C-Terminal and Internal Sequences of a Low Molecular Weight (LMW-s) Type of Glutenin Subunit

WILLIAM H. VENSEL,[1] GEORGE E. TARR,[2] and DONALD D. KASARDA[1]

Cereal Chem. 72(4):356–359

ABSTRACT

Amino acid sequences have been determined for a 25-residue C-terminal peptide and for two 9-residue peptides, all derived from cyanogen bromide fragmentation of a low molecular weight glutenin subunit (LMW-GS) with the N-terminal sequence: serine-histidine-isoleucine-proline-glycine- (LMW-s type). Previously, only N-terminal amino acid sequences have been available for LMW-s types, which appear to be the predominant LMW-GS type in hexaploid and tetraploid wheats. N-terminal sequences and complete sequences based on DNA sequences have been available, however, for subunits of the LMW-m type. These have the N-terminal sequence: methionine-glutamic acid-threonine-serine-cysteine-. All three peptides prepared from the LMW-s type subunit showed strong sequence similarities to LMW-m type subunits from hexaploid and tetraploid bread wheats. The 25-residue C-terminal fragment was 80% identical to DNA-derived sequences of specific LMW-m types. Two of the peptides contained cysteine and showed homology around these cysteines with LMW-m type sequences. The results support the basic similarity between LMW-s and LMW-m glutenin subunits.

Direct sequence analysis of the low molecular weight glutenin subunit (LMW-GS) proteins of hexaploid and tetraploid wheats indicates that the two main types are coded at the *Glu-3* loci, with some minor types similar to α- and γ-gliadins coded at the *Gli-1* and *Gli-2* loci (Lew et al 1992). The predominant subunit type coded at *Glu-3* has an N-terminal sequence beginning with: serine-histidine-isoleucine-proline-glycine- (Shewry et al 1986, Tao and Kasarda 1989, Wieser et al 1990, Lew et al 1992, Masci et al 1995). No DNA sequences have been reported for this subunit type, designated as LMW-s; the only sequence information available are N-terminal amino acid sequences ranging up to 21 residues in length. Complete DNA sequences have been available only for the other type of subunit, designated LMW-m, that has an N-terminal sequence beginning with: methionine-glutamic acid-threonine-serine-cysteine- (Lew et al 1992).

When the N-terminal amino acid sequence of an LMW-s type subunit (Fig. 1, peak 10; Lew et al 1992) and an LMW-m type amino acid sequence are aligned to maximize homologies, displacing the N-termini by three residues, there is a 76% identity between the aligned regions. Notable in this region is the presence of a cysteine residue at position 5 in some LMW-m types and its absence in the LMW-s type. Quality characteristics of the wheat glutenin proteins are related to the number and arrangement of cysteine residues in their primary structures, which determine their ability to form polymers of protein subunits through intermolecular disulfide bond formation. The absence of a cysteine residue at position 5 in the LMW-s type sequence (with consequent implications for formation of disulfide polymers) and the failure to clone genes corresponding to the LMW-s type sequence seemed sufficient reason to make further structural comparisons between the LMW-s and the LMW-m type subunits.

Additionally, because it would be desirable to have a complete sequence of the LMW-s type subunit, we thought that if sufficient differences at the C-termini were present, this would make it possible to synthesize DNA primers that could be used for PCR amplification of the LMW-s type gene. Accordingly, we purified the protein corresponding to peak 10 (LMW-s type) described by Lew et al (1992) and conducted cyanogen bromide (CNBr) fragmentation. We determined the amino acid sequence of the C-terminal peptides and two smaller peptides, and compared them with published DNA-derived sequences of LMW-m type glutenin subunits.

