Theor Appl Genet (1991) 81:653–660

THEORETICAL AND APPLIED GENETICS

Molecular characterization of a low-molecular-weight glutenin cDNA clone from *Triticum durum*

B. G. Cassidy * and J. Dvorak

Department of Agronomy and Range Science, University of California, Davis, CA 95616, USA

Received April 24, 1990; Accepted September 19, 1990
Communicated by A. L. Kahler

Summary. A full-length, low-molecular-weight (LMW) glutenin cDNA clone, pTdUCD1, has been isolated from a *Triticum durum* cv 'Mexicali' wheat cDNA library. The complete sequence was determined and compared to the LMW glutenin genes that have been isolated from hexaploid wheat, *Triticum aestivum*. This cDNA codes for a protein of 295 amino acids (33,414 daltons) including a 20-amino acid signal peptide as deduced from the DNA sequence. Northern analysis showed that this cDNA hybridizes to a family of related sequences ranging in length from 1,200 to 1,000 nucleotides. This gene is similar but not identical to previously published LMW glutenin gene sequences. The most striking characteristic of all cloned LMW glutenin genes is the conservation of eight cysteine residues, which could be involved in potential secondary or tertiary structure, disulfide bond interactions. This paper presents a structural map defining distinct regions of the LMW glutenin gene family.

Key words: Low-molecular-weight glutenin subunit – Wheat seed storage proteins

Introduction

Wheat seed storage proteins are important for their effect on the physical properties of dough and on end product quality. In durum wheat they have an effect on pasta firmness and surface properties.

Wheat seed storage proteins are classified as prolamins because of their high proline and glutamine amino acid composition. Prolamins have been subdivided into two groups, gliadins and glutenins, based on their differential solubilities in aqueous alcohols (Kreis et al. 1985).

In their native state, glutenins from large aggregates bonded together by disulfide linkages. This property is an integral part of the viscoelastic property of dough (Wall 1979). Glutenins have been categoried by their mobility following their reduction in SDS-PAGE into high-molecular-weight (HMW) and low-molecular-weight (LMW) subunits. Correlations between specific HMW glutenins and baking quality have been found in common bread wheat (Payne et al. 1979, 1981; Branlard and Dardevet 1985; Moonen and Zeven 1985; Lagudah et al. 1987). However, correlations between baking quality and LMW glutenins have been much more difficult to identify in hexaploid wheat (Payne 1987; Gupta and Shepherd 1988; Gupta et al. 1989). This is mainly due to the difficulty associated with identifying specific LMW glutenin subunits within the complex pattern of proteins expressed in hexaploid wheat and the close linkage of genes that encode them to other seed storage protein genes on homologous group 1 chromosomes (Shepherd 1988).

Similarities have been recently reported at the nucleotide level between all of the wheat seed storage protein genes analyzed to date (Colot et al. 1989). The relatedness and evolutionary divergence between the groups of seed storage proteins at the nucleotide and deduced amino acid level result in the difficulty associated with separating the closely related gene sequences into their proper classes.

The DNA sequence of five LMW glutenin genes, all from bread wheats (*T. aestivum*), have been reported (Bartels and Thompson 1983; Okita 1984; Okita et al. 1985; Pitts et al. 1988; Colot et al. 1989). This paper reports the first characterization of a full-length cDNA clone categorized as a LMW glutenin in durum wheat. By comparing this sequence with the other LMW

* *Present address:* The Samuel Roberts Noble Foundation, P.O. Box 2180, Ardmore, Ok 73402, USA

glutenin gene family members, we have been able to describe the fine structure of this gene family. This should facilitate the future assignment of seed storage protein gene sequences to this family.

Materials and methods

Materials

Restriction endonucleases and other DNA modifying enzymes were purchased primarily from Pharmacia or Promega Biotec and used according to the manufacturer's specifications. Basic cloning procedures were taken from Ausubel et al. (1987).

Cloning and sequencing

A cDNA library [a gift from Dr. S. Thomas (Plant Cell Research Institute, Dublin/CA) and sent by Dr. Natasha Raikhel] was constructed from mRNA from immature seeds of durum wheat (cv Mexicali) and cloned into pARC7 (Alexander 1987). An aliquot was plated onto bacterial media plates and 1,000 colonies were picked and plated onto master plates. The colonies were transferred to nitrocellulose (Schleicher and Schuell) and lysed in situ (Grunstein and Hogness 1975).

Prehybridization and hybridization were carried out according to the manufacturer's recommendations. The colonies were hybridized to a kinase-labeled, 30-mer synthetic oligonucleotide corresponding to nucleotides 629 to 659 from the partial cDNA sequence (Bartels and Thompson 1983).

