Identification of a new class of recombinant prolamin genes in wheat

Ila J. Nagy, Imre Takács, Angéla Juhász, László Tamás, and Zoltán Bedő

Abstract: A novel storage protein gene with obvious wheat chimeric structure was isolated from an immature kernel-specific cDNA library prepared from the old Hungarian variety, *Bánkúti 1201*. This clone contains γ-gliadin sequences in the 5′ region and LMW-glutenin sequences on the 3′ end. A frameshift mutation was also introduced by the putative recombination event. Hence, the amino acid sequence of the C-terminal region was transformed to a completely new polypeptide. Based on this finding, 7 additional recombinant prolamin genes of similar structure were isolated with specific PCR primers. The 8 chimeric clones seem to be derived from 4 individual γ-gliadin and 3 LMW-glutenin sequences. These genes show remarkable diversity in size, gliadin:glutenin ratio, frameshift mutations, and sulphur content. The putative functional characteristics of the chimeric polypeptides and problems related to the origin of the encoding genes are discussed.

Key words: prolamin, chimeric genes, recombination, wheat cDNA library.

Résumé : Un nouveau gène codant pour une protéine de réserve et montrant une structure chimérique évidente a été isolé au sein d'une banque d'ADNc spécifique du grain immature préparée à partir d'une vieille variété hongroise, *Bánkúti 1201*. Ce clone comprend des séquences de type γ-gliadine en 5′ et des séquences de type gluténine LMW en 3′. Un changement de cadre de lecture a également été introduit suite au possible événement de recombinaison. Ainsi, la séquence peptidique de la région C-terminale est complètement changée. À partir de cette observation, 7 gènes additionnels codant pour des prolamines recombinées de structure semblable ont été isolés à l'aide d'amorces PCR spécifiques. Les 8 clones chimériques semblaient dérivés de 4 γ-gliadines et de 3 gluténines LMW différentes. Ces gènes montrent une diversité remarquable en matière de taille, de ratio gliadine/gluténine, de changements de cadre de lecture et de contenu en soufre. Les caractéristiques fonctionnelles potentielles de ces polypeptides chimériques et les problèmes liés à la genèse des gènes correspondants sont discutés.

Mots clés : prolamine, gènes chimériques, recombinaison, banque d'ADNc de blé.

[Traduit par la Rédaction]

Introduction

Storage proteins are accumulated in the endosperm of the angiosperm species during seed development. During germination they are mobilized and provide an amino acid supply for the synthesis of macromolecules in the seedling. Apart from their basic biological function, wheat storage proteins (also called prolamins) have enormous economic significance: they are indispensable for the preparation of bread. The scientific investigation of flour goes back as far as 1745, when Beccari separated it into 2 components: amyloid and glutinin (Beccari 1745). Interestingly, his pioneering work has preserved its relevance through the centuries, since he succeeded in identifying the 2 main constituents of flour: starch and gluten. The wheat prolamin gene family represents immense genetic diversity, consisting of several hundred genes with an unknown number of alleles. Despite their structural similarity, they can be divided into 2 functionally different groups: gliadins and glutenins. The majority of the prolamins contain several cysteine residues, which are able to form disulphide bonds under oxidative conditions. In gliadins, these bonds are generally formed inside the molecules, whereas the glutenins are capable of polymerization via intermolecular bridges. When the flour is mixed with water, the glutenin subunits form giant macropolymers that basically determine the reheological parameters of the dough. Giant macropolymers that basically determine the rheological parameters of the dough. Thus, the breadmaking quality of wheat varieties depends mainly on their glutenin allele combination (Payne 1987). Gliadins can be classified into 3 groups: alpha, gamma, and omega. The latter have no cysteine residues in their polypeptide chain (Müller and Wieser 1995, 1997). Glutenin subunits can also be divided into 2 groups: low-molecular-weight (LMW) and high-molecular-weight (HMW) (Payne 1987). Gliadins prevail in the prolamin family with an estimate of more than 200 genes (Sabelli and

Received 18 August 2004. Accepted 29 March 2005. Published on the NRC Research Press Web site at http://genome.nrc.ca on 19 October 2005.

Corresponding Editor: R. Appels.

I.J. Nagy,[1] I. Takács,[1,2] A. Juhász, and Z. Bedő. Agricultural Research Institute of the Hungarian Academy of Sciences, Brunszvik u. 2., H-2462, Martonvásár, Hungary. **L. Tamás.** Department of Plant Physiology, Eötvös Loránd University, H-1445, Budapest, Hungary, POB 330.

[1]Corresponding author (e–mail: taki@mail.mgki.hu).

[2]These authors are listed alphabetically and contributed equally to this work.

 doi: 10.1139/G05-042

Shewry 1991; Anderson et al. 1997). The exact number of LMW glutenins is not yet known, but it may be approximately 40 (Cassidy et al. 1998). High-molecular-weight glutenin subunits are represented by a total of 6 genes, but only 3–5 of these are translated to protein in bread wheat (Harberd et al. 1986).

