Characterization of Low Molecular Weight Glutenin Subunits by Reversed-Phase High-Performance Liquid Chromatography, Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis, and N-Terminal Amino Acid Sequencing

ELLEN J.-L. LEW,[1] DONALD D. KUZMICKY,[1] and DONALD D. KASARDA[1,2]

ABSTRACT

Cereal Chem. 69(5):508–515

Ion-exchange-purified glutenin was prepared from the wheat cultivar Yecora Rojo, and reduced-alkylated protein subunits were purified from the mixture by reversed-phase high-performance liquid chromatography. N-terminal amino acid sequencing was performed on the fractions that corresponded to the B-type (mol wt 40,000–50,000) and C-type (mol wt 30,000–40,000) low molecular weight (LMW) glutenin subunits. The main B-type N-terminal sequence (NH₂-Ser-His-Ile-Pro-Gly-. . .) lacks cysteine and does not correspond exactly to any known DNA-based sequences for LMW-glutenin subunits. Also present were sequences that showed a close correspondence to those derived from DNA sequences (NH₂-Met-Glu-Thr-Ser-Cys-. . .), although the characteristic cysteine was not always present at position 5. C-type N-terminal sequences corresponded mainly to those associated with monomeric α-type and γ-type gliadins. These C-type LMW-glutenin sequences are likely to correspond to mutant forms that have acquired a single additional cysteine residue. They may act to terminate chain extension during formation of glutenin polymers.

Wheat (*Triticum aestivum* L.) gluten proteins may be classified as gliadins, monomeric proteins having only intramolecular disulfide bonds (or none), and glutenins, polymers of protein subunits linked by intermolecular disulfide bonds (Kasarda 1989). The glutenin fraction of the gluten proteins is most important in producing the unique viscoelastic properties of wheat flour doughs and probably in determining variations in the bread-making quality of different wheat cultivars (Bietz et al 1973, Payne et al 1984, Hoseney and Rogers 1990). The mechanisms responsible are not completely understood at the molecular level, but the joining together of protein subunits into polymers by means of intermolecular disulfide cross-linkages is almost certainly an essential feature (Ewart 1968, 1990). Glutenin polymers incorporate from two to many protein subunits so that molecular weights range from about 80,000 to several million (Kasarda 1989, Shewry et al 1989).

Glutenin subunits have been classified on the basis of their mobilities in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions into two main groups—high molecular weight glutenin subunits (HMW-GS) and low molecular weight glutenin subunits (LMW-GS). According to the scheme of Payne and Corfield (1979), HMW-GS were termed A subunits and LMW-GS were further broken down into two subgroups, B subunits (slower moving) and C subunits (faster moving). The HMW-GS have a molecular weight range of about 80,000–120,000 by SDS-PAGE (MacRitchie et al 1991). The LMW-GS have molecular weight ranges of about 40,000–50,000 for B and about 30,000–40,000 for C subunits.

To date, the HMW-GS have been moderately well characterized. Complete amino acid sequences are available for representatives of the main types as a result of sequencing of the genes coding for these proteins (Shewry et al 1989, Anderson et al 1990). The amino acid sequences of LMW-GS are less well characterized. Okita et al (1985) published a cDNA-based sequence that appears to correspond to a LMW-GS on the basis of sequence similarities to an N-terminal amino acid sequence reported by Shewry et al (1983). Sequences virtually identical with the sequence of Okita

et al (1985) and also derived from DNA clones have been reported by Colot et al (1989) and by Cassidy and Dvorak (1991). However, these DNA-derived sequences are slightly different from the N-terminal sequence of the "aggregated gliadin" subunit (LMW-GS) described by Shewry et al (1983) and the N-terminal sequence of a major LMW-GS reported by Kasarda et al (1988). Tao and Kasarda (1989) found this latter sequence to be the major type for the B-subunit group of LMW-GS; they also found an N-terminal sequence corresponding to the sequence of Okita et al (1985), but only in a minor amount.

Tao and Kasarda (1989) also demonstrated that N-terminal sequences of the faster-moving C group of LMW-GS corresponded mainly to γ-type and α-type gliadins. These gliadinlike C subunits presumably resulted from gliadin genes that acquired an extra cysteine residue due to a point mutation; the resulting odd number of cysteines necessitates the incorporation of the protein into glutenin by means of an intermolecular disulfide bond (Kasarda et al 1987, Kasarda 1989). Support for the presence of α-type gliadins in glutenin has been provided by Gupta (1989), who noted that some glutenin C subunits were coded by genes on group 6 chromosomes, which code for α-type gliadins, and by Skerritt and Robson (1990), who noted that monoclonal antibodies specific for α-gliadins reacted with some C subunits of glutenin. Sequences of the various types of B and C subunits also can be elucidated from composite sequences obtained by Bietz et al (1980), although the requisite sequences for interpretation of the results had not then been defined. Sequences of LMW-GS similar to those of Tao and Kasarda (1989) also were reported by Wieser et al (1990).

In this study, we reduced an ion-exchange (IE)-purified glutenin preparation from the cultivar Yecora Rojo and fractionated the reduced proteins by reversed-phase high-performance liquid chromatography (RP-HPLC). The protein (sometimes proteins) corresponding to each peak in the LMW-GS region of the chromatogram was reduced and alkylated with 4-vinylpyridine. The resulting pyridylethylated (PE) protein or proteins were refractionated by RP-HPLC. The PE-protein fractions (again derived from single peaks) from this last separation were characterized by SDS-PAGE and by N-terminal amino acid sequencing. Although this approach resulted in somewhat less pure fractions than those that were obtained by the spot sequencing approach (Tao and Kasarda 1989), the larger amounts of both major and minor components from RP-HPLC allowed us to obtain considerable new information about the types of LMW-GS present in wheat glutenin. The use of Yecora Rojo, which has good mixing and baking quality, provided a comparison with Chinese Spring, the poor-quality cultivar used in the study of Tao and Kasarda (1989).

MATERIALS AND METHODS

IE-Purified Glutenin

Wheat grain of the cultivar Yecora Rojo, lot CWC-141, was obtained from the California Wheat Commission. IE-purified glutenin was prepared from the endosperm fraction. Flour was milled in a Brabender Quadrumat Senior mill (C. W. Brabender, South Hackensack, NJ), and flour protein was 11.0% on a 13% moisture basis. Mixing and baking quality was as described by MacRitchie et al (1991). The glutenin preparation procedure will be described in detail elsewhere (D. D. Kasarda, N. F. Laird, M. D. Dietler, A. E. Adalsteins, and E. J.-L. Lew, *unpublished results*). In brief, the procedure was as follows. Flour was extracted with $0.1M$ acetic acid, which extracted about 70% of the protein as determined by Kjeldahl nitrogen analysis (nitrogen-to-protein conversion factor of 5.7). The acid-solubilized gliadin and glutenin proteins were precipitated by the addition of $0.25M$ NaCl to separate them from albumins and globulins. The resulting gliadin and glutenin mixture was fractionated first by size-exclusion chromatography on BioGel P-100 (Bio-Rad, Richmond, CA) in $0.1M$ acetic acid. Proteins that eluted at and shortly after the void volume included all glutenins and some ω-gliadins but excluded α-, β-, and γ-gliadins. The glutenin and ω-gliadin mixture was separated further by IE chromatography on carboxymethylcellulose (Whatman CM-32) using the method of Shewry et al (1981) with minor differences; mainly, $6M$ urea was included in all buffers to enhance glutenin solubility. This fractionation yielded glutenin polymers composed of HMW- and LMW-GS and free of ω-gliadins and β-amylases (Gupta et al 1991). The glutenin from this procedure was used for preparation of reduced and reduced-alkylated glutenin subunits.

