Theor Appl Genet (2007) 115:77–86
DOI 10.1007/s00122-007-0542-0

ORIGINAL PAPER

D hordeins of *Hordeum chilense*: a novel source of variation for improvement of wheat

Fernando Pistón · Peter R. Shewry · Francisco Barro

Received: 5 February 2007 / Accepted: 22 March 2007 / Published online: 26 April 2007

Abstract The high molecular weight subunits of wheat (*Triticum aestivum* L.) glutenin (HMW-GS) are important in determining the bread-making quality of flour and dough. There is therefore interest in transferring orthologous HMW-GS present in other grass species into wheat by wide crossing in order to extend the range of end use properties. In this work, we have isolated and characterized two genes encoding D hordeins from *Hordeum chilense* (Roem. et Schult.) lines H1 and H7, representing two ecotypes. The fragments were 4,305 bp for line H1 and 4,227 for line H7 and contained the promoter, coding and terminator regions. Both sequences differ in the presence of single base changes (SNPs) and insertions/deletions in the open reading frame (ORF). The encoded proteins comprise 870 and 896 amino acids for lines H1 and H7, respectively. The primary structure is similar to those of D hordeins of cultivated barley (*H. vulgare* L.) and HMW-GS from wheat. However, the D hordeins from *H. chilense* are significantly larger than those from cultivated barley due to the presence of longer repetitive regions. The *H. chilense* D hordeins also differ from those of cultivated barley in the distribution of the cysteine residues: whereas the D hordeins of cultivated barley contain ten cysteines with four in the repetitive domain, only nine are present in the *H. chilense* proteins with two in the repetitive domain. As in the HMW-GS, the central part of the D hordein proteins comprises repeated sequences based on short peptide motifs. The repetitive domain is divided in three regions named as R1 (N-terminal repeats), R2 (central degenerate repeats) and R3 (C-terminal repeats). Hexapeptide motifs are present throughout the repetitive domains of D hordeins with a consensus motif of PFQGQQ in R1 and R2 and PHQGQQ in R3. In addition, the tetrapeptide motif TTVS, which is characteristic of D hordeins of cultivated barley is present in the repetitive domain close to the protein C-terminus.

Communicated by B. Friebe.

F. Pistón · F. Barro (✉)
Instituto de Agricultura Sostenible, CSIC,
Apdo. 4084, 14080 Córdoba, Spain
e-mail: ge1balof@uco.es

P. R. Shewry
Rothamsted Research, Harpenden,
Hertfordshire AL5 2JQ, UK

Introduction

Prolamins are the major storage proteins synthesized in the seeds of cereals and other grasses (Lawrence and Shepherd 1981). Although prolamins are highly polymorphic both within and between species, those present in wheat (*Triticum aestivum* L.), barley (*Hordeum vulgare* L.) and other members of the Triticeae can be divided into three families based on their amino acid sequences (Miflin et al. 1983). The S-rich prolamins comprise the α gliadins, γ gliadins and low molecular weight glutenin subunits (LMW-GS) of wheat and the B and γ hordeins of barley; the S-poor prolamins comprise the ω gliadins and C hordeins of wheat and barley, respectively, while the high molecular weight prolamins comprise the high molecular weight glutenin subunits (HMW-GS) of wheat and D hordein of barley.

The HMW-GS of wheat have been studied in detail because of their importance in determining the bread-making quality of dough (Payne et al. 1987; Shewry et al. 1992). Amino acid sequence comparisons show that all HMW-GS share a conserved structure comprising a signal peptide (which is removed co-translationally), and short N-terminal and C-terminal domains flanking a central

repetitive domain (Shewry et al. 1995). The repetitive domain comprises repeats based on two or more short peptide sequences, with the degree of reiteration of the repeats largely accounting for the differences in length between the HMW-GS (Shewry et al. 1995). A range of studies including analyses of cultivars and crosses have shown that HMW subunits differ in their impact on bread-making performance with subunits 1Ax1 and 1Ax2* and subunits 1Dx5 and 1Dy10 (encoded by chromosomes 1A and 1D, respectively) in particular being associated with high dough strength and good bread-making quality (reviewed by Payne 1987; Shewry et al. 2003). These HMW-GS have therefore been used to alter wheat grain quality by genetic transformation (Altpeter et al. 1996; Barro et al. 1997; Blechl and Anderson 1996).

In contrast to our detailed knowledge of the HMW subunits of wheat, we know little about the orthologous subunits present in other grass species, which could be incorporated into wheat by wide crossing. For example, the D hordeins in barley are encoded by the *Hor3* locus on chromosome 5 (1H) have similar structures to the HMW-GS of wheat but have subtle differences, for example in the number and distribution of cysteine residues (Halford et al. 1992; Gu et al. 2003). These proteins may therefore confer useful properties if transferred to wheat.

Cultivated barley is not readily crossed with wheat and hence gene transfer between these species is difficult to achieve. In contrast, the related wild species, *H. chilense* (Roem. et Schult.) can be hybridized with both tetraploid durum wheat (*T. turgidum* spp. *durum* (Desf.) Husn) and hexaploid bread wheat (*T. aestivum* L.) and such hybrids have been used to develop hexaploid and octaploid forms of a novel cereal, Tritordeum (Martin and Sanchez-Monge 1982; Alvarez et al. 1992; Martin et al. 1999). The hexaploid Tritordeum forms derived from durum wheat are of particular interest as they combine high-nitrogen uptake and water use efficiency with improved bread-making quality, indicating that the H^{ch} genome can substitute to some extent for the D genome of bread wheat.

Accessions of *H. chilense* show considerable variation in their patterns of hordein storage proteins, including D hordein (Alvarez et al. 2001), and may therefore have different impacts on bread-making and malting quality when used to make Tritordeums (Alvarez et al. 1992; Martin et al. 1999). In particular, variation in the D hordein subunits contributed by the *H. chilense* parent has been shown to affect the dough strength in addition to the HMW subunits contributed by the durum wheat parent (Alvarez et al. 1999).

