Journal of Experimental Botany, Vol. 53, No. 370,
Inorganic Nitrogen Assimilation Special Issue, pp. 947–958, April 2002

Journal of Experimental Botany

Cereal seed storage proteins: structures, properties and role in grain utilization

Peter R. Shewry[1] and Nigel G. Halford

IACR-Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Long Ashton, Bristol BS41 9AF, UK

Received 18 July 2001; Accepted 8 October 2001

Abstract

Storage proteins account for about 50% of the total protein in mature cereal grains and have important impacts on their nutritional quality for humans and livestock and on their functional properties in food processing. Current knowledge of the structures and properties of the prolamin and globulin storage proteins of cereals and their mechanisms of synthesis, trafficking and deposition in the developing grain is briefly reviewed here. The role of the gluten proteins of wheat in determining the quality of the grain for breadmaking and how their amount and composition can be manipulated leading to changes in dough mixing properties is also discussed.

Key words: Cereals, functional properties, gene expression, genetic engineering, protein bodies, proteins.

Introduction

Cereals are the most important crops in the world, with total annual grain yields exceeding 2000 million tonnes (mt), compared with less than 250 mt for legume seeds (including pulses, soybean and groundnut) (FAO, 1999). Although a number of cereal species are grown, three (maize (604 mt in 1998), wheat (589 mt in 1998) and rice (563 mt in 1998)) together account for over 70% of the total production. Other cereals include barley, sorghum, millets (which comprise a number of small-seeded tropical species), oats and rye, in order of decreasing total production.

Cereal grains contain relatively little protein compared to legume seeds, with an average of about 10–12% dry wt. Nevertheless, they provide over 200 mt of protein for the nutrition of humans and livestock, which is about three times the amount derived from the more protein-rich (20–40%) legume seeds. In addition to their nutritional importance, cereal seed proteins also influence the utilization of the grain in food processing. This is particularly important in wheat, which is largely consumed by humans after processing into bread and other foods. It is not surprising, therefore, that cereal seed proteins have been a major topic of research for many years, with the aim of understanding their structures, control of synthesis and role in grain utilization.

Cereal seed storage proteins

The scientific study of cereal grain proteins extends back for over 250 years, with the isolation of wheat gluten first being described in 1745 (Beccari, 1745). Since then more systematic studies have been carried out, notably by TB Osborne (1859–1929) who can be regarded as the father of plant protein chemistry. Osborne developed a classification of plant proteins based on their solubility in a series of solvents, for example, albumins in water, globulins in dilute saline. Although 'Osborne fractionation' is still widely used, it is more usual today to classify seed proteins into three groups: storage proteins, structural and metabolic proteins, and protective proteins. Seed storage proteins fall into three different Osborne fractions and occur in three different tissues of the grain.

Storage globulins

The embryo and outer aleurone layer of the endosperm contain globulin storage proteins, and those from maize embryos have been characterized in some detail (Kriz, 1989, 1999; Kriz and Schwartz, 1986; Kriz and Wallace, 1991; Wallace and Kriz, 1991). These proteins are readily soluble in dilute salt solution and have sedimentation coefficients of about 7. They have limited sequence

[1] To whom correspondence should be addressed. Fax: +44 (0)1275 394299. E-mail: peter.shewry@bbsrc.ac.uk

similarity with, and may be homologous to, the 7S vicilins of legumes and other dicotyledonous plants; they also have similar structures and properties (Kriz, 1999). Related proteins have been found in embryos and/or aleurone layers of wheat, barley and oats (Burgess and Shewry, 1986; Yupsanis *et al*., 1990; Heck *et al*., 1993). 7S globulins from rice embryos have also been characterized (Horikoshi and Morita, 1975), but their relationships to other plant 7S globulins have not been established. The 7S globulins are stored in protein bodies and appear to function solely as storage proteins. However, they do not appear to be absolutely required for normal seed function, at least in maize, where a null mutant behaves normally in terms of development and germination (Kriz and Wallace, 1991). Furthermore, although the aleurone and embryo are rich in proteins compared with the starchy endosperm, the globulins in these tissues have limited impact on the end use properties of the grain. In small grained cereals, such as wheat, the aleurone and embryo account only for about 10% of the grain dry weight and are usually removed by milling (wheat), polishing (rice), pearling (barley) or decortication (sorghum), before human consumption. By contrast, the embryo of maize accounts for 10–11% of the grain and its high contents of protein and oil are important for livestock nutrition.

Storage globulins of 11–12S located in the starchy endosperm are also present in at least some cereal grains. In fact, in oats and rice these proteins form the major endosperm storage protein fraction, accounting for about 70–80% of the total protein. It is now known that these proteins are related to the widely distributed 'legumin' type globulins which occur in most dicotyledonous species (Casey, 1999). The rice proteins are not readily soluble in dilute salt solutions and hence are classically defined as glutelins, but they clearly belong to the 11–12S globulin family. They comprize subunits of M_r approx. 55 000 that are post-translationally cleaved to give acidic (M_r approx. 33 000 in oats, 28–31 000 in rice) and basic (M_r approx. 23 000 and 20–22 000, respectively) polypeptide chains linked by a single disulphide bond (Shotwell, 1999; Takaiwa *et al*., 1999). The oat globulin also resembles the legumins in forming a hexameric structure with a sedimentation coefficient of about 12.

Proteins related to legumins, called 'triticins', are present in starchy endosperm of wheat, although they account only for about 5% of the total seed protein (Singh *et al*., 1988). Triticins consist of large (M_r about 40 000) and small (M_r about 22–23 000) polypeptide chains, but the subunits appear to form dimeric structures rather than the typical legumin hexamers (Singh *et al*., 1988, 1993; Singh and Shepherd, 1985).

The high content of globulin storage proteins in oat grain may contribute to the high nutritional value when compared with other cereals, such as barley and wheat, an important factor in view of the widespread use of oats for livestock feed (Lockhard and Hurt, 1986; Cuddeford, 1995).

Prolamin storage proteins: general properties

With the exceptions of oats and rice, the major endosperm storage proteins of all cereal grains are prolamins. This name was originally based on the observation that they are generally rich in proline and amide nitrogen derived from glutamine, but it is now known that the combined proportions of these amino acids actually vary from about 30–70% of the total among different cereals and protein groups. Similarly, although prolamins were originally defined as soluble in alcohol/water mixtures (e.g. 60–70% (v/v) ethanol, 50–55% (v/v) propan-1-ol or propan-2-ol), some occur in alcohol-insoluble polymers. Nevertheless, all individual prolamin polypeptides are alcohol-soluble in the reduced state. The prolamins vary greatly, from about 10 000 to almost 100 000, in their molecular masses.

It is clear, therefore, that prolamin storage proteins are much more variable in structure than the 7S and 11/12S globulins, and it is possible that the major groups of prolamins in the Triticeae (wheat, barley, rye) and the Panicoideae (maize, sorghum, millets) have separate evolutionary origins. Nevertheless, most prolamins share two common structural features. The first is the presence of distinct regions, or domains, which adopt different structures to each other and may have different origins. The second is the presence of amino acid sequences consisting of repeated blocks based on one or more short peptide motifs, or enriched in specific amino acid residues, such as methionine. These features are responsible for the high proportions of glutamine, proline and other specific amino acids (e.g. histidine, glycine, methionine, phenylalanine) in some prolamin groups.

The prolamin superfamily

Discussions of prolamin structure and properties can be confusing for the non-expert because of the complexity of the fractions and their specialized nomenclature. However, the availability of complete amino acid sequences of representatives of all the major prolamin groups has allowed the redefinition of their classification in relation to structural and evolutionary relationships (Shewry and Tatham, 1990).