Plasmids were isolated from selected colonies (Holmes and Quigley 1981). The plasmid DNA was digested with restriction endonucleases to release the cDNA insert. The size of the insert was determined by electrophoresis in agarose gels. The longest cDNA insert, pTdUCD1, was completely sequenced. Unidirectional deletions were constructed from both ends of the insert (Henikoff 1984). A series of overlapping fragments were sequenced by the dideoxynucleotide method (Sanger et al. 1977) utilizing double-stranded DNA, primers specific to the 5′ or 3′ region of the vector adjacent to the cDNA insert, and the sequenase sequencing kit (US Biochemicals). Both strands were entirely sequenced.

Computer analysis

The analysis of this sequence and comparisons to it were carried out on a DEC Vax 11/785 and the sequence analysis package from the Wisconsin genetics computer group (Devereux et al. 1984). Sequences from the NIH-Genbank (TM) Genetic Sequence Databank, release 59, were used in the comparisons between pTdUCD1 and the α/β-gliadins, γ-gliadins, and HMW glutenin subunits.

RNA isolation and Northern analysis

RNA was isolated according to a procedure used by Dr. A. Blechl (personal communication) with minor modifications. The starting material was 10 g of immature seeds 15 days postanthesis. The seeds were ground to a powder in a mortar with liquid nitrogen. Thirty milliliters of NTES (10 m*M* NaCl, 10 m*M* TRIS, pH 8.5, 1 m*M* EDTA, 1.0% SDS) and 20 ml of PCI (phenol:chloroform:isoamyl alcohol; 25:24:1) were mixed together and added to the frozen powder. The phenol had previously been equilibrated with 0.1 *M* TRIS, pH 8.5, 0.1% 8-hydroxyquinolin, and 0.3% β-mercaptoethanol. The emulsion froze upon addition and was slowly ground until it had thawed completely. The emulsion was shaken gently in a 50-ml polycarbonate centrifuge tube for 5 min. The samples were centrifuged for 10 min at $10{,}000 \times g$ at 4 °C in a Sorvall centrifuge. The aqueous layer was transferred to a clean centrifuge tube and reextracted at least three times with an equal volume of PCI each time.

The final aqueous phase was collected and an equal volume of CsCl solution (1.0 g/ml) was added. This was layered over a 13-ml shelf of 5.7 *M* CsCl in a 39 ml quick-seal, polyallomer ultracentrifuge tube. The tubes were centrifuged at 15 °C in a Ti70 rotor at 55,000 rpm for 5.5 h. Following centrifugation, the tops of the tubes were cut off and the CsCl solutions were carefully pipetted off. The RNA pellet was rinsed twice with ice-cold 70% ethanol. The pellet was resuspended in sterile, RNAse-free water, transferred to a clean tube, and reprecipitated by adding 1/10 vol. of 2 *M* NaOAc (pH 5.5) and 2.5 vol. of 100% ethanol. The precipitated RNA was stored at −20 °C until needed. An aliquot of the precipitated RNA was removed, centrifuged at 10,000 rpm in a microcentrifuge for 10 min, and resuspended in sterile water to measure the optical density and prepare it for electrophoresis on a denaturing agarose gel. This procedure yielded approximately 250 µg of total RNA per gram of seeds.

RNA for Northern analysis was separated by electrophoresis on a 1.4% agarose/formaldehyde gel (Ausubel et al. 1987). The RNA was transferred to Hybond membrane (Amersham) and UV-crosslinked by exposure to a Fotodyne UV light box for 3 min. Prehybridization and hybridizations were carried out as suggested by Amersham. The cDNA insert of TdUCD1 was labeled by the specific primer labeling reaction (similar to Feinberg and Vogelstein 1983) to be used as a probe. Briefly, the clone containing the insert was linearized at the 3′ end of the insert by digestion with a restriction endonuclease. The linear plasmid was denatured by boiling for 10 min or alkaline treatment for 5 min at room temperature followed by precipitation. The denatured DNA (25–100 ng) was reannealed with a primer (40-ng, 17-mer oligonucleotide) specific for the 5′ end of the insert. The primer was elongated utilizing a Klenow fragment of DNA polymerase and ^{32}P-αdATP. The specific activity of the probe was between 1×10^9 and 1×10^{10} cpm/µg.