MATERIALS AND METHODS

Glutenin Preparation

Wheat grain of the *Triticum aestivum* cultivar Yecora Rojo, lot CWC-141, was obtained from the California Wheat Commission. Milling was performed using a Brabender Quadrumat Senior mill (C. W. Brabender, South Hackensack, NJ). Protein content was 11%, based on a 13% moisture content. The white flour fraction (endosperm) was used for protein preparations. Glutenins were prepared after the method of Singh et al (1991). Briefly, 50 mg of flour was placed in a 1.5-ml Eppendorf tube. Gliadins were extracted for 30 min at 65°C with 1 ml of 2-propanol and water (1:1). Samples were vigorously mixed at 10-min intervals. The tube was then centrifuged for 10 min at 5,000 rpm in an Eppendorf centrifuge. The gliadin-containing supernatant solution was removed by aspiration and discarded. Extraction and centrifugation were repeated three times. The final centrifugation was for 15 min at 14,000 rpm.

The gliadin-free pellet was suspended in 1 ml of pH 8.1, 0.1*M* Tris-HCl 2-propanol and water (1:1). Four mg of dithiothreitol was added, the pellet was vortexed, and reduction was performed at 65°C for 30 min. During that interval, the pellet was vortexed three times. After reduction, the pH was lowered by the addition of 100 μl of 88% formic acid, and the suspension was centrifuged for 15 min at 14,000 rpm. The supernatant solution was transferred to a clean Eppendorf tube and centrifuged as before. In preparation for reverse-phase high-performance liquid chromatography (RP-HPLC) purification of the glutenin proteins, 0.5 ml of the centrifuged solution was diluted to 2 ml with 0.1% trifluoroacetic acid (TFA) and loaded into a 2-ml HPLC injection loop.

HPLC Purification

The sample was injected onto an analytical (4.6- × 250-mm) Vydac C-4 reverse-phase column equilibrated with a mixture of 70% of the aqueous solvent (A) and 30% of the organic solvent (B). Five minutes after injection, a linear gradient of 110-min duration, from 30 to 65% solvent B was used to elute the glutenins. Solvent A was 0.1% TFA in water. Solvent B was 0.1% TFA in acetonitrile (MeCN), 2-propanol, and H_2O (6.75:2.25:1) (Vensel et al 1989). The elution volume of the peak corresponding to

[1]U.S. Department of Agriculture, Agricultural Research Service, Western Regional Research Center, 800 Buchanan Street, Albany CA 94710.

Mention of a particular product or instrument is for purposes of information only and is not meant to imply endorsement by the USDA over others that may be suitable.

[2]Department of Cardiology, Children's Hospital, Boston MA. Present address: PerSeptive Biosystems, Inc., 500 Old Connecticut Path, Framingham, MA 01701.

peak 10 of Lew et al (1992) was collected and dried (Savant Speed Vac) in preparation for reduction and alkylation.

Reduction and Alkylation

The peak 10 material was dissolved in 1 ml of 6*M* GuHCl (pH 8.2) 0.1*M* in Tris-HCl. Then 3.2 mg of dithiothreitol was added, and reduction was allowed to proceed for 1 hr at 65°C. The tube was cooled to room temperature, 6.4 μl of 4-vinylpyridine was added, and the contents of the tube was thoroughly mixed. Pyridylethylation was allowed to proceed at room temperature for 30 min. After the addition of 100 μl of 88% formic acid to the reaction mixture, the tube was vortexed and the pyridylethylated proteins were immediately purified by HPLC.

CNBr Fragmentation

CNBr fragmentation of the material corresponding to peak 10 (Lew et al 1992) was performed by dissolving it in 50 μl of 70% formic acid, adding 4 μl of 4*M* CNBr (in MeCN), and allowing the reaction to proceed for 22 hr, at room temperature, in the dark. The reaction mixture was dried, treated with 2 μl of neat aminoethanol (Tarr and Crabb 1983) for 20 min at room temperature, then the aminoethanol was removed under vacuum.

Separation of CNBr Fragments

The dried CNBr fragmentation mixture was dissolved in 20 μl of 0.1% TFA in water and injected onto a 2- × 250-mm Vydac C-4 column. A linear gradient of 30-min duration, starting from 100% of solvent A (0.1% TFA and water) and extending to 60% of solvent B (0.09% TFA and MeCN) at a flow rate of 250 μl/min, was delivered using a high-pressure liquid chromatograph (Hewlett Packard HP1090); detection was made with a Hewlett Packard diode array detector at wavelengths of 205, 215, 275, and 285 nm.