This new system of classification assigns all of the prolamins of the Triticeae (wheat, barley and rye) to three broad groups: sulphur-rich (S-rich), sulphur-poor (S-poor) and high molecular weight (HMW) prolamins, with several subgroups within the S-rich group (Table 1). These groups do not correspond directly to the polymeric and monomeric fractions in wheat (glutenins and gliadins, respectively) recognized by cereal chemists, as both

Table 1. *Summary of the types and characteritics of wheat grain prolamins (gluten proteins)*

Components	M_r (% total)	Polymers or monomers	Partial amino acid composition (mol%)
HMW prolamins			
HMW subunits of glutenin	65–90 000 (6–10%)	Polymers	30–35% Gly, 10–16% Pro, 15–20% Gly, 0.5–1.5% Cys, 0.7–1.4% Lys
S-rich prolamins			
γ-gliadins	30–45 000 (70–80%)	Monomers	30–40% Gln, 15–20% Pro, 2–3% Cys, <1.0% Lys
α-gliadins		Monomers	
B- and C-type LMW subunits of glutenin[a]		Polymers	
S-poor prolamins			
ω-gliadins	30–75 000 (10–20%)	Monomers	40–50% Gln, 20–30% Pro, 8–9% Phe, 0–0.5% Lys, 0–<0.5% Cys[b]
D-type LMW subunits of glutenin[a]		Polymers	

[a]C-type LMW subunits are essentially polymeric forms of α- and γ-gliadins and D-type LMW subunits polymeric ω-gliadins. The B-type LMW subunits constitute a discrete group of S-rich prolamins.
[b]Cys is present in D-type LMW subunits, but not ω-gliadins.

monomeric and polymeric forms of S-rich and S-poor prolamins occur.

Structures of typical S-rich, S-poor and HMW prolamins of wheat are summarized in Fig. 1. They all contain extensive repeated sequences based on proline-rich and glutamine-rich motifs with the repeat motifs of the S-rich and S-poor groups being clearly related. Similarly, sequence similarity is clearly present between the non-repetitive domains of the S-rich and HMW prolamins, particularly in positions of conserved cysteine residues and amino acid residues adjacent to these. On the basis of such comparisons, it can be concluded that the S-rich, S-poor and HMW prolamins have a common evolutionary origin. Wider comparisons show further evolutionary and structural relationships to several groups of zein proteins (see below), the prolamins of oats and rice, 2S albumin storage proteins of dicotyledonous seeds, cereal seed inhibitors of α-amylase and trypsin and a range of low M_r cysteine-rich plant proteins including lipid transfer proteins and cereal grain puroindolines. These proteins are, therefore, together defined as the Cereal Prolamin Superfamily of plant proteins (Kreis *et al.*, 1985).

In wheat, the prolamins form the major components of the gluten protein fraction which forms a viscoelastic network in doughs and is largely responsible for the ability to process wheat to form bread, pasta and many other food products.

The prolamins of maize

The prolamins of maize (called zeins) and of other panicoid cereals (e.g. sorghum and many millets) are comprised of one major group of proteins (α-zeins) and several minor groups (β, γ, δ-zeins) (Coleman and Larkins, 1999; Leite *et al.*, 1999) (Fig. 2). Amino acid sequence comparisons demonstrate that the β, γ and δ-zeins are all members of the prolamin superfamily, but only the

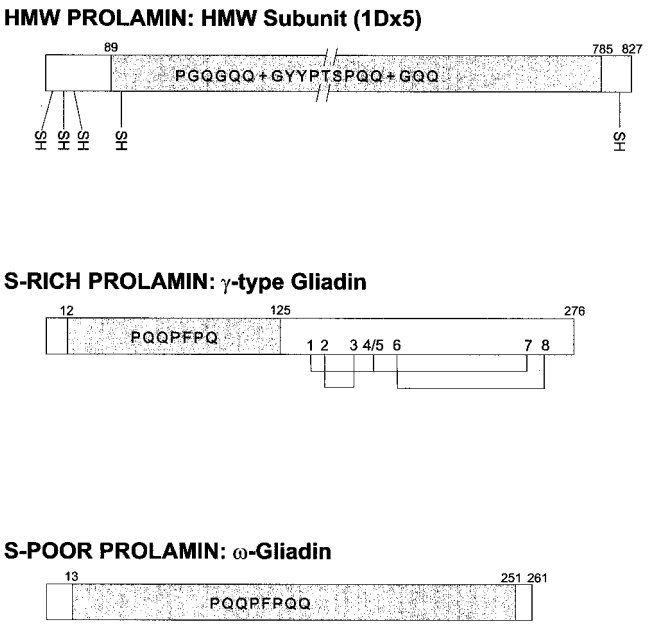

Fig. 1. Schematic structures of typical HMW, S-rich and S-poor prolamins (based on sequences in Anderson *et al.*, 1989; Bartels *et al.*, 1986; Hsia and Anderson, 2001). Repetitive sequences are shaded and disulphide bonds between conserved cysteine residues (1–8) in the γ-gliadin shown as lines. SH denotes the positions of cysteine residues in the HMW prolamins.

γ-zeins contain repeated amino acid sequences (either two or eight tandem repeats of Pro-Pro-Pro-Val-His-Leu). The β-zeins and δ-zeins are both rich in methionine with these residues being clustered in a region close to the *C*-terminus in the former.

By contrast, the α-zeins do not appear to be related to any other prolamins except the α-type prolamins of other panicoid cereals. They consist of two major subclasses called the 19K and 22K zeins based on their M_r determined by SDS-PAGE although they have true molecular masses of 23–24 000 and 26 500–27 000, respectively. Both subclasses contain degenerate repeats of about 20 amino

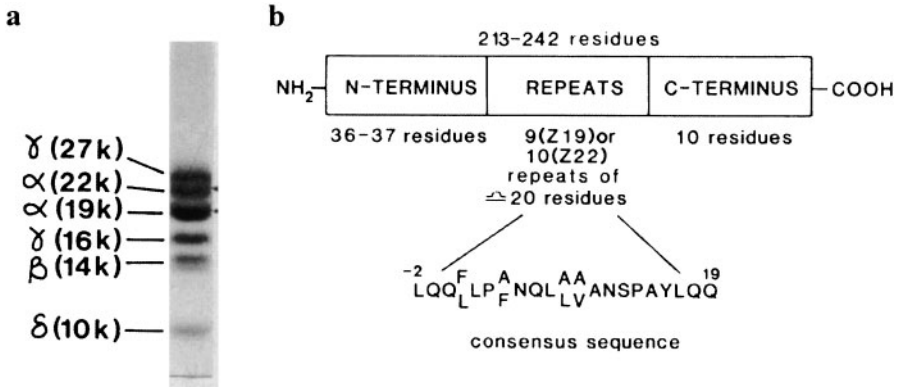

Fig. 2. (a) One-dimensional SDS-PAGE of total zeins of maize. (b) Schematic structures of the M_r 19 000 (Z19) and M_r 22 000 (Z22) α-zeins of maize. (Taken from Shewry and Tatham, 1990, with permission.)

acid residues, with nine such blocks present in the Z19 and ten in the Z22 zeins (Fig. 2).

The α-zeins contain only one or two cysteine residues per molecule and are present in the grain as monomers or oligomers, while the β-, γ- and δ-zeins are all richer in cysteine and form polymers.