Results

cDNA selection and sequencing

Forty-three cDNA clones were selected from the library by virtue of their hybridization to the oligonucleotide probe. The size of each of the cDNA inserts was determined and ten of the longest were sequenced beginning at their 5′ end. Only one of these cDNAs appeared to contain the entire coding region of a LMW glutenin. The insert from pTdUCD1 was presumed to be full length, because it contains a single open reading frame that begins with a methionine and a characteristically hydrophobic signal sequence. The deduced protein sequence from the 5′ end of pTdUCD1 is highly homologous to the amino terminal sequence of a LMW glutenin protein determined previously (Kasarda et al. 1988). Additionally, the comparison to other wheat seed storage protein sequences matches most closely the sequences from the LMW glutenin gene family that have been published (described below).

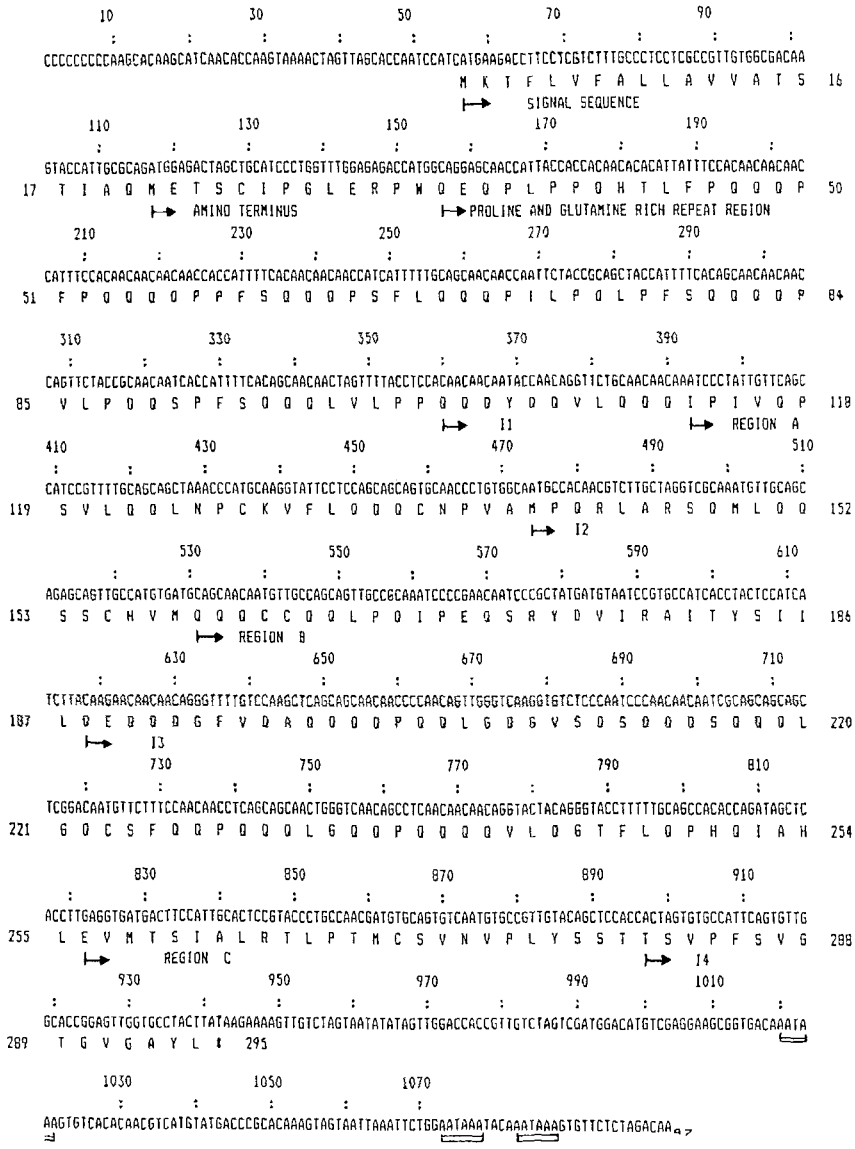

Fig. 1. The entire nucleotide sequence of the cDNA, pTdUCD1. The signal peptide and the start of the mature protein are designated. The termination codon of the protein coding region is marked by an asterisk at amino acid position 296. The consensus polyadenylation signals have been doubly underlined. The translation of the open reading frame is represented by the single letter amino acid codes beneath the DNA sequence. The borders of the subregions of the coding sequence have been designated with *arrows* and *labeled*. The numbers above the sequence are for the DNA sequence and the numbers at the ends of the rows are for the amino acid sequence