Gliadins and LMW glutenins are encoded by genes located on the short arm of group-1 homeologous chromosomes, except for alpha-gliadins (group-6 chromosomes). The HMW genes are located on the long arm of group-1 chromosomes (Payne et al. 1987). Recently, several LMW-gliadin sequences have also been identified, which are linked to group-4 and group-7 chromosomes (Clarke et al. 2002; Anderson et al. 2001*a*). Several other minor loci for prolamins have also been identified on various chromosomes (reviewed by Gianibelli et al. 2001).

The origin of storage protein genes was investigated by Kreis et al. (1985), based on the structure of the C-terminal region in S-rich prolamins (α-gliadins, γ-gliadins, and LMW glutenins). Three sequence motifs, A, B, and C, were identified in this region, separated and flanked by insertions. The limited homology among the 3 domains suggests the triplication of an ancestral gene coding for a protein of about 30 residues. The repetitive domain and the short N-terminal sequence were probably added to the gene later. A comparison of the repeat motifs of different prolamin genes reveals a clear relationship between gliadins and LMW glutenins, whereas HMW glutenins represent a separate lineage. Sequence domains similar to A, B, and C have also been found in a large number of monocot and dicot seed proteins (enzyme inhibitors), and even in some fern spores, extending the "prolamin superfamily" to a wide spectrum of taxa (Shewry and Tatham 1999). The divergence of the storage protein genes was presumably accelerated by the addition of the repetitive domain. Unequal crossing overs and (or) replication slip-mismatching mechanisms introduced a considerable amount of length polymorphism into this region. D'Ovidio et al. (1996) identified a long insertion in the repetitive domain of a high-molecular-weight glutenin subunit (HMW-GS) gene at the *Glu-D1* locus resulting from an unequal crossing-over event.

A large number of sulphur-rich LMW-gliadin sequences have been isolated by Anderson et al. (2001*a*) and Clarke et al. (2002), without any apparent repetitive domain. These molecules contain a very large number of cysteines, and the even number of these residues (14) may result in a compact structure stabilized by 7 intramolecular disulphide bonds. The extended repetitive domain, which has been regarded so far as a general structural unit in wheat storage proteins, is missing from the LMW-gliadins. Only a short proline-glutamine repeat motif can be recognized in the centre of the molecules. It is not clear yet whether this is the vestige of a once existing repetitive region, or whether such a domain has never been acquired. This special structural organization may impart outstanding importance to these genes in the evolution of the wheat prolamins.

Extremely large size differences were described in barley C-hordein genes by Sayanova et al. (1994), caused by variation in the repetitive region. Significant sequence variations and length polymorphisms were also detected in ω-gliadins (Hsia and Anderson 2001), in the ω-secalin genes of rye (Clarke and Appels 1999), and in LMW-GS (Ikeda et al. 2000).

In this paper, we describe several chimeric genes resulting from fusions between γ-gliadin and LMW glutenin sequences. The 8 chimeric clones exhibit remarkable variability in size, structure, and sulphur content. The recombination mechanism that gave rise to these sequences has yet to be elucidated.

Materials and Methods

cDNA cloning

The cDNA library was prepared, amplified, and screened for LMW-GS sequences as previously published by Nagy et al. (2003).

PCR amplification and cloning

Fifty thousand colonies from the cDNA library were grown up on Luria–Bertani–ampicilin plates. The colonies were washed off from the plates with 0.9% NaCl solution, and the plasmid DNA was extracted from the cells, according to the standard protocol. The chimera specific primers were as follows:

GliNF1: 5′-ATACAGGTCGACCCTAGTGG-3′, hybridizing to the N-terminal part of the gamma-gliadin gene *M16064* 58 bp from the start codon.

LMWCR2: 5′-TTATCAGTAGGCACCAACTC-3′ binding to the very end of the *LMW1D1* glutenin gene spanning the stop codons (Cassidy et al. 1998).

The PCR mixture contained 0.4 µmol primers (each)/L, 200 µmol dNTP-s/L, 1.5 µmol $MgCl_2$/L, 20 ng plasmid DNA, 2.5 U *Taq* polymerase (Promega, Madison, Wisconsin, USA) in 15 µL volume. The reactions were run through the following program: 94 °C for 4 min. (1 cycle), 95 °C for 30 s., 55 °C for 25 s., 72 °C for 1 min. (30 cycles), and 72 °C for 20 min.

In PCRs initiated from genomic DNA, the primer concentrations were doubled, the annealing temperature was increased to 60 °C, and the number of cycles was increased to 40. The reaction mixtures contained 100 ng genomic DNA (line B11310) in a total volume of 50 µL. The *Taq* polymerase was replaced by Herculase enhanced DNA polymerase (Stratagene, La Jolla, California, USA) to ensure high fidelity replication.

The amplification products were cloned with a PCR 2.1-TOPO TA Cloning kit (Invitrogen, Carlsbad, California, USA) according to the instructions of the manufacturer.