Synthesis and deposition of cereal seed storage proteins

Cereal seed storage proteins are produced by the secretory pathway and deposited in discrete protein bodies. However, the origins of the protein bodies and the mechanisms that determine the pathway of storage protein trafficking and deposition are still incompletely understood. The 7S and 11S storage globulins, which are present in the embryo and the aleurone layer and in the starchy endosperm, respectively, of some cereals are believed to follow the same route as the homologous proteins in dicotyledonous seeds. Thus, they are synthesized on rough endoplasmic reticulum (ER) membranes, transported co-translationally into the lumen and then pass via the Golgi apparatus into a specific population of protein storage vacuoles, which differ from the lytic vacuoles that are also present in developing seeds. The precise details of this process have been reviewed elsewhere (Kermode and Bewley, 1999). The precise sorting mechanisms are incompletely understood but physical aggregation within the Golgi appears to be important, leading to the formation of electron dense aggregates which form the contents of dense vesicles. Globulin storage proteins do not contain cleavable pro-domains which confer vacuolar targeting, but non-cleavable segments within the mature protein sequence may be important (Kermode and Bewley, 1999).

The mechanisms of prolamin transport and deposition are less well understood than those of globulins, but two pathways may occur. In maize, other panicoid cereals (sorghum, millets) and rice the prolamins appear to accumulate directly within the lumen of the ER, leading to the formation of discrete protein bodies surrounded by a membrane of ER origin (Coleman and Larkins, 1999; Muench *et al.*, 1999). In rice, this leads to the presence of two populations of protein bodies, PB-I which is of ER origin and contains prolamins and PB-II which is vacuolar in origin and contains globulins/glutelins (Fig. 3a) (Yamagata and Tanaka, 1986; Krishnan *et al.*, 1986). Okita and co-workers also provided evidence that prolamins and globulins/glutelins are synthesized in separate subdomains of the ER, with mRNAs for prolamins being targeted to the rough ER associated with the developing prolamin-containing protein bodies and globulin/glutelin mRNAs to more typical cysternal ER membranes (Li *et al.*, 1993*a*; Choi *et al.*, 2000). More recent work indicates that prolamin mRNA localization results from binding to a specific site on the tubulin and actin cytoskeleton (Muench *et al.*, 1998).

Oats resemble rice in having a high proportion of globulin-type storage proteins in starchy endosperm cells, but in this case the globulins and prolamins are co-located in the same protein bodies, with the prolamins present as lighter-staining inclusions (Fig. 3b) (Lending *et al.*, 1989). The authors suggested that this results from the fusion of two populations of protein bodies, of ER origin (containing prolamins) and of vacuolar origin (containing globulins).

Whereas the prolamins of maize, rice and probably also oats appear to accumulate directly within the ER with little or no evidence for transport to the vacuole, there is now overwhelming evidence that both routes of protein transport and protein body formation operate in wheat, barley and probably also rye. Evidence for this was reviewed elsewhere (Galili, 1997), but includes immunogold labelling of prolamins within Golgi complexes (see also Shewry, 1999), the observation of small protein bodies within or associated with the ER, the association of ER marker enzymes with protein bodies prepared by subcellular fractionation and the expression

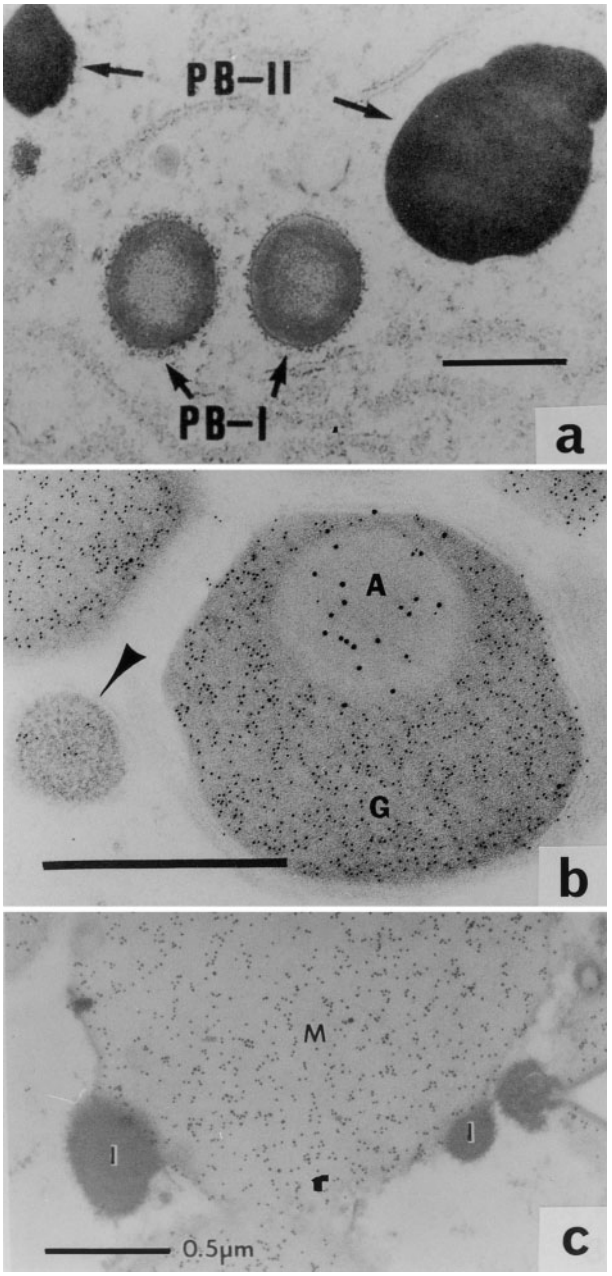

Fig. 3. Protein bodies in developing starchy endosperm cells of cereals. (a) Rice at 7 d after flowering, showing two populations of protein bodies. PB-I are spherical vesicles delimited by a single unit membrane derived from the ER and contain prolamins. PB-II are amorphous, derived from vacuolar deposition and contain globulins/glutelins. (Taken from Yamagata and Tanaka, 1986, with permission.) (b) Oats at 8 d after anthesis showing light-staining deposits of prolamin (labelled with 10 nm colloidal gold) as inclusions in protein bodies containing globulins (5 nm colloidal gold). (Taken from Lending *et al*., 1989, with permission.) (c) Wheat at 11 d after flowering showing inclusions of triticin (dark staining, labelled I) within a matrix of prolamin (M). (Taken from Bechtel *et al*., 1991, with permission.) The bars are 1.0 μM in (a) and 0.5 μM in (b) and (c).

of wild-type and mutant proteins in heterologous systems. The conclusion is that some prolamins, principally gliadins, are transported via the Golgi to the protein storage vacuole whereas others, principally glutenins, are retained within the ER. Galili and co-workers also suggested that the ER-derived protein bodies are subsequently absorbed by protein storage vacuoles in a process similar to autophagy. The protein bodies in developing wheat grains also contain dark-staining inclusions of the globulin storage protein triticin (Fig. 3c) (Bechtel *et al*., 1991), which is presumably transported via the Golgi to the vacuolar protein bodies.

The precise mechanism of protein body fusion in wheat remains to be resolved. However, the net result is the presence in mature, dry endosperm cells of a continuous matrix of proteins that surrounds starch granules and engulfs the remains of other cell structures. This matrix is the basis for the formation of the gluten network when wheat flour is mixed with water to form dough.

The mechanisms which determine whether prolamins are retained within the ER or transported via the Golgi to the vacuole are not known, and it is not possible to recognize either classical ER retention signals (i.e. the *C*-terminal tetrapeptides KDEL or HDEL) or vacuolar targeting sequences in these proteins. Expression of wild-type and mutant forms of γ-zein and γ-gliadin in heterologous systems has demonstrated that the proline-rich repetitive sequences are required for ER retention (Torrent *et al*., 1994; Geli *et al*., 1994; Altschuler *et al*., 1993; Altschuler and Galili, 1994), and it is possible that these regions form protein : protein interactions leading to the formation of insoluble accretions which accumulate directly in the ER rather than being transported to the Golgi and vacuole (Coleman and Larkins, 1999; Shewry, 1999).

