The Plant Cell, Vol. 7, 945–956, July 1995 © 1995 American Society of Plant Physiologists

Seed Storage Proteins: Structures and Biosynthesis

Peter R. Shewry,[1] Johnathan A. Napier, and Arthur S. Tatham

IACR–Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Long Ashton, Bristol BS18 9AF, United Kingdom

INTRODUCTION

The plant seed is not only an organ of propagation and dispersal but also the major plant tissue harvested by humankind. The amount of protein present in seeds varies from ~10% (in cereals) to ~40% (in certain legumes and oilseeds) of the dry weight, forming a major source of dietary protein. Although the vast majority of the individual proteins present in mature seeds have either metabolic or structural roles, all seeds also contain one or more groups of proteins that are present in high amounts and that serve to provide a store of amino acids for use during germination and seedling growth. These storage proteins are of particular importance because they determine not only the total protein content of the seed but also its quality for various end uses. For example, the low content of lysine, threonine, and tryptophan in various cereal seeds and of cysteine and methionine in legume seeds is due to the low proportions of these amino acids in the major storage proteins and may limit the nutritional quality of the seeds for monogastric animals. In the case of wheat, the storage proteins form the gluten fraction, whose properties are largely responsible for the ability to use wheat flour to make bread, other baked goods, and pasta. These properties are not shared by the storage proteins of other cereals.

In this article, we provide a broad overview of the structures and properties of the seed storage proteins of the major crop plants, emphasizing their biological roles, their evolutionary origins, and their modes of synthesis and deposition. Although some storage proteins may also play roles in defense or metabolism, we focus on those that function solely for storage.

Characteristics of Seed Storage Proteins

Despite wide variation in their detailed structures, all seed storage proteins have a number of common properties. First, they are synthesized at high levels in specific tissues and at certain stages of development. In fact, their synthesis is regulated by nutrition, and they act as a sink for surplus nitrogen. However, most also contain cysteine and methionine, and adequate sulfur is therefore also required for their synthesis. Many seeds contain separate groups of storage proteins, some of which are rich in sulfur amino acids and others of which are poor in them. The presence of these groups may allow the plant to maintain high levels of storage protein synthesis despite variations in sulfur availability. The strict tissue specificity of seed storage protein synthesis contrasts with that of tuber storage proteins, which may be synthesized in vegetative tissues under unusual conditions (for example, in vitro or after removal of tubers) (Shewry, 1995). A second common property of seed storage proteins is their presence in the mature seed in discrete deposits called protein bodies, whose origin has been the subject of some dispute and may in fact vary both between and within species. Finally, all storage protein fractions are mixtures of components that exhibit polymorphism both within single genotypes and among genotypes of the same species. This polymorphism arises from the presence of multigene families and, in some cases, proteolytic processing and glycosylation.

Classification of Seed Storage Proteins

Because of their abundance and economic importance, seed storage proteins were among the earliest of all proteins to be characterized. For example, wheat gluten was first isolated in 1745 (Beccari, 1745), and Brazil nut globulin was crystallized in 1859 (Maschke, 1859). However, the detailed study of seed storage proteins dates from the turn of the century, when Osborne (1924) classified them into groups on the basis of their extraction and solubility in water (albumins), dilute saline (globulins), alcohol/water mixtures (prolamins), and dilute acid or alkali (glutelins). The major seed storage proteins include albumins, globulins, and prolamins.

2S ALBUMIN STORAGE PROTEINS

The 2S albumins were initially defined as a group on the basis of their sedimentation coefficients ($S_{20,w}$) of ~2 (Youle and Huang, 1981). They are widely distributed in dicot seeds and have been most widely studied in the Cruciferae, notably oilseed rape (in which they are called napins) and Arabidopsis. The napins consist of two polypeptide chains with M_r values of ~9000 and 4000, which are linked by interchain disulfide

[1] To whom correspondence should be addressed.

