Theor Appl Genet (1985) 70:634–642

Structural homology of endosperm high molecular weight glutenin subunits of common wheat (*Triticum aestivum* L.)

G. Galili and M. Feldman

Department of Plant Genetics, The Weizmann Institute of Science, Rehovot, Israel

Accepted January 19, 1985
Communicated by J. MacKey

Summary. Several high molecular weight endosperm glutenin subunits, coded by genes located on chromosomes 1A, 1B and 1D of common wheat, *Triticum aestivum* L. em. Thell., were isolated from excised gel segments and subjected to amino acid analysis and peptide mapping; the latter was carried out following a limited digestion with trypsin, chymotrypsin or *Staphylococcus aureus* – V8 protease. Generally, all high molecular weight glutenins had a similar amino acid composition but several significant differences were observed in some of them. Both analyses revealed that the structural similarity among the various subunits was related to the homology of the genes coding them: subunits coded by homoalleles, i.e., different alleles of the same gene, were most similar; those coded by homoeoalleles, i.e., alleles of homoeologous genes, were less similar; whereas subunits coded either by alleles of different genes of the same gene cluster, or by non-homoeoalleles of homoeologous clusters, were the least similar. Several small peptides derived from protease digestion of various subunits had a higher than expected staining intensity indicating that small peptide repeats may be interspersed within the glutenin subunits. The evolutionary course of the high molecular weight glutenins is discussed.

Key words: Hexaploid wheat – *Triticum aestivum* – Glutenins – Electrophoresis – Amino acid analysis – Peptide mapping

Introduction

The fraction of endosperm high molecular weight (HMW) glutenins of common wheat, *Triticum aestivum*, noted for its contribution to the baking quality of wheat (Wall 1979), has been intensively studied in recent years. (Recent review see Payne et al. 1982.)

Using one-dimensional polyacrylamide gel electrophoresis (PAGE) containing sodium dodecyl sulphate (SDS), Bietz et al. (1975) detected in the common wheat cultivar 'Chinese Spring' four HMW glutenin subunits. Further analyses of various aneuploids derived from this cultivar revealed that the two central HMW glutenins are controlled by the long arm of chromosome 1B (1BL), while the rapidly and slowly migrating (outer) subunits are controlled by the chromosome arm 1DL (Bietz et al. 1975; Brown et al. 1979; Lawrence and Shepherd 1980; Galili and Feldman 1983a). Studies of many common wheat genotypes, as well as several intervarietal substitution lines have shown that in most cultivars chromosome 1B controls two to three HMW glutenin subunits, chromosome 1D – two subunits, and chromosome 1A – zero to two subunits (Lawrence and Shepherd 1980; Payne et al. 1981; Galili and Feldman 1983b). Shewry et al. (1984) concluded that differences between HMW glutenin subunits coded by one chromosome in a given cultivar were unlikely to be due to post-translational modifications of a single polypeptide chain; the presence of many amino-acid substitutions in these subunits indicate that they represent the products of more than one gene.

The HMW glutenin subunits controlled by each of the chromosomes 1A, 1B and 1D of hexaploid wheat were classified further into chromosomal subgroups on the basis of their occurrence in various cultivars (Galili and Feldman 1983b): subunits of a given subgroup were defined as those controlled by the same chromosome and never co-appear in a homozygous line; in hybrids these subunits appear in a ratio of 2:1, depending on the direction of the cross. Following the suggestion of Lawrence and Shepherd (1980) different subunits of the same subgroup were assumed to be coded by different alleles of the same gene. Accordingly, three genes were identified on chromosome 1B (*Glt-B1, Glt-B2,* and *Glt-B3*), two on 1D (*Glt-D1* and *Glt-D2*) and two on 1A (*Glt-A1* and *Glt-A2*).

Since the HMW glutenin genes are arranged in clusters (Brown et al. 1979; Lawrence and Shepherd 1980), the following four types of alleles controlling these proteins may be defined: (a) "Homoalleles" – alleles of homologous genes; (b) "Homoeoalleles" – alleles of homoeologous genes (homoeologous genes coding for subunits with a similar rate of migration); (c) "Paraalleles" – alleles of different genes of the same gene cluster [such genes were defined as paralogous genes (Fitch 1973)]; (d) "Parahomoeoalleles – alleles of paralogous genes of homoeologous gene clusters, which are not homoeoalleles. For example, alleles coding for subunits B2, B4 and B6 (Table 1) are homoalleles, those coding for B2, D1 and A1 are homoeoalleles, those coding for B2, B8 and B10 are paraalleles and those coding for B2 and D4 are parahomoeoalleles.

In the present work amino acid analyses and peptide mappings after limited digestion with either chymotrypsin, trypsin or V8 protease were used as probes to further assess the structural relationships among the various HMW glutenin subunits.

Materials and methods

Plant material

The following cultivars and lines of common wheat, *Triticum aestivum* L. em. Thell., were used: 'Chinese Spring', 'Atlas 66' (TAA 19), 'Timstein', 'Thatcher' and 'Sonora 66' (Table 1). The SDS PAGE pattern of the endosperm proteins of these genotypes was reported in a previous work (Galili and Feldman 1983 b; Fig. 1).

Protein extraction and electrophoresis

Extraction of total endosperm proteins, fractionation by SDS PAGE, staining and destaining of the gels and the use of MW markers were as previously described (Galili and Feldman 1983 a, b).