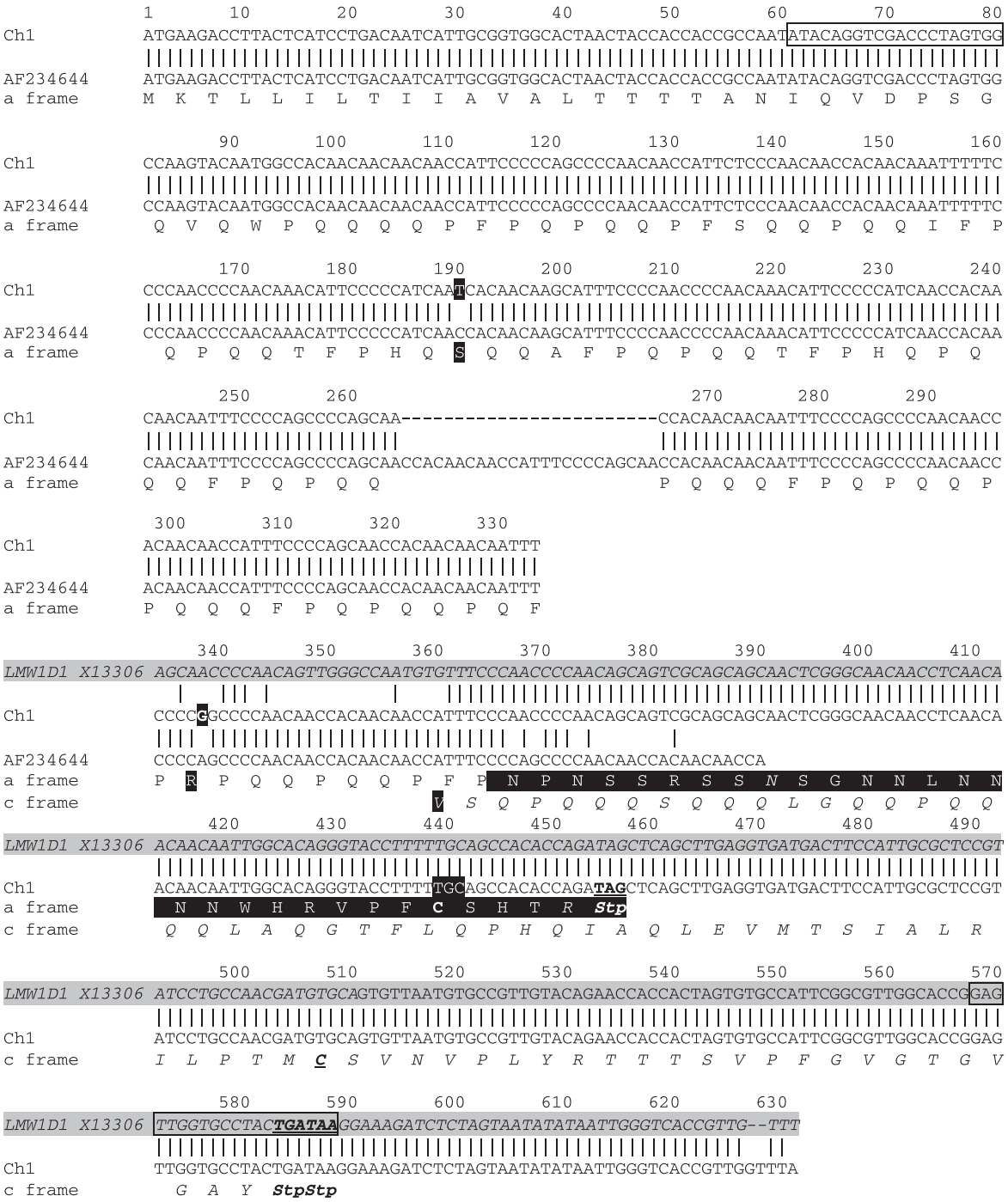
DNA electrophoresis

PCR products were analysed on a 1% agarose gel (Sigma, St. Louis, Missouri, USA) in Tris–borate–EDTA buffer (90 mmol Tris–borate/L, pH 8.0, and 2mmol EDTA/L). Half the volume of each sample was taken and run in the gel, along with a 100-bp ladder (Invitrogen, Carlsbad, California, USA).

Sequence analyses

High quality plasmid DNA was purified from the recombinant clones using a Qiagen Plasmid Kit (Qiagen, Hilden, Germany). The sequencing reactions were carried out using an ABI PRISM® BigDye® Terminator Cycle Sequencing

Fig. 1A. Nucleotide sequence alignment of the *Ch1* clone to the γ-gliadin (*af234644*) and LMW- glutenin (*X13306*) genes. PCR-primer binding sites are boxed. The 24-bp deletion is designated with a dashed line. Gliadin and glutenin sequences overlap at 360th–366th nucleotides. Mutations are indicated with shaded letters. Glutenin sequences are marked with italics in a grey background. Amino acid sequences are also presented. The "a frame" indicates the gliadin-specific and the frameshift modified amino acid sequences. The "c frame" represents the amino acid array of the glutenin fragment.



Ready Reaction Kit 3.1, and the samples were run on an ABI PRISM® 3100 automated sequencer (Applied Biosystems, Foster City, California, USA).

Sequence data were analysed using the GCG software package (Genetics Computer Group, Madison, Wisconsin, USA).