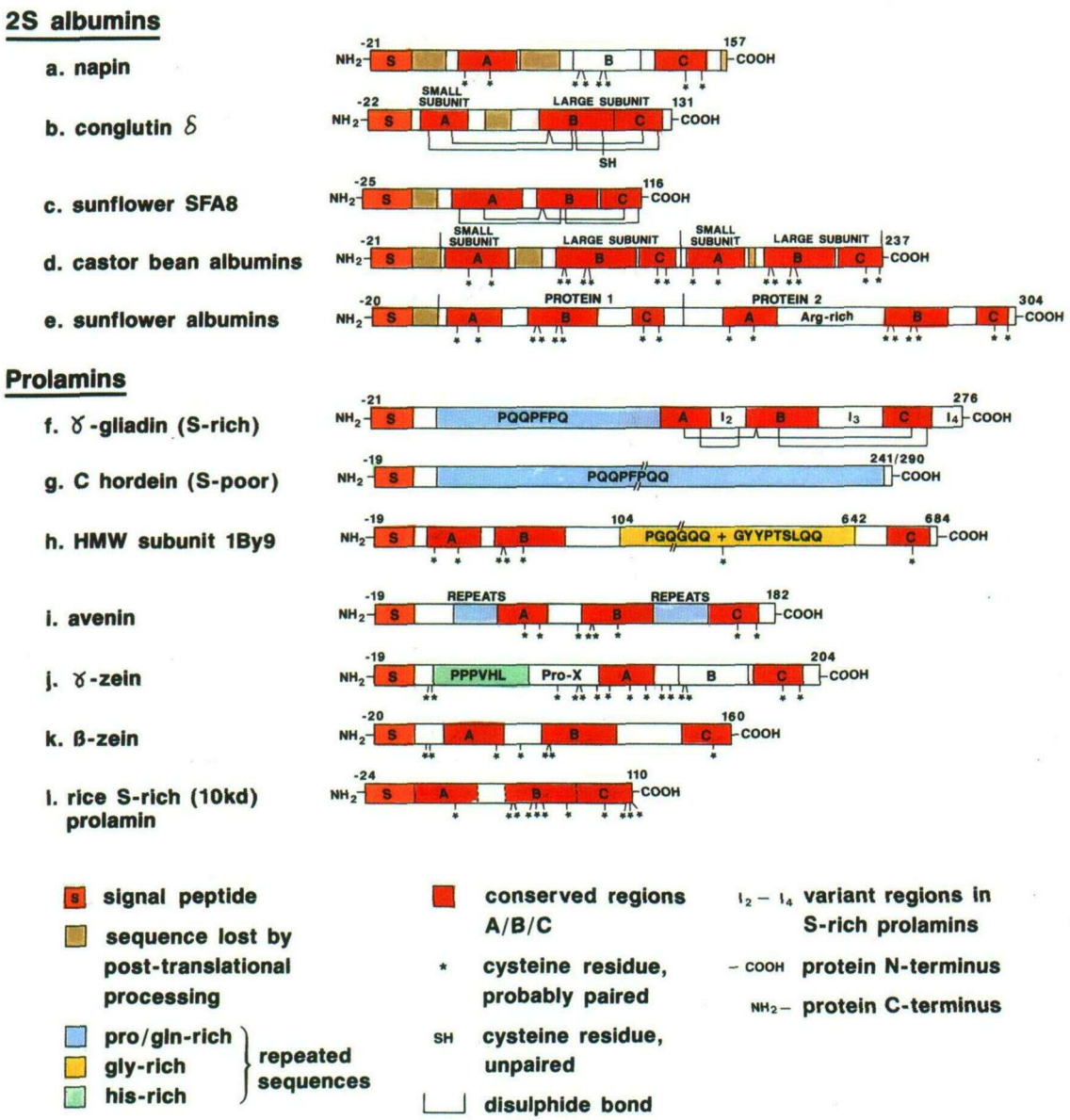

Figure 1. Schematic Structures of Members of the Cereal Prolamin Superfamily.

The cereal prolamin superfamily comprises the 2S albumins of dicots, the prolamins of the Triticeae, oats, and rice, and the β- and γ-zeins of maize. Three conserved regions (A, B, and C) are present in all except C hordein, although their boundaries are often poorly defined. These three regions also show homology with each other and contain cysteine residues that may be conserved within or between the different groups of proteins. For example, the 2S albumins shown all contain eight cysteine residues that are conserved in terms of context and position, including Cys-Cys and Cys-Xaa-Cys motifs, which are present in many of the other proteins. Based on data of Sharief and Li (1982), Crouch et al. (1983), Bartels et al. (1986), Boronat et al. (1986), Higgins et al. (1986), Lilley and Inglis (1986), Pedersen et al. (1986), Halford et al. (1987), Egorov (1988), Chestnut et al. (1989), Masumura et al. (1989), Gayler et al. (1990), Irwin et al. (1990), Entwistle et al. (1991), Kortt et al. (1991), and Shewry et al. (1993). The sunflower albumin structure and the disulfide structure of SFA8 are unpublished results of A. Tatham, J. Napier, R. Fido, P. Thoyts, T. Egorov, and P. Shewry.

bonds (Ericson et al., 1986). They are synthesized as single precursor proteins that are proteolytically cleaved with the loss of a linker peptide and short peptides from both the N and C termini (Figure 1a; Crouch et al., 1983; Ericson et al., 1986). This appears to be the most typical 2S albumin structure: similar heterodimeric proteins are present in species as diverse as pumpkin (Hara-Nishimura et al., 1993), cotton (Galau et al., 1992), castor bean (Sharief and Li, 1982), and lupin (Lilley and

Inglis, 1986). The presence of two interchain bonds has been directly demonstrated in 2S albumins from lupin (Figure 1b). Variant types of 2S albumin also occur. Those of pea appear to lack interchain disulfide bonds (Higgins et al., 1986), whereas the 2S albumins of sunflower remain uncleaved (Figures 1c and 1e; Kortt and Caldwell, 1990; Anisimova et al., 1994). In addition, in sunflower and castor bean, some mRNAs encode two mature albumin proteins, each consisting of one or two subunits (Figures 1d and 1e; Allen et al., 1987; Irwin et al., 1990; P. Thoyts, J. Napier, M. Millichip, T. Griffiths, A. Tatham, A. Stobart, and P. Shewry, unpublished results).

Despite differences in their subunit structure and synthesis, all the 2S albumins are compact globular proteins with conserved cysteine residues (Figure 1). Although little is known about the detailed three-dimensional structures of 2S albumins, that of yellow mustard has been reported to contain ~50% α-helix, with little or no β-sheet (Menéndez-Arias et al., 1987). The authors proposed a ring structure with tightly packed α-helices, as suggested for zeins (see later discussion), but there is no experimental evidence for this structure.

Much of the recent interest in 2S albumins has focused on their exploitation in genetic engineering. Most notably, Altenbach et al. (1987, 1992) have used the 2S albumin of Brazil nut, which is rich in methionine (Youle and Huang, 1981), to increase the methionine content of tobacco seeds by up to 30%, and Higgins and co-workers have used the methionine-rich sunflower 2S albumin SFA8 to increase the methionine content of forage grasses (Tabe et al., 1993). In addition, the 2S albumins of Arabidopsis have been used as "hosts" for the synthesis of biologically active peptides, including the pentapeptide Leu-enkephalin (Vandekerckhove et al., 1989) and a 28-residue antibacterial peptide from Xenopus (De Clercq et al., 1990; Krebbers et al., 1993). In this work, the peptide was expressed as an insert within a variable loop region of the 2S albumin and then isolated by enzymatic cleavage. Although yields in oilseed rape equivalent to 1 kg of a 25-residue peptide per hectare were achieved (Krebbers et al., 1993), the commercial viability of the work is uncertain.