Okita and co-workers have proposed the existence of a specific mechanism leading to the retention of prolamins in the ER of rice. This involves an interaction with the molecular chaperone BiP (binding protein) which may bind the nascent polypeptide and retain it in the ER until assembly into a protein body (Li *et al*., 1993*b*; Muench *et al*., 1999). This mechanism has not so far been reported for other cereals.

Organization of proteins within protein bodies

The clear separation of prolamin and globulin components in the biphasic protein bodies of oats (Fig. 3b) and wheat (Fig. 3c) may result from the initial deposition of these components in separate populations of protein bodies (as in rice), which subsequently fuse. However, it could also result from phase separation of the prolamin and globulin proteins due to their different structures and properties.

Evidence for the spatial separation of different types of prolamins within protein bodies is less clear cut as the different prolamin groups tend to show similar staining properties on preparation for electron microscopy. However, Lending and Larkins presented elegant studies

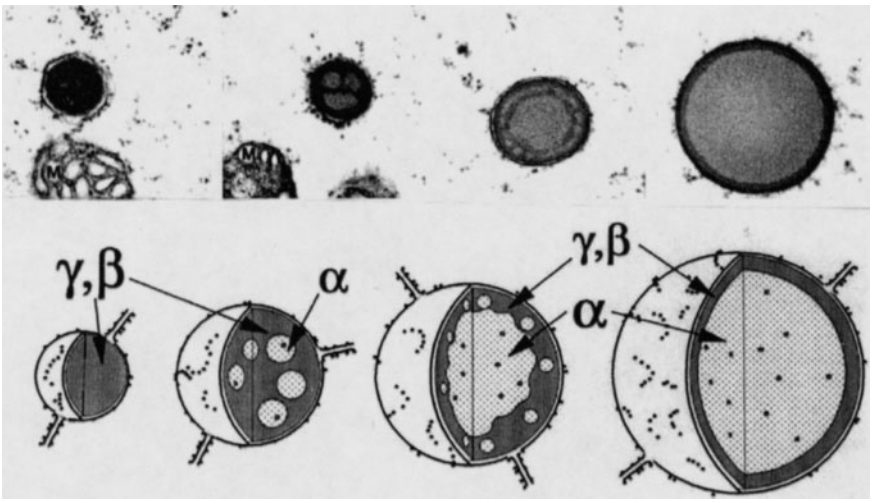

Fig. 4. Developmental pattern of protein body formation in maize endosperm. The most immature protein body is on the left and development is from left to right. Greek letter designations in the lower figure indicate the location of the corresponding zein classes as determined by immunolocalization. (Taken from Coleman *et al*., 1999, with permission.)

showing differences in the distribution of zeins within protein bodies related to their position in the developing endosperm (Lending and Larkins, 1989). Thus, the protein bodies in the youngest cells in the sub-aleurone region contain mainly β- and γ-zeins which are distributed throughout the protein body. Passing into the endosperm the protein bodies increased in size, with the appearance of centrally-located 'locules' containing α-zeins. Finally, these locules fused to form a continuous central region of α-zeins, with the β- and γ-zeins being located peripherally in the mature protein bodies of the central endosperm cells (Fig. 4).

The analysis of protein body development in maize was facilitated by the availability of highly specific antibodies for α, β and γ-zeins, and similar studies have not so far been reported for other cereals. However, recent work carried out with different gliadin fractions indicates that separation could occur, resulting in the formation of microphases. Purified α-gliadin and ω-gliadin fractions were dissolved in 70% (v/v) ethanol, mixed in ratios of 1 : 3, 1 : 1 and 3 : 1, applied to a mica surface and the solvent allowed to evaporate. Analysis of the surface properties of the dried films with atomic force microscopy showed two phases in proportions approximately related to the ratios of the components (McMaster *et al*., 1999) (Fig. 5). This suggests that gliadins and glutenins could partition into separate microphases in protein bodies, which would not be observed by conventional electron microscopy.

Expression in heterologous systems also indicated that a mixture of zein classes may be required for the formation of normal protein bodies. Thus, co-expression of γ-zein was required for the accumulation of α-zein in transgenic tobacco (Coleman *et al*., 1996) while β- and δ-zeins formed abnormal protein bodies when expressed

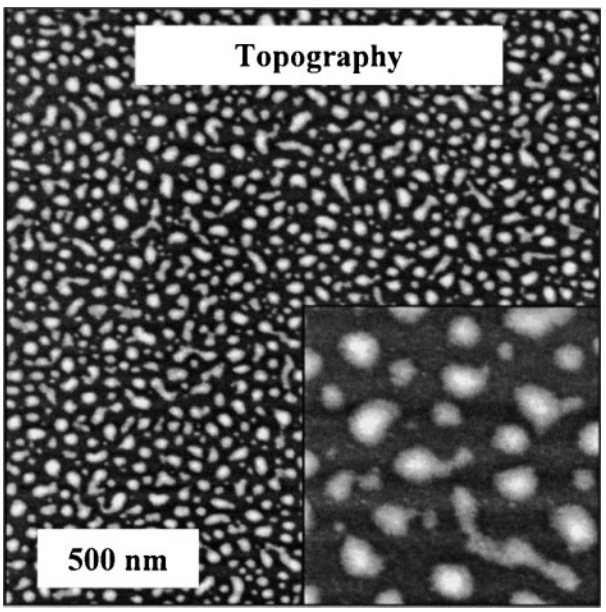

Fig. 5. Topographic image determined by scanning probe microscopy of a 1 : 1 mixture of purified α- and ω-gliadins deposited from solution on a mica surface. Note the presence of two clear phases. (Taken from McMaster *et al*., 1999, with permission.)

alone in tobacco, but apparently normal protein bodies when expressed together (Bagga *et al*., 1995, 1997). It is clear from these studies that there is still much to learn about the mechanisms of protein body formation in cereals and the roles and organization of the different protein groups.

Spatial distribution of prolamins within the starchy endosperm

Because the aleurone cells continue to divide periclinally in the developing cereal endosperm the youngest cells are

present in the sub-aleurone layer and the oldest cells in the central part of the endosperm. In both small grain cereals and panicoid species the sub-aleurone cells contain few starch granules, which tend to be smaller than those in the central endosperm cells. Consequently, these cells contain high proportions of proteins, although the total protein content per cell varies little across the whole wheat endosperm (Evers, 1970). In maize, the protein bodies in the sub-aleurone and outer parts of the endosperm are enriched in γ- and β-zeins and low in α-zeins, the latter being more uniformly distributed across the endosperm (Geetha *et al.*, 1991). This distribution is consistent with the ontogeny of maize protein bodies discussed above.

Differences in protein distribution throughout the endosperm also occur in barley (and probably also in wheat), although the developmental basis for this is unclear. The sub-aleurone cells of both species are rich in proteins, but immunocytochemical and pearling studies of barley show that they contain mainly S-rich and S-poor prolamins (principally B and C hordeins), with the HMW prolamin (D hordein) only occurring in significant amounts below the sub-aleurone (Shewry *et al.*, 1996; Tecsi *et al.*, 2000). This may have significance for the utilization of wheat and barley, as D hordeins are major components of the gel protein fraction that may limit the modification of barley during malting (Smith and Lister, 1983), while the homologous HMW subunits of wheat glutenin are major components of the elastomeric polymers that underpin breadmaking and other food uses (Shewry *et al.*, 1995).