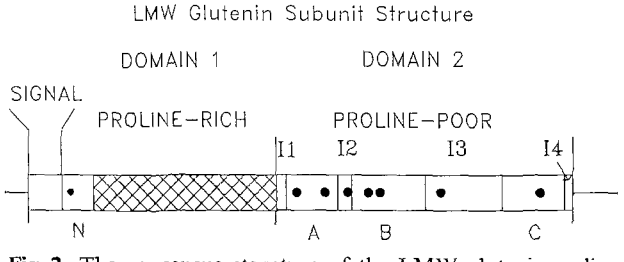

Fig. 2. The consensus structure of the LMW glutenin coding regions. The hatched area represents the repeat region. The *solid circles* represent the consensus position of the cysteine residues

The nucleotide sequence and gene structure of pTdUCD1

The cDNA is 1,103 nucleotides long and includes an open reading frame from nucleotide 57 to 942, which would be translated into a protein of 295 amino acids or 33,414 kDa. The sequence of pTdUCD1 (Fig. 1) appears to fit the general structure for LMW glutenins (aggregated gliadins) proposed by Kreis et al. (1985). There appear to be two major structural domains within the coding region of the gene: (1) domain 1: a repetitive proline and glutamine rich and cysteine poor region, and (2) domain 2: a nonrepetitive proline-poor and cysteine-rich C-terminal region (Fig. 2).

Each region could be subdivided further into distinct subdomains (Fig. 2). From the 5′ end, the gene begins with a hydrophobic signal sequence of 60 nucleotides (20 amino acids). The signal sequence is followed by a distinct amino terminus region of 39 nucleotides, which represents the amino terminus of the mature protein. The region encodes 13 amino acids, including a single cysteine residue at position 5. This region is followed by a proline- and glutamine-rich and cysteine-poor repeat region, which extends 204 nucleotides (68 amino acids). The 582

nucleotide long, nonrepetitive, proline-poor and cysteine-rich sequence (domain 2, Fig. 2) has been subdivided into three regions – A, B, and C – that are conserved among the LMW glutenin gene family, as well as the gliadins, and among seed storage proteins in rye (secalins) and barley (hordeins) (Kreis et al. 1985; Colot et al. 1989). These regions are separated and flanked by regions termed 'intermediate' (I) by Kreis et al. (1985). Regions I1 and I3 contain stretches of polyglutamines, making them higher in glutamine content than regions A, B, or C. Intermediate region I2 is conserved among the LMW glutenin genes but has diverged from the corresponding region of the gliadins, secalins, and hordeins (Kreis et al. 1985; Colot et al. 1989). Region I4 corresponds to the terminal 45 nucleotides (the carboxyl terminus of the protein), which again appear to be conserved within the LMW glutenin genes, but is different from the termini of other families of seed storage genes.

The 3′ nontranslated region of pTdUCD1 is 159 nucleotides long and contains 3 consensus polyadenylation coding sites, AATAAA (Fig. 1). The first signal is 76 nucleotides from the termination codon and the other two are 54 and 64 nucleotides further downstream. The last two signals are within 30 nucleotides of the polyadenosine cDNA tail.

Nucleotide sequence comparisons

The nucleotide sequence of the coding region of pTdUCD1 was compared to the nucleotide sequence of the coding regions of the LMW glutenin genes from *Triticum aestivum* reported to date. pTdUCD1 shares between 93 and 82% homology with the other LMW glutenin cDNA clones characterized (Table 1). The comparison to the two published LMW glutenin genomic sequences demonstrates a wider divergence between LMW glutenin alleles. LMWG-1D1, a Chinese Spring allele, had 92% homology to pTdUCD1, whereas PL1211, a Yamhill allele, had only 72% homology. The nucleotide sequence of the coding region of pTdUCD1 also was compared to the nucleotide sequence of the coding region of the other wheat seed storage proteins in Genbank release 59. Over the total length of the coding regions, the α/β-gliadins and γ-gliadins exhibit homologies on the average of 60 and 67%, respectively. However, stronger homologies exist between the central conserved regions of the nucleotide sequences (regions A, B, C). Only very minor homologies were detected between pTdUCD1 and any HMW glutenin gene.

Amino acid comparison

The amino acid sequences were deduced from the nucleotide sequences of the following LMW glutenin subunit cDNA clones: pTdUCD1, WHTGLIGBA, WHTGLIGBB, WHTGLIGBC, WHTGLG, LMWG-1D1, and LP1211 (Fig. 3A, B, C). The deduced amino acid sequences of the 5′ ends of the complete sequences and the N-terminal protein sequence of a LMW glutenin determined by direct sequencing of a protein isolated from a 2-dimensional gel (Kasarda et al. 1988) are compared in Fig. 3A. The amino termini include a 20-amino acid signal sequence that would not be found in the mature protein sequence reported by Kasarda et al. (1988) or in the partial cDNA sequences reported.