PROLAMIN STORAGE PROTEINS

Whereas the 2S albumin and globulin (see later discussion) storage proteins are widely distributed in flowering plants, the prolamins are restricted to one family, the grasses. These include the major cereals, in which prolamins usually account for approximately half of the total grain nitrogen. Exceptions to this general rule are oats and rice, in which the major storage proteins are 11S globulin–like and prolamins are present at low levels (~5 to 10% of the total grain protein).

Prolamins are traditionally recognized as a group on the basis of their solubility in alcohol/water mixtures (usually 60 to 70% [v/v] ethanol or 50% [v/v] propan-1-ol) and their high levels of glutamine and proline. However, comparisons of amino acid sequences have shown that this definition must be widened to include components that are insoluble in aqueous alcohols in the native state due to the presence of interchain disulfide bonds and to recognize that all prolamins, even those that are insoluble in aqueous alcohols, are related, except for the α-zeins of maize (and their homologs present in related Panicoid cereals). All other prolamins form a single group known as the prolamin superfamily.

The Prolamin Superfamily of the Triticeae

Our understanding of this storage protein family stemmed initially from studies of the temperate cereals of the tribe Triticeae: barley, wheat, and rye. The prolamins of all three species are highly polymorphic mixtures of components whose M_r values range from ~30,000 to 90,000. These prolamins are classified into three groups (Miflin et al., 1983)—the S-rich, S-poor, and high molecular weight (HMW) prolamins—based on their amino acid sequences. Typical structures of the three types of prolamin are summarized in Figures 1f, 1g, and 1h.

The S-rich prolamins are the quantitatively major prolamin group in all three species, accounting for ~80 to 90% of the total prolamin fractions. They include polymeric (that is, with interchain disulfide bonds) and monomeric (with intrachain disulfide bonds) components and consist of at least two families in each species: the B and γ-hordeins of barley; two types of γ-secalin of rye; and the α-gliadins, γ-gliadins, and low molecular weight (LMW) glutenin subunits of wheat. Their amino acid sequences consist of two separate domains: an N-terminal domain composed of repeated sequences, and a nonrepetitive C-terminal domain (see Figure 1f). The repetitive domain consists of tandem or interspersed repeats based on one or two short peptide motifs rich in proline and glutamine; this structure accounts for the high proportions of these two residues in the protein as a whole. For example, the repetitive domain of the γ-gliadin shown in Figure 1f is based on a Pro-Gln-Gln-Pro-Phe-Pro-Gln heptapeptide. This domain forms a secondary structure containing β-reverse turns and poly-L-proline II helix (Tatham et al., 1990), as discussed later for the S-poor prolamins. In contrast, the nonrepetitive domain appears to have a globular structure rich in α-helix. This domain also contains most or all of the cysteine residues. Eight cysteines are present in the monomeric γ-gliadin, which form four intrachain disulfide bonds (Figure 1f). Six of these cysteine residues are also present in the monomeric α-gliadins (based on sequence context); additional "unpaired" cysteine residues present in the polymeric LMW glutenin subunits may be responsible for polymer formation (see Shewry et al., 1993).

The S-poor prolamins include C hordein of barley (Figure 1g), the ω-secalins of rye, and the ω-gliadins of wheat (Kasarda et al., 1983). Several genes encoding ω-secalins and C hordeins have been isolated (Entwistle et al., 1991; Hull et al., 1991; Sayanova, 1993). In all cases, the encoded proteins consist almost entirely of repeats of the octapeptide motif Pro-Gln-Gln-Pro-Phe-Pro-Gln-Gln that are flanked at the N-terminal side by short unique sequences of 12 residues and at the C-terminal

side by short unique sequences of either six (C hordeins; see Figure 1g) or four (ω-secalin) residues. The S-poor prolamins generally lack cysteine residues and therefore cannot form oligomers or polymers. Structural studies of C hordein indicate that the highly conserved repetitive primary structure results in a similarly conserved supersecondary structure. This is a loose spiral based on elements of β-turn and poly-L-proline II helix, the whole molecule forming a "stiff worm-like coil" of ~70 nm in length (l'Anson et al., 1992).

The HMW prolamins are typified by the HMW subunits of wheat glutenin, which have been studied in detail because of their putative role in determining the elasticity, and hence the bread-making performance, of wheat doughs (Payne, 1987; Shewry et al., 1989, 1992). Extensive repeated sequences are present, flanked by nonrepetitive N- and C-terminal domains (Figure 1h). The repeated sequences are based on the motifs Gly-Tyr-Tyr-Pro-Thr-Ser-Pro or Leu-Gln-Gln, Pro-Gly-Gln-Gly-Gln-Gln, and, in some subunits only, Gly-Gln-Gln. Differences in the number of repeated peptides are largely responsible for variation in HMW subunit size.

Although the repeated sequences present in the HMW subunits are not related to those in the S-poor prolamins, they appear to adopt a similar spiral supersecondary structure, although one that is more compact because it includes β-turns but not poly-L-proline II structure. The net result is a rod-shaped molecule (Field et al., 1987), which has been imaged directly by scanning probe microscopy (Miles et al., 1991). As in the S-poor prolamins, cysteine residues are largely restricted to the nonrepetitive domains (Figure 1h). These domains appear to be globular (being rich in α-helix), with the cysteine residues allowing the formation of an elastic network stabilized by interchain disulfide bonds.

