpression' allele at the Glu-B1 high-molecular-weight glutenin locus of
938 wheat: Sequence characterisation and functional effects Theor. Appl. Genet.
939 124, 333-343.
940 Geng, Y., Pang, B., Hao, C., Tang, S., Zhang, X., Li, T., 2014. Expression of wheat
941 high molecular weight glutenin subunit 1Bx is affected by large insertions and
942 deletions located in the upstream flanking sequences. PLoS ONE 9, e105363.
943 Gu, Y.Q.,Wanjugi, H., Coleman-Derr D., Kong, X., Anderson, O.D., 2010. Conserved
944 globulin gene across eight grass genomes identify fundamental units of the
945 loci encoding seed storage proteins. Funct. Integr. Genomics 10, 111–122.

946 Gu, Y.Q., Coleman-Derr, D., Kong, X., Anderson, O.D., 2004. Rapid genome evolution
947 revealed by comparative sequence analysis of orthologous regions from four *Triticeae*
948 genomes. Plant Physiol. 135, 459-470.
949 Gupta, R.B.. Popineau, Y., Lefebvre, J., Cornect, M., Lawrence, G.J., MacRitchie, F.,
950 1995. Biochemical basis of flour properties in bread wheats. II. Changes in
951 polymeric protein formation and dough/gluten properties associated with the
952 loss of low M_r or high M_r glutenin subunits. J. Cereal Sci. 21, 103-116.
953 Gupta, R.B., Shepherd, K.W., 1993. Production of multiple wheat-rye 1RS
954 translocation stocks and genetic analysis of LMW subunits of glutenin and
955 gliadins in wheats using these stocks. Theor. Appl. Genet. 85, 719-728.
956 Gupta, R.B., MacRitchie, F., 1994. Allelic variation at glutenin subunit and gliadin loci,
957 *Glu-1*, *Glu-3*, and *Gli-1* of common wheats. II. Biochemical basis of the allelic
958 effects on dough properties. J. Cereal Sci. 19, 19-29.
959 Halford, N.G., Field, J.M., Blair, H., Urwin, P., Moore, K., Robert, L., Thompson, L.,
960 Flavell, R.B., Tatham, A.S., Shewry, P.R., 1992. Analysis of HMW glutenin
961 subunits encoded by chromosome 1A of bread wheat (*Triticum aestivum* L.)
962 indicates quantitative effects on grain quality. Theor. Appl. Genetics 83, 373-
963 378.
964 Ibba, M.I., Kiszonas, A.M., Morris, C.F., 2017a. Evidence of intralocus recombination
965 at the *Glu-3* loci in bread wheat (*Triticum aestivum* L.). Theor Appl. Genet.
966 130, 891-902.
967 Ibba, M.I., Kiszonas, A.M., Morris, C.F., 2017b. Influence of low-molecular-weight
968 glutenin subunit haplotypes on dough rheology in elite common wheat
969 varieties. Cereal Chem. 94, 1016-1027.