It is clear, therefore, that the D hordeins of *H. chilense* represent a novel type of functional protein, which can be transferred to bread or durum wheats by conventional crossing rather than by using GM technology, which would be required to transfer the orthologous D hordeins of cultivated barley. To explore this possibility we have isolated and characterized two genes encoding D hordeins from *H. chilense* and compared the sequences of the encoded proteins with those of D hordeins from *H. vulgare* and HWM-GS from wheat. This shows that they have differences in structure compared to D hordeins of cultivated barley, which could be exploited by transfer to conventional bread-making wheats as well as by expression in Tritordeums.

Materials and Methods

Plant material

The diploid accession lines of *H. chilense* (Roem. et Schulz) H1, H8, H7, H10, H11, H16, H47, H55, H56, H57, H74, H75, H204, H205, H210, H220, H252, H293, H297 and H303 were from the germplasm collection of the Instituto de Agricultura Sostenible, CSIC, Cordoba, Spain. Plants were grown in a greenhouse with supplementary lights providing a day/night regime of 12/12 h.

Amplification of short fragments of D hordeins

Genomic DNA from the *H. chilense* lines was isolated using a CTAB method (Stacey and Isaac 1994). Primers HoD*3 and HoD*4 were used to amplify short fragments encoding *H. chilense* D hordeins (Table 1). These primers were designed on the basis of the published sequence of the mRNA sequence of D hordein from *H. vulgare* (GenBank accession number D82941). All primer used in this work were synthesized by Sigma-Genosys, Pampisford, UK. The PCR conditions to amplify the short fragments of *H. chilense* D hordein were as in Piston et al. (2005). The resulting products were gel-purified, cloned in pGEMT-Easy vector (Promega, Madison, WI, USA), and introduced into competent *Escherichia coli* (DH5α) cells by cell transformation. The plasmid was isolated and purified using QIAprep® Spin Miniprep Kit from Qiagen, Valencia, CA, USA, and used as template for sequencing.

Amplification of cDNA ends (RACE PCR)

RNA isolation was carried out as described previously (Piston et al. 2005). Total RNA from the endosperm of *H. chilense* line H7 10 days after flowering was used to synthesize the anchored cDNA and the 5′and 3′ ends as described by the SMART™ RACE cDNA Amplification Kit from Clontech, Palo Alto, CA, USA. The specific primers HoD*5, HoD*6, HoD*7 and HoD*8 (Table 1) were designed using the short fragments previously sequenced to generate overlapping 5′and 3′ RACE products. The 5′and 3′ end products were size-fractionated by gel electrophoresis,

Table 1 Description of PCR primers for cloning and molecular characterization of the D hordeins of *H. chilense*

Primer	Description	Sequence (5′ → 3′)
HoD*3	Forward for D hordein	TGCAAGGGCAACAACAAGGACAAGGG
HoD*4	Reverse for D hordein	CCACTAGGCTGGTATCCTTGTTGTCCTTGTCCTG
HoD*5	External primer for 3′ RACE	CAAGAAGGATACTACCCAAGTGCAACTTTTCC
HoD*6	External primer for 5′ RACE	ACTGTTGCGGAGAAGTTGCACTTGGGTAG
HoD*7	Nested primer for 3′ RACE	CAACAGCCAGGACAAGGGCAAACAGG
HoD*8	Nested primer for 5′ RACE	AGTATCCTGTTTGCCCTTGTCCTGGCTG
HoD*13	Nested primer for 3′ genome walking	TGTATACAATAAGCGTGACGTGTGCTC
HoD*14	Nested primer for 5′ genome walking	GGTAGAAGGATCCTCCCTTGGATG
HoD*15	External primer for 3′ genome walking	TCCAACGAGTTGTACTATGACAGC
HoD*16	External primer for 5′ genome walking	ATAGGTTTCCCACCATCTTCCTTG
potHorD*1	Forward for full-length gene	CTCATTCGGGAAGCGGGAAAA
potHorD*2	Reverse for full-length gene	CCTCTTCGTCCTGCCGCAGAC
potHoD:1402-F	Forward for D hordein repetitive domain	GCAGCAAGGGTCCTATCCAAGC
potHoD:2688-R	Reverse for D hordein repetitive domain	CGGAGAAGTTGCAGTTGGGTAGTAGT

bands eluted using the Gel Extraction Kit from Qiagen, and cloned and sequenced as describe above.

Isolation of 5′and 3′ untranslated regions by genome walking

Genomic DNA was isolated from young leaves by the Plant DNAzol® reagent from Invitrogen, Carlsbad, CA, USA, following the manufacturer's instructions. The isolation of the promoter and the terminator region of the D hordeins was carried out from DNA of *H. chilense* lines H1 and H7 using the Universal Genome Walker Kit (Clontech) according to manufacturer's instructions and using the gene-specific nested primer pairs HoD*16/HoD*14 and HoD*15/HoD*13, respectively (Table 1). The primers were designed on the basis of the 5′ and 3′ RACE-PCR fragments obtained above.

The nested PCR fragments obtained from the Genome Walker 'library' were cloned into pGEMT-Easy vector (Promega) and used to transform the *E. coli* DH5α competent cells. The plasmids were isolated using QIAprep Spin Miniprep Kit from Qiagen, and sequenced.

Full-length gene isolation and repetitive region amplification

The sequences obtained from the 5′ and 3′ untranslated region were used to design the primers potHorD*1 and potHorD*2 (Table 1) to amplify the full D hordein gene from genomic DNA from *H. chilense* lines H1 and H7. The PCR amplification was carried out using the Advantage 2 Polymerase Mix (Clontech).