Reduction of IE-Purified Glutenin

The IE-purified glutenin was dissolved in 50 mM Tris-HCl buffer (pH 8) containing $4M$ urea. The solution was flushed with nitrogen, the reductant dithiothreitol was added at an approximate dithiothreitol-to-cysteine ratio of 20:1, and the solution was once again flushed with nitrogen. Reduction was allowed to proceed at room temperature for 2–4 hr, at which time glacial acetic acid was added to lower the pH and to end the reduction. The reduced glutenin was then filtered through a 0.45-μm-pore filter and desalted and separated by RP-HPLC.

RP-HPLC

RP-HPLC was performed on a Vydac C_{18} semipreparative column (Vydac, Hesperia, CA) and equilibrated at 50°C with a Spectra-Physics SP8700 solvent delivery system (Spectra-Physics, San Jose, CA), an Isco V4 detector (Isco, Lincoln, NE), and a SP4270 integrator (Spectra-Physics). Sample load ranged from 0.5 to 1.5 mg in volumes of 250–500 μl. The separation was achieved with a linear gradient of 28–57% aqueous acetonitrile (with 0.05% trifluoroacetic acid) in 55 min at a flow rate of 1.5 ml/min. The proteins were detected by UV absorbance at 210 nm. Entire peaks were collected. Each of the 10 peaks (fractions) in the LMW-GS region was collected and dried on a Savant Speed Vac concentrator (Savant, Farmingdale, NY). Each dried fraction was reduced by the above method and alkylated with 4-vinylpyridine by the method described in the following section.

Reduction and Pyridylethylation of Glutenin Fractions

After reduction, 4-vinylpyridine (Sigma, St. Louis, MO) was added to total thiol at an approximate ratio of 4:1, and the solution was flushed with nitrogen. Alkylation was allowed to proceed overnight at room temperature in the dark (Tarr 1986). Glacial acetic acid was added to quench the reaction and to lower the pH. The solution was filtered through a 0.45-μm-pore filter and desalted and separated by RP-HPLC as described above. The resulting PE-glutenin fractions were dried and further characterized by SDS-PAGE and by N-terminal amino acid sequencing.

Electrophoresis

One-dimensional SDS-PAGE of the PE-protein fractions was performed according to the procedure of MacRitchie et al (1991).

Amino Acid Sequencing

Sequencing was performed with an Applied Biosystems (Foster City, CA) model 477A pulsed liquid phase amino acid sequencer and a model 120A on-line phenylthiohydantoin (PTH)-amino acid analyzer. Each PE-glutenin fraction to be sequenced was dissolved in 50% aqueous acetonitrile (with 0.05% trifluoroacetic acid) and applied onto the filter pretreated with 3.0 mg of Biobrene Plus (Applied Biosystems). Analyses of the data were performed by the instrument's data analysis program and, independently, by the operator. Fractions that had important features, such as the presence or absence of a cysteine residue as compared with the DNA-based sequences, and fractions for which the results were not clear were subjected to two or three analyses. Multiple analyses were performed also to verify longer runs extending to more than 20 residues when a cysteine residue was to be analyzed (the PTH derivative of PE-cysteine gives a low yield in sequencing by the Edman degradation).

RESULTS

Sixteen peaks were chosen for analysis from the RP-HPLC separation of reduced IE-purified glutenin (Fig. 1). Peaks 1–6, in the HMW-GS region, were not examined in detail but apparently corresponded to HMW-GS, eluting in the order (numbering according to the system of Payne et al 1984) 10, 5, 18, 1, and 17 (peak 2 was not identified). Peaks 7–16, in the LMW-GS region, were collected separately, dried, reduced, and pyridylethylated. The RP-HPLC separation of the PE fractions 7–15 is shown in Figure 2. PE fraction 16 (not shown in Fig. 2), had a single major peak eluting at 39 min. With the exceptions of peaks 10, 14, and 16, all PE fractions separated into two or three peaks, although most did not give baseline separations. It appears that pyridylethylation of the cysteine residues often permitted further resolution of LMW-GS by RP-HPLC. This may be because the numbers of cysteines differed among the components or were arranged differently in the primary structures so that the changes in surface hydrophobicity resulting from added pyridylethyl groups caused the elution times to differ. Changes in elution times for various wheat proteins after alkylation have been noted by Marchylo et al (1988).

Figure 3 shows one-dimensional SDS-PAGE patterns of the PE fractions of Figure 2, the starting IE-purified glutenin (reduced), and the total seed proteins from Yecora Rojo. The LMW-glutenin subunits fall into two groups (B, mol wt 40,000–50,000; and C, mol wt 30,000–40,000) as classified by Payne and Corfield (1979). PE fractions 7A, 7B, 8A, 10, 11B, 11C, 13B, 14, 15A, 15B, and 16 appear each to have a single major band in SDS-PAGE, although sequencing indicated that
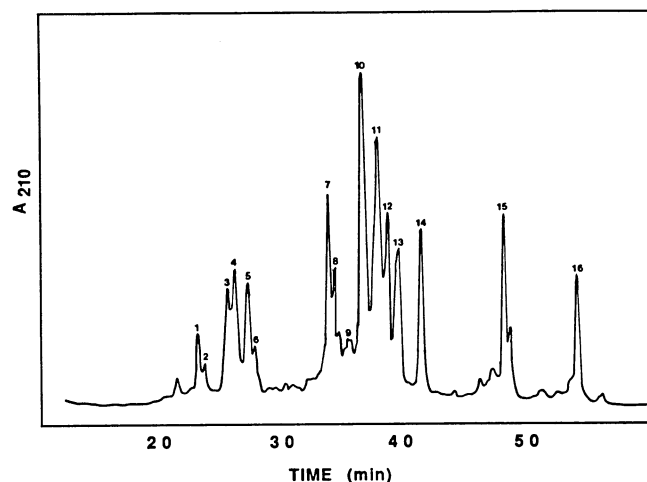