The deduced protein sequences from the LMW glutenin genes have been aligned in Fig. 3B for maximal homology at the amino acid level. There are eight cysteine residues in each of the complete sequences and seven in the partial sequences. In all cases except one, the position of the cysteine residues has been conserved (boxed in Fig. 3B). Six of the cysteine residues are 100% conserved and are clustered in the middle of the protein. One cysteine is found very close to the N-terminus, at amino acid position 5 in the mature protein sequence, in all of the sequences except LP1211, and the other is 25 amino acids from the carboxyl terminus. LP1211 does not have a cysteine at position 5, yet maintains the total number of eight by having a cysteine within region I2 at amino acid position 142 (boxed in Fig. 3B).

The proline- and glutamine-rich repeat region of pTdUCD1 appears to be more internally divergent than that found by others (Colot et al. 1989; Bartels and Thompson 1983; Okita 1984; Okita et al. 1985; Pitts et al. 1988). In pTdUCD1, the repeat region is divided into ten different groups of amino acids varying from pentamers

Table 1. Comparison of *Triticum aestivum* nucleotide sequences to *Triticum durum* cDNA, pTdUCD1

Original designation	Genbank code	Origin	Type	Reference	Homology to pTdUCD1
pB11-13	WHTGLIGBA	cv Cheyenne	cDNA:F	Okita et al. (1985)	93%
pB31	WHTGLIGBB	cv Cheyenne	cDNA:P	Okita et al. (1985)	83%
pB48	WHTGLIGBC	cv Cheyenne	cDNA:P	Okita (1984)	82%
pTAG544	WHTGLG	cv Chinese Spring	cDNA:P	Bartels and Thompson (1983)	87%
LMWG-1D1	not entered	cv Chinese Spring	Genomic	Colot et al. (1989)	92%
LP1211	not entered	cv Yamhill	Genomic	Pitts et al. (1988)	72%