The fragments amplified were cloned and sequenced as described above. Three clones for each line were sequenced in both forward and reverse orientation to ensure the sequence. To determine the complete DNA sequence of each D hordein the full-length nucleotide sequence was assembled with the sequence data from the genome walking.

The full-length D hordeins sequences from lines H1 and H7 were used as template to design the primer pair potHoD:1402-F/potHoD:2688-R (Table 1) to amplify part of the repetitive domain in 15 lines of *H. chilense*. Accessions numbers of the full-length genes are EF417988 and EF417989 for H1 and H7, respectively.

Bioinformatic analyses

All the bioinformatics design and analyses were performed using the Vector NTI 9.1.0 suite (Invitrogen). Sequence identity searches were performed at the National Center for Biotechnology Information (NCBI) http://www.ncbi.nlm.nih.gov. Sequences were aligned using the CLUSTALX (Version 1.81) (Thompson et al. 1997). ProML in Felstein's PHYLIP package (Felsenstein 1999) was used to generate a maximum Likelihood tree based on the Jone–Taylor–Thornton algorithm (Jones et al. 1992).

Results

Isolation and nucleotide sequence of full-length D hordein genes

Previous analyses of 38 lines of *H. chilense* showed ten allelic forms of D hordein with one form (subunit 1^{Hch}) being present in 11 of the lines (Alvarez et al. 2001). The gene encoding this subunit was isolated from *H. chilense*

line H1 and that encoding the less widespread subunit 3^{Hch} (present in five lines) from line H7, using the primers HoD*3 and HoD*4 (Table 1).

The fragments from H1 and H7 were confirmed as encoding D hordein by BLASTn search and were used to design the primer pairs HoD*5/HoD*7 and HoD*6/HoD*8 (Table 1) to amplify the 5′ and 3′ ends from cDNA synthesized from line H7 by RACE-PCR. Subsequently, the primer pairs HoD*13/HoD*15 and HoD*14/HoD*16 (Table 1) were designed, based on the 5′ and 3′ cDNA ends, to amplify the 5′ and 3′ untranslated regions by genome walking from lines H1 and H7. Finally, the primer pair potHorD*1/potHorD*2 designed based on the 5′ and 3′ untranslated sequences, was used to amplify the full-length D hordein genes. Enzymes with proofreading activity were used for the amplifications to minimize PCR errors. The full-length sequences were confirmed as encoding D hordeins by BLASTn searches. The fragments were 4,305 bp for line H1 and 4,227 for line H7 and contained the promoter, coding and terminator regions (Fig. 1a). The promoter regions contained all the regulatory elements and motifs that are characteristic of HMW-GS genes and are considered to confer endosperm-specific expression (Halford and Shewry 2007).

The open reading frames (ORFs) for H1 and H7 comprise 2,691 and 2,613 bp, respectively, this difference being due to the presence of insertions/deletions between both sequences. For example, the line H1 sequence has two insertions with respect to the line H7 sequence of 108 and 18 bp, respectively, whereas the line H7 sequence has an insertion, with respect to the H1 sequence, of 48 bp (Fig. 1a). These deletions/insertions occur only within the coding regions of the genes and do not interrupt the ORFs. Comparison of the sequences of H1 and H7 also shows the presence of single base changes (SNPs). In total, 54 SNPs were present and were more frequent in the coding regions (42 SNPs) than in the promoter (four SNPs) and terminator (eight SNPs) regions. Primers potHoD:1402-F/potHoD:2688-R (Table 1), flanking the positions of the insertions/deletions, were designed to test if differences in the length of the ORF also occurred in other lines of *H. chilense* (Fig. 1a, b). Of the 13 additional lines that were analysed, five (H8, H57, H75, H204 and H16) showed the same 1,312 bp band as H1, and eight (H10, H11, H74, H303, H47, H55, H56 and H205) the same 1,234 bp as H7 (Fig. 1b).

Amino acid sequence analysis

The D hordein ORFs encode proteins of 870 and 896 amino acid residues for lines H1 and H7, respectively. Analysis of the amino acid sequences showed that D hordeins from *H. chilense* have a primary structure similar to those of D hordeins of cultivated barley and HMW-GS from wheat (Fig. 2). This conserved structure consists of a 21 amino acid signal peptide and short N-terminal and C-terminal domains flanking a central repetitive domain (Table 2, Fig. 3). It should be noted that D hordeins from *H. chilense* are significantly larger than those from cultivated barley due to the presence of longer repetitive regions. The *H. chilense* D hordeins also resemble those from cultivated barley in having more cysteine residues than the HMW-GS but differ in their distribution: whereas the D hordeins of cultivated barley contain ten cysteines with four being present in the repetitive domain, only nine are present in the *H. chilense* proteins with two in the repetitive domain (Table 2).

The non-repetitive N- and C-terminal domains of the D hordeins from *H. chilense* and cultivated barley show high-

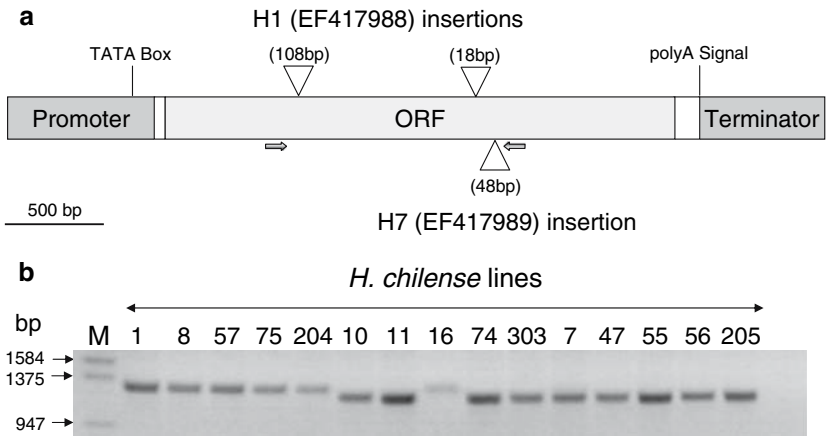

Fig. 1 **a** Nucleotide structure of the D hordeins from *H. chilense*, lines H1 and H7, accession numbers EF417988 and EF417989, respectively. The insertions present in H1 or H7 and not present in the *other line* are indicated. The relative positions of the primers that amplify the repetitive region containing all insertion are indicated by the *arrows*. **b**. Agarose gel of the fragment amplified using the primer arrowed in *panel* **a** and indicated in Table 1. Each number corresponds to an accession of *H. chilense*. *M* molecular marker