Prediction of HPLC Elution Positions

The elution positions of peptides separated on RP-HPLC are directly related to the peptide amino acid composition and can be predicted from published retention coefficients (Meek 1979). We used the known amino acid sequence of an LMW-m type glutenin subunit (Okita et al 1985) to predict the peptides expected upon CNBr cleavage. Determination of the expected peptides and their relative RP-HPLC elution positions was made with a program written by one of the authors (GET). The program predicted a set of CNBr fragments, their expected elution order at pH 2.1 from a reverse-phase silica-based column, and their UV absorption characteristics. Programs that provide similar information are readily available from commercial sources.

Sequence Determination

Sequence determination of peptides by the Edman degradation method was performed using standard conditions with a protein sequencer (Applied Biosystems 477A). The amount of each peptide loaded was estimated, on the basis of the absorption at 205 nm, to be ~0.15, 0.17, and 0.44 μg. The peptide sample to be sequenced was applied to a preconditioned Biobrene-coated glass fiber filter (Applied Biosystems).

RESULTS AND DISCUSSION

The results of the CNBr fragment separation are shown in Figure 2. The material eluting at ~5 min represents the breakthrough volume of the column. Based on the assumption that the elution pattern would be similar to that expected from peptides of an LMW-m type of glutenin, we predicted that the peaks labeled 1 and 2 (which showed almost no absorption at 275 or 285 nm) would correspond to internal fragments lacking aromatic residues. The peak eluting just after peak 2 was associated with aminoethanol. The UV absorption at 275 and 285 nm and its elution position suggested that peak 3 was the aromatic-containing C-terminal peptide. The large UV-absorbing cluster of peaks eluting between 24 and 28 min contained the remaining CNBr fragments, which were not examined in this work.

Sequence analysis indicated that the amount of protein recovered was 25, 28, and 20 pmol, respectively, for peaks 1, 2, and 3. Thus, based on an assumed initial sequence yield of 50%, at least 50 pmol of the LMW-s protein was obtained from 12 mg of Yecora Rojo flour. The results of the protein sequencing showed each sample to have a single sequence. The repetitive efficiencies were: peptide 1, 89% (glutamine 2 and glutamine 8); peptide 2, 85% (glutamine 2 and valine 9); peptide 3, 91% (valine 3 and valine 22).

The BLAST (Basic Local Alignment Search Tool) network service at the National Center for Biological Information (Altschul et al 1992) was used to conduct a database search of DNA-derived