Fig. 1. Reversed-phase high-performance liquid chromatographic separation of reduced ion-exchange-purified glutenin. Separation conditions are described in Materials and Methods. Peaks 7–16 in the low molecular weight region were examined in this study.
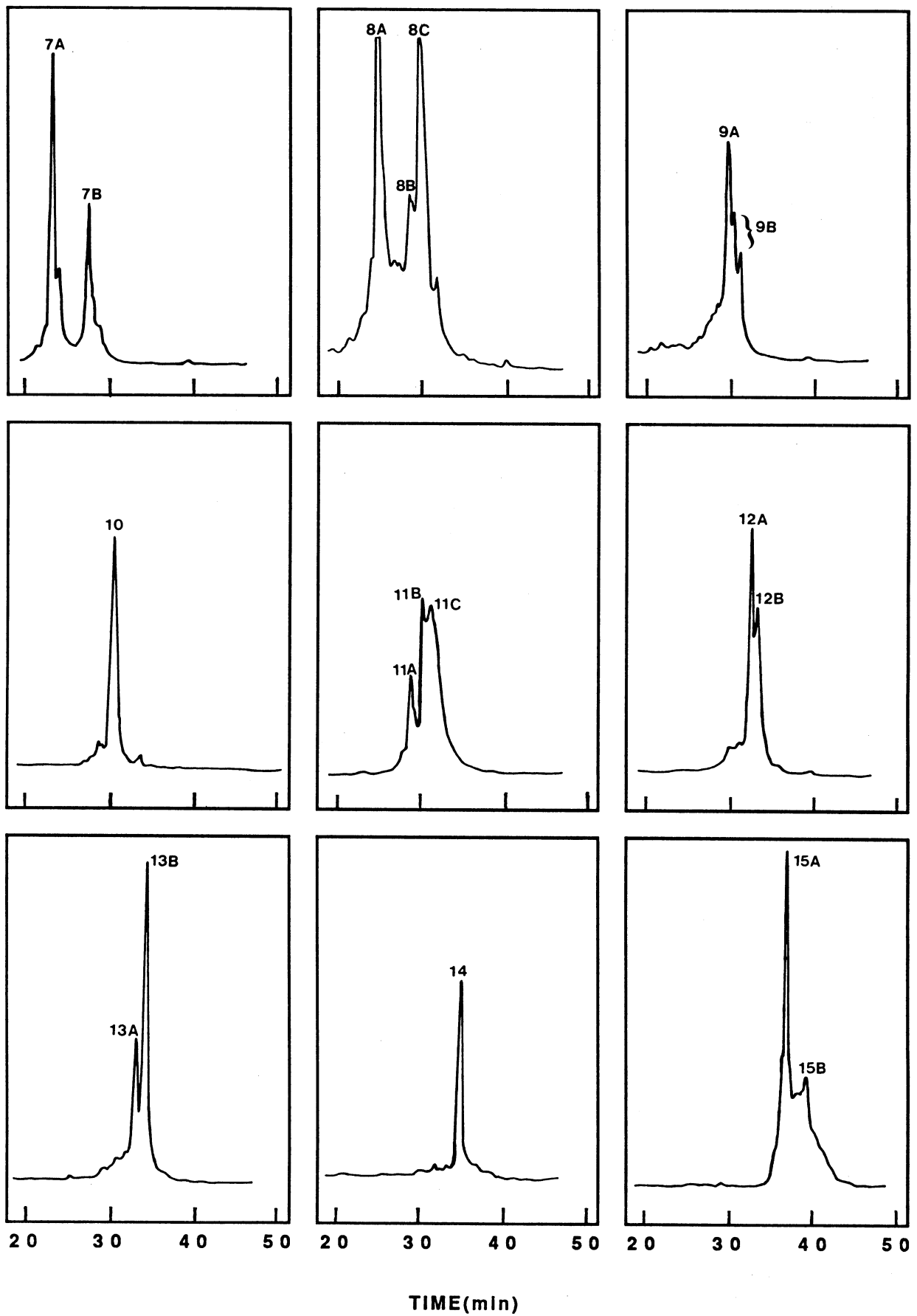
Fig. 2. Reversed-phase high-performance liquid chromatographic separation of reduced and pyridylethylated fractions 7–15. Separation conditions are described in Materials and Methods. PE-16, not shown, had a single major peak eluting at 39 min.

most of these fractions were mixtures, as will be discussed. Minor amounts of higher molecular weight bands appeared in the SDS-PAGE patterns of some fractions, especially fraction 10 (Fig. 3). These bands, which were noticeable only when the samples were heavily loaded, probably correspond to oligomers of the main monomeric component that are cross-linked in some way other than by disulfide cross-linking, as the samples had been reduced and alkylated, and, in addition, 5% mercaptoethanol had been included in the electrophoresis sample buffer. We cannot, however, readily explain the basis of such cross-linking. The second and minor bands (Fig. 3) in PE fractions 12A, 12B, and 13A may be due to overlapping of fractions because baseline separations were not achieved by RP-HPLC (Fig. 2). This was fairly typical of most fractions in our study, which emphasized mainly analysis of overlapping fractions corresponding to almost the entire area of the chromatogram (LMW-GS region). We did not attempt to resolve completely the individual subunits of this highly complex mixture.

N-terminal amino acid sequencing was performed for each of the PE fractions (Table I). The sequence type described by Okita et al (1985), Colot et al (1989), and Cassidy and Dvorak (1991), which has methinonine as the first amino acid in the sequence, was called LMW-m. The sequence type described by Shewry et al (1983), Kasarda et al (1988), and Tao and Kasarda (1989), which has serine as the first amino acid in the sequence, was named LMW-s.

Some fractions from the first chromatographic separation (protein reduced only, Fig. 1) gave only a single peak in the second fractionation as well (protein reduced and alkylated, Fig. 2); furthermore, the fractions had predominantly a single sequence (for example, see results for peak 10). Other fractions, for example, peak 11 (Fig. 1), gave more than a single peak after reduction and alkylation. Thus, the three peaks resulting from chromatography of peak 11 after it had been reduced and alkylated were designated peaks 11A, 11B, and 11C (Fig. 2 and Table I). Furthermore, sequencing of protein corresponding to each of the three peaks gave more than one amino acid at each cycle, indicating that a mixture of protein types was present in each fraction.

Because the main types of sequence for LMW-GS had been established in earlier work (Tao and Kasarda 1989), we were able to determine the approximate proportions of these sequence types in mixtures. This was done by comparing yields of the more stable PTH-amino acids found in the early cycles of each sequence type. The PTH-amino acids especially suitable for this purpose are valine, leucine, isoleucine, and methionine. To illustrate, consider the results for peak 11A. Of the sequence types known to occur in glutenins, only the α-type sequence has valine at positions 1, 3, and 5; thus, the yields of valine (301, 251, and 185 pmol) at these positions indicated the amount of α-type protein in the mixture. Next, only the LMW-s sequence has isoleucine at position 3 and leucine at position 6; the recovery of these high-yielding amino acids (55 and 49 pmol, respectively) indicated the amount of the LMW-s-type subunit present in the mixture. Finally, methionine at cycle 1 (16 pmol) is unique to the LMW-m type of sequence and provides a reasonable estimate of the amount of this sequence type present in the mixture, which was checked by the amount of isoleucine (18 pmol) at position 6. Because no other amino acids showed up in meaningful amounts in the early cycles that could correspond to sequences other than the three indicated, we concluded that fraction 11A contained only these subunits and designated the sequences 11Aa, 11Ab, and 11Ac (Table I) in order of decreasing amounts. Amino acids such as serine, threonine, histidine, and arginine that give a relatively poor yield in the Edman degradation were avoided as the basis for estimation of mixture proportions.