Amino acid analysis of individual protein bands

Individual bands were excised from preparative slab gels and homogenized in 5 ml of 1% sodium dodecyl phosphate (SDS) using a pestle and mortar. The gel slurry was shaken for 15–20 h at room temperature. Gel pieces were removed by centrifugation at room temperature for 15 min at 10,400 × g. The supernatant was brought to about 85% with cold acetone (acetone precipitation was performed twice) and stored for 24 h at −20 °C. Proteins precipitated by the acetone were centrifuged at 4 °C at 10,400 × g for 20 min and the pellet was resuspended in 0.3 ml of 6N HCl containing a few drops of phenol. The proteins were then hydrolized at 110 °C for 23 h in vacuum. Following hydrolysis, the HCl was removed by lyophilization and the remaining amino acid powder was resuspended in 0.1 ml of 0.2 M sodium citrate-HCl pH 2.2. Analysis of amino acids was performed by a dionex D-502 amino acid analyzer.

Proteolytic digestion by trypsin or chymotrypsin

Excision and extraction of individual bands from the gels as well as acetone precipitation were as above. The protein pellet was resuspended in 0.1–0.2 ml of 50 mM $(NH_4)_2CO_3$. A given amount of trypsin or chymotrypsin (Worthington) was added and digestion proceeded at 37 °C. Digestion was terminated by the addition of an equal volume of 3 × sample buffer and heating at 100 °C for 2 min. Trypsin cleaves specifically at lysine and arginine; chymotrypsin cleaves preferentially at tryptophane, phenyl-alanine and tyrosine and much slower at asparagine, glutamine, histidine, leucine, lysine, methionine, serine and threonine (Smythe 1967).

Protein digestion by Staphylococcus aureus V8 protease

Individual bands were excised from the gels and soaked for 30 min in excess of 1.0 mM ethylenediaminetetra-acetic acid (EDTA), 1% 2-mercaptoethanol (ME), 0.1% SDS and 0.125 M Tris-HCl pH 6.8. The bands were placed in wells of a second gel which were then filled with 10–30 µl of a solution consisting of 20% glycerol, 0.1% SDS, 0.1% bromophenol blue and 0.125 M Tris-HCl pH 6.8 followed by 10 µl of a solution of 10% glycerol, 0.1% SDS, 0.125 M Tris-HCl pH 6.8 and a given amount of *Staphylococcus aureus* – V8 protease (Miles). The enzyme cleaves specifically at glutamate and aspartate (Drapeau 1977).

Gels were allowed to run at 100 volts for about 2 h until the bromophenol blue migrated close to the end of the stacking gel. The current was then stopped for 30 min to complete the protease digestion and the running was continued at 150 volts until the dye reached the bottom of the gel. Staining and destaining were as mentioned above.

Results

Eleven HMW glutenin bands, previously assigned (Galili and Feldman 1983 b) to six different chromosomal subgroups (genes), were studied (Table 1). The extent of homology between these subunits was determined from their amino acid composition and peptide mapping after a limited digestion with trypsin, chymotrypsin or V8 protease.

Amino acid composition

Amino acid composition was determined for both HMW glutenin bands and interband gel segments. Analysis of the interband gel segments revealed a non-specific background of amino acids, accounting in molaric quantity for about 10% of the various glutenin bands. The amino acid compositions of the individual HMW glutenins, corrected for by this non-specific background, are shown in Table 2. On the whole, there was a great similarity in amino acid composition among all subunits. In fact, the mole percentage of aspartic acid + asparagine, glutamic acid + glutamine, glycine, phenyl alanine and lysine did not differ significantly among the various HMW glutenin subunits. On the other hand, histidine, alanine, isoleucine and arginine were the most variable. For the rest of the amino acids, slight though significant differences were observed. Comparisons of the relative amounts of each amino acid in the different subunits by the Duncan's Multiple Range

Table 1. The HMW glutenin subunits studied, their origin and assignment into chromosomal groups and subgroups[a]. Subunits of one subgroup are coded by different alleles of one gene; genes of the same chromosomal group comprise one gene cluster

Chromosomal group	Chromosomal subgroup	Band designation[a]	MW (KD)	Origin
Chromosome 1B	*Glt-B1*	B2	98	'Chinese Spring'
		B4	95	'Atlas 66'
		B6	92	'Timstein'
	Glt-B2	B8	90	'Atlas 66'
		B9	88	'Timstein'
	Glt-B3	B10	86	'Chinese Spring'
Chromosome 1D	*Glt-D1*	D1	108	'Chinese Spring'
		D3	106	'Thatcher'
	Glt-D2	D4	84	'Thatcher'
		D5	80	'Chinese Spring'
Chromosome 1A	*Glt-A1*	A1	114	'Sonora 66'

[a] According to Galili and Feldman (1983b)

Table 2. Amino acid composition of individual HMW glutenin subunits isolated from excised gel segments (values are expressed as mole % and represent the average of at least three different determinations)[a, b]