Results and discussion

Identification of the *Ch1* gene

A cDNA library was prepared from the immature kernels (18 DPA) of the isogenic line B11310 isolated from variety *Bánkúti 1201*, to isolate novel storage protein alleles. After

Fig. 1B. Amino acid sequence of the Ch1 chimeric protein. The amino acid sequence of the chimeric protein deduced from the *Ch1* gene is aligned to demonstrate the consensus sequences and repeat motifs characteristic of prolamin proteins. The cysteine residue is underlined and a larger bold letter is used.

```
MKTLLILTII AVALTTTTA                Signal peptide
NIQVDPSGQVQ W                       N-terminal region
PQQQQ                               Start of repetitive region
PFPQ PQQ
PFSQ Q PQQ
IFPQ PQQ
TFPH Q SQQ                          point mutation (P replaced by S)
AFPQ PQQ
TFPH Q PQQ
QFPQ PQQ PQQ                        deleted from Ch1 protein
PF   PQQ PQQ
QFPQ PQQ PQQ
PFPQ Q PQQ
QFPR PQQ PQQ                        point mutation (Q replaced by R)
PFP
NPNSSRSS NSGNNLNNNN WHRVPFCSHTR     C-terminal region
```

consensus sequence for gamma gliadin
PFPQ (Q) $(PQQ)_{1\text{-}2}$

screening the library with a probe specific to the C-terminal part of the *LMW1D1* gene (Ciaffi et al. 1999), a large number of positive clones were isolated and sequenced (Nagy et al. 2003). Among the regular LMW-GS type sequences, a recombinant gene with several unique characteristics was identified. Although it was isolated with an LMW-GS probe, only the C-terminal part of the gene contained an LMW1D1-specific sequence. The N-terminal part of the clone showed almost complete homology to a gamma-gliadin gene, *af234644* (Anderson et al. 2001*a*). A 24-nucleotide sequence was, however, missing near the 3′ end of the gliadin fragment (see Fig. 1). The fusion between the 2 different genes resulted in a relatively short chimeric (recombinant) gene. The open reading frame (ORF) of the glutenin fragment was changed because of the recombination, and the frameshift mutation truncated the coding sequence even shorter (from 193 to 151 aa.). The recombination process, together with the frameshift, gave rise to major rearrangements in the C-terminal domain. The amino acid sequence changed to a completely new array, representing a unique peptide motif without significant homology to any known storage proteins. The only cysteine residue that could potentially be obtained with the LMW-GS C-terminal region was also eliminated from the putative protein, and a new cysteine occurred at the end of the molecule, in position 147. This gene was denominated as *Ch1*.

Isolation of additional chimeric genes

After the identification of the *Ch1* gene, the cDNA library was screened with specific primers to isolate further chimeric clones. Several bands of various sizes were detected, all of which were cloned and sequenced. The sequence analysis identified a total of 7 additional chimeric clones. All these genes resembled the chimeric gene *Ch1*, consisting of different proportions of gamma-gliadin and LMW-GS sequences (see Table 1).

To confirm the presence of these genes in the genome and exclude the impact of any posttranscriptional modification, along with the potential cDNA cloning artefacts, PCRs were also carried out on genomic DNA. Identical genomic copies of 4 previously described cDNA clones, *Ch1*, *Ch5*, *Ch6*, and *Ch8* were amplified in this way (see Fig. 2). *Ch1* was amplified and cloned in 3 independent experiments with the same results.

Despite the low number of chimeric genes isolated by genomic PCR, it is assumed that the other chimeric cDNA clones are also present in the genome. The low efficiency of genomic PCR may be attributed to the very high complexity of the template DNA. Further experimental efforts are needed to solve this problem.

Structural and functional analysis of the chimeric sequences

The most important features of the chimeric genes are summarized in Table 1. These sequences show very significant variation not only in size but also in the relative ratio of gliadin:glutenin sequences and also in the origin of the parental genes. Calculations on the coding sequences indicate that the shortest molecule is *Ch6*, with no more than 82 residues. *Ch2* and *Ch4* are almost 4 times longer than *Ch6* and similar in size to the parental genes (see Fig. 3). As shown in Table 1, the length of the proteins for *Ch2* and *Ch4* are 318 and 307 amino acids, respectively.

The gliadin:glutenin sequence ratio also varies over a wide range from 63%:37% (*Ch7*) to 11%:89% (*Ch4*), but interestingly the breakpoint never goes beyond the 420th bp of the gliadin parent, so the cysteine residues always come from the glutenin ancestor.

Many of the chimeric genes contain deletions of various sizes without modifying the original ORF (see Table 1). The majority of them occur in the gliadin fragments, near to the gliadin/glutenin border. The short sequence repetitions observed upstream of the borders of the deleted fragments suggest in vivo recombination, via either unequal crossovers or reciprocal recombination. The longest deletion can be detected in the glutenin domain of *Ch3*. This rearrangement is of special interest, as *Ch3* seems to have been derived from *Ch4* via a 597-bp deletion, which eliminated the major part of the glutenin sequence. Such descendants cannot be observed among the other chimeric genes cloned in this work. This means that the 8 chimeric clones were presumably generated by 7 independent recombination events. A different type of relationship can be observed between *Ch1* and *Ch2*. Both genes carry an identical, 24-bp deletion starting at the 216th nucleotide of the γ-gliadin fragment. In this case, the deletion probably took place in the gliadin ancestor prior to recombination with the glutenin progenitor.