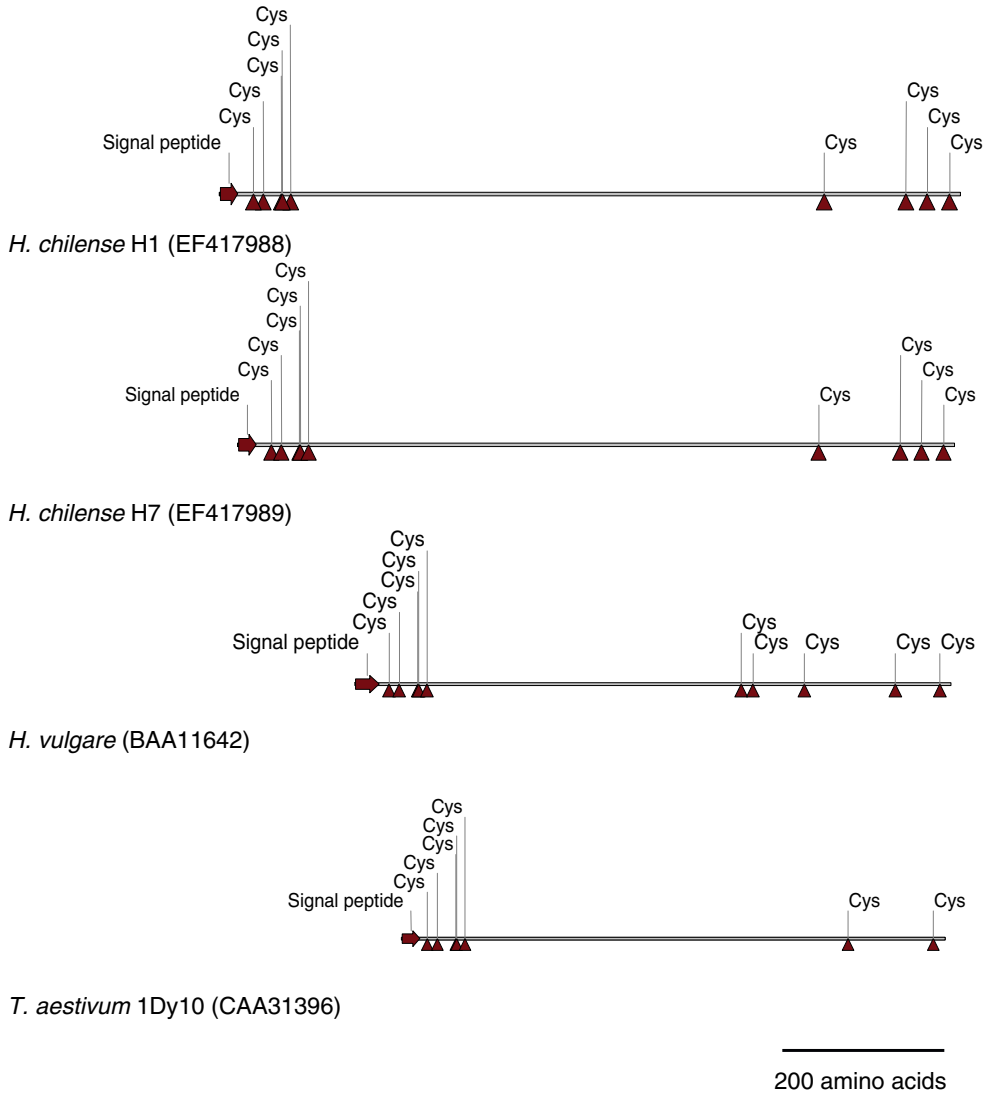

Fig. 2 Model of D hordeins from *H. chilense* lines H1 and H7 and comparison with D hordein from *H. vulgare* and the HMW-GS 1Dy10 from wheat. Signal peptide and the position of the *cysteine* residues are indicated

sequence identity, as demonstrated by the alignment shown in Fig. 3. The alignment also shows that the D hordeins from *H. chilense* are similar to the *x*- and *y*-type HMW-GS, with the sequence of the N-terminal domain being more similar to those of *y*-type HMW-GS and the C-terminal domain to those of *x*-type HMW-GS, respectively. When the combined sequences of these domains are used to construct a phylogenetic tree using the neighbour-joining method (Fig. 5) the D hordeins of *H. chilense* and *H. vulgare* group together and closer to the *y*-type HMW subunit: this reflects the fact that the 110 residue N-terminal domain has a greater effect on the analysis than the 42 residue C-terminal domain.

As in the wheat HMW-GS, the central part of the D hordein proteins comprises repeated sequences based on short peptide motifs. Gu et al. (2003) defined two blocks of repeats separated by a non-repetitive region containing several cysteine residues in D hordeins of barley. However, manual alignment of the sequences (Fig. 4) shows that this region is also repetitive but that the structure is degenerate with frequent mutations and insertions/deletions. We will therefore refer to these three regions of the repetitive domain as R1 (N-terminal repeats), R2 (central degenerate repeats) and R3 (C-terminal repeats) (Fig. 5).