Gene	*Glt-A1*	*Glt-B1*			*Glt-D1*		*Glt-B2*		*Glt-B3*	*Glt-D2*		SE
Subunit	A1	B2	B4	B6	D1	D3	B8	B9	B10	D4	D5	
Asp[c]	1.14	0.74	1.33	1.15	0.90	0.91	1.16	1.45	1.15	1.76	1.33	0.42
Thr*	3.00	3.39	3.30	3.39	2.86	2.81	3.42	3.30	3.43	3.59	3.39	0.14
Ser*	6.68	7.98	7.69	8.25	5.48	5.92	6.51	6.80	7.51	6.59	6.16	0.32
Glu[d]	39.46	36.50	36.98	36.41	38.46	37.38	36.28	35.48	35.15	34.62	36.95	1.32
Pro*	13.36	12.89	12.22	12.24	14.64	13.77	9.73	10.96	10.84	11.42	12.29	0.42
Gly	18.88	19.64	18.76	18.18	20.73	20.92	19.98	18.74	19.33	19.38	18.95	0.74
Ala*	2.48	3.14	3.52	3.73	3.17	3.13	4.58	4.13	4.74	4.53	3.99	0.29
Val*	0.90	1.20	1.69	1.58	1.38	1.46	2.68	2.28	1.97	2.21	1.67	0.29
Met*	0.40	0.36	0.23	0.31	0.34	0.23	0.44	0.44	0.19	0.80	0.63	0.10
Ile*	0.29	0.58	0.54	0.81	0.30	0.47	1.30	1.28	1.12	0.84	0.91	0.14
Leu*	4.44	3.01	2.73	3.35	4.42	4.49	4.34	4.37	4.17	4.27	3.79	0.20
Tyr*	4.26	6.50	6.66	5.92	4.64	4.94	5.14	5.17	5.01	4.86	5.05	0.48
Phe	0.23	0.35	0.15	0.44	0.25	0.39	0.58	0.50	0.61	0.77	0.42	0.41
His*	0.69	0.61	0.28	0.65	0.60	0.52	1.22	1.57	1.31	1.78	2.00	0.11
Lys	0.93	0.64	1.42	0.92	0.91	1.05	1.48	1.36	1.42	0.80	1.15	0.25
Arg*	2.68	2.11	2.25	2.50	0.73	1.26	2.02	2.40	1.93	2.06	1.43	0.22

[a] The amino acids cystein and tryptophan were not determined
[b] The values of amino acids marked by an asterisk (*) differed significantly among various subunits at the 5% level (F test)
[c] Aspartic acid + asparagine
[d] Glutamic acid + glutamine

Test (Table 3), showed that subunits coded by homoalleles were most similar to one another; some of those coded by homoeoalleles were less similar and subunits coded by the other homoeoalleles as well as those coded by paraalleles and parahomoeoalleles were least similar; the latter two types of alleles differed from each other only slightly. These comparisons also indicated that the similarity between subunits coded by the genes *Glt-A1* and *Glt-D1* was greater than between those coded by *Glt-B1* and either *Glt-A1* or *Glt-D1*; subunits coded by *Glt-B2* were very similar to that coded by *Glt-B3*; the latter, however, was more similar to those of the homoeologous gene *Glt-D2* than the subunits of *Glt-B2*.

Several amino acids were either particularly low or high in subunits of certain genes: serine and arginine were low while proline was high in subunits of *Glt-D1*; serine and tyrosine were high while leucine low in sub-

Table 3. Comparison between subunits produced by homoalleles, homoeoalleles, paraalleles and parahomoeoalleles as to differences in their amino acids mole % (based on values of Table 2); the symbol (+) designates significant difference (at the 0.05 level) between at least one pair of subunits controlled by the indicated alleles as determined by the Duncan's Multiple Range Test

Amino acid	Subunits produced by homoalleles of the genes			Subunits produced by homoeoalleles of the genes				Subunits produced by paraalleles of the genes				Subunits produced by parahomoeoalleles of the genes								Total
	Glt-B1	*Glt-B2*	*Glt-D1*	*Glt-D2*	*Glt-A1* and *Glt-B1*	*Glt-A1* and *Glt-D1*	*Glt-B1* and *Glt-D1*	*Glt-B3* and *Glt-D2*	*Glt-B1* and *Glt-B2*	*Glt-B1* and *Glt-B3*	*Glt-B2* and *Glt-B3*	*Glt-D1* and *Glt-D2*	*Glt-A1* and *Glt-B2*	*Glt-A1* and *Glt-B3*	*Glt-A1* and *Glt-D2*	*Glt-B1* and *Glt-D2*	*Glt-D1* and *Glt-B2*	*Glt-D1* and *Glt-B3*	*Glt-B2* and *Glt-D2*	
Asp[a]																				0
Thr												+			+		+	+		4
Ser					+	+	+		+							+	+	+		7
Glu[b]															+					1
Pro							+	+	+	+		+	+	+	+		+	+	+	11
Gly							+													1
Ala					+		+		+	+		+	+	+	+	+	+	+		11
Val									+			+	+	+	+	+	+		+	7
Met								+				+	+	+	+				+	5
Ile					+				+	+		+	+	+	+		+	+	+	11
Leu					+		+		+	+		+				+				6
Tyr					+		+			+						+				4
Phe										+		+	+			+				4
His	+								+	+		+	+	+	+	+	+		+	14
Lys		+			+			+					+							1
Arg						+	+								+		+	+	+	8
Total	1	1	0	0	6	2	8	3	8	7	0	9	5	5	10	9	8	7	6	

[a] Aspartic acid + asparagine [b] Glutamic acid + glutamine

units of *Glt-B1*; methionine and histidine were high in subunits of *Glt-D2*; isoleucine was high while proline low in subunits of *Glt-B2* and *Glt-B3*; tyrosine was low in subunits of *Glt-A1*.