The repetitive region is the most variable part of the prolamin genes. This structural unit exhibits remarkable size polymorphism in the chimeric clones, too. *Ch2* and *Ch4* have extensive composite (gliadin/glutenin) repetitive domains (115 and 105 aa, respectively), closely resembling the parental genes. *Ch1* and *Ch7* also have long gliadin-derived repetitive fragments of comparable size (91 and 109 aa). All the other molecules contain only rudimentary repetitive regions ranging from 19 (*Ch6*) to 28 residues (*Ch8*). In this respect, they resemble the LMW-gliadin genes isolated by Clarke et al. (2002).

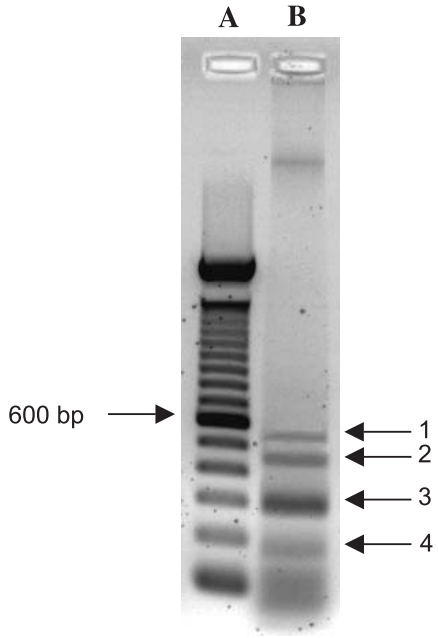
Considering their structural characteristics, the recombinant genes can be grouped in various ways. Genes *Ch1*,

Table 1. Main structural characteristics of the isolated chimeric genes.

Designation	Length of isolated fragment (nt)	Gliadin–Glutenin border (bp)[b]	Gliadin (%)[c]	Deletions (nt)[d]	Length of the protein (aa)[e]	No. of Cys res. in the protein	Homologues	
							Gliadin part	Glutenin part
Ch1[a]	529	302–308 (+)	63	24 (206th)	151 (42)	1	*af234644*	*X13306*
Ch2	902	253 (–)	33	24 (206th)	318	7	*af234644*	*X13306*
Ch3	272	41–43 (–)	31	597 (99th)	108	1	*af234644*	*X13306*
Ch4	869	41–43 (–)	11	—	307	7	*af234644*	*X13306*
Ch5	287	101–113 (–)	50	—	113	1	*M16064*	*X13306*
Ch6	322	92 (+)	40	—	82 (42)	1	*M16064*	*X13306*
Ch7	594	359–361 (+)	65	69 (157th)	176 (40)	1	*af234644*	*ab062863*
Ch8	438	122 (+)	37	93 (46th)	122 (40)	1	*af144104*	*ab062851*

[a]*Ch1* is represented by the PCR-amplified clones, for easier comparison.
[b]+, frameshift present; –, frameshift absent.
[c]The original coding sequences of the parental genes are calculated, regardless of the frameshift mutations.
[d]The position of the deletion is indicated in parentheses.
[e]The number of residues eliminated by the frameshift is shown in parentheses. Clones isolated both from the cDNA library and from the genome are indicated in boldface.

Fig. 2. Agarose gel separation of amplified fragments produced by chimera specific primers using genomic DNA as template. Lane A: 100-bp ladder. Lane B: fragments 1 and 2 represent genes *Ch1* and *Ch8*, respectively. Fragment 3 contains *Ch5* and *Ch6*, whereas fragment 4 represents a product with almost complete primer binding sites, but the rest of the fragment does not show significant homology to any known storage protein genes.



Ch6, *Ch7*, and Ch8 carry frameshift mutations at the junction points. Consequently, their protein products cannot be regarded as gliadin–glutenin hybrids, or even as true prolamins, compared with the rest of the clones with original ORFs. There is a great difference in the number of cysteines too. *Ch2* and *Ch4* contain 7 cysteine residues in the C-terminal glutenin domain, owing to the N-terminal position of the breakpoint and the preservation of the original ORF. The majority of the clones carry a single cysteine in the C-terminal domain, either left in the original position (*Ch3* and *Ch5*) or newly created by a frameshift (*Ch1*, *Ch6*, *Ch7*, and *Ch8*).