970 Jackson, E.A., Holt, L.M., Payne, P.I., 1983. Characterisation of high molecular weight
971 gliadin and low-molecular-weight glutenin subunits of wheat endosperm by
972 two-dimensional electrophoresis and the chromosomal localisation of their
973 controlling genes. Theor. Appl. Genet. 66, 29-37.
974 Jiang, Q-T., Wei, Y-M, Wang, F., Wang, J-R., Yan, Z-H., Zheng, Y-L., 2009.
975 Characterization and comparative analysis of HMW glutenin 1Ay alleles with
976 differential expressions. BMC Plant Biol. 9, 16.
977 Jang, Y.-R., Kim, S., Sim, J.-R., Lee, S-B., Lim, S-H., Kang,C-S., Choi, C., Goo, T-W.,
978 Lee, J-Y., 2021. High-throughput analysis of high-molecular weight glutenin
979 subunits in 665 wheat genotypes using an optimized MALDI-TOF–MS
980 method. 3 Biotech 11, 92.
981 Jiang, Q.T., Zhang, X.W., Ma, J.,Wei, L., Zhao, S., Zhao, Q-Z., Qi, P-F., Lu, Z-X.,
982 Zheng, Y-L., Wei, Y-M., 2013. Characterization of high-molecular-weight
983 glutenin subunits from *Eremopyrum bonaepartis* and identification of a novel
984 variant with unusual high molecular weight and altered cysteine residues.
985 Planta 239, 865–875.
986 Juhász, A., Gianibelli, M.C., 2006. Low-molecular-weight glutenin subunits: insights
987 into this abundant subunit group present in glutenin polymers. In: Wrigley, C.,
988 Bekes, F., Bushuk, W. (Eds.), Gliadin and Glutenin: the Unique Balance of
989 Wheat Quality. St. Paul, American Association of Cereal Chemistry, pp. 171–
990 212.
991 Juhász, A., Larroque, O.R., Tamás, L., Hsam, S.L.K., F. Zeller F.J., Békés, F., Bedő,
992 Z., 2003. Bánkúti 1201 -An old Hungarian wheat variety with special storage
993 protein composition. Theor.Appl. Genet. 107, 697-704.

994 Kasarda, D.D., 1989. Glutenin structure in relation to wheat quality. In Wheat is unique,
995 Y. Pomeranz ed, AACC, pp. 277-302.
996 Kong, X-Y., Gu, Y.Q., You, F.M., Dubcovsky, J., Anderson, O.D., 2004. Dynamics of
997 the evolution of orthologous and paralogous portions of a complex locus
998 region in two genomes of allopolyploid wheat. Plant Mol. Biol. 54, 55-69.
999 Lawrence, G.J., MacRitchie, F., Wrigley, C.W., 1988. Dough and baking quality of
1000 wheat lines deficient in glutenin subunits controlled by the Glu-A1, Glu-B1 and
1001 Glu-D1 loci. J. Cereal Sci. 7, 109-112.
1002 Lew, E-J.L., Kuzmicky, D.D., Kasarda, D.D., 1992., Characterization of low molecular
1003 weight glutenin subunits by reversed-phase high-performance liquid
1004 chromatography, sodium dodecyl sulfate-polyacrylamide gel electrophoresis,
1005 and N-Terminal Amino Acid Sequencing. Cereal Chem. 69, 508-515.
1006 Li, W., Wan, Y., Liu, Z., Liu, K., Liu, X., Li, B., Li, Z., Zhang, X., Dong, Y., Wang, D.,
1007 2004. Molecular characterization of HMW glutenin subunit allele 1Bx14:
1008 further insights into the evolution of *Glu-B1-1* alleles in wheat and related
1009 species. Theor. Appl. Genet. 109, 1093-1104
1010 Li, Y. An, X., Yang, R., Yue, G., Fan, R., Li, B., L, Z., Zhang, K., Dong, Z., Zhang, L.,
1011 Wang, J., Jia, Zhang, A., Zhang, X., Wang, D. 2015. Dissecting and enhancing
1012 the contributions of high-molecular-weight glutenin subunits to dough
1013 functionality and bread quality. Mol. Plant, 8, 332-334.
1014 Lukaszevski A., 2003. Registration of Six Germplasms of Durum Wheats- with
1015 Introgressions of the Glu-D1 Locus Crop Sci. 43, 1137–1138.
1016 Lukow, O.M., Forsyth, A.S., Payne, P.I., 1992. Over-production of HMW glutenin
1017 subunits coded on chromosome IB in common wheat. *Triticum aestivum*. J.
1018 Genet. Breed. 46, 187-192.