Hexapeptide motifs are present throughout the repetitive domains of D hordeins (54 in H1, 56 in H7, 48 and 43 in *H. vulgare*) (Table 3). These have a consensus motif of PFQGQQ (which is identical to that present in the HMW-GS) in R1 and R2 and PHQGQQ in R3. In the R1 and R2 regions these hexapeptides occur tandemly or interspersed with 11 residue repeats (consensus GYYPSATSPQQ) which are related to the nonapeptides present in the HMW-GS (consensus GYYPTSP or LQQ). A total of 31 residue repeats of these are present in H1, 30 in H7 but only 19 in the *H. vulgare* proteins (Table 3). Finally, a small number of

Table 2 Characteristics of the primary structure of the D hordeins from *H. chilense* lines H1 and H7 in comparison with those D hordeins from *H. vulgare* and the HMW-GS from wheat (*T. aestivum*). The D hordeins from *H. vulgare* are indicated by their accession number

Subunit	Number of amino acid residues				Number of cysteine residues			
	N-terminal domain	C-terminal domain	Repetitive domain	Total	N-terminal domain	C-terminal domain	Repetitive domain	Total
T. aestivum								
1Ax1	86	42	681	809	3	1	0	4
1Ax2	86	42	666	794	3	1	0	4
1Bx7	81	42	645	768	3	1	0	4
1By9	104	42	538	684	5	1	1	7
1Dx2	88	42	687	817	3	1	0	4
1Dx5	89	42	687	818	3	1	1	5
1Dy10	104	42	481	627	5	1	1	7
1Dy12	104	42	493	639	5	1	1	7
H. vulgare								
AAP31051	110	42	584	736	5	1	4	10
BAA11642	110	42	534	686	5	1	4	10
H. chilense								
H1	110	42	723	875	5	2	2	9
H7	110	42	697	849	5	2	2	9

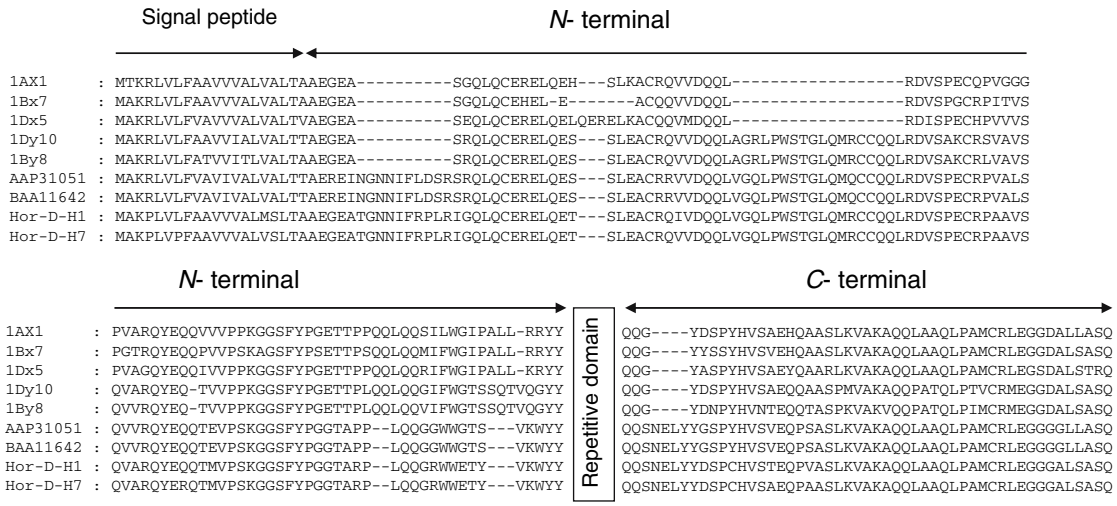

Fig. 3 Amino acid sequence alignment of the N-terminal and C-terminal regions of representative HMW-GS *x* and *y* types from wheat, D hordeins from *H. vulgare* and D hordeins from *H. chilense*. The *block* between the N- and C-terminal regions represents the repetitive domain

tripeptides (five to seven) are present either between blocks of hexapeptides or between hexapeptides and 11 residue peptides. The hexapeptides in the R3 region are mainly interspersed but in this case with a novel tetrapeptide motif (TTVS) which has no counterpart in the HMW-GS (Fig. 4). It is notable that only four copies of this motif are present in H1 and H7 of which only two matches the consensus. In contrast, this region is expanded in the *H. vulgare* AAP31051 protein to comprise 16 copies, 12 of which are adjacent to consensus hexapeptides motifs (PHQGQQ). In the second *H. vulgare* protein (BAA11642) the R3 region is truncated and only 11 hexapeptide + tetrapeptide repeats are present. The numbers of repeat motifs in the four D hordein proteins are summarized in Table 3.

Discussion

The HMW-GS of wheat are major determinants of dough strength and hence processing quality in bread and durum