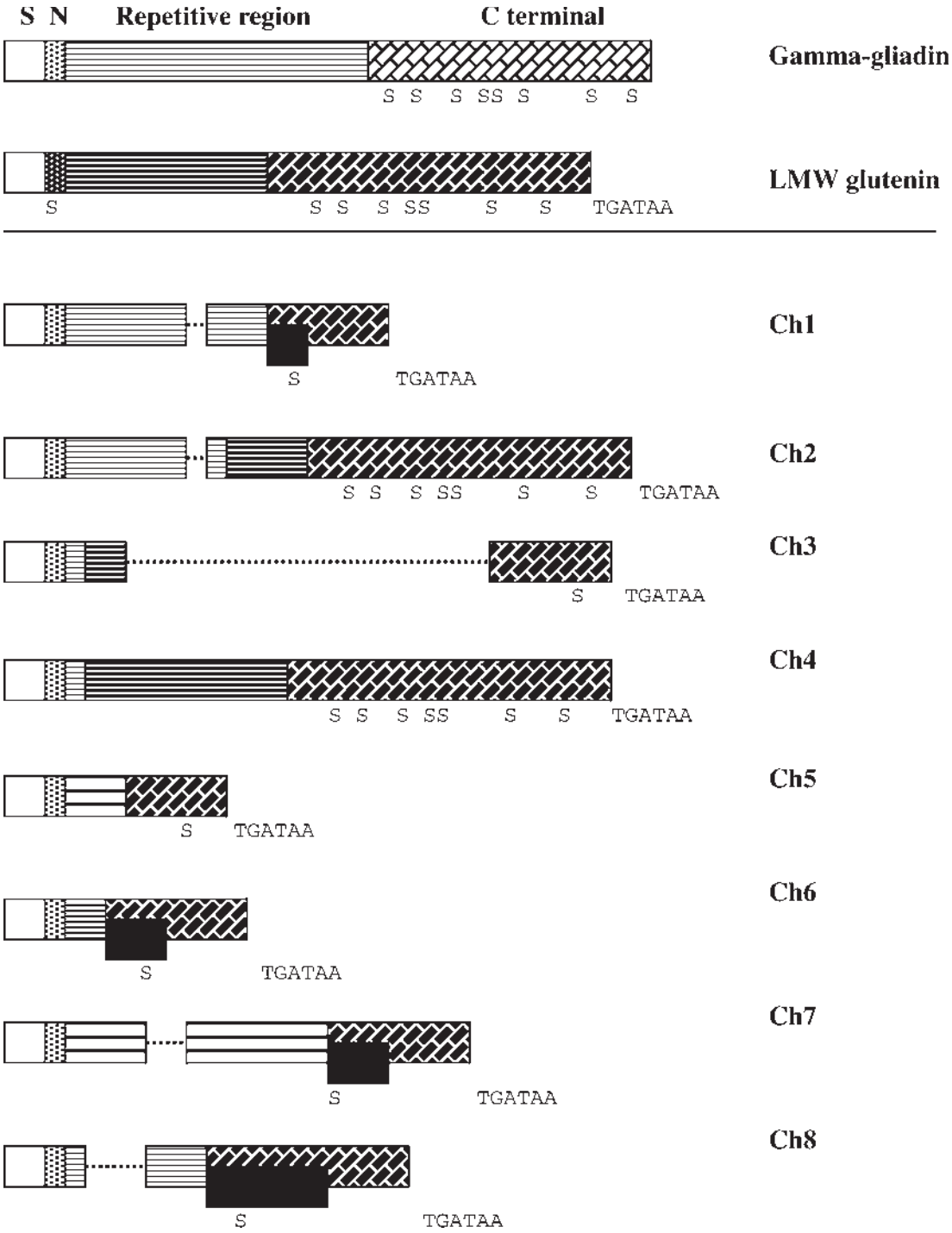
Although we have no evidence that the chimeric prolamin transcripts are translated to polypeptides, this putative family of proteins may form a new class of wheat prolamins. To the best of our knowledge, no storage proteins of such chimeric structure have been identified so far. In the absence of functional data, it is difficult to determine how these proteins contribute to the quality characteristics of wheat flour. It appears from the deduced amino acid sequences that proteins without a frameshift may form a distinct group as putative chain terminators. The cysteine residues present in the LMW-GS subunits have different functional values. Only 2 of them, at distinct positions, are capable of intermolecular disulphide bonding (Köhler et al. 1993). The 2 longest chimeras, with 7 cysteines, preserved 1 of these residues; hence, they are probably capable of incorporating the gluten polymer and acting as chain terminators. This is also the most likely explanation of the functional behavior of the single cysteine molecules in this group. The incorporation of prolamin polypeptides with a single cysteine into glutenin polymers was described by Tamás et al. (1998).

The 4 polypeptides containing frameshifts are quite different from the others. Although they also have single cysteines, this residue is embedded in a completely different amino acid sequence; therefore, its functional behavior is difficult to predict. Nevertheless, some of these proteins may also incorporate the gluten matrix.

Origin of the chimeric genes

The chimeric genes were apparently created by recombination among γ-gliadin and LMW-glutenin sequences, most probably via crossing over between the *Gli1* and *Glu3* loci. Other alternatives, involving recombination among in vivo reverse transcripts of prolamin mRNAs can be ruled out, because their expression is strictly endosperm-specific (Colot et al. 1987). Therefore, the inheritance of such products by subsequent generations is very unlikely (Dubcovsky et al. 1997). At the same time, there are 2 circumstances that question the possibility of regular chiasmatic recombination:

Fig. 3. Schematic representation of 8 recombinant storage protein genes. Gliadin-specific and glutenin-specific sequences are indicated by a white or dark background, respectively. Black rectangles symbolize frame shift mutated regions. The letter "S" marks cysteine residues. "TGATAA" designates the stop codons. Open area, signal peptides; dotted area, N-terminal domains; horizontal lines, repetitive domains; dotted lines, deletions; "bricks", C-terminal domains.