Regulation of prolamin gene expression

Prolamin genes are subject to tissue-specific and developmental regulation, being expressed exclusively in the starchy endosperm during mid- and late-development, and nutritional regulation, responding sensitively to the availability of nitrogen and sulphur in the grain (Duffus and Cochrane, 1992; Giese and Hopp, 1984). This control of gene expression is exerted primarily at the transcriptional level (Bartels and Thompson, 1986; Sørensen *et al.*, 1989).

Little, if anything, is known about the mechanisms by which the expression of prolamin genes responds to sulphur. However, a motif involved in the response of S-poor and S-rich prolamin genes of wheat, barley and rye to nitrogen has been identified. This motif (the N motif or nitrogen element) (Hammond-Kosack *et al.*, 1993; Muller and Knudsen, 1993) is present within a highly conserved sequence called the prolamin box. The prolamin box (sometimes called the endosperm element) was the first prolamin gene regulatory sequence to be reported and was identified through a comparison of the promoters of several gliadin and hordein genes (Forde *et al.*, 1985). This revealed the presence of the conserved sequence, which is approximately 30 bp long, around 300 bp upstream of the transcription start site (it was first called the −300 element). The consensus sequence for the element is: 5′-TGACATGTAA AGTGAATAAG ATGAGTCATG.

The N motif is at the 3′ end of the box and has the consensus sequence G(A/G)TGAGTCAT in S-rich prolamin genes. It is present in the reverse orientation within the prolamin box of S-poor prolamin genes (Shewry *et al.*, 1999). It has similarity with the binding site of the GCN4 transcription factor, which is a component of the nitrogen signalling pathway in yeast, and is sometimes called the GCN4-like motif (GLM). The prolamin box contains a second highly conserved motif, sequence TGTAAAGT, that has been called the endosperm or E motif (Hammond-Kosack *et al.*, 1993).

Promoter regions containing the prolamin box have been shown to be functional by introducing promoter/chloramphenicol acetyl transferase (CAT) reporter gene constructs into transgenic tobacco (Colot *et al.*, 1987; Marris *et al.*, 1988). A regulatory role for the prolamin box itself has been established previously (Müller and Knudsen, 1993; Hammond-Kosack *et al.*, 1993). Müller and Knudsen used an homologous, transient expression system, involving particle bombardment of cultured barley endosperms with C hordein promoter/β-glucuronidase (GUS) constructs. These experiments confirmed that the E and N motifs are separate elements and showed that the N motif exerts a negative effect on gene expression at low nitrogen levels and interacts with the E motif and other upstream elements to give high expression when nitrogen levels are adequate.

Hammond-Kosack *et al.* used *in vivo* footprinting and gel retardation assays to show that E motifs within the prolamin box and further upstream in the promoter of a wheat low molecular weight (LMW) subunit gene bound a putative transcription factor, ESBF-1 (Hammond-Kosack *et al.*, 1993). A second putative transcription factor, ESBF-II, bound the N motif prior to maximum expression of the gene. A third putative transcription factor, SPA, has since been shown to recognize the N motif (Albani *et al.*, 1997).

Together, the results of these experiments suggest that the N motif is an important component of the nitrogen regulatory mechanism for S-rich and S-poor prolamin genes. Its function requires interaction with the E motif, the two motifs together making up the prolamin box. However, the prolamin box is not present in all prolamin genes. Zein gene promoters, for example, contain a highly conserved 15 bp element that has been suggested to act as a tissue-specific enhancer (Quayle and Feix, 1992). It contains the sequence TGTAAAG, which resembles the E motif, but the N motif is absent (Coleman

and Larkins, 1999). The N motif is present in γ-zein promoters, but it is separated from the E motif and its function, if any, has not been investigated (Coleman and Larkins, 1999).

The prolamin box in its entirety is also not present in HMW prolamin gene promoters (Shewry *et al*., 1999). Instead, HMW prolamin promoters contain a major regulatory element (identified by Thomas and Flavell, 1990) which is located in a 38 bp sequence with the consensus: 5′-GTTTTGCAAA GCTCCAATTG CTCCTTGCTT ATCCAGCT.

The location of this sequence is highly conserved in all HMW prolamin promoters, beginning at position −185 to −189 (Shewry *et al*., 1999). The element contains the sequence TGCAAAG, which is similar to the E motif sequence TGTAAAG that is also present in zein genes, but it does not contain anything resembling the N motif.

Sequences corresponding to parts of the N and E motifs are present in HMW prolamin promoters upstream from the major enhancer (Lamacchia *et al*., 2001). However, deletion of these sequences does not appear to affect activity of the promoter, at least when driving reporter gene expression in transgenic tobacco (Halford *et al*., 1989; Thomas and Flavell, 1990).

Cereal seed proteins and grain utilization

The total protein contents of cereal seeds vary from about 10–15% of the grain dry weight, with about half of the total being storage proteins. Nevertheless, proteins have major impacts on the end use properties of the grain. The prolamins, which form the major storage protein fraction in all the major cereals except oats and rice, are deficient in the essential amino acids lysine and in threonine and tryptophan (particularly in maize). This results in nutritional deficiencies in these amino acids when the whole grain are fed to monogastric livestock such as pigs and poultry. It is therefore usual to combine cereals with other sources of these amino acids for animal feed, for example, legume seeds (notably soybean), oilseed cake, fish meal or synthetic amino acids. The combination of cereals and legume seeds is particularly favoured, as these two types of seeds are essentially complementary in their compositions of essential amino acids: cereals tend to be rich in sulphur-containing amino acids and low in lysine and legume seeds vice versa.

The nutritional quality of cereals is not generally an important consideration for human diets in the developed world, although it is still important in some developing countries. The major consideration is the impact of the grain proteins on functional properties for food processing, since the bulk of all cereals, except rice, are consumed in processed foods. Processing quality is particularly important for wheat where the gluten proteins are the major determinant of end use quality.

Gluten proteins and wheat grain quality

Most of the wheat consumed by humans is processed from white flour, which is produced by milling to remove the germ (embryo) and bran (pericarp, testa, nucellar layer, and aleurone layer). It therefore corresponds to the starchy endosperm cells and contains high proportions of starch and gluten.

The gluten proteins form a continuous matrix in the mature dry endosperm cells, as discussed above. When flour is mixed with water to form a dough, the protein matrices in the individual cells are brought together to form a continuous network. This confers visco-elastic properties that allow the dough to be expanded by fermentation and baked into leavened bread or processed into pasta, noodles and a range of other foods.

The molecular basis for the visco-elastic properties of wheat gluten has fascinated cereal scientists for many years as a phenomenon of fundamental interest as well as in relation to improving the end use properties of wheat flours. Of particular importance are the glutenin polymers, and it is well established that strong (i.e. highly visco-elastic) doughs contain high proportions of high molecular mass glutenin polymers (Field *et al*., 1983*b*). However, a breakthrough in understanding came about 20 years ago when Payne and co-workers demonstrated that allelic variation in the composition of the HMW prolamins (the HMW subunits of glutenin) was strongly correlated with differences in the breadmaking quality of European bread wheats (Payne, 1987). This association has been confirmed in many laboratories worldwide and has led to detailed studies of HMW subunit structure and properties.

Cultivars of hexaploid bread wheat have six HMW subunit genes, two each at the *Glu-1* loci on the long arms of the group 1 chromosomes (1A, 1B, 1D). Each of these loci encodes one x-type and one y-type subunit. However, differences in gene expression result in the presence of only three, four or five HMW subunit proteins, with 1Dx, 1Dy and 1Bx subunits being present in all cultivars and 1Ax and/or 1By subunits only in some cultivars. Good breadmaking quality is particularly associated with the presence of a 1Ax subunit (compared with the silent or null allele) and the chromosome 1D-encoded subunit pair 1Dx5 + 1Dy10 (compared with the allelic subunit pairs 1Dx2 + 1Dy12, 1Dx3 + 1Dy12 and 1Dx4 + 1Dy12).