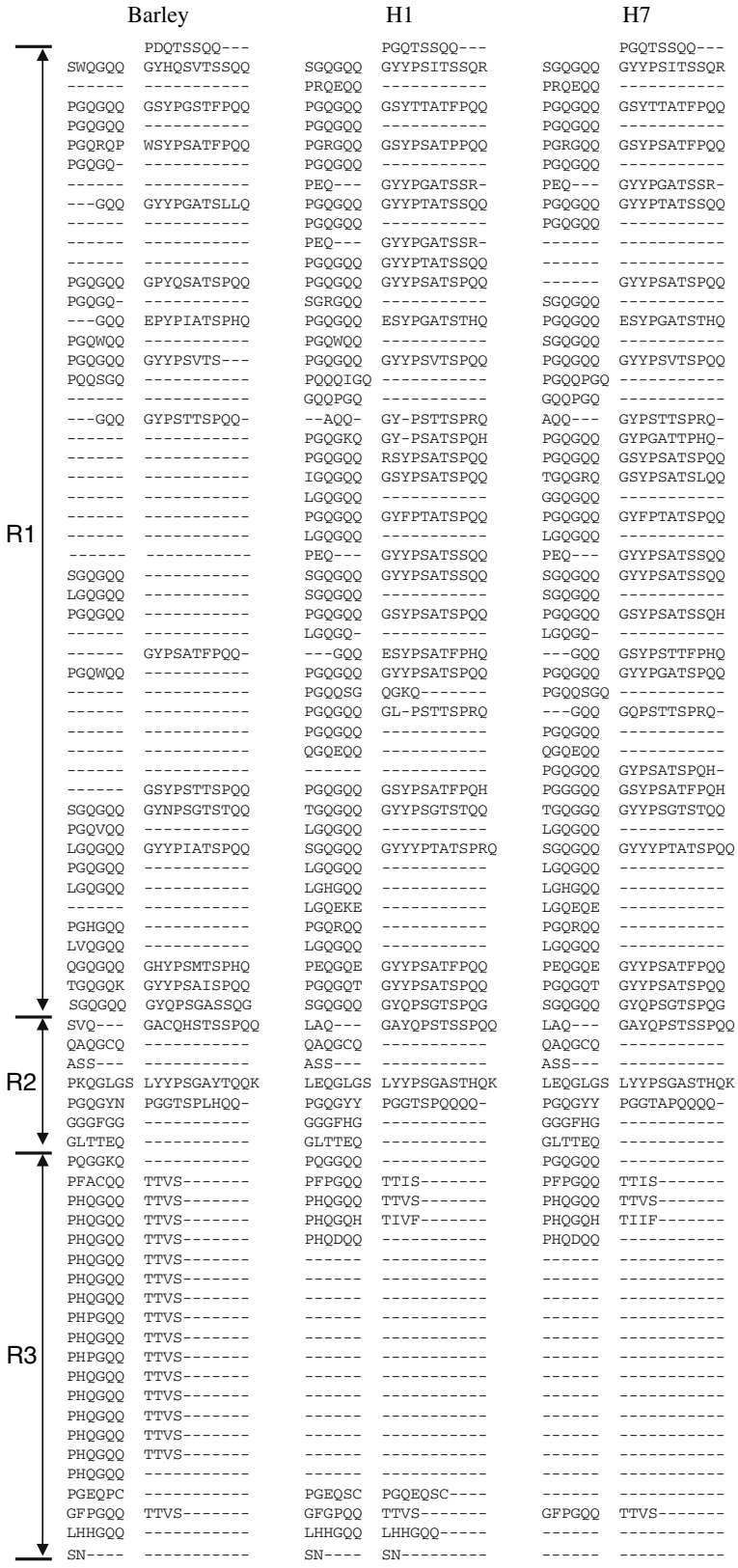

Fig. 4 Manual alignment of the repetitive domains of D hordeins from *H. chilense* (H1, H7) and *H. vulgare* (AAP31051). The repetitive domain is divided in three regions named as R1 (N-terminal repeats), R2 (central degenerate repeats) and R3 (C-terminal repeats)

wheats. They occur in a range of allelic forms and specific alleles are associated with strong or weak dough properties (reviewed by Payne 1987; Shewry et al. 2003). However, despite this polymorphism there is limited variation in their detailed structures, the most striking being differences in the length of their repetitive domains. A number of studies

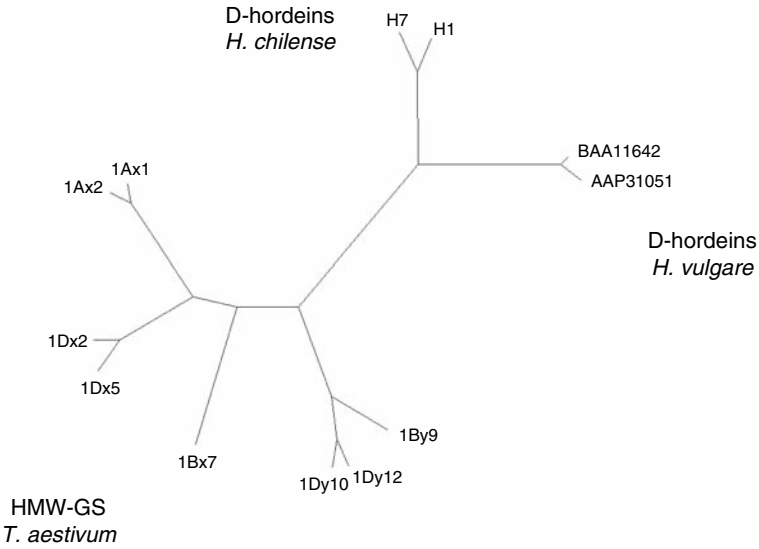

Fig. 5 Maximum likelihood tree obtained with the alignment of the full amino acid sequences show in the figure. The amino acid sequences were compared by Clustal analysis with minor modifications. The D hordeins from *H. vulgare* are indicated by their corresponding accession numbers

Table 3 Repeat peptide compositions of the R1, R2 and R3 regions of the repetitive domains of D hordeins from *H. chilense* (H1, H7) and *H. vulgare* (AAP31051, BAA11642)

	Number of repeated peptides											
	H1			H7			AAP31051			BAA11642		
	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2	R3
Tripeptide	5	2	0	5	2	0	3	2	0	3	2	0
Tetra	0	0	4	0	0	4	0	0	16	0	0	11
Hexa	43	5	8	41	5	8	24	5	19	24	5	14
11 Residue	28	3	0	27	3	0	16	3	0	16	3	0

have therefore been directed towards the identification of new HMW-GS alleles in exotic germplasm and related wild species of *Triticum* and *Aegilops*, as reviewed by Shewry et al. (2006). Nevertheless, the extent of variation in HMW-GS structure in these lines is still limited, with the most successful approach being to increase the number of expressed genes by introducing *Glu-A1* loci expressing two subunits compared to the single subunit expressed by this locus in bread wheat cultivars (Ciaffi et al. 1995; Rogers et al. 1997). The D hordeins of cultivated barley and *H. chilense* therefore provide a wider range of variation which, in the latter species, can be readily incorporated into wheat by wide crossing.