Peptide mapping

The protease digestion of each subunit yielded a unique and fairly repeatable peptide map. Trypsin and chymotrypsin gave complex maps of many peptides which in most cases revealed some similarity only between subunits of homoalleles (Figs. 1 and 2). Digestion with V8 protease yielded less complex maps with considerable similarity between subunits of homoalleles and some similarity between subunits of several homoeoalleles (Figs. 3–6). In most cases, several peptides appeared with a higher than expected staining intensity, indicating that small peptide repeats might be interspersed within the HMW glutenin subunits (Figs. 1–7).

Homology between subunits coded by homoalleles

Maps of subunits coded by homoalleles showed several peptides with identical MW as well as several different peptides, indicating a considerable structural homology besides various degrees of amino acids alterations in the sequence of these subunits. Maps of several subunits, especially those coded by the homoalleles of the genes *Glt-D1* and *Glt-D2* (Figs. 5 and 6), showed, in several peptides, a similar size variation to that revealed by the original undigested bands from which they were derived. This indicates that duplication or insertion of small peptides may have also contributed to the size variation between these subunits.

Glt-B1. Generally, subunits B2, B4 and B6 of this gene (Fig. 1, left) showed a partial homology in the trypsin peptide maps (Fig. 2) and a considerable homology in the chymotrypsin and V8 protease peptide maps (Figs. 1 and 3). The trypsin maps showed that subunits B2 and B6 were more similar to each other than to subunit B4. Despite the considerable homology, several unique peptides were noted in each of the peptide maps derived from the three proteases.

Glt-B2. Subunits B8 and B9 of this gene (Fig. 1) showed a partial homology in the chymotrypsin and trypsin derived peptide maps (Figs. 1 and 2) and almost full homology following V8 protease digestion (Fig. 4). The latter protease cleaved B8 and B9 into a few peptides only.

Glt-D1. Peptide maps of subunits D1 and D3 of this gene (Fig. 1) revealed a few homologous peptides with identical MW following digestion with either one of the

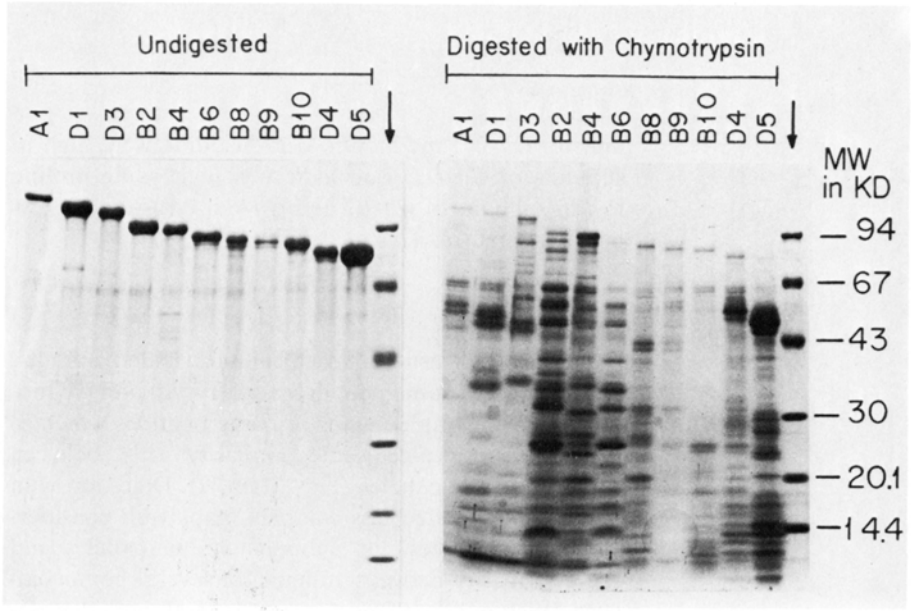

Fig. 1. Peptide maps of excised HMW glutenin bands: undigested (*left*) and digested with 10 µg/ml chymotrypsin (*right*) for 100 min at 37 °C. *Arrows* indicate MW markers