Results from a range of studies are consistent with the hypothesis that the HMW subunits form an elastomeric polymer network which provides a ‘backbone’ for interactions with other glutenin subunits and with gliadins. There is little doubt that this network is stabilized by inter-chain disulphide bonds (Shewry and Tatham, 1997), but Belton has proposed that inter-chain hydrogen bonds,

formed in particular between glutamine residues present in the repetitive domains (Fig. 1), are also important in conferring elasticity (Belton, 1999).

The biological significance of the biophysical properties of gluten is unknown, as the proteins function essentially as grain storage proteins and have no known biological requirement to exhibit visco-elasticity. However, it is probable that the molecular interactions which determine these properties are initially established in the developing grain, because protein bodies isolated during the mid-development stage contain disulphide-linked glutenin polymers (Field *et al*., 1983*a*) which can be demonstrated to exhibit visco-elasticity (authors' unpublished observations). The precise roles of the enzyme, protein disulphide isomerase, in catalysing disulphide bond formation between gluten proteins in the ER and of BiP in establishing other protein : protein interactions remain uncertain (Grimwade *et al*., 1996; DuPont *et al*., 1998; Galili, 1997; Shewry, 1999).

Manipulating HMW subunit composition and wheat grain quality

A number of factors have contributed to the choice of the HMW subunits as an early target for genetic engineering to improve grain quality. They show clear associations with grain quality, with both quantitative effects related to gene expression (e.g. 1Ax subunits) (Halford *et al*., 1992) and qualitative effects associated with allelic differences in subunit structure and properties (e.g. subunits 1Dx5 + 1Dy10). Furthermore, a number of genes encoding subunits associated with good and poor breadmaking quality are available, and the transgene products are readily identified by simple SDS-PAGE of total seed proteins.

Four laboratories have so far reported the expression of HMW subunit transgenes in bread wheat, in each case using either the gene's own promoter or that of another HMW subunit gene (Blechl and Anderson, 1996; Altpeter *et al*., 1996; Barro *et al*., 1997; Alvarez *et al*., 2000). All reported expression of the transgenes at levels up to or exceeding those of the endogenous HMW subunit genes. Transgene expression also appeared to be restricted to starchy endosperm cells, and this was confirmed by the analysis of transgenic wheat lines in which the HMW subunit 1Dx5 gene promoter was used to drive the *Uid*A reporter gene encoding β-glucuronidase (Gus) (Lamacchia *et al*., 2001).

The impact of the HMW subunit transgenes on dough strength has been determined on selected lines grown in the glasshouse (Barro *et al*., 1997; Rooke *et al*., 1999) and field plots (Popineau *et al*., 2001) using the Mixograph which measures the energy input during dough mixing. This showed that the expression of the HMW subunit 1Ax1 transgene in a poor quality background did indeed result in the expected improvement in dough strength. However, when the HMW subunit 1Dx5 gene was highly expressed in either the same background or in a good breadmaking quality background, it had unexpected effects. Flours milled from these lines failed to form a normal dough on hydration and mixing, resulting in decreases in the Mixograph parameters that are usually associated with dough strength. Rheological analysis of glutens isolated from these lines indicated that the connectivity of the gluten network had been greatly increased, by about 100× and 10× in the poor and good quality backgrounds, respectively (Popineau *et al*., 2001), resulting in similar properties to those of gluten modified by transglutaminase (Larre *et al*., 2000).

An increase in the degree of cross-linking between gluten proteins in the transgenic lines may have resulted from the presence of a cysteine residue towards the *N*-terminal end of the repetitive domain of subunit 1Dx5 (Fig. 1) because cysteine residues are not present in equivalent positions in other 1Dx subunits. However, there is also evidence that allelic pairs of HMW subunits such as 1Dx5 + 1Dy10 exist as dimers in the glutenin polymers (Shewry and Tatham, 1997). The expression of high levels of subunit 1Dx5 in the absence of equivalent amounts of subunit 1Dy10 could, therefore, result in a drastic restructuring of the glutenin polymers.

It is clear from the preliminary results discussed above that it is feasible to manipulate the structure and properties of wheat gluten by genetic engineering, although the current knowledge of gluten structure and functionality is still insufficient to predict the results reliably. Bread wheat is currently being transformed with a range of wild-type and mutant genes encoding HMW subunits and other gluten proteins. This should allow the determination of the precise roles of specific proteins and structural features as well as the production of lines with improved properties suitable for incorporation into plant breeding programmes.

Conclusions

The storage proteins of cereals are of immense importance in determining the quality and end use properties of the grain. Understanding the structures of these proteins, their biophysical and functional properties, and the biological mechanisms which determine their synthesis, trafficking and deposition in the grain is important to underpin future attempts to improve the end use quality of grain by genetic engineering.

Acknowledgement

IACR receives grant-aided support from the Biotechnology and Biological Sciences Research Council of the United Kingdom.

References

Albani D, Hammond-Kosack MCU, Smith C, Conlan S, Colot V, Holdsworth M, Bevan MW. 1997. The wheat transcriptional activator SPA: a seed-specific bZIP protein that recognizes the GCN4-like motif in the bifactorial endosperm box of prolamin genes. *The Plant Cell* **9,** 171–184.

Altpeter E, Vasil V, Srivastava V, Vasil IK. 1996. Integration and expression of the high-molecular-weight glutenin subunit 1Ax1 gene into wheat. *Nature Biotechnology* **14,** 1151–1159.

Altschuler Y, Galili G. 1994. Role of conserved cysteines of a wheat gliadin in its transport and assembly into protein bodies in *Xenopus* oocytes. *Journal of Biological Chemistry* **269,** 6677–6682.

Altschuler Y, Harel R, Galili G. 1993. Role of the *N*- and *C*-terminal regions of a wheat γ-gliadin in its transport via the endoplasmic reticulum of *Xenopus* oocytes. *The Plant Cell* **5,** 443–450.

Alvarez ML, Guelman S, Halford NG, Lustig S, Reggiardo MI, Ryabushkina N, Shewry P, Stein J, Vallejos RH. 2000. Silencing of HMW glutenins in transgenic wheat expressing extra HMW subunits. *Theoretical and Applied Genetics* **100,** 319–327.

Anderson OD, Greene FC, Yip RE, Halford NG, Shewry PR, Malpica-Romero J-M. 1989. Nucleotide sequences of the two high-molecular-weight glutenin genes from the D-genome of a hexaploid bread wheat, *Triticum aestivum* L. cv. Cheyenne. *Nucleic Acids Research* **17,** 461–462.

Bagga S, Adams H, Kemp JD, Sengupta-Gopalan C. 1995. Accumulation of 15-kilodalton zein in novel protein bodies in transgenic tobacco. *Plant Physiology* **107,** 13–23.

Bagga S, Adams H, Rodriquez FD, Kemp JD, Sengupta-Gopalan C. 1997. Co-expression of the maize δ- and β-zein genes results in stable accumulation of δ-zein in ER-derived protein bodies formed by β-zein. *The Plant Cell* **9,** 1683–1696.

Barro F, Rooke L, Békés F, Gras P, Tatham AS, Fido R, Lazzeri PA, Shewry PR, Barceó P. 1997. Transformation of wheat with high molecular weight subunit genes results in improved functional properties. *Nature Biotechnology* **15,** 1295–1299.