Evolutionary Relationships among the S-Rich, S-Poor, and HMW Prolamins

The three groups of prolamins present in the Triticeae all consist of at least two discrete domains, one of which is based on repeated sequences. More detailed comparisons show that the prolamins are likely to have evolved from a single ancestral protein. Comparisons of the nonrepetitive domains of a range of S-rich prolamins (Kreis et al., 1985a, 1985b; Kreis and Shewry, 1989) show that all contain three conserved regions of between 20 and 30 residues. These regions, designated A, B, and C (Figure 1f), contain most of the conserved cysteine residues and are also related to each other, indicating that they are likely to have originated from the triplication of a short ancestral domain. Insertion of additional variable regions (I_1 to I_4) and of repeated sequences at the N-terminal side of region I_1 would have given rise to the range of present-day S-rich prolamins. Short regions related to A, B, and C are also present in the HMW prolamins, although in this case regions A and B are in the N-terminal domain and region C is in the C-terminal domain (Figure 1h). Therefore, these proteins are likely to have evolved from the same ancestor as did the S-rich prolamins, although unrelated repeated sequences have been inserted between regions B and C.

The S-poor prolamins are also clearly related to the S-rich prolamins in that their repetitive sequences are based on similar proline- and glutamine-rich peptide motifs. For example, the heptapeptide and octapeptide motifs present in γ-gliadin and C hordein (Figure 1) differ in only a single glutamine residue. The S-poor group is hypothesized to have evolved from the S-rich prolamins by further amplification of the repeated sequences and deletion of most of the nonrepetitive domain that contains regions A, B, and C (Kreis et al., 1985a, 1985b; Kreis and Shewry, 1989).

The Prolamin Superfamily in Other Species

Prolamins related to those present in the Triticeae are also present in a range of other cereals. These include oats, in which the avenins contain regions A, B, and C together with two blocks of repeats rich in proline and glutamine (Figure 1i; Egorov, 1988; Chesnut et al., 1989), and rice. The prolamins of rice consist of three groups of small proteins (Kim and Okita, 1988a, 1988b; Masumura et al., 1989, 1990). Although these do not contain repeated sequences, they appear to be related to one another and to the prolamins of the Triticeae. For example, the sulfur-rich M_r 10,000 prolamins shown in Figure 1l appear to contain vestiges of regions A, B, and C.

The prolamins of maize, known as the zeins, and of related Panicoid cereals such as sorghum, pearl millet, and Job's tears, fall into four groups, three of which belong to the prolamin superfamily. In maize, these are the β-, γ-, and δ-zeins. The β-zeins (Pedersen et al., 1986) and γ-zeins (Boronat et al., 1986) both contain regions related to A, B, and C (Figures 1j and 1k). The δ-zeins do not contain repeats or any other distinguishing features, but homology with the prolamin superfamily can be inferred from some sequence identity with the 2S albumin of Brazil nut (Kirihara et al., 1988; see later discussion). All three of these groups of zeins are rich in cysteine and/or methionine, residues deficient in the α-zeins.

The 2S Albumins Are Also Related to the Prolamin Superfamily

The 2S albumins also contain three conserved regions related to regions A, B, and C. These regions contain the eight conserved cysteine residues present in most 2S albumins (see Figures 1a to 1e), with region A and regions B and C corresponding to the small and large subunits, respectively, of the heterodimeric 2S albumins (for example, napin and conglutin δ; Figures 1a and 1b). The absence of repeated sequences and the widespread distribution of 2S albumins in dicots (and even in ferns; Rodin and Rask, 1990) may indicate that they are similar to the ancestral protein of the prolamin superfamily, although this would have lacked the proteolysis site between regions A and B.

The α-Zeins of Maize

The α-zeins account for ~75 to 80% of the total prolamins in maize and are classified into two groups with slightly different M_r (~19,000 and ~22,000). They have similar structures, consisting of unique N- and C-terminal domains flanking repeated sequences (Marks and Larkins, 1982; Pedersen et al., 1982; Marks et al., 1985). Although the latter are generally considered to contain blocks of ~20 residues, they are highly degenerate, with no clear consensus motif. There is no evidence of homology with the repeated sequences present in other prolamins, and the unique N- and C-terminal sequences do not appear to be related to any other protein. The size difference between the M_r 19,000 and M_r 22,000 zeins may result from variation in the number of blocks present in the repetitive domains (nine and 10, respectively) or from the insertion of a loop region of ~20 residues in the C-terminal domain of the M_r 22,000 proteins.

The precise structure adopted by the α-zeins is still uncertain. Whereas a range of biophysical studies demonstrates that they have extended conformations when in solution (Tatham et al., 1993), they may adopt a more compact conformation when present in the hydrated solid state, that is, in protein bodies (see later discussion). For example, Argos et al. (1982) proposed that α-zeins form an antiparallel ring of nine α-helices, facilitating packaging in the protein bodies.

GLOBULIN STORAGE PROTEINS

The globulins are the most widely distributed group of storage proteins; they are present not only in dicots but also in monocots (including cereals and palms) and fern spores (Templeman et al., 1987). They can be divided into two groups based on their sedimentation coefficients ($S_{20,w}$): the 7S vicilin-type globulins and the 11S legumin-type globulins. Both groups show considerable variation in their structures, which results partly from post-translational processing. In addition, both have nutritional significance in that they are deficient in cysteine and methionine, although 11S globulins generally contain slightly higher levels of these amino acids. The globulin storage proteins have been studied in most detail in legumes, notably peas, soybean, broad bean (*Vicia faba*), and French bean (*Phaseolus vulgaris*).

The 11S Globulins

The 11S legumins are the major storage proteins not only in most legumes but also in many other dicots (for example, brassicas, composits, and cucurbits) and some cereals (oats and rice). The mature proteins consist of six subunit pairs that interact noncovalently. Each of these subunit pairs consists in turn of an acidic subunit of M_r ~40,000 and a basic subunit of M_r ~20,000, linked by a single disulfide bond. Each subunit pair is synthesized as a precursor protein that is proteolytically cleaved after disulfide bond formation. Legumins are not usually glycosylated, an exception being the 12S globulin of lupin (Duranti et al., 1988). This contrasts with the 7S globulins (see later discussion).