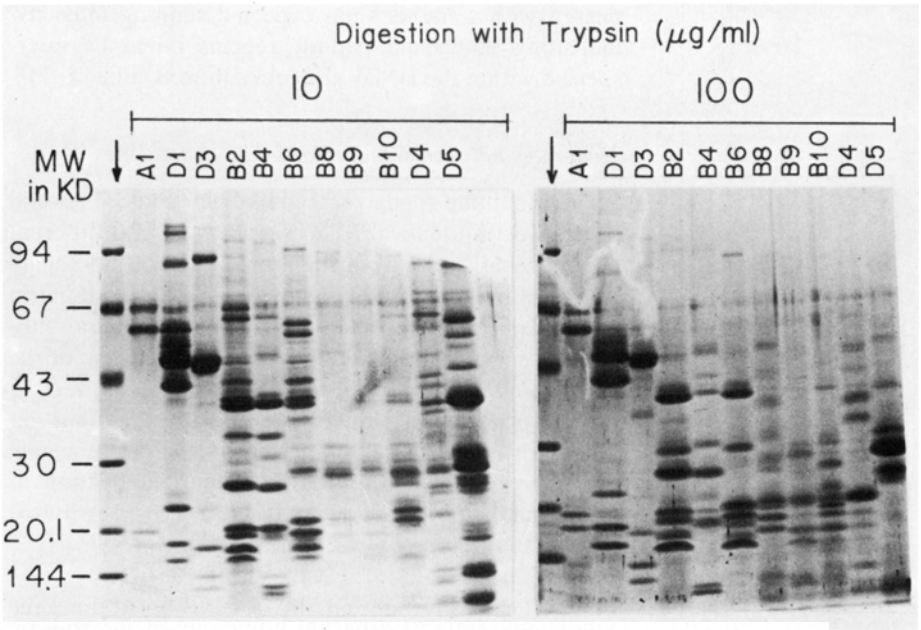

Fig. 2. Peptide maps of excised HMW glutenin bands digested with 10 or 100 µg/ml trypsin for 100 min at 37°C. *Arrows* indicate MW markers

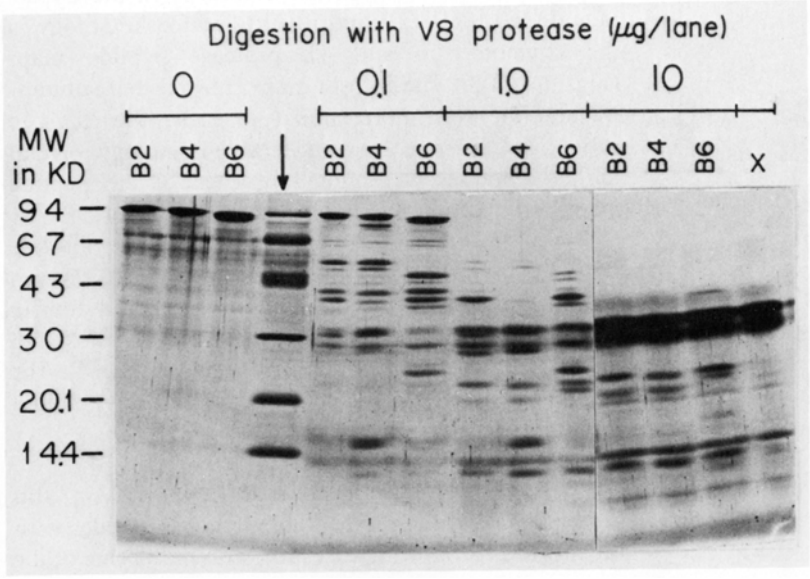

Fig. 3. Peptide maps of excised HMW glutenin bands of *Glt-B1* digested with 0, 0.1, 1.0 or 10 µg/lane with V8 protease. (X) digestion products of 10 µg/lane V8 protease. *Arrow* indicates MW markers

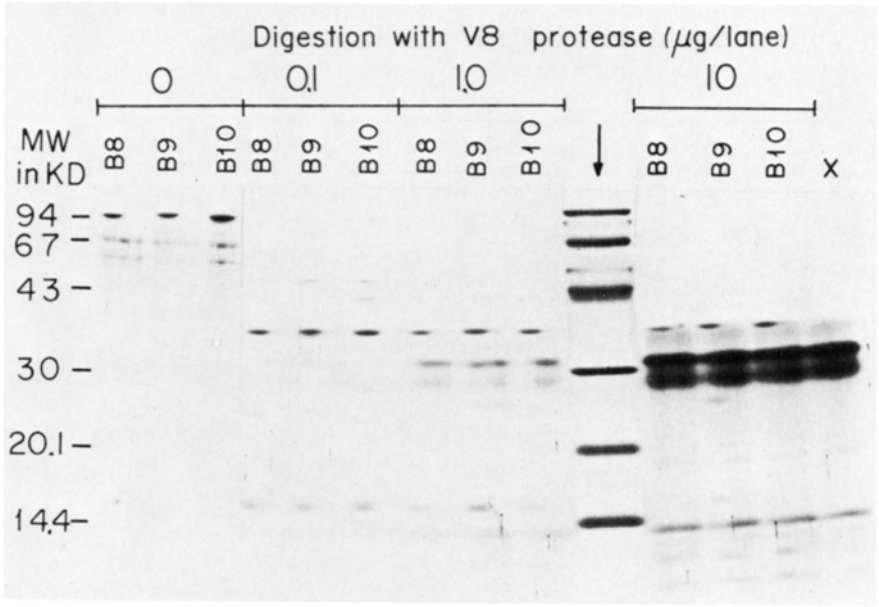

Fig. 4. Peptide maps of excised HMW glutenin bands of *Glt-B2* and *Glt-B3* digested with 0, 0.1, 1.0 or 10 μg/lane V8 protease. (X) digestion products of 10 μg/lane V8 protease. Fractionated MW markers are indicated by *arrow*