Bartels D, Thompson RD. 1986. Synthesis of messenger-RNAs coding for abundant endosperm proteins during wheat grain development. *Plant Science* **46,** 117–125.

Bartels D, Altosaar I, Harberd NP, Barker RF, Thompson RD. 1986. Molecular analysis of γ-gliadin gene families at the complex *Gli-1* locus of bread wheat (*T. aestivum* L.). *Theoretical and Applied Genetics* **72,** 845–853.

Beccari. 1745. *De Frumento*. De Bononiensi Scientiarum et Artium Instituto atque Academia Commentarii, II. Part I., 122–127.

Bechtel DB, Wilson JD, Shewry PR. 1991. Immunocytochemical localization of the wheat storage protein triticin in developing endosperm tissue. *Cereal Chemistry* **68,** 573–577.

Belton PS. 1999. On the elasticity of wheat gluten. *Journal of Cereal Science* **29,** 103–107.

Blechl AE, Anderson OD. 1996. Expression of a novel high-molecular-weight glutenin subunit gene in transgenic wheat. *Nature Biotechnology* **14,** 875–879.

Burgess SR, Shewry PR. 1986. Identification of homologous globulins from embryos of wheat, barley, rye and oats. *Journal of Experimental Botany* **37,** 1863–1871.

Casey R. 1999. Distribution and some properties of seed globulins. In: Shewry PR, Casey R, eds. *Seed proteins*. Dordrecht: Kluwer Academic Publishers, 159–169.

Choi S-B, Wang C, Muench DG, Ozawa K, Franceschi VR, Wu Y, Okita TW. 2000. Messenger RNA targeting of rice seed storage proteins to specific ER subdomains. *Nature* **407,** 765–767.

Coleman CE, Larkins BA. 1999. The prolamins of maize. In: Shewry PR, Casey R, eds. *Seed proteins*. Dordrecht: Kluwer Academic Publishers, 109–139.

Coleman CE, Herman EM, Takasaki K, Larkins BA. 1996. The maize γ-zein sequesters α-zein and stabilizes its accumulation in protein bodies of transgenic tobacco endosperm. *The Plant Cell* **8,** 2335–2345.

Colot V, Robert LS, Kavanagh TA, Bevan MW, Thompson RD. 1987. Localization of sequences in wheat endosperm protein genes which confer tissue-specific expression in tobacco. *The EMBO Journal* **6,** 3559–3564.

Cuddeford D. 1995. Oats for animal feed. In: Welch RW, ed. *The oat crop: production and utilization*. London: Chapman & Hall, 321–368.

Duffus CM, Cochrane MP. 1992. Grain structure and composition. In Shewry PR, ed. *Barley: genetics, biochemistry, molecular biology and biotechnology*. Wallingford: CAB International, 291–317.

DuPont FM, Hurkman WJ, Tanaka CK, Chan R. 1998. BiP, HSP70, NDK and PDI in wheat endosperm. I. Accumulation of MRNA and protein during grain development. *Physiologia Plantarum* **103,** 70–79.

Evers AD. 1970. Development of the endosperm of wheat. *Annals of Botany* **34,** 547–555.

FAO. 1999. *FAO Yearbook Production* **53**.

Field JM, Shewry PR, Burgess SR, Forde J, Parmar S, Miflin BJ. 1983*a*. The presence of high molecular weight aggregates in the protein bodies of developing endosperms of wheat and other cereals. *Journal of Cereal Science* **1,** 33–41.

Field JM, Shewry PR, Miflin BJ. 1983*b*. Solubilization and characterization of wheat gluten proteins; correlations between the amount of aggregated proteins and baking quality. *Journal of the Science of Food and Agriculture* **34,** 370–377.

Forde BG, Heyworth A, Pywell J, Kreis M. 1985. Nucleotide sequence of a B1 hordein gene and the identification of possible upstream regulatory elements in endosperm storage protein genes from barley, wheat and maize. *Nucleic Acids Research* **13,** 7327–7339.

Galili G. 1997. The prolamin storage proteins of wheat and its relatives. In: Larkins BA, Vasil IK, eds. *Cellular and molecular biology of plant seed development*. The Netherlands: Kluwer Academic Publishers, 221–256.

Geetha KB, Lending CR, Lopes MA, Wallace JC, Larkins BA. 1991. Opaque-2 modifiers increase γ-zein synthesis and alter its spatial distribution in maize endosperm. *The Plant Cell* **3,** 1207–1219.

Geli MI, Torrent M, Ludevid D. 1994. Two structural domains mediate two sequential events in γ-zein targeting: protein endoplasmic reticulum retention and protein body formation. *The Plant Cell* **6,** 1911–1922.

Giese H, Hopp E. 1984. Influence of nitrogen nutrition on the amount of hordein, protein Z and β-amylase messenger RNA in developing endosperms of barley. *Carlsberg Research Communications* **49,** 365–383.

Grimwade B, Tatham AS, Freedman RB, Shewry PR, Napier JA. 1996. Comparison of the expression patterns of wheat gluten proteins and proteins involved in the secretory pathway in developing caryopses of wheat. *Plant Molecular Biology* **30,** 1067–1073.

Halford NG, Forde J, Shewry PR, Kreis M. 1989. Functional analysis of the upstream regions of a silent and an expressed

member of a family of wheat seed protein genes in transgenic tobacco. *Plant Science* **62,** 207–216.

Halford NG, Field JM, Blair H, Urwin P, Moore K, Robert L, Thompson R, Flavell RB, Tatham AS, Shewry PR. 1992. Analysis of HMW glutenin subunits encoded by chromosome 1A of bread wheat *Triticum aestivum* L. indicates quantitative effects on grain quality. *Theoretical and Applied Genetics* **83,** 373–378.

Hammond-Kosack MCU, Holdsworth MJ, Bevan MW. 1993. *In vivo* footprinting of a low molecular weight glutenin gene (LMWG-1D1) in wheat endosperm. *The EMBO Journal* **12,** 545–554.

Heck GR, Chamberlain AK, Ho DT-H. 1993. Barley embryo globulin 1 gene, *Beg1*: characterization of cDNA chromosome mapping and regulation of expression. *Molecular and General Genetics* **239,** 209–218.

Horikoshi M, Morita Y. 1975. Localization of γ-globulin in rice seed and changes in γ-globulin content during seed development and germination. *Agricultural and Biological Chemistry* **39,** 2309–2314.

Hsia CC, Anderson OD. 2001. Isolation and characterization of wheat ω-gliadin genes. *Theoretical and Applied Genetics* **103,** 37–44.

Kermode AR, Bewley JD. 1999. Synthesis, processing and deposition of seed proteins: The pathway of protein synthesis and deposition of the cell. In: Shewry PR, Casey R, eds. *Seed proteins*. Dordrecht: Kluwer Academic Publishers, 807–841.

Kreis M, Forde BG, Rahman S, Miflin BJ, Shewry PR. 1985. Molecular evolution of the seed storage proteins of barley, rye and wheat. *Journal of Molecular Biology* **183,** 499–502.

Krishnan HB, Franceschi VR, Okita TW. 1986. Immunochemical studies on the role of the Golgi complex in protein body formation in rice seeds *Planta* **169,** 471–480.

Kriz AL. 1989. Characterization of embryo globulins encoded by the maize *Glb* genes. *Biochemical Genetics* **27,** 239–251.

Kriz AL. 1999. 7S globulins of cereals. In: Shewry PR, Casey R, eds. *Seed proteins*. Dordrecht: Kluwer Academic Publishers, 477–498.

Kriz AL, Schwartz D. 1986. Synthesis of globulins in maize embryos. *Plant Physiology* **82,** 1069–1075.