Although the 11S globulin of Brazil nut was one of the first proteins to be crystallized (Maschke, 1859), the crystals of this and other 11S globulins have generally been small and disordered and have failed to provide any details of protein structure. However, a recent study of edestin, an 11S globulin from hempseed, is more promising. Although the crystals showed some disorder, they exhibited enough symmetry so that some measurements could be made. These indicated that the subunits are arranged in an open ring structure, oriented alternately up and down, in a disk whose diameter is 145 Å and whose thickness is ~90 Å (Patel et al., 1994).

The 7S Globulins

7S globulins are typically trimeric proteins of M_r ~150,000 to 190,000 that lack cysteine residues and hence cannot form disulfide bonds. Their detailed subunit compositions vary considerably, mainly because of differences in the extent of post-translational processing (proteolysis and glycosylation). For example, the vicilin subunits of pea are initially synthesized as groups of polypeptides of M_r ~47,000 and ~50,000, but post-translational proteolysis and glycosylation then give rise to subunits with M_r values between 12,500 and 33,000 (Figure 2; Gatehouse et al., 1984; Casey et al., 1986, 1993). These subunits are difficult to purify and characterize, but molecular cloning allowed their origins and the sites of proteolytic cleavage and glycosylation to be identified.

The 7S globulins of *P. vulgaris* and soybean differ from those of pea and *V. faba* in that glycosylation is more extensive but proteolysis does not occur. For example, the 7S phaseolin of

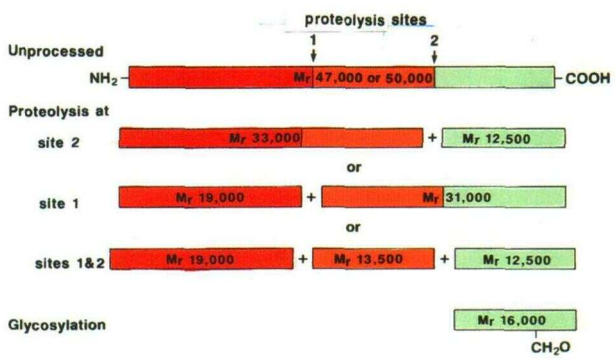

Figure 2. Schematic Diagram Showing the Origin of the Pea Vicilin Subunits.

Based on Gatehouse et al. (1984) and Casey et al. (1993). The *Mr* 19,000, *Mr* 13,500, and *Mr* 12,500/16,000 subunits are shown in red, orange, and green, respectively.

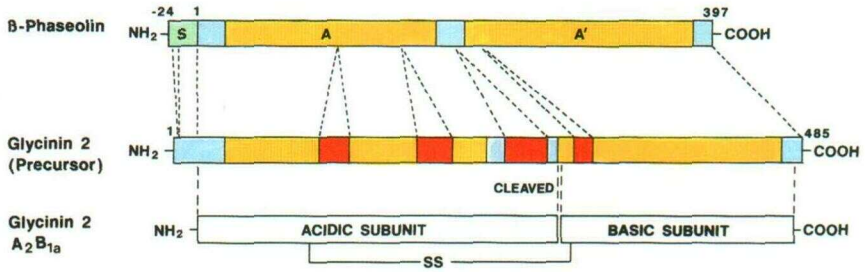

Figure 3. Alignment of Phaseolin and Glycinin 2 Subunit Sequences.

The alignment of β-phaseolin (French bean 7S globulin) and glycinin 2 (soybean 11S globulin) is based on the structure of phaseolin (Lawrence et al., 1994). Two structurally similar units, A and A′ (shown in yellow), have been defined by x-ray crystallography. Each unit consists of a β-barrel with a "jelly roll" motif followed by an α-helical domain comprising three helices. Unit A is located in the acidic subunit of the 11S protein and unit A′ in the basic subunit. The blue areas flanking and separating regions A and A′ correspond to regions of sequence homology that are not reflected in the secondary structures as determined by crystallography. The distribution of globally conserved residues in the 7S/11S alignment indicates the presence of four major insertions in the 11S protein (shown in red). Two of these are in unit A and one is in unit A′, all falling within loop regions connecting structural elements. The fourth insertion is in the region of sequence homology between A and A′, at the C-terminal end of the acidic subunit. Insertions and deletions of less than six residues are not shown. The signal peptide of the β-phaseolin, shown in green, has limited sequence homology with the N-terminal region of the mature glycinin 2 precursor. Based on data of Lawrence et al. (1994).

P. vulgaris consists of glycosylated subunits with M_r values between ~43,000 and 53,000 (Hall et al., 1977; Bollini and Chrispeels, 1978).

The three-dimensional structures of several 7S globulins have recently been determined using x-ray crystallography (Lawrence et al., 1990, 1994; Ko et al., 1993). These show that the trimeric proteins are disk shaped, with diameters of ~90 Å and thicknesses of 30 to 40 Å.

11S and 7S Globulins Are Related

Although the 7S and 11S globulins show no obvious sequence similarities, they do have similar properties, including the ability to form both trimeric and hexameric structures. In the case of the 7S globulins, the mature protein is trimeric, but it may undergo reversible aggregation into hexamers, depending on the ionic strength (Thanh and Shibasaki, 1979). The mature 11S globulin, by contrast, is hexameric but is initially assembled and transported through the secretory system as an intermediate trimer (Gatehouse et al., 1984; Müntz et al., 1993). Therefore, it is not surprising that more sophisticated comparisons have shown that the 11S and 7S globulin subunits are related in structure (Gibbs et al., 1989; Lawrence et al., 1994). Such comparisons indicate that the basic (C-terminal) chain of the 11S legumins is related to the C-terminal region of the 7S vicilins. Lawrence et al. (1994) determined the x-ray crystal structure of 7S phaseolin and established sequence homologies between 7S/11S proteins. The homologies showed that the 11S sequences, with four major insertions of sequence, can be aligned with 7S sequences. The authors further proposed that the 11S legumins have a tertiary structure similar to that of the 7S vicilins and concluded that the 7S and 11S proteins evolved from a common ancestral protein (Figure 3).