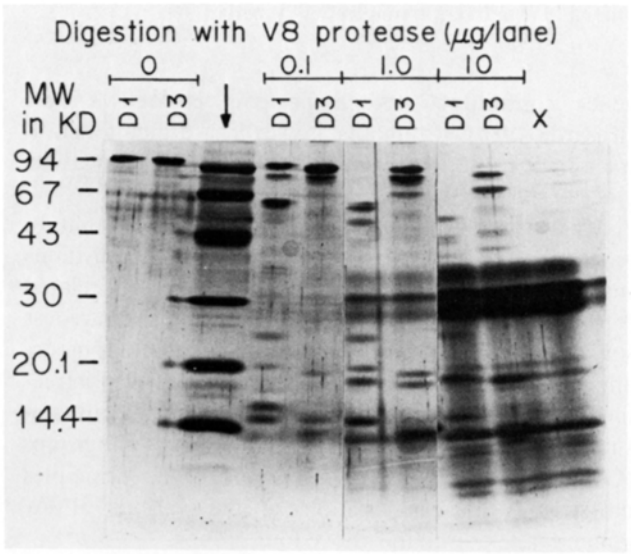

Fig. 5. Peptide maps of excised HMW glutenin bands of *Glt-D1* digested with V8 protease. (X) digestion products of 10 μg/lane of V8 protease. Fractionated MW protein markers are indicated by *arrow*

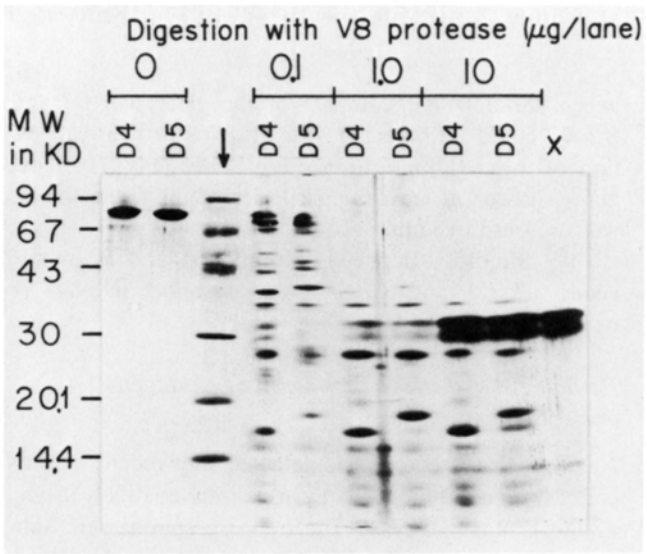

Fig. 6. Peptide maps of excised HMW glutenin bands of *Glt-D2* with 0, 0.1, 1.0 or 10 μg/lane V8 protease. (X) digestion products of 10 μg/lane V8 protease. Fractionated MW markers are indicated by *arrow*

three proteases (Figs. 1, 2 and 5). In these maps however, several subunits could be detected with size variation of about 2 KD – the size difference between D1 and D3 (Table 1). Each map also possessed several unique peptides.

Glt-D2. Subunits D4 and D5 of this gene showed a considerable number of homologous peptides after digestion with chymotrypsin (Fig. 1) and a smaller number of homologous peptides in maps of trypsin or V8 protease digest (Figs. 2 and 6). Several peptides which could account for the size difference of 4 KD between D4 and D5 (Table 1), were seen in all protease digests. In the peptide maps derived from the higher concentration of trypsin, a considerable number of unique peptides was also seen (Fig. 2).

Homology between subunits coded by homoeoalleles

Subunits coded by homoeoalleles revealed in their chymotrypsin and trypsin maps a low degree of homology. However, some degree of homology be-

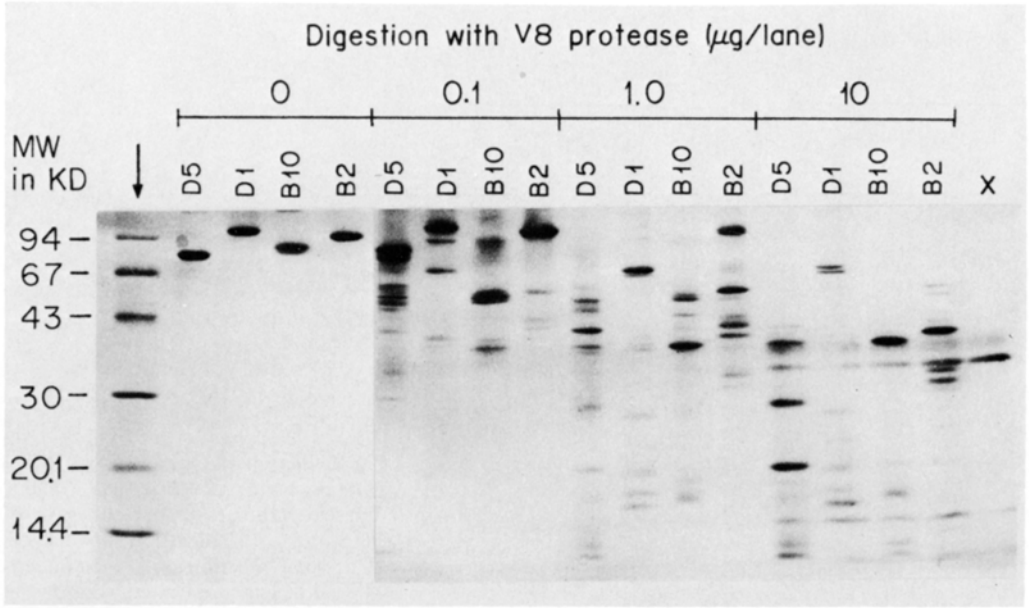

Fig. 7. Peptide maps of excised HMW glutenin bands of *Glt-B1, Glt-B3, Glt-D1* and *Glt-D2* digested with 0, 0.1, 1.0 or 10 μg/lane V8 protease. (X) digestion products of 10 μg/lane V8 protease. Fractionated MW markers are indicated by *arrow*

tween subunits B10 and D5 of the genes *Glt-B3* and *Glt-D2* could be detected following V8 protease digestion of these subunits in parallel lanes (Fig. 7). These maps showed at least five peptides of the same size as well as about five non-related peptides.