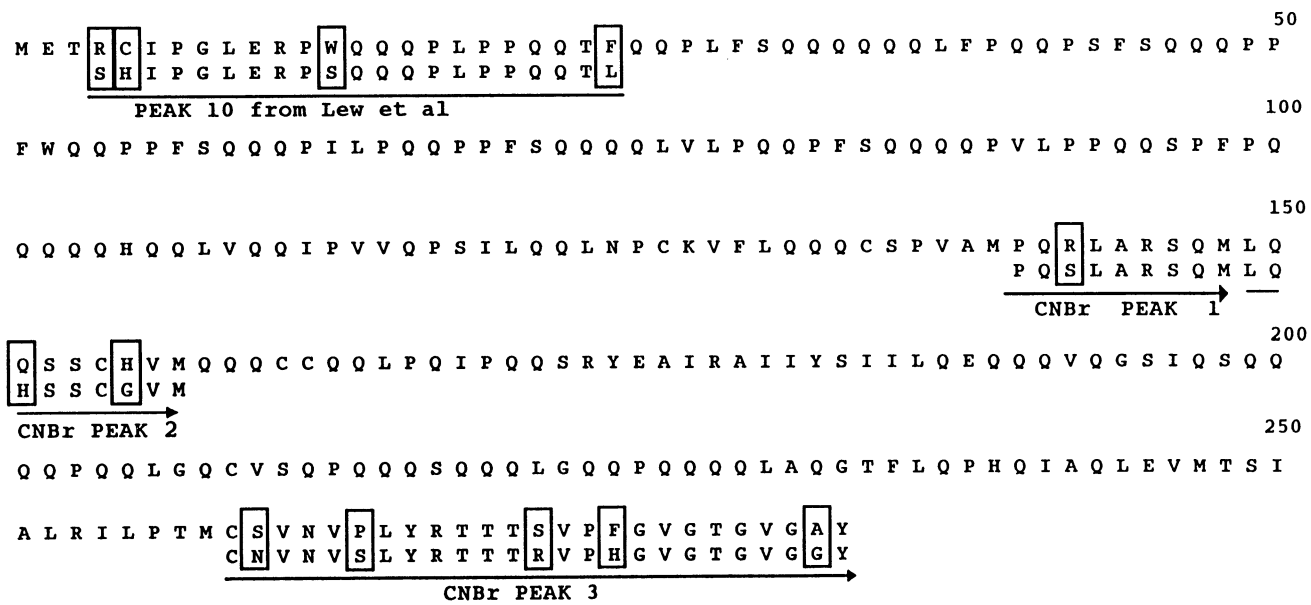

Fig. 1. DNA-derived amino acid sequence of a low molecular weight glutenin subunit (LMW-m) type glutenin (Colot et al 1989), along with the amino acid sequence of the three cyanogen bromide (CNBr) peptides and the N-terminal sequence from the LMW-s glutenin subunit (Lew et al 1992). CNBr fragments obtained in this study (CNBr peaks 1–3) are underscored with an arrow (→), the N-terminal sequence is underscored with a line (—), and positions where the LMW-m and LMW-s sequences differ are boxed. Single letter nomenclature for amino acids: C, cysteine; D, aspartic acid; E, glutamic acid; 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.

protein sequences for comparison with the amino acid sequences of the peptide fragments shown in Figure 1. Some of the results are shown in Figure 3. The DNA-derived sequence with the greatest similarity to the three LMW-s type peptide sequences was the LMW-m type sequence of Colot et al (1989). The sequences of peptides 1, 2, and 3, along with the N-terminal sequence of peak 10 (Lew et al 1992), are shown aligned on the LMW-s sequence in Figure 1. Peptides 1 and 2, which terminated with PTH-homoserine (a derivative of methionine that results from CNBr cleavage at methionine residues), are indicated with the C-terminal residue as methionine. Except for the substitution of serine (S) for arginine (R) at the third position, the peak 1 fragment is identical to a 9-amino acid fragment from the LMW-GS gene of *Triticum aestivum* designated TAGLU1D1 (Colot et al 1989).

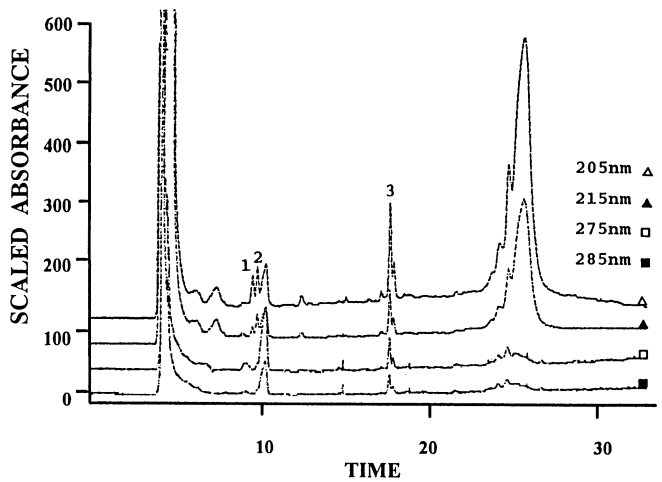

Fig. 2. Elution profile for the cyanogen bromide (CNBr) fragments of the reduced and alkylated peak 10 protein. Large peak with a retention time of ~5 min represents the column breakthrough. Peaks labeled 1, 2, and 3 were analyzed by N-terminal protein sequencing. Four different traces shown in Figure 1 represent detection at four wavelengths (205, 215, 275, and 285 nm). UV absorbancies at each wavelength were adjusted (scaled absorbance) so that the chromatograms could all be fit into the same figure. Chromatographic conditions for the separation are described in the text.