(*i*) the distance between the 2 loci and (*ii*) the limited sequence homology between the parental genes. Although the *Gli1* and *Glu3* loci are closely linked, interlocus recombination has been detected in all 3 genomes in various *Triticum* species (Dubcovsky et al. 1997). In spite of the lack of direct evidence, there are strong indications that the gliadin/LMW-glutenin recombinant genes are located in the D genome. Seven of the 8 chimeric clones contain D genome-specific LMW-glutenin sequences (see Table 1). Only the chromosomal position of the LMW-glutenin parent of the *Ch7* gene is unknown (D'Ovidio and Masci 2004). Recently, a physical map has been made for *Triticum tauschii*, the D-genome progenitor of wheat. This map shows a distance of approximately 100 kb between the 2 complex loci

(Spielmeyer et al. 2000). Chromosomal mispairing over a very long range would be needed to bridge this distance, unless intrachromosomal recombination shortened the interlocus region to a large extent (Peterhans et al. 1990; Gal et al. 1991).

Despite the fact that γ-gliadins and LMW-glutenins are closely related gene families, the nucleotide sequence homology between the parental genes hardly exceeds 60% over long sections. There are only short continuously homologous stretches in these sequences, and very few of the fusion breakpoints fall into these sites. The only significant overlap of 12 bp is found in *Ch5*. At the same time, short overlaps can be detected in the parental genes (see Table 1) adjacent to the fusion sites.

In recent years, the large-scale sequencing and comparative analysis of various Triticeae genomes have revealed major rearrangements in the intergenic space, including deletions, insertions, duplications, and inversions (Wicker et al. 2003). It was also demonstrated that a large deletion resulted in the truncation of the *TdLRR-1* pseudogene. Because of these findings, the possibility that the chimeric prolamin genes were created by large-scale deletions can also be considered. The clear advantage of this model is that it suggests a simple intrachromosomal mechanism for gene fusion, instead of an intrachromosomal recombination.

Chimeric sequences of unrevealed origin have been previously identified in various plant species: a kinase subunit resulting from a gene fusion has been described in maize (Lumbreras et al. 2002), and a chimeric ORF has been identified in the mitochondrial DNA of several species in the Triticeae family (Hedgcoth et al. 2001). Nevertheless, little information on the origin of the prolamin chimeras is provided by these findings. In contrast with storage protein genes, kinase subunit sequences have introns, which may facilitate recombination between different genes, whereas the mitochondria contain a prokaryotic type of genome and recombination system.

It is clear from the above that the exact recombination mechanism responsible for the chimeric prolamin genes has yet to be clarified.

Acknowledgements

This research was funded by the Biotechnology 2000 project of the National Committee for Technological Development, No. 2578. The authors are grateful to Prof. László Patthy, Dr. Renato D'Ovidio, and Dr. Bryan C. Clarke for the valuable discussions and useful advice.

References

Anderson, O.D., Litts, J.C., and Greene, F.C. 1997. The α-gliadin gene family. I. Characterization of ten new wheat α-gliadin genomic clones, evidence for limited sequence conservation of flanking DNA, and southern analysis of the gene family. Theor. Appl. Genet. 95: 50–58.

Anderson, O.D., Hsia, C.C., Adalstein, A.E., Lew, E.J.–L, and Kasarda, D.D. 2001*a*. Identification of several new classes of low-molecular-weight wheat gliadin-related proteins and genes. Theor. Appl. Genet. **103**: 307–315.

Anderson, O.D., Hsia, C.C., and Torres, V. 2001*b*. The wheat γ-gliadin genes: characterization of ten new sequences and further understanding of γ-gliadin gene-family structure. Theor. Appl. Genet. **103**: 323–330.

Beccari, J.B. 1745. De frumento. De bononiensi scientarium et artium. Instituto atque academia commentarii. Vol. 2. Bologna. pp. 122–127. [In Italian.]

Cassidy, B.G., Dvorak, J., and Anderson, O.D. 1998. The wheat low-molecular-weight glutenin genes: characterization of six new genes and progress in understanding gene family structure. Theor. Appl. Genet. **96**: 743–750.

Ciaffi, M., Lee, Y.K., Tamás, L., Gupta, R., Skerritt, J., and Appels, R. 1999. The low-molecular-weight glutenin subunit proteins of primitive wheats. III. The genes from the D-genome species. Theor. Appl. Genet. **98**: 135–148.

Clarke, B.C., and Appels, R. 1999. Sequence variation at the *Sec-1* locus of rye, *Secale cereale* (Poaceae). Plant Syst. Evol. **214**: 1–14.

Clarke, B.C., Phongkham, T., Gianibelli, M.C., Beasley, H., and Bekes, F. 2002. The characterization and mapping of a family of LMW-gliadin genes: effects on dough properties and bread volume. Theor. Appl. Genet. **106**(4): 629–635.

Colot, V., Robert, L.S., Kavanagh, T.A., Bevan, M.W., and Thompson, R.D. 1987. Localization of sequences in wheat endosperm protein genes which confer tissue-specific expression in tobacco. EMBO (Eur. Mol. Biol. Organ) J. **6**: 3559–3564.