No homology between subunits D1 and B2 of the genes *Glt-D1* and *Glt-B1* was detected in the V8 protease digests (Fig. 7).

Homology between subunits coded by paraalleles or by parahomoeoalleles

No homology could be detected between subunits coded by paraalleles and by parahomoeoalleles in any of the protease digests. The only exception were subunits B8 and B9 of the *Glt-B2* gene and subunit B10 of the *Glt-B3* gene. These subunits showed a partial homology in the maps of chymotrypsin and trypsin digests (Figs. 1 and 2) and almost a full homology in the maps of the V8 protease (Fig. 4).

Discussion

The relative similarity among the various HMW glutenin subunits, as revealed both by amino-acid analyses and peptide mappings, strongly indicates a common ancestral origin of the genes coding them. This is in accord with the conclusion of Shewry et al. 1984, that the various wheat genes for HMW glutenin subunits evolved from a single ancestral gene. The conservative nature of genes controlling storage proteins is well known (Kasarda 1980). It remains, however, a difficult task to trace their evolution through the analysis of the various types of the HMW glutenins.

The data obtained in this work revealed that in terms of their structure and amino acid compositions subunits coded by homoalleles are most closely related to each other; subunits coded by homoeoalleles are less similar and those coded by paraalleles and parahomoeoalleles are least similar. This finding, which is in agreement with the previous assignment of the HMW glutenin subunits into chromosomal groups and subgroups (Galili and Feldman 1983 b), supports the principles underlying the classification of the various HMW glutenin alleles into the above four groups.

Peptide mapping revealed that subunits controlled by homoalleles of the B genome genes were more closely related than those of the D genome. The highest degree of similarity was found among subunits of the *Glt-B1* gene. This is in contrast to the finding that *Glt-B1* is the most polymorphic gene, having the widest qualitative and quantitative variation among all HMW glutenin genes in hexaploid wheat (Galili and Feldman 1983 b). In addition, peptide maps of subunits coded by homoalleles of the genes *Glt-D1* and *Glt-D2* showed a considerable number of small peptides with size differences similar to the original undigested subunits. Apparently, insertion of small peptides has contributed to this size variation among subunits coded by the homoalleles of these two genes. The differences between subunits controlled by homoalleles of the B and

D genomes may reflect different patterns of divergence within the HMW glutenin genes of these genomes.

The relative similarity between subunits coded by homoeoalleles as compared with that of subunits coded by paraalleles or parahomoeoalleles, is in accord with the classification of the genes *Glt-A1, Glt-B1* and *Glt-D1* into one homoeologous group coding for the slowly migrating subunits, and *Glt-B3* and *Glt-D2* into a second homoeologous group coding for the rapidly migrating subunits. In addition, both amino acid analyses and peptide mapping show that subunits B10 and D5 of the homoeologous genes *Glt-B3* and *Glt-D2,* respectively, are closer to one another than subunits B2 and D1 of the homoeologous genes *Glt-B1* and *Glt-D1.* This may indicate that the homoeologous genes coding for the slowly migrating subunits have further differentiated from one another than those coding for the rapidly migrating subunits. Thus, the differentiation rate of different genes in one gene cluster may be different and independent of each other.

The second homoeologous group of genes, i.e., that coding for the rapidly migrating subunits, lacks a representative of chromosome 1A. As was suggested by Payne et al. (1981) this gene is either absent or inactive in hexaploid wheat. It is, however, present in the diploid wheat, *T. monococcum,* the donor of the A genome of polyploid wheats, as well as in the wild tetraploid wheat *T. turgidum* var. 'dicoccoides' (Galili and Feldman 1983c).

Subunits coded by paraalleles or by parahomoeoalleles showed conspicuous differences from one another in their amino acid composition and peptide mapping. These differences strongly indicate that such subunits represent products of different genes. This is in agreement with the conclusion of Shewry et al. 1984 that the differences between subunits coded by paralogous genes are unlikely to be due to posttranslational modifications of a single polypeptide chain, because of the presence of amino acid substitutions and deletions in these subunits. Exceptionally, subunits of the paralogous genes *Glt-B2* and *Glt-B3* of chromosome 1B were found to be very similar to each other and therefore, very closely related. Since a gene homoeologous to *Glt-B2* is absent from chromosomes 1A and 1D of common wheat, as well as from the genomes of *Aegilops* species of section Sitopsis (closely related to the donor of the B genome; Galili and Feldman, unpublished), the duplication and divergence that gave rise to the *Glt-B2* occurred at the polyploid level and is a relatively recent event.

That divergence of the ancestral HMW glutenin gene followed gene duplication is suggested from the presence of several small peptides with higher than expected staining intensity in most peptide maps. It seems that repeats of small peptides might be interspersed among the HMW glutenin subunits. This is also supported by the presence of DNA sequences coding for basic repeats, of six, nine and 21 amino acids, recently shown in two cDNA clones produced from mRNAs coding for HMW glutenins in common wheat (Forde et al. 1983; Thompson et al. 1983). Repeated amino acid sequences were also shown in wheat gliadins by Kasarda (1980); Bartels and Thompson (1983) and in corn zeins by Geraghty et al. (1981) and Pedersen et al. (1982). Argos et al. (1982) postulated that amino acid repeats are important for the compact organization of zeins in protein bodies.