Because of the complexity of some fractions, proportions we determined for those fractions are meant only to be semiquantitative. On the basis of long experience with sequencing mixtures of gluten proteins (see, for example, Autran et al 1979), we suggest that the proportions given in Table I are accurate within a range of about ±10%. A similar approach to analysis of mixture sequences was taken by Falk et al (1991).

DISCUSSION

LMW-s was the major LMW-GS sequence type in our glutenin preparation, on the basis of its occurrence in the various fractions described in Table I together with consideration of the corresponding areas of peaks in Figure 1. The LMW-m type was the next most abundant. It appears that the characterization of LMW-GS based on recombinant DNA approaches (Okita et al 1985, Colot et al 1989, Cassidy and Dvorak 1991) is incomplete at this time as no clone has been described that fits exactly the N-terminal sequence of the predominant LMW-s type, although clear sequence similarities exist between the LMW-m and LMW-s
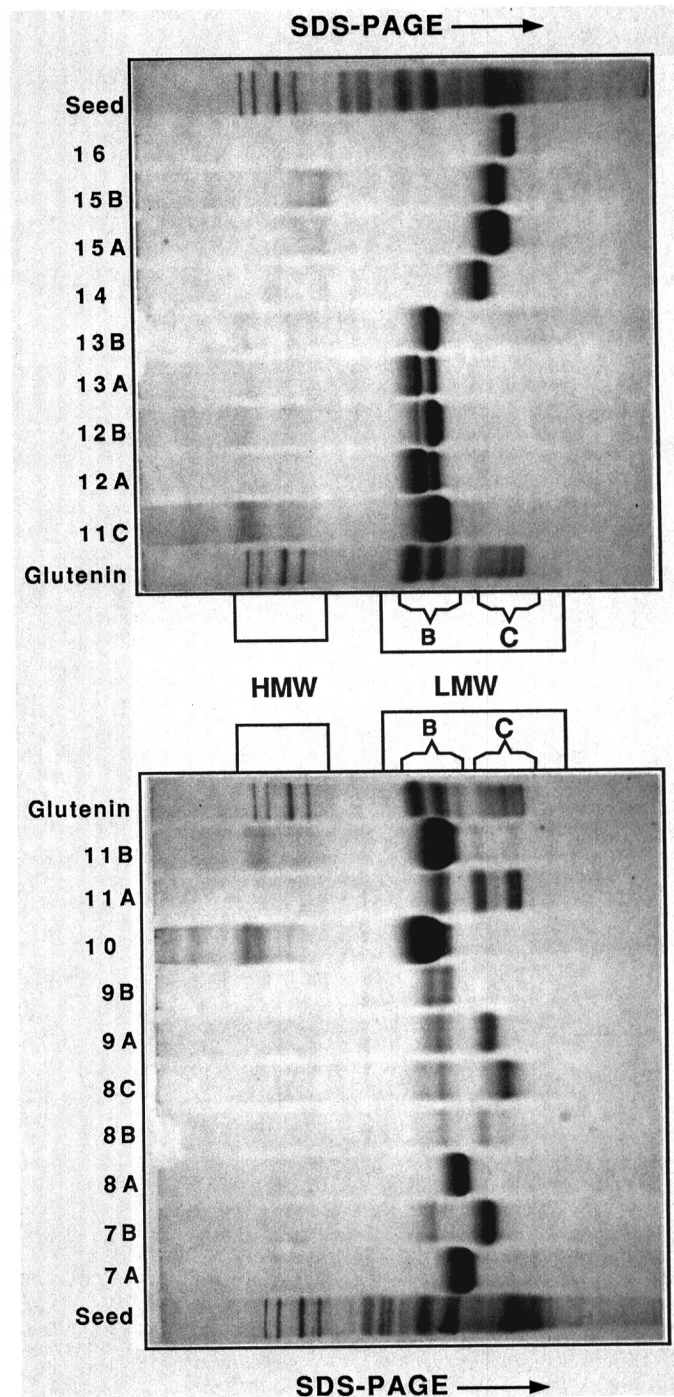

Fig. 3. One-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) of pyridylethylated fractions of Figure 2, the starting ion-exchange-purified glutenin (reduced), and the total seed proteins from Yecora Rojo. HMW = high molecular weight, LMW = low molecular weight.