SYNTHESIS, ASSEMBLY, AND DEPOSITION OF SEED STORAGE PROTEINS

All of the seed storage proteins discussed here are secretory proteins synthesized with a signal peptide that is cleaved as the protein is translocated into the lumen of the endoplasmic reticulum (ER). The subsequent events in storage protein processing are less clearly understood and may vary not only between different species but also within the same species, depending on the protein type and stage of development. The events that occur in the different compartments of the secretory system are discussed later and summarized in Table 1.

Storage Protein Folding and Assembly in the ER

Secretory proteins assume their folded conformations within the lumen of the ER, which is also the site of disulfide bond formation. Studies of other systems demonstrate that three types of ER lumenal proteins may assist in these processes. Molecular chaperones of the HSP70/BiP family may facilitate folding by binding transiently to the nascent polypeptides and may also prevent the formation of incorrect inter- or intramolecular interactions. BiP-related proteins are present in developing endosperms of cereals such as rice (Li et al., 1993b), wheat (Giorini and Galili, 1991), and maize (Boston et al., 1991), and they accumulate in higher than normal levels in high-lysine maize mutants (Boston et al., 1991), possibly due to the presence of incorrectly folded zeins.

A second group of proteins, the peptidyl-prolyl *cis–trans* isomerases (PPI), or cyclophilins, of which one subclass (the S-cyclophilins) is resident in the ER lumen, may also assist folding by accelerating the isomerization of Xaa-Pro bonds,

Table 1. Processing and Assembly of Storage Proteins in the Secretory System

Compartment	Event	2S Albumins	Prolamins	7S Globulins	11S Globulins
ER	Co-translational insertion	+[a]	+	+	+
	Signal peptide cleavage	+	+	+	+
	Chaperone-mediated folding (BiP)[b]	NA	NA	NA	NA
	S–S bond formation (PDI)	NA	+[c]	–	+
	Isomeration of Xaa-Pro bonds (PPI)[d]	NA	NA	NA	NA
	N-Glycosylation	–	–	+	–
Golgi	Complex glycan addition	–	–	+	–
Vacuole	Propeptide processing (cleavage at asparagine residues by proteases)	+[e]	–	NA	+[f]

[a] +, the process has been experimentally observed; –, the process has been looked for but not detected; NA, the process has not been examined.

[b] Although BiP is likely to interact transiently with every elongating nascent chain translocating across the ER membrane, its subsequent role in folding is likely to be protein dependent.

[c] Because the majority of storage proteins form disulfide bonds, it is assumed that this reaction is catalyzed by protein disulfide isomerase (PDI) in the ER lumen.

[d] Peptidyl-prolyl *cis–trans* isomerase (PPI) may be required for the folding in the ER of proline-rich proteins (Stamnes et al., 1992) such as the prolamins.

[e] Aspartic and thiol proteases have been characterized from *B. napus* and castor bean, respectively.

[f] Thiol proteases have been characterized from soybean and castor bean.

which is a rate-limiting step in the folding of some proteins. The repetitive domains of cereal prolamins contain high levels of proline, and isomerization of Xaa-Pro bonds might therefore be expected to limit their folding. This does not, however, appear to be the case, at least in vitro. For example, C hordein contains ~30 mol % proline residues, all of which appear to be in the *trans* configuration (Tatham et al., 1985). Nevertheless, it folds readily in vitro (Tamas et al., 1994). Our preliminary studies also show that the levels of cyclophilin-related transcripts decrease during the period of gluten protein synthesis in developing wheat seeds (B. Grimwade, R. Freedman, P. Shewry, and J. Napier, unpublished results). 7S and 11S globulin subunits are also assembled in the ER, with the 7S globulins forming the mature trimers (Ceriotti et al., 1991).

Whether protein disulfide isomerase (PDI) catalyzes disulfide bond formation in storage proteins also remains to be established, although Bulleid and Freedman (1988) showed that depletion of PDI from dog pancreas microsomes resulted in defective synthesis of disulfide bonds in a γ-gliadin synthesized in vitro. PDI has also been shown to be associated with the ER in developing wheat endosperms (Roden et al., 1982), although the levels of PDI transcripts peak somewhat earlier than those of gluten proteins (B. Grimwade, R. Freedman, P. Shewry, and J. Napier, unpublished results). The assembly of some prolamins into disulfide-stabilized polymers presumably also takes place in the ER, although there is no information available on how this occurs.

N-linked glycosylation of the 7S phaseolin subunits also occurs in the ER lumen, probably as a cotranslational event (Bollini et al., 1983; Vitale et al., 1993). Phaseolin has two consensus N-glycosylation sites, one of which is always used, whereas the other, located closer to the C terminus, is used less frequently (Ceriotti et al., 1991). Wild-type phaseolin assembles into trimers in the ER, but trimerization is prevented if a C-terminal sequence of 59–amino acid residues is deleted. In this case, the protein monomer remains in the ER and becomes glycosylated at the second site. Moreover, this assembly-defective protein interacts with BiP in an ATP-dependent manner, highlighting the role of BiP in binding to malfolded proteins (Pedrazzini et al., 1994).

Storage Protein Transport and Protein Body Formation in Cereals

Two routes of protein body formation appear to operate in developing cereal endosperms, in one of which the protein body forms from the vacuole and in the other of which it forms from the ER. For example, the major storage proteins in oats and rice are related to the 11S globulins of dicots and appear to be transported from the ER lumen via the Golgi apparatus to the vacuole. The protein bodies then appear to form by fragmentation of the vacuole. In contrast, the prolamins of rice (Krishnan et al., 1986) and maize (Larkins and Hurkman, 1978) appear to be retained within the lumen of the ER, which becomes distended to form protein bodies. Thus, rice endosperm cells contain two populations of protein bodies, some of vacuolar origin (containing glutenins) and others of ER origin (containing prolamins).