F: Full length
P: Partial

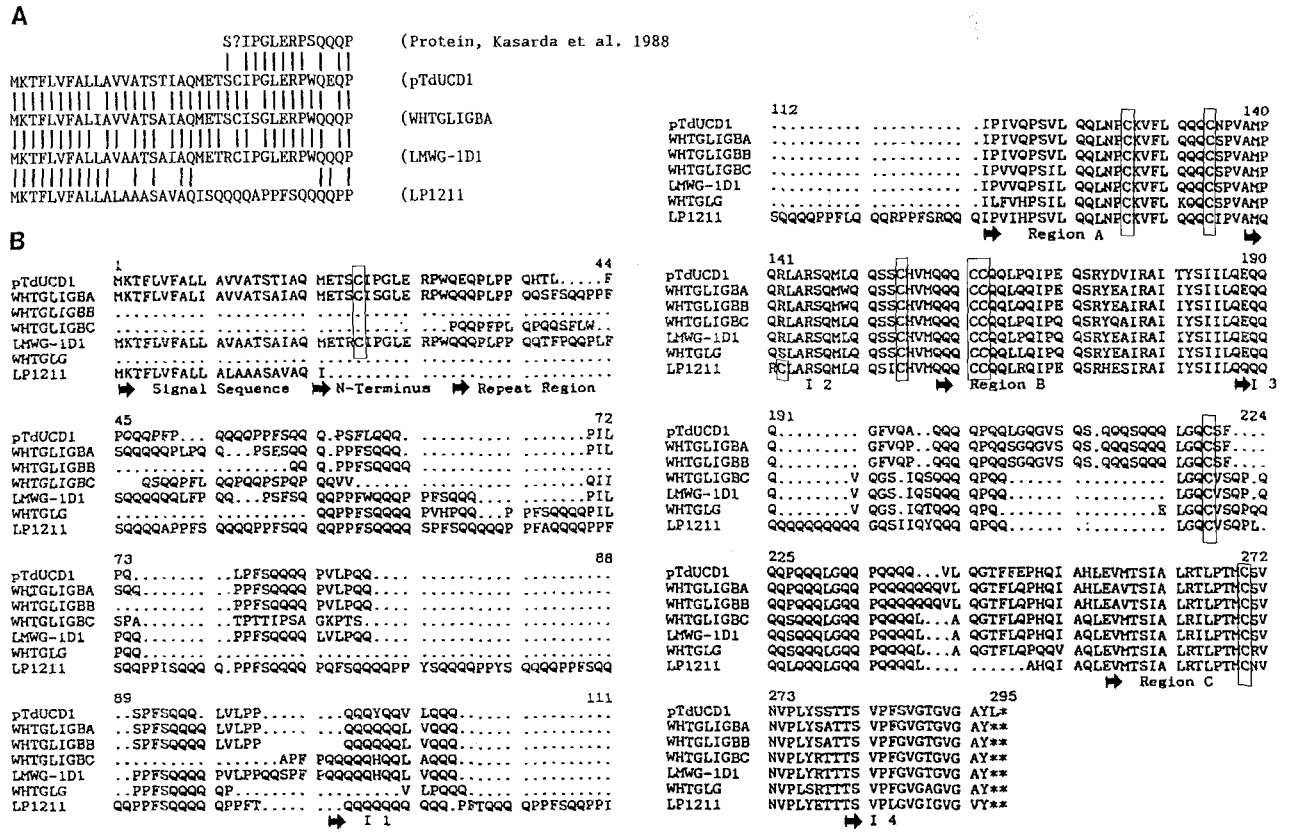

Fig. 3 A–C. Comparison of the deduced amino acid sequences. The standard single-letter amino acid code has been used. **A** The deduced signal peptide and N-terminus amino acid sequences of pTdUCD1, WHTGLIGBA, LMWG-1D1, LP1211, and the N-terminal protein sequence from Kasarda et al. 1988 have been compared. Horizontal lines indicate homology to pTduCD1. **B** the total deduced amino acid sequences of pTdUCD1, WHTGLIGBA, WHTGLIGBB, WHTGLIGBC, LMWG-1D1, WHTGLG, and LP1211 are aligned for maximal homology. *Periods* indicate gaps inserted for maximal homology. The cysteine residues have been *boxed* and the subdomains have been *labeled* below the sequences and delineated with *arrows*. The numbering corresponds to the amino acid sequence of pTdUCD1. **C** A compilation is shown of the repeats from the repeat region of pTdUCD1, WHTGLIGBA, LMWG-1D1, LP1211, WHTGLIGBB, WHTGLIGBC, and WHTGLG

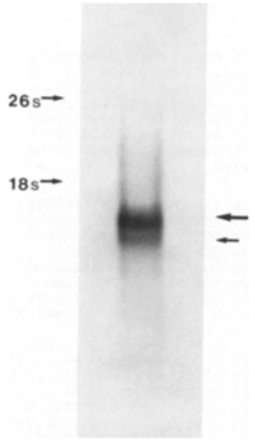

Fig. 4. Northern analysis of total RNA from *Triticum durum* cv. Mexicali seeds. The position of 18s and 26s ribosomal RNA has been labeled. The major hybridizing RNA population is designated with a *bold arrow* and the minor population with a *small arrow*

to an 11-mer, each ending in a string of glutamines. In contrast, WHTGLIGBA, WHTGLIGBB, WHTGLG, LMWG-1D1, and LP1211 consist largely of repeats containing heptamers to nanomers, with a strong homology to $PPFSQQQ_n$ in every other repeat (Fig. 3C). The repeat region of the partial cDNA of WHTGLIGBC does not appear to fit this general model yet is still proline- and glutamine-rich.

mRNA analysis

Northern analysis was carried out to determine the size distribution of the LMW glutenin mRNA (Fig. 4) in a durum wheat. pTdUCD1 hybridized to two distinct RNA populations from immature seeds of cv Mexicali: one RNA population of about 1,200 nucleotides and a second, less abundant, population approximately 1,000 nucleotides long. The size distribution of the mRNA from immature Mexicali seeds agrees with the expected size deduced from the cDNA sequence (about 1,140 nucleotides). The less abundant RNA population detected in the Northern blot may represent a smaller, less efficiently transcribed member of the gene family or an alternative polyadenylation site.

Discussion

The cDNA clone that we characterized from durum wheat cv Mexicali constitutes a full-length member of the LMW glutenin gene family. This was established by the following criteria: (1) it contains one long, open-reading frame that begins with a methionine, (2) it hybridizes to a mRNA population consistent with it being a full-length clone, (3) it contains a hydrophobic signal sequence at its 5′ end, (4) it shares strong homology (11 out of 14 amino acids) to the N-terminal amino acid sequence of a mature LMW glutenin protein from hexaploid wheat (Kasarda et al. 1988). By comparison at the nucleotide or deduced amino acid sequence level, pTdUCD1 is most homologous to other LMW glutenin gene sequences. Finally, 12 amino termini of LMW glutenin proteins from hexaploid wheat have been sequenced from protein spots from a 2-dimensional gel (Tao and Kasarda 1989). One of the protein spots contained an amino terminus identical to that of pTdUCD1.