Kriz AL, Wallace NH. 1991. Characterization of the maize *Globulin-2* gene and analysis of two null alleles. *Biochemical Genetics* **29,** 241–254.

Lamacchia C, Shewry PR, Di Fonzo N, Forsyth JL, Harris N, Lazzeri PA, Napier JA, Halford NG, Barcelo P. 2001. Endosperm-specific activity of a storage protein gene promoter in transgenic wheat seed. *Journal of Experimental Botany* **52,** 243–250.

Larré C, Denery-Papini S, Popineau Y, Deshayes G, Lefebvre J. 2000. Biochemical analysis and rheological properties of gluten modified by transglutaminase. *Cereal Chemistry* **77,** 121–127.

Leite A, Neto GC, Vettore AL, Yunes JA, Arruda P. 1999. The prolamins of sorghum, *Coix* and millets. In: Shewry PR, Casey R, eds. *Seed proteins*. Dordrecht: Kluwer Academic Publishers, 141–157.

Lending CR, Larkins BA. 1989. Changes in the zein composition of protein bodies during maize endosperm development. *The Plant Cell* **1,** 1011–1023.

Lending CR, Chesnut RS, Shaw KL, Larkins BA. 1989. Immunolocalization of avenin and globulin storage proteins in developing endosperm of *Avena sativa* L. *Planta* **178,** 315–324.

Li X, Franceschi V, Okita TW. 1993*a*. Segregation of storage protein mRNAs on the rough endoplasmic reticulum membranes of rice endosperm membrane. *Cell* **72,** 869–879.

Li X, Wu Y, Zhang D-Z, Gillikin JW, Boston RS, Franceschi VR, Okita TW. 1993*b*. Rice prolamin protein body biogenesis: a BiP-mediated process. *Science* **262,** 1054–1056.

Lockhart HB, Hurt HD. 1986. Nutrition of oats. In: Webster FH, ed. *Oats: chemistry and technology*. St Paul, Minnesota, USA: American Association of Cereal Chemists, Inc., 297–308.

Marris C, Gallois P, Copley J, Kreis M. 1988. The 5′ flanking region of a barley B hordein gene controls tissue and developmental specific CAT expression in tobacco plants. *Plant Molecular Biology* **10,** 359–366.

McMaster TJ, Miles MJ, Wannerberger L, Eliasson A-C, Shewry PR, Tatham AS. 1999. The identification of microphases in mixed α- and γ-gliadin protein films investigated by atomic force microscopy. *Journal of Agricultural and Food Chemistry* **47,** 5093–5099.

Muench DG, Wu Y, Coughlan SJ, Okita TW. 1998. Evidence for a cytoskeleton-associated binding site involved in prolamine mRNA localization to the protein bodies in rice endosperm tissue. *Plant Physiology* **116,** 559–569.

Muench DG, Ogawa M, Okita TW. 1999. The prolamins of rice. In: Shewry PR, Casey R, eds. *Seed proteins*. Dordrecht: Kluwer Academic Publishers, 93–108.

Müller M, Knudsen S. 1993. The nitrogen response of a barley C-hordein promoter is controlled by positive and negative regulation of the GCN4 and endosperm box. *The Plant Journal* **4,** 343–355.

Payne PI. 1987. Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. *Annual Review of Plant Physiology* **38,** 141–153.

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

Quayle T, Feix G. 1992. Functional analysis of the −300 region of maize zein genes. *Molecular and General Genetics* **231,** 369–374.

Rooke L, Békés F, Fido R, Barro F, Gras P, Tatham AS, Barcelo P, Lazzeri P, Shewry PR. 1999. Overexpression of a gluten protein in transgenic wheat results in highly elastic dough. *Journal of Cereal Science* **30,** 115–120.

Shewry PR. 1999. The synthesis, processing and deposition of gluten proteins in the developing wheat grain. *Cereal Foods World* **44,** 587–589.

Shewry PR, Tatham AS. 1990. The prolamin storage proteins of cereal seeds: structure and evolution. *Biochemical Journal* **267,** 1–12.

Shewry PR, Tatham AS. 1997. Disulphide bonds in wheat gluten proteins. *Journal of Cereal Science* **25,** 207–227.

Shewry PR, Tatham AS, Barro F, Barcelo P, Lazzeri P. 1995. Biotechnology of breadmaking: unravelling and manipulating the multi-protein gluten complex. *Bio/Technology* **13,** 1185–1190.

Shewry PR, Brennan C, Tatham AS, Warburton T, Fido R, Smith D, Griggs D, Cantrell I, Harris N. 1996. The development, structure and composition of the barley grain in relation to its end use properties. In: *Cereals 96*. Proceedings of the 46th Australian Cereal Chemistry Conference, Sydney, September 1996, 158–162.

Shewry PR, Tatham AS, Halford NG. 1999. The prolamins of the Triticeae. In: Shewry PR, Casey R, eds. *Seed proteins*. Dordrecht: Kluwer Academic Publishers, 35–78.

Shotwell MA. 1999. Oat globulins. In: Shewry PR, Casey R, eds. *Seed proteins*. Dordrecht: Kluwer Academic Publishers, 389–400.

Singh NK, Shepherd KW. 1985. The structure and genetic control of a new class of disulphide-linked proteins in wheat endosperm. *Theoretical and Applied Genetics* **71,** 79–92.

Singh NK, Shepherd KW, Langridge P, Gruen LC, Skerritt JH, Wrigley CW. 1988. Identification of legumin-like proteins in wheat. *Plant Molecular Biology* **11,** 633–639.

Singh NK, Donovan GR, Carpenter HC, Skerritt JH, Langridge P. 1993. Isolation and characterization of wheat triticin cDNA revealing a unique lysing-rich repetitive domain. *Plant Molecular Biology* **22,** 227–237.

Smith DB, Lister PR. 1983. Gel forming proteins in barley grain and their relationship with malting quality. *Journal of Cereal Science* **1,** 229–239.

Sørensen MB, Cameron-Mills V, Brandt A. 1989. Transcriptional and post-transcriptional regulation of gene expression in developing barley endosperm. *Molecular and General Genetics* **217,** 195–201.

Takaiwa F, Ogawa M, Okita TW. 1999. Rich glutelins. In: Shewry PR, Casey R, eds. *Seed proteins*. Dordrecht: Kluwer Academic Publishers, 401–425.

Tecsi L, Darlington HF, Harris N, Shewry PR. 2000. Patterns of protein deposition and distribution in developing and mature barley grain. In: *Barley genetics*, VIII. Proceedings of the 8th International Barley Genetics Symposium, Adelaide, Australia, Vol. 2, 266–268.

Thomas MS, Flavell RB. 1990. Identification of an enhancer element for the endosperm-specific expression of high molecular weight glutenin. *The Plant Cell* **2,** 1171–1180.

Torrent M, Geli MI, Ruiz-Avila L, Canals J, Puigdomènech P, Ludevid MD. 1994. Role of structural domains for maize γ-zein retention in *Xenopus* oocytes. *Planta* **192,** 512–518.

Wallace NH, Kriz AL. 1991. Nucleotide sequence of a cDNA clone corresponding to the maize *Globulin-2* gene. *Plant Physiology* **95,** 973–975.

Yamagata H, Tanaka K. 1986. The site of synthesis and accumulation of rice storage proteins. *Plant and Cell Physiology* **27,** 135–145.

Yupsanis T, Burgess SR, Jackson PJ, Shewry PR. 1990. Characterization of the major protein component from aleurone cells of barley *Hordeum vulgare* L. *Journal of Experimental Botany* **41,** 385–392.