The peak 2 fragment consisted of a 9-residue peptide, seven residues of which are identical to the fragment adjacent to the peak 2 fragment from the same clone. As expected, the peak 3 peptide corresponded to the C-terminal sequence; it showed an 80% identity with the C-terminus of the TAGLU1D1 clone (Colot et al 1989) and with the GDB3_WHEAT sequence of Okita (1984). The C-terminal LMW-m type glutenin sequences from *Triticum durum* reported by D'Ovidio et al (1992) and by Cassidy and Dvorak (1991), and the sequence from *Triticum aestivum* reported by Pitts et al (1988) showed lesser degrees of similarity.

The number of methionines in the TAGLU1D1 clone (Fig. 1) suggests that there should have been a third small peptide of 11 residues (Fig. 1, residues 248–258) with a hydrophobicity that would have caused it to elute between peptides 2 and 3. The number of small peptides released by CNBr thus suggests that the peak 10 LMW-s type subunit we sequenced has a different number or distribution of methionine residues than are found in the TAGLU1D1 sequence, or that the hydrophobicity of the expected peptide was markedly different from that expected from the TAGLU1D1 sequence.

The results presented here demonstrate that there is a high degree of correspondence between C-terminal and N-terminal regions of the LMW-s and the LMW-m type of glutenin subunits and in sequences surrounding two of their cysteine residues (both located in the C-terminal half of the molecule). The absence of cysteine near the region of the N-terminus suggests that the LMW-s types might possess an odd number of cysteine residues, and thus could function as chain terminators. Yecora Rojo, however, has strong mixing characteristics, and the predominance of the LMW-s type sequence (Lew et al 1992) in this variety argues against its function as a terminator of glutenin polymers. Alternatively, there may be an extra cysteine, as yet unidentified, in the LMW-s sequence that would cause it to behave as a chain extender (Masci et al 1995) by providing two cysteine residues available for intermolecular disulfide bond formation.

The high degree of correspondence between both the C-terminal and N-terminal regions of the LMW-s and the LMW-m glutenin subunits suggests that probes for PCR amplification based on these sequences do not seem to be a likely basis for discriminating LMW-s and LMW-m genes. The failure of molecular biologists to find any clones that have an N-terminal sequence corresponding to the LMW-s type is puzzling, but thus may simply be the result

PEPTIDE / LOCUS	SEQUENCE	IDENTITY	Species	Type	ACCESSION NO.	REFERENCE
PEAK 3	CNVNVSLYRTTTRVPHGVGTGVGGY					
	N S R H G					
TAGLU1D1_5	CSVNVPLYRTTTSVPFGVGTGVGAY	80%	T. ae.	LMW-m	X13306	Colot et al, 1989
GDB3_WHEAT	CSVNVPLYRTTTSVPFGVGTGVGAY	80%	T. ae.	LMW-m	P04730	Okita, 1984
TAGLU1D1_3	CSVNVPLYRTTTSVPFGVGTGVGAY	80%	T. ae.	LMW-m	X13306	Colot et al, 1989
GLTB_WHEAT	CSVNVPLYRTTTSVPFGVGTGVGAY	80%	T. ae.	LMW-m	P10386	Colot et al, 1989.
	N S RT R H G					
GDB1_WHEAT	CSVNVPLYSATTSVPFGVGTGVGAY	72%	T. ae.	gamma	P04729	Okita et al, 1985
WHTGLIGBB_1	CSVNVPLYSATTSVPFGVGTGVGAY	72%	T. ae.	gamma	M11336	Okita et al, 1985
	S R R H T G					
GLTA_WHEAT	CNVNVPLYETTTSVPLGVGIGVGVY	76%	T. ae.	LMW-	P10385	Pitts et al, 1988
	N S RT R HG G					
TDGLUT_3	CSVNVPLYSSTTSVPFSVGTGVGAY	68%	T. d.	LMW-m	X51759	Cassidy and Dvorak, 1991
GLTC_WHEAT	CSVNVPLYSSTTSVPFSVGTGVGAY	68%	T. d.	LMW-m	P16315	Cassidy and Dvorak, 1991
	N S RT R H					
TDLMWG_1	CSVNVPLYSSTTSVPFGV	63%	T. d.	LMW-m	X62588	D'Ovidio et al, 1992