With the expansion of the sequence data base for the LMW glutenins, it is now possible to construct a more accurate map of the coding region. This map should facilitate the future assignment of sequences to the LMW glutenin subunit gene family. There are several elements which appear to distinguish the LMW glutenins from other seed storage proteins that can be used to identify the type of seed storage protein directly from the nucleotide sequence. The signal sequences and the amino terminus regions appear to be conserved between all the LMW glutenins except for the amino terminus of LP1211 (Fig. 3A). Our determination of the signal sequence length differs slightly from the assignment by Kreis et al. (1985) and Colot et al. (1989), because our determination was based on the results of direct sequencing of a protein whose amino terminus began METS . . . (Tao and Kasarda 1989). This would place the cleavage site between the glutamine and methionine, one amino acid beyond the point proposed by Kreis et al. and three amino acids before that proposed by Colot et al. This is also the point at which LP1211 diverges from the rest of the reported LMW glutenin sequences (Fig. 3A), supporting our assignment of the signal sequence cleavage point.

The proline- and glutamine-rich repeat region (domain I) is clearly the most divergent region among the LMW glutenins. This region is diverging the fastest because of its repetitive structure, which allows for slipping and duplicating or deleting sequences during replication. Despite this heterogeneity, the structure of the repeat region can also be used to distinguish the LMW glutenins from other seed storage proteins as discussed by Colot et al. (1989).

The existence of multiple polyadenylation sites appears to be a common phenomenon among the LMW glutenins. Multiple signals prior to the polyadenosine tail in the cDNAs sequenced occur in all but WHTGLIGBA. pTdUCD1 has three polyadenylation signals, two of which are within 30 nucleotides of the polyadenosine tail. The first polyadenylation signal must be read through to produce the message we have isolated (1,140 nucleotides). However, by Northern analysis, there is a small population of mRNAs which hybridize to pTdUCD1

that are approximately 100 nucleotides smaller. This population could correspond to messages that have terminated at the first polyadenylation signal.

The high degree of sequence homology between the LMW glutenins pTdUCD1, LMWG-1D1, and WHT-GLG may be due in part to the method used to isolate these genes. pTdUCD1 and LMWG-1D1 were both isolated using probes derived from the sequence of WHT-GLG, which may have excluded the isolation of less homologous members of the LMW glutenin gene family. These might include those genes whose amino terminus matches the protein sequence determined by Kasarda et al. (1988) and the genomic clone, LP1211, from Pitts et al. (1988). However, the other LMW glutenin genes cloned were selected randomly from a cDNA library and probably represent the most abundantly expressed genes in cv Cheyenne (Okita et al. 1985). The sequencing of the amino terminus of full-length, LMW glutenin subunit genes reported here, except for the genomic clone LP1211, is highly homologous to the amino terminus protein sequence determined by Kasarda et al. (1988). The different amino terminus sequence and the long polyglutamine stretches appear to set LP1211 apart from the other LMW glutenin genes. It is possible that LP1211 is a member of a second family of LMW glutenins or perhaps a different type of seed storage protein altogether.

Domain 2 contains the most highly conserved regions among all of the LMW glutenin subunits. As pointed out by Kreis et al. (1985) and Colot et al. (1989), regions A, B (subdomain I by Colot et al.), and C (subdomain II by Colot et al.) of domain 2 are highly conserved not only between the wheat seed storage proteins (α/β- and γ-gliadins) but also between γ-secalins and β-hordeins. These regions may represent portions of an ancestral gene that has been conserved to maintain a specific structure or function. These regions contain six of the eight cysteine residues in the LMW glutenin subunits that have the potential for forming inter- and intramolecular disulfide cross linkages, implying that these cysteine residues have been conserved due to some functional role.