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

D'Ovidio, R., Lafiandra, D., and Porceddu, E. 1996. Identification and molecular characterization of a large insertion within the repetitive domain of a high-molecular-weight glutenin subunit gene from hexaploid wheat. Theor. Appl. Genet. **93**: 1048–1053.

Dubcovsky, J., Echaide, M., Giancola, S., Rousset, M., Luo, M.C., Joppa, L.R., et al. 1997. Seed-storage-protein loci in RFLP maps of diploid, tetraploid, and hexaploid wheat. Theor. Appl. Genet. **95**: 1169–1180.

Gal, S., Pisan, B., Hohn, T., Grimsley, N., and Hohn, B. 1991. Genomic homologous recombination in planta. EMBO (Eur. Mol. Biol. Organ) J. **10**: 1571–1578.

Gianibelli, M.C., Larroque, O.R., MacRitchie, F., and Wrigley, C.W. 2001 Biochemical, genetic, and molecular characterization of wheat glutenin and its component subunits. Cereal Chem. **78**: 635–646.

Harberd, N.P., Bartels, D., and Thompson, R.D. 1986. DNA restriction-fragment variation in the gene family encoding high-molecular-weight (HMW) glutenin subunits of wheat. Biochem. Genet. **24**: 579–595.

Hedgcoth, C., El–Shehawi, A.M., Wei, P., Clarkson, M., and Tamalis, D. 2002. A chimeric open reading frame associated with cytoplasmic male sterility in alloplasmic wheat with *Triticum timopheevii* mitochondria is present in several *Triticum* and *Aegilops* species, barley, and rye. Curr. Genet. **41**: 357–365.

Hsia, C.C., and Anderson, O.D. 2001. Isolation and characterization of wheat ω-gliadin genes. Theor. Appl. Genet. **103**(1): 37–44.

Ikeda, T.M., Nagamine, T., Fukuoka, H., and Yano, H. 2002. Identification of new low-molecular-weight glutenin subunit genes in wheat. Theor. Appl. Genet. **104**: 680–687.

Köhler, P., Belitz, H.D., and Wieser, H. 1993. Disulphide bonds in wheat gluten: further cysteine peptides from high-molecular-weight (HMW) and low-molecular-weight (LMW) subunits of glutenin and from γ-gliadins. Z. Lebensm-Unters-Forsch. **196**: 239–247.

Kreis, M., Forde, B.G., Rahman, S., Miflin, B.J., and Shewry, P.R. 1985. Molecular evolution of the seed storage proteins of barley, rye and wheat. J. Mol. Biol. **183**: 499–502.

Lumbreras, V., Alba, M.M., Kleinow, T., Koncz, C., and Pages, M.

2001. Domain fusion between SNF1-related kinase subunits during plant evolution. EMBO. Rep. **2**(1): 55–60.

Müller, S., and Wieser, H. 1995. Disulphide bonds of α-type gliadins. J. Cereal Sci. **22**: 21–27.

Müller, S., and Wieser, H. 1997. The location of disulphide bonds in monomeric γ-gliadins. J. Cereal Sci. **26**: 169–176.

Nagy, I.J., Takács, I., Tamas, L., and Bedő, Z. 2003. Molecular cloning and characterization of low molecular-weight glutenin sequences from an old Hungarian wheat variety, *Bánkúti 1201*. Cereal Res. Comm. **31**(1–2): 25–31.

Payne, P.I. 1987. Genetics of wheat storage proteins and the effect of allelic variation on bread making quality. Annu. Rev. Plant Physiol. **38**: 141–153.

Peterhans, A., Schlüpman, H., Basse, C., and Paszkovsky, J. 1990. Intrachromosomal recombination in plants. EMBO. (Eur. Mol. Biol. Organ) J. **9**: 3437–3445.

Sabelli, P., and Shewry, P.M. 1991. Characterization and organisation of gene families at the *Gli-1* loci of bread and durum wheats by restriction fragment analysis. Theor. Appl. Genet. **83**: 209–216.

Sayanova, O.V., Mekhedov, S.L., Burkitbaev, E.M., Shchipkova, N.I., and Ananev, E.V. 1994. Structural organization of barley C-hordein genes. Genetika, **30**(6): 749–755.

Shewry, P.M., and Tatham, A.S. 1999. The characteristics, structures and evolutionary relationships of prolamins. *In* Seed proteins. *Edited by* P.R. Shewry and R Casey. Kluwer Academic Publishers, Dordrecht, Netherlands. pp. 11–33.

Spielmeyer, W., Moullet, O., Laroche, A., and Lagudah, E.S. 2000. Highly recombinogenic regions at seed storage protein loci on chromosome 1DS of *Aegilops tauschii*, the D-genome donor of wheat. Genetics, **155**: 361–367.

Tamás, L., Bekes, F., Greenfield, J., Tatham, A.S., Gras, P.W., Shewry, P.R., et al. 1998. Heterologous expression and dough mixing studies of wild-type and mutant c hordeins. J. Cereal Sci. **27**:15–22.

Wicker, T., Yahiaoui, N., Guyot, R., Schlagenhauf, E., Liu, Z.-D., Dubcovsky, J., et al. 2003 rapid genome divergence at orthologous low molecular weight glutenin loci of the A and A^m genomes of wheat. Plant Cell. **15**: 1186–1197.