Fig. 3. C-terminal amino acid sequence comparison between cyanogen bromide (CNBr) peak 3 fragment of the Yecora Rojo low molecular weight (LMW-s) type glutenin subunit and the 10 top-scoring wheat DNA clone alignments. Differences between groupings (clones with the same degree of identity) and the peak 10 C-terminal fragment are shown above each grouping. Gene locus and the accession number are shown for each fragment. Species abbreviations: T. ae. = *Triticum aestivum;* T. d. = *Triticum durum.* Single letter nomenclature for amino acids as in Figure 1. Database used in this search was the nonredundant PDB+SwissProt+SPupdate+PIR+GenPept (translated GenBank)+GPupdate (cumulative daily updates). Accession numbers starting with X or M refer to the GenBank database; those starting with P refer to the SwissProt database.

of chance. Until such clones are available, more extensive direct sequencing is likely to be the best way to obtain information about the structural differences between LMW-s and LMW-m types of glutenin subunits.

LITERATURE CITED

ALTSCHUL, S. F. W., GISH, W., MILLER, W., MYERS, E.W., and LIPMAN, D. J. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403-410.

CASSIDY, B. G., and DVORAK, J. 1991. The molecular characterization of a low molecular weight glutenin protein 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 prolamine genes. Mol. Gen. Genet. 216:81-90.

D'OVIDIO, R., TANZARELLA, O. A., and PORCEDDU, E. 1992. Nucleotide sequence of a low-molecular-weight glutenin from *Triticum durum*. Plant Mol. Biol. 18:781-784.

LEW, E. J.-L., KUZMICKY, D. D., and KASARDA, D. D. 1992. Characterization of low molecular weight glutenin subunits by reverse-phase high-performance liquid chromatography, sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and N-terminal amino acid sequencing. Cereal Chem. 69:508-515.

MASCI, S., LEW, E. J.-L., LAFIANDRA, D., PORCEDDU, E., and KASARDA, D. D. 1995. Characterization of low-molecular-weight glutenin subunits in durum wheat by RP-HPLC and N-terminal sequencing. Cereal Chem. 72:100-104.

MEEK, J. L. 1979. Prediction of peptide retention times in high-pressure liquid chromatography on the basis of the amino acid composition. Proc. Natl. Acad. Sci. USA 77:1632-1636.

OKITA, T. W. 1984. Identification and DNA sequence analysis of a gamma-type gliadin cDNA plasmid from winter wheat. Plant Mol. Biol. 3:325-332.

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

PITTS, E. G., RAFALSKI, J. A., and HEDGCOTH, C. 1988. Nucleotide sequence and encoded amino acid sequence of a genomic gene region for a low molecular weight glutenin. Nucleic Acids Res. 16:11376.

SHEWRY, P .R., MIFFLIN, B. J., LEW, E. J.-L., and KASARDA, D. D. 1986. The preparation and characterization of an aggregated gliadin fraction from wheat. J. Exp. Bot. 34:205-209.

SINGH, N. K., SHEPHERD, K. W., and CORNISH, G. B. 1991. A simplified SDS-PAGE procedure for separating LMW submits of glutenin. J. Cereal Sci. 14:203-208.

TARR, G. E., and CRABB, J. W. 1983. Reverse-phase high-performance liquid chromatography of hydrophobic proteins and fragments thereof. Anal. Biochem. 131:99-107.

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.

VENSEL, W. H., LAFIANDRA, D., and KASARDA, D. D. 1989. The effect of an organic eluent modifier and pH on the separation of wheat-storage proteins: Application to the purification of γ-gliadins of *Triticum monococcum* L. Chromatographia 28:133-138.

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.

[Received December 12, 1994. Accepted March 22, 1995]