TABLE I
N-Terminal Sequences of Reduced-Pyridylethylated Fractions

Fraction Number	Sequence Type[a]	Percentage of Fraction	Sequence[b]
			1 10 15
7Aa	α	54	V- R- V- P- V- P- Q- L- Q- P- Q- N- P- S- Q-
7Ab	γ	23	N- M- Q- V- D- P- S- G- Q- V- Q- W- P- Q- Q-
7Ac	LMW-s	16	S- X- I- P- G- L- E- R- P- S- Q- Q- Q- P- L-
7Ad	LMW-m	7	M- E- T- S- (C/H)- I- P- G- L- E- R- P- S- Q- Q-
7Ba	α	31	V- R- V- P- V- P- Q- L- Q- P-
7Bb	γ	28	N- M- Q- V- D- P- S- G- Q- V-
7Bc	LMW-s	25	S- H- I- P- G- L- E- R- P- S-
7Bd	LMW-m	15	M- E- T- S- H- I- P- G- L- E-
8A			Blocked? Hint of α type
8Ba	α	58	V- R- V- P- V- P- Q- L- Q- P-
8Bb	LMW-s	19	S- H- I- P- G- L- E- (R/K)- P- S-
8Bc	γ	15	N- M- Q- V- D- P- S- G- Q- V-
8Bd	LMW-m	8	M- E- T- S- H- I- P- G- L- E-
8Ca	LMW-s	48	S- X- I- P- G- L- E- X- P- S-
8Cb	LMW-m	25	M- E- T- S- X- I- P- G- L- E-
8Cc	α	15	V- X- V- P- V- P- Q- L- Q- P-
8Cd	γ	12	N- M- Q- V- D- P- S- G- Q- V-
9Aa	α	33	V- R- V- P- V- P- Q- L- Q- P-
9Ab	LMW-s	32	S- H- I- P- G- L- E- (R/K)- P- S-
9Ac	LMW-m	29	M- E- T- S- C- I- P- G- L- E-
9Ad	γ	6	N- M- Q- V- D- P- S- G- Q- V-
9Ba	LMW-s	81	S- H- I- P- G- L- E- (R/K)- P- S-
9Bb	LMW-m	19	M- E- T- S- H- I- X- G- L- E-
10	LMW-s	100	S- H- I- P- G- L- E- (R/K)- P- S- Q- Q- Q- P- L- P- P- Q- Q- T- L- X- X- X- (Q/P)- X- Q- Q- P- (I/P)-
11Aa	α	81	V- R- V- P- V- P- Q- L- Q- P- Q- N- P- S- Q- Q- Q- P- Q- (K/E)- Q- V-
11Ab	LMW-s	15	S- H- I- P- G- L- E- (K/R)- P- S- Q- Q- Q- P- L-
11Ac	LMW-m	4	M- E- T- S- (R/H)- I- P- G- L- E-
11Ba	LMW-s	80	S- H- I- P- G- L- E- (K/R)- P- S- Q- Q- Q- P- L- P- L- X- Q- X- L-
11Bb	LMW-m	20	M- E- T- S- H- I- P- G- L- E- K- P- (L/S)- Q- Q- Q- P- X- P- X- Q-
11Ca	LMW-s	62	S- H- I- P- G- L- E- (K/R)- P-
11Cb	LMW-m	38	M- E- T- S- C- I- P- G- L-
12Aa	LMW-s	64	S- H- I- P- G- L- E- (K/R)- P- S-
12Ab	LMW-m	36	M- E- T- S- H- I- P- G- L- E-
12B	LMW-m	100	M- E- T- S- (H/R)- (I/V)- P- G- L- E-
13Aa	LMW-m	55	M- E- T- S- H- (I/V)- P- G- L- E-
13Ab	LMW-s	45	S- X- I- P- G- L- E- R- P- S-
13B	LMW-m	100	M- E- T- S- R- V- P- G- L- E- K- P- S- Q- Q-
14	LMW-m	100	M- E- T- R- C- I- P- G- L- E- R- P- X- Q- Q- Q- P- L- X- (P/Q)-
15A+B	γ	100	N- M- Q- V- D- P- S- G- Q- V- Q- W- P- Q- Q- P- F- P- Q- P- Q- P- F- C- Q- P- Q- R- T- I- (26)
16a	γ	74	N- M- Q- V- D- P- S- G- Q- V- Q- W- P- Q- Q- P- F- P- Q- P- Q- P- F- C- Q- Q-
16b	HMW-1	14	G- G- S- F- Y- P- G- Q- T- T- P- P- Q- Q- L- Q- Q-
16c	HMW-2	12	G- Q- Q- P- G- Q- G- Q- P- G- Q- G- Q- G- Q- P- G- Q- G- Q- Q- G- Q- Q- P-

[a] LMW-m, the sequence type described by Okita et al (1985), Colot et al (1989), and Cassidy and Dvorak (1991) (m represents methionine, the first amino acid of the sequence); LMW-s, the sequence type described by Shewry et al (1983), Kasarda et al (1988), and Tao and Kasarda (1989) (s represents serine, the first amino acid of the sequence); α, the α-type sequence; γ, the γ-type sequence (see Kasarda et al 1983 for α- and γ-type gliadin nomenclature).

[b] The single-letter nomenclature for amino acids is as follows: A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; X, unidentified; Y, tyrosine. More than one letter in a position indicates heterogeneity, with the letter on top representing the major amino acid.

types. Neither type has ever been reported as characteristic of monomeric gliadins.

Furthermore, all DNA clones characterized so far yield a protein sequence with cysteine in the fifth position (Met-Glu-Thr-Ser-Cys- or Met-Glu-Thr-Arg-Cys-). We found this cysteine in some of our LMW-m-type sequences, which we designated LMW-mc5, but sequences having histidine or arginine at the fifth position were about equally common (LMW-mh5 or LMW-mr5). The LMW-s type has no cysteine in the first 20 residues of the N-terminal sequence (see sequence of fraction 10 in Table I).

The locations of cysteine residues in LMW-GS are likely to be of importance in understanding the contributions of these subunits to glutenin properties. These cysteines must participate in either inter- or intramolecular disulfide cross-links. There is no evidence for free sulfhydryls in glutenin (Ewart 1985). To the extent that the sequences of LMW-GS may be compared with those of γ-gliadins, which are known to be monomeric and have only intramolecular disulfide bonds (with exceptions discussed later), the location of a cysteine residue near the N terminus of the LMW-mc5 molecule is anomalous. This cysteine, separated from the other cysteines of the molecule that are located in the C-terminal half of the molecule by repeating sequences (composed mainly of glutamine, proline, and phenylalanine residues), is more likely to be involved in inter- rather than intramolecular disulfide bonding.

Figure 4 illustrates schematically the expected locations of cysteine residues in the LMW-mc5-type subunit. This is based on the identity of our N-terminal sequence with that obtained by Okita et al (1985) from a cDNA clone, B11-33, and an assumption that the remainder of the sequence is similar. This is probably a good assumption for LMW-m-type sequences in general. We compare these with the locations of cysteines in a typical γ-gliadin, represented by a sequence derived from the clone pTag 1436 of Bartels et al (1986) (see also Rafalski 1986).

Cysteines designated 1–5 and 7 in the LMW-mc5 sequence appear to be homologous with the equivalents in γ-gliadins. These cysteines of LMW-mc5 might form intramolecular disulfide bonds, although this is, of course, speculative because the only disulfide bond locations in glutenin known at this time are those described by Wieser et al (1990), which link two y-type HMW-GS.

Cysteine 8 of the γ-gliadins is missing from the LMW-mc5-type subunits, and the cysteine at position 210 of the LMW-mc5-type subunits does not appear to be homologous with cysteine 6 of γ-gliadins. Accordingly, it seems reasonable to suggest that cysteine 210 and the cysteine near the N terminus are involved in intermolecular cross-linking. If this assumption is correct, then subunits of the LMW-mc5 type should correspond to chain extenders (Kasarda 1989) in having at least two cysteines available to form intermolecular disulfide cross-links with other protein molecules.

What then of the LMW-mr- and LMW-mh-type subunits we have found that have no cysteine at position 5 and the LMW-s-type subunits that also have no cysteine near the N terminus? Do these serve as chain terminators (Kasarda 1989) in having only one cysteine available for intermolecular disulfide bond formation, or do they have differences in the number or arrangement of cysteines in parts of the molecule yet to be sequenced that make them chain extenders? This question can only be answered by future work, because almost no information about the arrangement of disulfide linkages in the gluten proteins exists despite the key role they play in determining the properties and the varietal and environmental differences in quality of wheat flour doughs.