1019 Lutz, E., Wieser, H., Koehler, P., 2012. Identification of disulfide bonds in wheat gluten
1020 proteins by means of mass spectrometry/electron transfer dissociation. J.
1021 Agric. Food Chem. 60, 3708-3716.
1022 Margiotta, B., Urbano, M., Colaprico, G., Johansson, E., Buonocore, F., D'Ovidio, R.,
1023 Lafiandra, D., 1996. Detection of y-type subunit at the *Glu-A1* locus in some
1024 Swedish bread wheat lines. J. Cereal Sci., 23, 203-211.
1025 Margiotta, B., Colaprico, G., D'ovidio, R., Lafiandra, D., 1993. Characterization of high
1026 m_r subunits of glutenin by combined chromatographic (RP-HPLC) and
1027 electrophoretic separations and restriction fragment length polymorphism
1028 (RFLP) analyses of their encoding genes. J. Cereal Sci. 17, 221-236.
1029 Margiotta, B., Pfluger, L., Roth, M.R., MacRitchie F., Lafiandra D., 2000. Isogenic bread
1030 wheat lines differing in number and type of high Mr glutenin subunits. In:
1031 Shewry, P.R., Tatham, A.S. (Eds.), Wheat Gluten. Royal Society of Chemistry,
1032 London, pp. 29–33.
1033 Markgren, J., Hedenqvist, M., Rasheed, F., Skepö, M., Johansson, E., 2020. Glutenin
1034 and gliadin, a piece in the puzzle of their structural properties in the cell
1035 described through monte carlo simulations Biomolecules, 10, 1095.
1036 Masci, S., D'Ovidio, R., Lafiandra, D., Kasarda, D.D., 1998. Characterization of a low-
1037 molecular-weight glutenin subunit gene from bread wheat and the
1038 corresponding protein that represents a major subunit of the glutenin polymer.
1039 Plant Physiol., 118, 1147-1158.
1040 Masci, S., Egorov, T.A., Ronchi, C., Kuzmicky, D.D., Kasarda, D.D., Lafiandra, D.,
1041 1999. Evidence for the presence of only one cysteine residue in the D-type
1042 low molecular weight subunits of wheat glutenin. J. Cereal Sci. 29, 17-25.

1043 Masci, S., Rovelli, L., Kasarda, D.D., Vensel, W.H., Lafiandra, D., 2002.
1044 Characterisation and chromosomal localisation of C-type low-molecular-
1045 weight glutenin subunits in the bread wheat cultivar Chinese Spring. Theor.
1046 Appl. Genet. 104, 422-428.
1047 Muccilli V, Cunsolo V, Saletti R, Foti, S., Masci, S., Lafiandra, D., 2005.
1048 Characterization of B- and C-type low molecular weight glutenin subunits by
1049 electrospray ionization mass spectrometry and matrix-assisted laser
1050 desorption/ionization mass spectrometry. Proteomics 5,719–728.
1051 McIntosh, R.A., Yamazaki, Y., Dubcovsky, J., Rogers, J., Morris, C., Appels, R., Xia,
1052 X.C., 2013. Catalogue of Gene Symbols for Wheat. In: 12th International
1053 Wheat Genetics Symposium, Yokohama, Japan
1054 Müntz, K., Shutov, A.D., 2002. Legumains and their functions in plants. Trends Plant
1055 Sci. 7, 340-344.
1056 Nunes-Miranda, J.D., Bancel, E, Viala, D.,Chambon, C., Capelo, J.L., Branlard, G.,
1057 Ravel, C., Igrejas, G., 2017. Wheat glutenin: the “tail” of the 1By protein
1058 subunits. J. Proteome 169:136–142.
1059 Payne, P., I., 1987. Genetics of wheat seed storage proteins and the effect of allelic
1060 variation on bread-making quality. Ann. Rev. Plant Physiol. 38, 141-53.
1061 Payne, P.I., Holt, L.M., Jackson, E.A., Law, C.N., 1984.Wheat storage proteins: their
1062 genetics and their potential for manipulation by plant breeding. Phil. Trans. R.
1063 Soc. Lond. B 304, 359-371.
1064 Payne, P.I., Holt, L.M., Krattiger, A.F., Carrillo, J.M., 1988. Relationships between
1065 seed quality characteristics and HMW glutenin subunit composition
1066 determined using wheats grown in Spain. J. Cereal Sci. 7, 229–235.