The two *H. chilense* D hordein genes reported here have highly conserved structures, showing 94% sequence identity and differing only in SNPs and small insertions/deletions. These differences are concentrated in the coding regions of the genes, with 1 SNP per 68 bp, with the 5′and 3′ regions being highly conserved. In fact, the major differences between the two genes were insertions and deletions within the sequences encoding the repetitive domains of the proteins, which is consistent with the proposed mechanism of evolution of these domains by mispairing and unequal crossing over of repetitive nucleotide sequences (D'Ovidio et al. 1996; Wan et al. 2005).

Analysis of 13 additional lines showed only restriction fragments corresponding to those present in the H1 and H7 genotypes, which agrees with the suggestion of Vaz Patto et al. (2001) that *H. chilense* comprises two major ecotypes (represented in this study by H1 and H7) with a third group of lines being hybrids between the two.

The proteins encoded by the H1 and H7 genes differ from each other in amino acid substitutions but also in insertion/deletions, which in most cases reflect the block structure. Thus, the H7 protein differs from that encoded by H1 in deletions of 36 residues (comprising one tripeptide, two hexapeptides and two blocks of 10 and 11 residues) and seven residues (a degenerate block of four and half of a hexapeptide) and by the insertion of a hexapeptide and an adjacent block of 10 residues (indicated in Fig. 4). This results in the size difference between the two mature proteins, which comprise 875 (H1) and 849 (H7) residues.

A higher degree of re-iteration of the repetitive domain also accounts for the difference in size between the D hordeins of *H. chilense* and barley but in this case there are also differences in the degree of re-iteration of the different motifs. In particular, the extensive reiteration of the hexapeptide and tetrapeptide motif in the R3 region is reduced from 16 blocks in the protein encoded by cultivated barley line AAP31051 to only four blocks in the H1 and H7

proteins. However, the R1 region is shorter in the barley protein and the whole protein is only 584 residues.

In wheat the length and degree of conservation of the repetitive sequences present in the HMW-GS are both considered to contribute to gluten strength due to their impact on the interactions of the subunits (Belton 1999, 2005). According to this model, subunits with long and highly conserved repetitive domains will form more elastic dough due to the formation of extensive arrays of inter-chain hydrogen bonds when the repetitive domains become aligned during mixing. This model is supported by spectroscopic and biomechanical measurements of gluten fractions and recombinant peptides (Wellner et al. 2005, 2006). The higher degree of re-iteration of the repeats in the *H. chilense* D hordeins and their greater similarity to those present in the HMW-GS of wheat (which lack the TTVS motif) may therefore result in an ability to form more stable interactions with the HMW-GS in dough than those formed by D hordeins of cultivated barley.

In addition, the D hordeins from the two *Hordeum* species differ from each other and from the HMW-GS in another potentially important respect: the number and distribution of cysteine residues. Thus, four cysteines are present in the repetitive domain of the *H. vulgare* proteins and two in *H. chilense*, compared with zero or one in the HMW-GS. In addition, either 5 + 1 or 5 + 2 cysteines are present in the N-terminal + C-terminal domains of the *H. vulgare* and *H. chilense* proteins, respectively, compared with 3 + 1 or 5 + 1 in the *x*-type and *y*-type HMW-GS, respectively (Table 2). Disulphide bonds between cysteine residues are considered to be essential to stabilize the glutenin polymers (as discussed by Shewry et al. 2003) but there is also evidence from the over-expression of subunit 1Dx5 in transgenic wheat that excessive cross-linking can have a negative impact leading to polymers which are too stiff and unable to hydrate fully on dough mixing (Popineau et al. 2001). Consequently it would be expected that the lower cysteine content of the *H. chilense* D hordein would result in more acceptable properties when transferred to durum or bread wheat than their homeologous from cultivated barley.

Conclusions

Two D hordeins from *H. chilense* have been isolated and characterized. These show differences in SNPs and insertions/deletions, which result in differences in the size of the encoded proteins. Comparisons with D hordeins of cultivated barley show that those from the *H. chilense* have longer repetitive region but fewer and differently distributed cysteine residues. The D hordeins of *H. chilense* represent a novel type of functional protein which can be transferred to bread or durum wheats by conventional crossing to widen the properties for conventional (breadmaking or pastamaking) or novel uses.

Acknowledgements The authors acknowledge funding by the Spanish Comisión Interministerial de Ciencia y Tecnología (C.I.C.Y.T.), project AGL2004-03361-C02-2. Technical assistance by Ana García is also acknowledged. Rothamsted Research receives grant-aided support from the Biotechnology and Biological Sciences Research Council (BBSRC) of UK.