In subunits with two cysteines available for intermolecular disulfide bond formation, location of one near the N terminus and the other somewhere in the C-terminal part of the protein molecule might have a different effect on the properties of the resulting glutenin polymers than would location of both in the
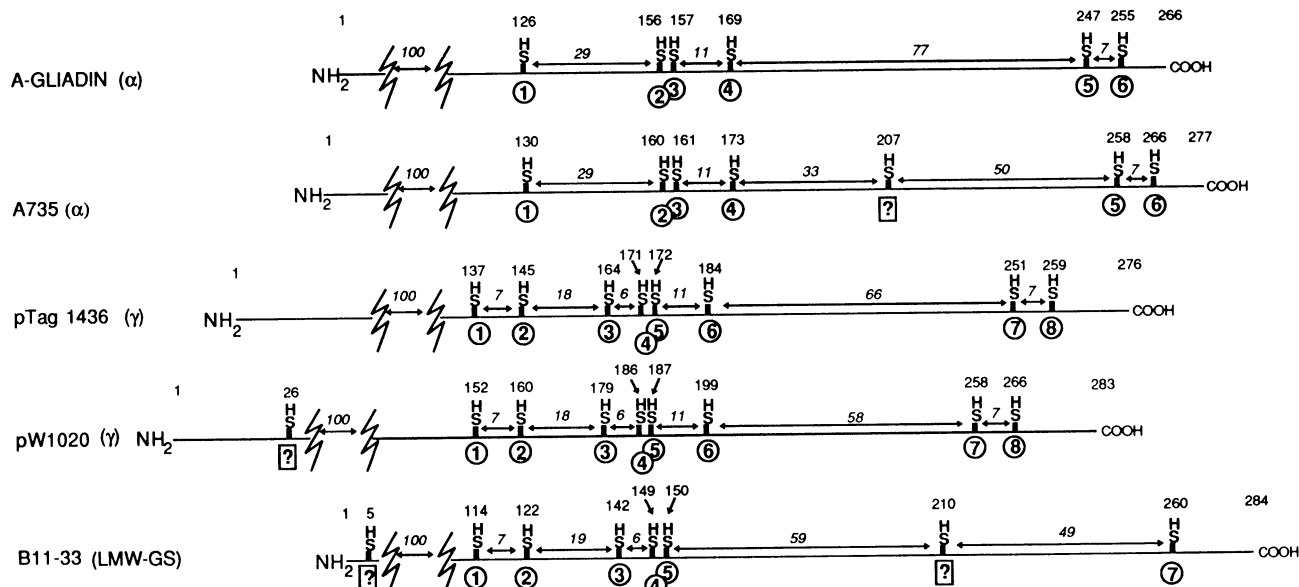

Fig. 4. Schematic depiction of cysteine residue arrangements in the polypeptide chains of an α-type gliadin (A-gliadin) (Kasarda et al 1984); an α-type gliadin, clone A735 (Okita et al 1985); a γ-type gliadin, clone pTag 1436 (Bartels et al 1986); a γ-type gliadin, clone pW1020 (Scheets and Hedgcoth 1988); and a low molecular weight glutenin subunit (LMW-GS), clone B11-33 (Okita et al 1985, see also Cassidy and Dvorak 1991). The A-gliadin sequence was obtained directly; the others are based on DNA sequences. The lengths of the horizontal lines representing the number of amino acid residues in each polypeptide chain are approximately in scale relative to one another, as are the positions of the cysteine residues (-SH) in the chains. A break equivalent to 100 amino acids is inserted in the N-terminal region of each polypeptide chain for scaling purposes. (This N-terminal region corresponds largely to repeating sequences consisting mainly of glutamine, proline, and tyrosine or phenylalanine residues.) Cysteines of normal gliadins are indicated by circled numbers below the line representing the polypeptide chain (numbered 1–8 for γ types as represented here by pTag 1436; 1–6 for α types as represented here by the A-gliadin sequence). Cysteines proposed to be involved in intermolecular cross-linking are not numbered but are indicated by boxed question marks since these have no equivalents in normal gliadins. The sequence positions of cysteine residues are indicated above the -SH symbols representing them, and the number of the C-terminal amino acid in each chain is also indicated. Lengths of intervening sequences between cysteine residues are indicated by italicized numbers and double-arrowed lines. NH₂ represents the amino terminus and COOH represents the carboxyl terminus of each polypeptide chain. Cysteine 2 of the two γ types and the B-type LMW-GS are aligned with one another and with cysteine 1 of each of the two α types on the basis of some apparent similarities in sequences near these cysteines. Additionally, cysteines 4, 5, and 7 of γ types are possibly homologous with cysteines 2, 3, and 5, respectively, of the α types on the same basis.

C-terminal half. For example, the former arrangement might restrict the interactions of the repeating sequence region located in the N-terminal half of the molecule, whereas the latter arrangement might leave this region available to interact with the equivalent region of other molecules.

In agreement with previous work (Tao and Kasarda 1989), we have found clear evidence for the presence in glutenin of subunits with N-terminal sequences corresponding to α-type and γ-type gliadins (Table I, especially fractions 7Aa, 7Ba, 8Ba, 9Aa, 11Aa, 7Ab, 7Bb, 15, and 16a). It has been suggested (Kasarda et al 1987, Kasarda 1989) that these glutenin subunits are related to proteins corresponding to particular DNA clones described in the literature as α-type and γ-type gliadins but that differ from normal α-type and γ-type gliadins in that a serine residue has been mutated to a cysteine residue (Okita et al 1985, Scheets and Hedgcoth 1988). These mutations result in proteins with an odd number of cysteine residues (Fig. 4, sequences based on clones A735 and pW1020), and the extra cysteine will almost certainly have to form an intermolecular disulfide bond, resulting in incorporation of the polypeptide chain into the glutenin fraction. Furthermore, the presence of a single extra cysteine is likely to make these proteins chain terminators unless the extra cysteine prevents formation of one or more intramolecular disulfide bonds.

Fractions 15A+B and 16 had clean γ-type sequences with a cysteine in position 26. This sequence corresponds closely to the DNA-based (Fig. 4, pW1020) γ-type gliadin sequence of Scheets and Hedgcoth (1988). The finding of this sequence provides evidence that the pW1020 γ-type gliadin, which has nine cysteines, is actually a γ-type LMW-GS. Cysteine 26 is likely to form a chain-terminating intermolecular disulfide bond because the other cysteines in the polypeptide chain are similar to those of normal γ-type gliadins (represented by pTag 1436 in Fig. 4) and are likely to form intramolecular disulfide bonds.

Among the fractions sequenced, 11Aa had the most predominant α-type sequence, corresponding to about 80% of the fraction. We hypothesized that if we found a lysine at position 20 of this sequence, it would provide evidence that our α-type LMW-GS corresponded to a clone (A735) described by Okita et al (1985). This clone of Okita et al (1985) corresponded to a protein described as an α/β-type gliadin, but it had an extra cysteine at position 207 and a lysine at position 20. Both of these substitutions were characteristic only of clone A735 and have not been reported in sequences derived from other α-type gliadin clones. Indeed, we found lysine as the predominant amino acid at position 20 of fraction 11Aa, whereas normally glutamic acid would be the only amino acid at position 20 in α-type gliadins. This indicated that the 11Aa α-type gliadin sequence is likely to be a glutenin subunit corresponding to the A735 clone of Okita et al (1985). Some glutamic acid was also present at position 20, which may indicate that the A735 clone is not the only α-type protein capable of becoming part of the glutenin complex.