The situation appears to be more complicated in barley, wheat, and rye, with prolamins present in both ER-derived and vacuolar protein bodies. In the case of wheat, these may differ in their protein content, as Rubin et al. (1992) suggested when they reported that glutenins are retained predominantly in ER-derived protein bodies, whereas gliadins are present in both types of protein body. In addition, Levanony et al. (1992) proposed that ER-derived protein bodies may subsequently fuse with the vacuoles, bypassing the Golgi apparatus.

The mechanisms that determine whether a prolamin is transported to the vacuole or retained in the ER are not known, because neither vacuolar targeting nor ER retention sequences have been identified. Li et al. (1993b) have suggested that rice prolamins are retained in the ER by interaction with BiP, which itself has a C-terminal ER retention sequence. Although the work of Li et al. (1993b) implies a "once only" binding of BiP to the emerging nascent polypeptide chain, it is now known that BiP binds to such chains in a sequential manner, "pulling" the protein into the ER lumen (Wickner, 1994). This indicates that a stable interaction of BiP with a nascent prolamin chain is very unlikely. In fact, the only clear examples of BiP binding to storage proteins are to malfolded or assembly-defective forms (D'Amico et al., 1992; Zhang and Boston, 1992; Pedrazzini et al., 1994). The expression of BiP-related transcripts is not coordinated with prolamin gene expression in developing endosperms of wheat (B. Grimwade, R. Freedman, P. Shewry, and J. Napier, unpublished results), and expression of wheat γ-gliadin in seed of transgenic tobacco plants does not alter the level of BiP transcripts (G. Richard, M. Turner, J. Napier, and P. Shewry, unpublished results), which suggests that BiP is unlikely to be involved in prolamin retention in the ER.

Studies of γ-gliadin transport and retention in *Xenopus* oocytes seem to indicate that prolamin accumulation in the ER does not require any plant-specific factors (Altschuler et al., 1993). Whereas a truncated form of the protein corresponding to the N-terminal domain accumulated in the ER to form protein body–like structures, a truncated form containing the C-terminal domain was secreted. The intact wild-type protein was also secreted, but at a lower rate than was the C-terminal domain. Although the prolamin repetitive domains could be responsible for ER retention by interacting with ER components, a simpler model is that interactions between the individual prolamin molecules result in the formation of insoluble aggregates that are not readily transported from the ER lumen. Such a model is supported by the observation of Li et al. (1993a) that rice prolamin mRNAs are segregated to a distinct region of the rough ER. Such segregation could allow aggregation of the prolamins to occur in localized parts of the ER, preventing widespread effects on ER integrity.

Storage Protein Transport and Protein Body Formation in Dicots

The 2S albumins and 7S/11S globulins of legumes and other dicots are transported via the Golgi apparatus to the vacuole, which fragments to form protein bodies. Despite several attempts, specific vacuolar targeting sequences have not been identified in these proteins (Müntz, 1989; Chrispeels, 1991; Saalbach et al., 1991; Chrispeels and Raikhel, 1992; Müntz et al., 1993). Instead, it is probable that one or more exposed regions of the correctly folded protein are recognized by the sorting machinery.

The assembly of the 11S globulins appears to be a highly regulated event. The monomeric proteins are initially assembled in the lumen of the ER into trimers that are then transported from the ER to the storage vacuole, where they are assembled into their final hexameric form. This assembly process requires specific proteolytic cleavage of the subunits present in the trimers (Dickinson et al., 1989). Uncleaved trimers cannot assemble into hexamers in vitro unless they have been treated with papain. This cleavage does not cause the trimers to disassemble but may result in a conformational change that favors assembly into hexamers. The 11S globulin vacuolar processing protease has been characterized from several species and shown to recognize asparagine processing sites specifically. Scott et al. (1992) purified a soybean protease that cleaves the trimeric 11S globulin proproteins. Hara-Nishimura et al. (1993) also purified an 11S globulin processing peptidase from castor bean that displays similar processing specificity and also appears to be a thiol protease but is unglycosylated.

The 11S globulin processing peptidase of Hara-Nishimura et al. (1993) is also able to cleave 2S albumin precursors in vitro at their asparagine processing sites. A 2S albumin processing protease that cleaves 2S albumin proproteins from Arabidopsis in vitro has also been characterized (D'Hondt et al., 1993). Although this enzyme has the same specificity as that of the one purified by Hara-Nishimura et al. (1993), it is an aspartic protease rather than a thiol protease. The processing of 2S albumins does not appear to be required for their assembly, in contrast with the case of the 11S globulins.

Storage Protein Packaging

Little is known about how storage proteins are organized within protein bodies, although this organization may well be important in ensuring efficient use of storage space and facilitating mobilization of storage proteins during germination. Whereas prolamin and globulin storage proteins are present in separate protein bodies in rice (see previous discussion), they are located within the same protein bodies (although in separate phases of them) in other cereals. Prolamin inclusions are present in a globulin matrix in oats (Lending et al., 1989), and globulin (triticin) inclusions are present within a prolamin matrix in wheat (Bechtel et al., 1991).

In leguminous plants, the 7S and 11S globulins appear to be in the same protein bodies with no spatial separation (see Harris et al., 1993). As discussed previously, their structures may facilitate efficient packaging. In many other dicots, such as pumpkin, sunflower, brassicas, and castor bean, 2S albumins

are stored together with 11S globulins, but how these distinct types of storage protein are organized in the protein bodies is not known. In contrast, there is evidence that the different types of prolamins are spatially separated in the protein bodies of cereal endosperms. This is most clear in maize, where immunogold labeling has shown that the α-zeins form the protein body core, with β- and γ-zeins at the periphery (Lending et al., 1992). The δ-zeins may also be present in the protein body core (Esen and Stetler, 1992). The evidence for spatial separation of prolamins in protein bodies of the Triticeae is less convincing, but Rechinger et al. (1993) proposed that the quantitatively minor S-rich γ_1- and γ_2-hordeins form a peripheral layer surrounding a core of B hordeins (S-rich) and C hordeins (S-poor) in barley. The separation of different proteins in cereal protein bodies could result from the properties of the proteins themselves (for example, their ability to separate into separate phases) or from different patterns of deposition during protein body biogenesis.