1067 Pirozi, M.R., Margiotta, B., Lafiandra, D., MacRitchie, F., 2008. Composition of
1068 polymeric proteins and bread-making quality of wheat lines with allelic HMW-
1069 GS differing in number of cysteines. J. Cereal Sci. 48, 117–122.
1070 Popineau, Y., Cornec, M., Lefebvre, J., Marchylo, B., 1994. Influence of high M_r
1071 glutenin subunits on glutenin polymers and rheological properties of glutens
1072 and gluten subfractions of near-isogenic lines of wheat Sicco. J. Cereal Sci.
1073 19, 231-241.
1074 Qi, P.F., Wei, Y.-M., Ouellet, T., Chen, Q., Tan, X., Zeng, Y-L., 2009. The γ-gliadin
1075 multigene family in common wheat (*Triticum aestivum*) and its closely related
1076 species. BMC genomics 10, 168.
1077 Ragupathy, R., Naeem, H.A., Reimer, E., Lukow, O.M., Sapirstein, H.D., Cloutier, S.,
1078 2008. Evolutionary origin of the segmental duplication encompassing the
1079 wheat GLU-B1 locus encoding the overexpressed Bx7 ($Bx7^{OE}$) high molecular
1080 weight glutenin subunit. Theor. Appl. Genet. 116, 283-296.
1081 Rogers, W.J., Miller, T.E., Payne, P.I., Seekings, J.A., Sayers, E.J., Holt, L.M., Law,
1082 C.N., 1997. Introduction to bread wheat (*Triticum aestivum* L.) and
1083 assessment for bread-making quality of alleles from *T. boeoticum* Boiss. ssp.
1084 *thaoudar* at *Glu-A1* encoding two high-molecular-weight subunits of glutenin.
1085 Euphytica 93, 19-29.
1086 Rogers, W.J., Payne, P.I., Seekings, J.A., Sayers, E.J., 1991. Effect on breadmaking
1087 quality of x-type and y-type high molecular weight subunits of glutenin. J.
1088 Cereal Sci. 14, 209-221.
1089 Roy, N., Islam, S., Al-Habbar, Z., Yu, Z., Liu, H., Lafiandra, D., Masci, S., Lu, M.,
1090 Sultana, N., Ma. W., 2021. Contribution to Breadmaking Performance of Two

1091 Different HMW Glutenin 1Ay Alleles Expressed in Hexaploid Wheat. J. Agric.
1092 Food Chem. 69, 36-44.
1093 Schmid, M., Wieser, H., Koehler, P., 2016. Isolation and characterization of high-
1094 molecular-weight (HMW) gliadins from wheat flour. Cereal Chem.93, 536-542.
1095 Schmid, M., Wieser, H., Koehler, P., 2017a. Alkylation of free thiol groups during
1096 extraction: Effects on the Osborne fractions of wheat flour. Cereal Chem. 94,
1097 546-553.
1098 Schmid, M., Wieser, H., Koehler, P., 2017b. Disulphide structure of high-molecular-
1099 weight (HMW-) gliadins as affected by terminators. J. Cereal Sci. 78, 66-74.
1100 Sears, E.R. 1976. Genetic control of chromosome-pairing in wheat. Ann. Rev. Genet.
1101 10, 31–51.
1102 Shen, L., Luo, G., Song, Y., Song, S., Li, Y., Yang, W., Li, X., Sun, J., Liu, D., Zhang,
1103 A., Low molecular weight glutenin subunit gene composition at *Glu-D3* loci of
1104 *Aegilops tauschii* and common wheat and a further view of wheat evolution.
1105 Theor. Appl. Genet. 131, 2745–2763.
1106 Shewry, P.R., D'Ovidio, R., Lafiandra, D., Jenkins, J.A, Mills, E.N.C., Békés, F., 2009.
1107 Wheat grain proteins. In: Khan, K. and Peter, R. Shewry (Eds.). St. Paul
1108 American Association of Cereal Chemistry, pp. 223–298.
1109 Shewry, P., Gilbert, S., Savage, A., A S Tatham, Y-F Wan, P S Belton, N Wellner, R
1110 D'Ovidio, F Békés, N G Halford 2003. Sequence and properties of HMW
1111 subunit 1Bx20 from pasta wheat (*Triticum durum*) which is associated with
1112 poor end use properties. Theor. Appl. Genet. 106, 744–750.
1113 Shewry, P.R., Halford, N.G., Lafiandra, D., 2006. The High-Molecular-Weight subunits
1114 of glutenin. In: Wrigley, C., Bekes, F., Bushuk, W. (Eds.), Gliadin and Glutenin:

the Unique Balance of Wheat Quality. St. Paul American Association of Cereal Chemistry, pp. 143–169.

Shewry, P.R., Miflin, B.J., Lew, E.J.-L., Kasarda, D.D., 1983. The preparation and characterization of an aggregated gliadin fraction from wheat. J. Exp. Bot. 34, 1403-1410.

Shewry P.R., Tatham A.S., 1997. Disulphide Bonds in Wheat Gluten Proteins. J. Cereal Sci. 25, 207-227.

Tahir, M., Pavoni, A., Tucci, G.F., Turchetta, T., Lafiandra, D., 1996. Detection and characterization of a glutenin subunit with unusual high Mr at the *Glu-A1* locus in hexaploid wheat. Theor.Appl. Genet. 92, 654-659.

Tamás, L., Gras, P.W., Solomon, R.G., Morell, M.K., Appels, R., Békés, F., 2002. Chain extension and termination as a function of cysteine content and the length of the central repetitive domain in storage proteins. J. Cereal Sci. 3, 313-325.

Tang, H., Du, X. Y., Wang, H.W., Ma, X., Bo, C., Li, A., Kong, L., 2018. Cloning and characterization of a novel low-molecular-weight glutenin subunit gene with an unusual molecular structure of *Aegilops uniaristata*. J. Genet. 97, 1463–1468**.**

Tao, H.P., Adalsteins, A.E., Kasarda, D.D., 1992. Intermolecular disulfide bonds link specific high-molecular-weight glutenin subunits in wheat endosperm. Biochim. Biophys. Acta 1159, 13-21.

Tao, H.P., Kasarda, D.D., 1989. Two-dimensional gel mapping and N-terminal sequencing of LMW-glutenin subunits. J. Exp. Botany. 40, 1015-1020.

1138 Uthayakumaran, S., Batey, I.L., Wrigley, C.W., 2005. On-the-spot identification of
1139 grain variety and wheat-quality type by Lab-on-a-chip capillary
1140 electrophoresis. J. Cereal Sci. 41, 371-374.
1141 Vensel, W.H., Tanaka, C.K., Altenbach, S.B., 2014. Protein composition of wheat
1142 gluten polymer fractions determined by quantitative two-dimensional gel
1143 electrophoresis and tandem mass spectrometry. Proteome Sci. 12, 8.
1144 Wang, D., Li, F., Cao, S., Zhang, K., 2020. Genomic and functional genomics analyses
1145 of gluten proteins and prospect for simultaneous improvement of end-use and
1146 health-related traits in wheat. Theor. Appl. Genet. 133, 1521-1539.
1147 Wang, Z., Li, Y., Yang, Y., Liu, X., Qin, H., Dong, Z., Zheng, S., Zhang, K., Wang, D.,
1148 2017. New insight into the function of wheat glutenin proteins as investigated
1149 with two series of genetic mutants. Sci. Rep. 7**,** 3428.
1150 Wang, Y., Chen, Q., Li, Y., Guo, Z., Liu, C., Wan, Y., Hawkesford, M., Zhu, J., Wu, W.,
1151 Wei, M., Zhao, K., Jiang, Y., Zhang, Y., Xu, Q., Kong, L., Pu, Z., Deng, M.,
1152 Jiang, Q., Lan, X., Wang, J., Chen, G., Ma, J., Zheng, Y., Wei, Y., Qi, P.,
1153 2021. Post-translational cleavage of HMW-GS Dy10 allele improves the
1154 cookie-making quality in common wheat (*Triticum aestivum).,* 2021. Mol.
1155 Breed. 41, 49.
1156 Wieser, H., Bushuk, W., MacRitchie, F., 2006. The polymeric glutenins. In: Wrigley,
1157 C., Bekes, F., Bushuk, W. (Eds.), Gliadin and Glutenin: the Unique Balance of
1158 Wheat Quality. St. Paul American Association of Cereal Chemistry, pp. 213–
1159 240.
1160 Wieser, H., Kieffer, R., 2001. Correlations of the amount of gluten protein types to the
1161 technological properties of wheat flours determined on a micro-scale. J.
1162 Cereal Sci. 34, 19-27.