References

Altpeter F, Vasil V, Srivastava V, Vasil IK (1996) Integration and expression of the high-molecular-weight glutenin subunit 1Ax1 gene into wheat. Nat Biotechnol 14:1155–1159

Alvarez JB, Ballesteros J, Sillero JA, Martin LM (1992) Tritordeum: a new crop of potential importance in the food industry. Hereditas 116:1–2

Alvarez JB, Campos LAC, Martín A, Sillero JA, Martín LM (1999) Genetic analysis of prolamins synthesised by the H^{ch} genome and their effects on gluten strength in hexaploid tritordeum. Euphytica 107:177–184

Alvarez JB, Martin A, Martin LM (2001) Variation in the high-molecular-weight glutenin subunits coded at the *Glu-*H^{ch1} locus in *Hordeum chilense*. Theor Appl Genet 102:134–137

Barro F, Rooke L, Bekes F, Gras P, Tatham AS, Fido R, Lazzeri PA, Shewry PR, Barcelo P (1997) Transformation of wheat with high-molecular-weight subunit genes results in improved functional-properties. Nat Biotechnol 15:1295–1299

Belton PS (1999) On the elasticity of wheat gluten. J Cereal Sci 29:103–107

Belton PS (2005) New approaches to study the molecular basis of the mechanical properties of gluten. J Cereal Sci 41:203–211

Blechl AE, Anderson OD (1996) Expression of a novel high-molecular-weight glutenin subunit gene in transgenic wheat. Nat Biotechnol 14:875–879

Ciaffi M, Lafiandra D, Turchetta T, Ravaglia S, Bariana H, Gupta R, MacRitchie F (1995) Breadmaking potential of durum wheat lines expressing both x and y type subunits at the *Glu-A1* locus. Cereal Chem 72:465–469

D'Ovidio R, Lafiandra D, Porceddu E (1996) Identification and molecular characterisation 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

Felsenstein J (1999) Coalescents, phylogenies, and likelihoods. Biol Bull 196:343–344

Halford NG, Tatham AS, Sui E, Daroda L, Dreyer T, Shewry PR (1992) Identification of a novel beta-turn-rich repeat motif in the D-hordeins of barley. BBA 1122:118–122

Halford NG, Shewry PR (2007) The structure and expression of cereal storage protein genes. In: Olsen O-A (ed) Plant cell monographs: endosperm developmental and molecular biology. Springer, Berlin, Germany (in press)

Gu YQ, Anderson OD, Londeore CF, Kong X, Chibbar RN, Lazo GR (2003) Structural organisation of the barley D-hordein locus in comparison with its orthologous regions of wheat genomes. Genome 46:1084–1097

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275–282

Lawrence GJ, Shepherd KW (1981) Chromosomal location of genes-controlling seed proteins in species related to wheat. Theor Appl Genet 59:25–31

Martin A, Alvarez JB, Martin LM, Barro F, Ballesteros J (1999) The development of tritordeum–a novel cereal for food- processing. J Cereal Sci 30:85–95

Martin A, Sanchez-Monge E (1982) Cytology and morphology of the amphiploid *Hordeum chilense* x *Triticum turgidum* conv *durum*. Euphytica 31:261–267

Miflin BJ, Field JM, Shewry PR (1983) Cereal storage proteins and their effects on technological properties. In: Daussant J, Mossé J, Vaughan J (eds) Seed Proteins. Academic, London, pp 255–319

Payne PI (1987) Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. Annu Rev Plant Physiol 38:141–153

Payne PI, Nightingale MA, Krattiger AF, Holt LM (1987) The relationship between HMW glutenin subunit composition and the breadmaking quality of british grown wheat varieties. J Sci Food Agri 40:51–65

Piston F, Martin A, Dorado G, Barro F (2005) Cloning and molecular characterization of B-hordeins from *Hordeum chilense* (Roem. et Schult.). Theor Appl Genet 111:551–560

Popineau Y, Deshayes G, Lefebvre J, Fido RJ, Tatham AS, Shewry PR (2001) Prolamin aggregation, gluten viscoelasticity and mixing properties of transgenic wheat lines expressing 1Ax and 1Dx HMW glutenin subunit transgenes. J Agric Food Chem 49:395–401

Rogers WJ, Miller TE, Payne PI, Seekings JA, Holt LM, Law CN (1997) Introduction to bread wheat (*Triticum aestivum* L.) and assessment for breadmaking quality of alleles from *T. boeoticum* Boiss ssp *thaoudar* at *Glu-A1* encoding two high molecular weight subunits of glutenin. Euphytica 93:19–29

Shewry PR, Halford NG, Tatham AS (1992) The high molecular weight subunits of wheat glutenin. J Cereal Sci 15:105–120

Shewry PR, Napier JA, Tatham AS (1995) Seed storage proteins: structures and biosynthesis. Plant Cell 7:945–956

Shewry PR, Halford NG, Tatham AS, Popineau Y, Lafiandra D, Belton PS (2003) The high molecular weight subunits of wheat glutenin and their role in determining wheat processing properties. Adv Food Nutr Res 45:221–302

Shewry PR, Halford NG, Lafiandra D (2006) Chapter 5: the high-molecular-weight subunits of Glutenin. In: Wrigley C, Bushuk W, Bekes F (eds) Gliadins and Glutenins. AACC, St. Paul, MN, USA

Stacey J, Isaac P (1994) Isolation of DNA from plants. In: Isaac PG (ed) Methods in molecular biology: protocols for nucleic acid analysis by non-radiactive probes. Humana Press, Totawa, NJ, pp 9–15

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The clustal(x windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Vaz Patto MC, Aardse A, Buntjer J, Rubiales D, Martín M, Niks RE (2001) Morphology and AFLP markers suggest three *Hordeum chilense* ecotypes that differ in avoidance to rust fungi. Can J Bot 79:204–213

Wan Y, Yan Z, Liu K, Zheng Y, d'Ovidio R, Shewry PR, Halford NG, Wang D (2005) Comparative analysis of the D genome-encoded high-molecular weight subunits of glutenin. Theor Appl Genet 111:1183–1190

Wellner N, Mills ENC, Brownsey G, Wilson RH, Brown N, Freeman J, Halford NG, Shewry PR, Belton PS (2005) Changes in prolamin secondary structure during gluten deformation studied by dynamic FT-IR spectroscopy. Biomacromolecules 6:255–261

Wellner N, Marsh J, Ssavag A, Halford NG, Shewry PR, Mills ENC, Belton PS (2006) Structure-function relationships in an elastomeric protein of wheat: synthesis, expression and characterisation of repetitive peptides based on sequence variants present in the HMW subunits. Biomacromolecules 7:1096–1103