The protein corresponding to clone A735 (Fig. 4) has the normal complement of six cysteines characteristic of and apparently homologous to those of normal α-gliadins (represented in Fig. 4 by the A-gliadin sequence). Normal α-gliadins are monomeric and have three intramolecular disulfide bonds. The cysteine at position 207 in the A735 sequence contributes a seventh cysteine with no apparent sequence homology to any of the usual α-type gliadin cysteines and, like the cysteine at position 26 of the pW1020 sequence, may form an intermolecular disulfide bond. Again, it seems possible that this subunit will act as a chain terminator, but the possibility remains that the extra cysteine might result in a failure of other cysteines to link up intramolecularly so that they become available for intermolecular bond formation. In the latter case, the subunit would become a branching chain extender.

If there is slight overlap in the time frames for intra- and intermolecular disulfide bond formation during and after biosynthesis of wheat gluten proteins, small amounts of normal gliadins, which mostly form monomers, could conceivably be trapped into the glutenin network through interactions with cysteines of glutenin subunits. Some intermolecular disulfide bonds, such as those linking the x-y-type HMW-GS dimers

(Werner et al 1992) are relatively stable to reduction and may form very rapidly during biosynthesis. At present, however, we have no information about the rapidity with which either intra- or intermolecular bonds form during in vivo synthesis and little understanding of whether or not other proteins that catalyze or guide the process participate in folding and disulfide bond formation.

It seems reasonable to assume that γ-type LMW-GS are coded at the Gli-1 locus on the short arms of group 1 homoeologous chromosomes and that α-type LMW-GS are coded at the Gli-2 locus on the short arms of group 6 chromosomes, because these are where the loci coding for the structurally equivalent gliadin genes can be found (Payne et al 1984). LMW-GS have been assumed to be coded by the Glu-3 locus, which is very closely linked to the Gli-1 locus. Recombination between the two loci is rare, but it seems fairly certain, however, that two distinct loci are involved (Singh and Shepherd 1988). In this study, we found that γ-type and α-type subunits made up about 20 mol % of LMW-GS, and in other work (D. D. Kasarda and E. J.-L. Lew, unpublished results), they made up as much as 40 mol % of the LMW-GS. Accordingly, we suggest that it is not strictly correct to consider LMW-GS as being controlled only by Glu-3. Glu-3 seems likely to code only for B-type LMW-GS (LMW-s and LMW-m). Furthermore, correlations observed between quality test results and proteins coded at the Gli-1 and Gli-2 loci (Metakovsky et al 1990) might have their origins in LMW-GS coded by these loci. The correlations with quality (Gupta and Shepherd 1990) that have been reported for particular combinations of LMW-GS, as defined by their SDS-PAGE patterns, may relate to the structural types and relative proportions of the LMW-GS subunits present.

The subunits with LMW-s-type sequences corresponded *entirely* to B subunits in their SDS-PAGE mobilities, whereas the LMW-m types corresponded *largely* to B subunits in their mobilities. Some LMW-m types corresponded to C subunits in their mobilities. Glutenin subunits of the α-type and γ-type corresponded largely to C subunits in their mobilities, but fraction 7A, which definitely exhibited α-type and γ-type sequences (Table I), fell within the range assigned to B subunits (Fig. 3). As is frequently the case in dealing with the complexity of the wheat gluten proteins, terminology becomes a problem. We tentatively suggest that LMW-s and LMW-m types might be referred to generically as B-type LMW-GS (divorcing the name somewhat from its electrophoretic origin) and the remaining types as α-type or γ-type LMW-GS.

Two minor sequences of fraction 16 (Table I) appeared to correspond to two truncated HMW-GS sequences that could be recognized in the repeat regions of the normal HMW-GS 1Dx2 and 1Dx5. We cannot explain the occurrence of these sequences, which were clearly identified in repeat experiments despite their low proportions in the fraction. They might be the result of some residual proteolytic activity in our preparations or might be the products of truncated genes.

Fraction 8A, which corresponded to only a minor peak in the total preparation, appeared to have a blocked N terminus as it gave only a hint of the α-type sequence when analyzed. In our experiments, we found little indication other than this of blocked components, in contrast to the studies with Chinese Spring LMW-GS (Tao and Kasarda 1989). This might result from the different fractionation approaches used, or it is possible that we might not have recognized that some fractions contained components with blocked N termini. There were general similarities in the types of LMW-GS found in this study for Yecora Rojo in comparison with those found for Chinese Spring by Tao and Kasarda (1989). The essentially nonquantitative nature of the spot-sequencing approach of the latter article made quantitative comparisons difficult.

CONCLUSIONS

We have defined the following main types of LMW-GS in Yecora Rojo on the basis of N-terminal sequences: LMW-s,

LMW-mc5, LMW-mh5, LMW-mr5, γ, and α. The LMW-s type was clearly predominant. We have provided additional evidence for the possibility that γ types and α types are incorporated because of mutations that affect their ability to form only intramolecular disulfide bonds and for the possibility that they might act as chain terminators during formation of glutenin polymers. Furthermore, we suggest that the *Gli-1* and *Gli-2* loci may make significant contributions to glutenin, along with the *Glu-3* loci and, of course, the *Glu-1* loci.

LITERATURE CITED

ANDERSON, O. D., CASSIDY, B., STEFFEN, J., DVORAK, J., and GREENE, F. C. 1991. Structure of the high- and low-molecular weight gene families of the homoeologous group 1 chromosomes of the hexaploid bread wheat cultivar Cheyenne. Pages 512-519 in: Gluten Proteins 1990. W. Bushuk and R. Tkachuk, eds. Am. Assoc. Cereal Chem.: St. Paul, MN.

AUTRAN, J.-C., LEW, E. J.-L., NIMMO, C. C., and KASARDA, D. D. 1979. N-terminal amino acid sequencing of prolamins from wheat and related species. Nature 282:527-529.

BARTELS, D., ALTOSAAR, I., HARBERD, N. P., BARKER, R. F., and THOMPSON, R. D. 1986. Molecular analysis of γ-gliadin gene families at the complex *Gli-1* locus of bread wheat (*T. aestivum* L.). Theor. Appl. Genet. 72:845-853.

BIETZ, J. A., and WALL, J. S. 1980. Identity of high molecular weight gliadin and ethanol-soluble glutenin subunits of wheat: Relation to gluten structure. Cereal Chem. 57:415-421.

BIETZ, J. A., HUEBNER, F. R., and WALL, J. S. 1973. Glutenin, the strength protein of wheat flour. Baker's Dig. 47(1):26-31, 34-35, 67.

CASSIDY, B. G., and DVORAK, J. 1991. Molecular characterization of a low-molecular-weight glutenin cDNA clone from *Triticum durum*. Theor. Appl. Genet. 81:653-660.