The repeated amino acid sequences in storage proteins suggest that even the ancestral storage protein genes evolved by duplication and divergence of small basic DNA sequences (Kasarda 1980). Later during evolution, when these repeats were divergent enough, duplication of larger DNA sequences might have also occurred. This may explain the fact that several groups of repeated DNA sequences of different sizes may be observed within one storage protein gene. It might also partially explain the wide size variation of storage proteins. Amplification or deletion of repeated DNA sequences can be carried out between genes by unequal crossing over at meiosis or even within a gene by intrastrand recombination (Kasarda 1980; Thompson et al. 1983).

The fact that subunits coded by paraalleles showed a greater diversity than subunits coded by homoeoalleles suggests that duplication of the ancestral gene preceded the speciation of the ancestral wheat plant. This is in agreement with the finding that most lines of diploid wheat and *Aegilops* possess more than one HMW glutenin band (Galili and Feldman, unpublished). It appears that the formation of paralogous genes for storage proteins occurred very early in the development of the grass family since the loci for Hordein in barley and Zein in corn are also known to be comprised of several paralogous genes.

References

Argos P, Pedersen K, Marks MD, Larkins BA (1982) A structural model for maize zein proteins. J Biol Chem 257: 9984–9990

Bartels D, Thompson RD (1983) The characterization of cDNA clones coding for wheat storage proteins. Nucleic Acids Res 11:2961–2977

Bietz JA, Shepherd KW, Wall JS (1975) Single-kernel analysis of glutenin: use in wheat genetics and breeding. Cereal Chem 52:513–532

Brown JWS, Kemble RJ, Law CN, Flavell RB (1979) Control of endosperm protein in *Triticum aestivum* (var. 'Chinese Spring') and *Aegilops umbellulata* by homoeologous group 1 chromosomes. Genetics 93: 189–200

Drapeau GR (1977) Cleavage at glutamic acid with Staphylococcal protease. In: Hirs CHW, Timascheff SN (eds) Methods in enzymology, vol 47. Academic Press, New York, pp 189–191

Fitch WM (1973) Aspects of molecular evolution. Annu Rev Genet 7:343–380

Forde J, Forde BG, Fry RP, Kreis M, Shewry PR, Miflin BJ (1983) Identification of barley and wheat cDNA clones related to the high-MW polypeptides of wheat glutenin. FEBS Lett 162:360–366

Galili G, Feldman M (1983a) Genetic control of endosperm proteins in wheat. 1. The use of high resolution one-

dimensional gel electrophoresis for the allocation of genes coding for endosperm protein subunits in the common wheat cultivar 'Chinese Spring'. Theor Appl Genet 64: 97–101

Galili G, Feldman M (1983 b) Genetic control of endosperm proteins in wheat. 2. Variation in high molecular weight glutenin and gliadin subunits of *Triticum aestivum*. Theor Appl Genet 66:77–86

Galili G, Feldman M (1983 c) Diploidization of endosperm protein genes in polyploid wheats. In: Proc 6th Int Wheat Genet Symp, Kyoto, Japan, pp 1119–1123

Geraghty D, Peifer MA, Rubenstein I, Messing J (1981) The primary structure of a plant storage protein: zein. Nucleic Acids Res 9:5163–5174

Kasarda DD (1980) Structure and properties of α-gliadins. Ann Technol Agric 29:151–173

Lawrence GJ, Shepherd KW (1980) Variation in glutenin protein subunits of wheat. Aust J Biol Sci 33:221–233

Payne PI, Holt LM, Law CN (1981) Structural and genetical studies on the high-molecular-weight subunits of wheat glutenin. 1. Allelic variation in subunits amongst varieties of wheat (*Triticum aestivum*). Theor Appl Genet 60: 229–236

Payne PI, Holt LM, Lawrence GJ, Law CN (1982) The genetics of gliadin and glutenin, the major storage proteins of the wheat endosperm. Qual Plant Foods Hum Nutr 31: 229–241

Pedersen K, Devereux J, Wilson DR, Sheldon E, Larkins BA (1982) Cloning and sequence analysis reveal structural variation among related zein genes in maize. Cell 29: 1015–1026

Shewry PR, Field JM, Faulks AJ, Parmar S, Miflin BJ, Dieter MD, Lew EJ-L, Kasarda DD (1984) The purification and N-terminal amino acid sequence analysis of the high molecular weight gluten polypeptides of wheat. Biochim Biophys Acta 788:23–34

Smythe DG (1967) Techniques in enzymic hydrolysis. In: Hirs CHW (ed) Methods in enzymology, vol 11. Academic Press, New York, pp 214–231

Thompson RD, Bartels D, Harberd NP, Flavell RB (1983) Characterization of the multigene family coding for HMW glutenin subunits in wheat using cDNA clones. Theor Appl Genet 67:87–96

Wall JS (1979) The role of wheat proteins in determining baking quality. In: Laidman DL, Wyn Jones RG (eds) Recent advances in the biochemistry of cereals. Academic Press, New York, pp 275–311