Wieser, H., Zimmermann, G., 2000. Importance of amounts and proportions of high molecular weight subunits of glutenin for wheat quality. Eur. Food Res. Technol. 210, 324-330.

Yang, Y., Li, S., Zhang, K., 2014. Efficient isolation of ion beam-induced mutants for homoeologous loci in common wheat and comparison of the contributions of *Glu-1* loci to gluten functionality Theor. Appl. Genet. 127, 359–372.

Yu, Z., Peng, Y., Islam, M.S., She, M., Lu, M., Lafiandra, D., Roy, N., Juhasz, J., Yan, G., Ma, W., 2019. Molecular characterization and phylogenetic analysis of active y-type high molecular weight glutenin subunit genes at *Glu-A1* locus in wheat. J. Cereal. Sci. 86, 9–14.

Zhang, X., Liu, D., Zhang, J., Jiang, W., Luo, G., Yang, W., Sun, J., Tong, Y., Cui, D., Zhang, A., 2013. Novel insights into the composition, variation, organization and expression of the low-molecular-weight glutenin subunit gene family in common wheat. J. Exp. Bot. 64, 2027-2040.

Figure Legends

Figure. 1. SDS PAGE of HMW-GS present in durum and bread wheat showing different alleles present at the three *Glu-1 loci* (Lanes 1, 6 *Glu-A1*; lanes 7-14 *Glu-D1*, lanes 15-20 *Glu-B1*). In lane 6, 7 and 8 are shown subunits with very high molecular size, associated at the *Glu-A1* (2.1*) and *Glu-D1* (2.2 and 2.2*), generated as result of insertion of large fragments in the repetitive domain of corresponding genes.

Fig. 2. Electrophoretic separation of HMW-GS present in wild and cultivated wheat species. From left: 1. bread wheat variety Torim, 2. durum wheat variety Duramba,

3-5. *T. araraticum*, 6. *T. timopheevii*, 7-9. *T. urartu*, 10-11. *T. boeoticum*, 12-14. *T. dicoccoides*, 15. bread wheat line with six HMW-GS.
x and y-type subunits associated with the A genome are indicated with empty and filled triangles. (taken From Shewry et al., 2006)

Fig. 3. Structures of B-type LMW-GS (A) LMW-m and LMW-s type (B) LMW-i type. Signal peptide (S); N-terminal region (N); repetitive domain REP); C-Terminal Region (C-TER I, II, III). The cysteines involved in inter-molecular disulphide bonds are indicated with dotted lines (modified by D'Ovidio and Masci 2004).

Table 1. Role of different number of cysteine residues present in HMW-GS in influencing dough characteristics. The improvement associated to the presence of the 1Bx7 ^{OE} result of a quantitative effect rather than a different number of cysteine residues.

Table 2. Composition of B- and C-Type subunits separated by RP-HPLC subunits deduced by N-Terminal analyses (From Masci et al., 2002).