The LMW glutenin subunits and the gliadins share several structural features yet have distinct properties. Their amino acid content is very similar and within the A, B, and C regions of domain 2 they are approximately 80% homologous. This homology includes four cysteine residues (nos. 2, 4, 5, and 8 of pTdUCD1) that are 100% conserved. However, unlike the LMW glutenins, the gliadins do not appear to form intermolecular disulfide bonds. The γ/β-gliadins have a total of six cysteines, all within domain 2, whereas the LMW glutenin subunits have seven cysteines within domain 2; and except for the genomic clone LP1211, which has two additional cysteines in domain 2, all have an additional cysteine at position 5 of the mature protein. These two additional cysteine residues could cause a change in the secondary structure of the protein by stabilizing an alternative secondary conformation, which could allow some cysteines to form intermolecular disulfide cross linkages. Alternatively, the additional cysteine residues may themselves form disulfide cross links with other proteins. In either case, some cysteine residue(s) are capable of forming the intermolecular disulfide cross links distinguishing the LMW glutenin subunits from the gliadins. To better understand this relationship, it will be necessary to determine which of the cysteine residues are involved in the formation of the intermolecular crosslinks and which are not involved.

Acknowledgements. The authors thank Dr. D.D. Kasarda for his help regarding the analysis of the protein sequence. They also thank Dr. J. Simmonds for his help in preparing the manuscript and H. Hakim for his help with the computer analysis. This work was supported by a gift from the Anheuser-Busch Company. The cDNA sequence has been given the accession number X-51759 by the EMBL library.

References

Alexander DC (1987) An efficient vector-primer cDNA cloning system. In: Wu R, Grossman L (eds) Methods in enzymology, vol 154. Academic Press, New York, pp 44–64

Ausubel FM, Brent R, Kingston RE, Moore DD, Siedman JG, Smith JA, Struhl K (1987) Current methods in molecular biology. Greene Publishing Associates and Wiley Interscience, New York

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

Branlard G, Dardevet M (1985) Diversity of grain proteins and bread wheat quality. II. Correlation between high-molecular-weight subunits of glutenin and flour quality characteristics. J Cereal Sci 3:345–354

Colot V, Bartels D, Thompson R, 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

Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12:387–395

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Grunstein M, Hogness DS (1975) Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc Natl Acad Sci USA 72:3961–3965

Gupta RB, Shepherd KW (1988) Low-molecular-weight glutenin subunits in wheat: their variation, inheritance, and association with bread-making quality. In: TE Miller, RMD Koebner (eds) 7th Int Wheat Genet Symp Inst Plant Sci Res, pp 943–949

Gupta RB, Singh NK, Shepherd KW (1989) The cumulative effect of allelic variation in LMW and HMW glutenin subunits on dough properties in the progeny of two bread wheats. Theor Appl Genet 77:57–64

Henikoff S (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28:351–359

Holmes DS, Quigley M (1981) A rapid boiling method for the preparation of bacterial plasmids. Anal Biochem 114:193–197

Kasarda DD, Tao HP, Evans PK, Adalsteins AE, Yuen SW (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

Kreis M, Shewry PR, Forde BG, Forde J, Miflin BJ (1985) Structure and evolution of seed storage proteins and their genes with particular reference to those of wheat, barley, and rye. Oxford Surv Plant Mol Cell Biol 2:253–317

Lagudah S, MacRitchie F, Halloran GM (1987) The influence of high-molecular-weight subunits of glutenin from *Triticum tauschii* on flour quality of synthetic hexaploid wheat. J Cereal Sci 5:129–138

Moonen JHE, Zeven AC (1985) Association between high-molecular-weight subunits of glutenin and bread-baking quality in wheat lines derived from backcrosses between *Triticum aestivum* and *Triticum speltoides*. J Cereal Sci 3:97–101

Okita TW (1984) Identification and DNA sequence analysis of a gamma-type gliadin cDNA plasmid from winter wheat. Plant Mol Biol 3:325–332

Okita TW, Chessbrough V, Reeves CD (1985) Evolution and heterogeneity of the α/β-type sequences and γ-type gliadin DNA sequences. J Biol Chem 260:8203–8213

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

Payne PI, Corfield KG, Blackman JA (1979) Identification of a high-molecular-weight subunit of glutenin whose presence correlates with bread-making quality in wheats of related pedigree. Theor Appl Genet 55:153–159

Payne PI, Corfield KG, Holt LM, Blackman JA (1981) Correlations between the inheritance of certain high-molecular-weight subunits of glutenin and bread-making quality in progenies of six crosses of bread wheat. J Sci Food Agric 32:51–60

Pitts EG, Rafalski JA, 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

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Shepherd KW (1988) Genetics of wheat endosperm proteins – in retrospect and prospect. In: Miller TE, Koebner RMD (eds) 7th Int Wheat Genet Symp Inst Plant Sci Res, pp 919–931

Tao P, Kasarda DD (1989) 2-dimensional gel mapping and N-terminal sequencing of LMW glutenin subunits. J Exp Bot 40:1015–1020

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, pp 273–311