COLOT, V., BARTELS, D., THOMPSON, R., and FLAVELL, R. 1989. Molecular characterization of an active wheat LMW glutenin gene and its relation to other wheat and barley prolamin genes. Mol. & Gen. Genet. 216:81-90.

EWART, J. A. D. 1968. A hypothesis for the structure and rheology of glutenin. J. Sci. Food Agric. 19:617-623.

EWART, J. A. D. 1985. Blocked thiols in glutenin and protein quality. J. Sci. Food Agric. 36:101-112.

EWART, J. A. D. 1990. Comments on recent hypothesis for glutenin. Food Chem. 38:159-169.

FALK, K., ROTZSCHKE, O., STEVANOVIC, S., JUNG, G., and RAMMENSEE, H.-G. 1991. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. Nature 351:290-296.

GUPTA, R. B. 1989. Low-molecular-weight subunits of glutenin in wheat and related species: Their characterization, genetics, and relation to bread-making quality. Ph.D. thesis. University of Adelaide: Adelaide, Australia.

GUPTA, R. B., and SHEPHERD, K. W. 1990. Two-step one-dimensional SDS-PAGE analysis of LMW subunits of glutelin. Theor. Appl. Genet. 80:65-74.

GUPTA, R. B., SHEPHERD, K. W., and MacRITCHIE, F. 1991. Genetic control and biochemical properties of some high molecular weight albumins in bread wheat. J. Cereal Sci. 13:221-236.

HOSENEY, R. C., and ROGERS, D. E. 1990. The formation and properties of wheat flour doughs. CRC Crit. Rev. Food Sci. Nutr. 29(2):73-93.

KASARDA, D. D. 1989. Glutenin structure in relation to wheat quality. Pages 277-302 in: Wheat Is Unique. Y. Pomeranz, ed. Am. Assoc. Cereal Chem.: St. Paul, MN.

KASARDA, D. D., AUTRAN, J.-C., LEW, E. J.-L., NIMMO, C. C., and SHEWRY, P. R. 1983. N-terminal amino acid sequences of ω-gliadins and ω-secalins. Implications for the evolution of prolamin genes. Biochim. Biophys. Acta 747:138-150.

KASARDA, D. D., OKITA, T. W., BERNARDIN, J. E., BAECKER, P. A., NIMMO, C. C., LEW, E. J.-L., DIETLER, M. D., and GREENE, F. 1984. Nucleic acid (cDNA) and amino acid sequences of α-type gliadins from wheat (*Triticum aestivum*). Proc. Natl. Acad. Sci. USA 81:4712-4716.

KASARDA, D. D., ADALSTEINS, A. E., and LAIRD, N. F. 1987. γ-Gliadins with α-type structure coded on chromosome 6B of the wheat (*Triticum aestivum* L.) cultivar "Chinese Spring." Pages 20-29 in: Proc. Int. Workshop Gluten Proteins, 3rd. A. Laztity and F. Bekes, eds. World Scientific: Singapore.

KASARDA, D. D., TAO, H. P., EVANS, P. K., ADALSTEINS, A. E., and YUEN, S. W. 1988. Sequencing of a protein from a single spot of a 2-D gel pattern: N-terminal sequence of a major wheat LMW-glutenin subunit. J. Exp. Bot. 39:899-906.

MacRITCHIE, F., KASARDA, D. D., and KUZMICKY, D. D. 1991. Characterization of wheat protein fractions differing in contributions to breadmaking quality. Cereal Chem. 68:122-130.

MARCHYLO, B. A., HATCHER, D. W., and KRUGER, J. E. 1988. Identification of wheat cultivars by reversed-phase high-performance liquid chromatography of storage proteins. Cereal Chem. 65:28-40.

METAKOVSKY, E. V., WRIGLEY, C. W., BEKES, F., and GUPTA, R. B. 1990. Gluten polypeptides as useful genetic markers of dough quality in Australian wheats. Aust. J. Agric. Res. 41:289-306.

OKITA, T. W., CHEESBROUGH, V., and REEVES, C. D. 1985. Evolution and heterogeneity of the α/β-type and γ-type gliadin DNA sequences. J. Biol. Chem. 260:8203-8213.

PAYNE, P. I., and CORFIELD, K. G. 1979. Subunit composition of wheat glutenin proteins isolated by gel filtration in a dissociating medium. Planta 145:83-88.

PAYNE, P. I., HOLT, L. M., JACKSON, E. A., and LAW, C. N. 1984. Wheat storage proteins: Their genetics and their potential for manipulation by plant breeding. Philos. Trans. R. Soc. Lond. B 304:359-371.

RAFALSKI, J. A. 1986. Structure of wheat gamma-gliadin genes. Gene 43:221-229.

SCHEETS, K., and HEDGCOTH, C. 1988. Nucleotide sequence of a γ gliadin gene: Comparisons with other γ gliadin sequences show the structure of γ gliadin genes and the general primary structure of γ gliadins. Plant Sci. 57:141-150.

SHEWRY, P. R., LEW, E. J.-L., and KASARDA, D. D. 1981. Structural homology of storage proteins coded by the *Hor-1* locus of barley (*Hordeum vulgare* L.). Planta 153:246-253.

SHEWRY, P. R., MIFLIN, B. J., LEW, E. J.-L., and KASARDA, D.D. 1983. The preparation and characterization of an aggregated gliadin fraction from wheat. J. Exp. Bot. 34:1403-1410.

SHEWRY, P. R., HALFORD, N. G., and TATHAM, A. S. 1989. The high molecular weight subunits of wheat, barley, and rye: Genetics, molecular biology, chemistry, and role in wheat gluten structure and functionality. Oxf. Surv. Plant Mol. Cell Biol. 6:163-219.

SINGH, N. K., and SHEPHERD, K. W. 1988. Linkage mapping of genes controlling endosperm storage proteins in wheat. 1. Genes on the short arms of group 1 chromosomes. Theor. Appl. Genet. 75:628-641.

SKERRITT, J. H., and ROBSON, L. G. 1990. Wheat low molecular weight glutenin subunits—Structural relationship to other gluten proteins analyzed using specific antibodies. Cereal Chem. 67:250-257.

TAO, H. P., and KASARDA, D. D. 1989. Two-dimensional gel mapping and N-terminal sequencing of LMW-glutenin subunits. J. Exp. Bot. 40:1015-1020.

TARR, G. E. 1986. Manual Edman sequencing system. Pages 155-194 in: Methods of Protein Microcharacterization: A Practical Handbook. J. E. Shively, ed. Humana Press: Clifton, NJ.

WERNER, W. E., ADALSTEINS, A. E., and KASARDA, D. D. 1992. Composition of high-molecular-weight glutenin subunit dimers formed by partial reduction of residue glutenin. Cereal Chem. 69:535-541.

WIESER, H., SEILMEIER, W., and BELITZ, H.-D. 1990. Characterization of ethanol-extractable reduced subunits of glutenin separated by reversed-phase high-performance liquid chromatography. J. Cereal Sci. 12:63-71.