FUTURE DIRECTIONS

Much of the recent work on seed storage proteins was performed to provide a basis for improving the nutritional and processing properties of crops using genetic engineering. The recent development of reliable transformation procedures for maize, small grain cereals (wheat and barley), and grain legumes means that this is now possible.

A detailed understanding of storage protein structure and diversity is an important prerequisite for attempts to manipulate quality because it indicates the extent to which the structure of the proteins can be manipulated without affecting their biological properties. For example, much of the work on engineering 2S albumins has been based on manipulation of a variable "loop" region identified by sequence comparisons (see previous discussion), whereas Wallace et al. (1988) used a structural mode for zein (Argos et al., 1982) to identify sites for the addition of lysine residues. We are using a similar approach to develop novel wheats with improved bread-making quality, based on our understanding of the structures and properties of individual gluten proteins.

Transgenic plants are being used to develop improved lines for incorporation into plant breeding programs, but equally important is their use as tools to explore aspects of seed protein folding, assembly, transport, and deposition. There are still many gaps in our knowledge of these processes. Current evidence indicates that ER lumenal proteins such as PDI and BiP may play a role in storage protein folding and assembly in vitro and under abnormal circumstances (for example, in high-lysine mutants of maize), but we do not know whether they are required for storage protein accumulation under normal conditions.

The mechanisms of protein targeting and protein body formation are also obscure. Although it is well established that the 2S albumin, 7S globulin, and 11S globulin storage proteins are transported to the vacuole, none of these proteins appears to contain a cleavage-targeting peptide like the N- or C-terminal peptides present on other proteins known to be transported to plant vacuoles (Chrispeels and Raikhel, 1992). Defining the precise targeting mechanisms for these proteins may be difficult if the sorting machinery in the Golgi apparatus recognizes one or more short peptide sequences exposed on the surface of the correctly folded proteins. However, identification of these sorting determinants is necessary for identification and isolation of the receptors within the secretory system that interact with them.

Prolamin targeting and deposition in cereals are less well understood, with some components apparently retained in the ER and others transported to the vacuole. Expression of epitope-tagged mutant proteins will allow the products of transgenes to be followed in a homologous background, whereas analysis of lines with increased or decreased levels of ER lumenal proteins and components of the sorting machinery (for example, small GTP binding proteins and vesicle coat proteins, or COPs; Pelham, 1994) will undoubtedly add to our knowledge of prolamin targeting. However, one additional factor needs to be considered—the physical properties of the folded proteins (and in particular those that form disulfide-bonded polymers), which may result in an ER retention system based on solubility rather than specific recognition processes. The segregation of prolamin mRNAs to specific regions of the rough ER may ensure that this retention does not cause widespread disruption of ER integrity and function.

ACKNOWLEDGMENTS

The authors are grateful to Linda Castle (IACR–Rothamsted, Harpenden, UK) for preparing the figures.

REFERENCES

Allen, R.D., Cohen, E.A., Vonder Haar, R.A., Adams, C.A., Ma, D.P., Nessler, C.L., and Thomas, T.L. (1987). Sequence and expression of a gene encoding an albumin storage protein in sunflower. Mol. Gen. Genet. **210,** 211–218.

Altenbach, S.B., Pearson, K.W., Leung, F.W., and Sun, S.S.M. (1987). Cloning and sequence analysis of a cDNA encoding a Brazil nut protein exceptionally rich in methionine. Plant Mol. Biol. **8,** 239–250.

Altenbach, S.B., Kuo, C.-C., Staraci, L.C., Pearson, K.W., Wainwright, C., Georgescu, A., and Townsend, J. (1992). Accumulation of a Brazil nut albumin in seeds of transgenic canola results in enhanced levels of seed protein methionine. Plant Mol. Biol. **18,** 235–245.

Altschuler, Y., Rosenburg, N., Harel, R., and Galili, G. (1993). The N- and C-terminal regions regulate the transport of wheat γ-gliadin through the endoplasmic reticulum in Xenopus oocytes. Plant Cell **5,** 443–450.

Anisimova, I., Fido, R.J., Tatham, A.S., and Shewry, P.R. (1994). Heterogeneity and polymorphism of sunflower seed 2S albumins. Biotech. Biotechnol. Equip. **7,** 63–69.

Argos, P., Pedersen, K., Marks, M.D., and Larkins, B.A. (1982). A structural model for maize zein proteins. J. Biol. Chem. **257,** 9984–9990.

Bartels, D., Altosaar, I., Harberd, N.P., Barker, R.F., and Thompson, R.D. (1986). Molecular analysis of γ-gliadin gene families at the complex *Gli-1* locus of bread wheat (*T. aestivum* L.). Theor. Appl. Genet. **72,** 845–853.

Beccari, J.B. (1745). De frumento. De bononiensi scientarium et artium. Instituto atque Academia Commentarii, Bologna **2,** 122–127.

Bechtel, D.B., Wilson, J.D., and Shewry, P.R. (1991). Immunocytochemical localization of the wheat storage protein triticin in developing endosperm tissue. Cereal Chem. **68,** 573–577.

Bollini, R., and Chrispeels, M.J. (1978). Characterization and subcellular localization of vicilin and phytohemagglutinin, the two major reserve proteins of *Phaseolus vulgaris* L. Planta **142,** 291–298.

Bollini, R., Vitale, A., and Chrispeels, M.J. (1983). In vivo and in vitro processing of seed reserve protein in the endoplasmic reticulum: Evidence for two glycosylation steps. J. Cell Biol. **96,** 999–1007.

Boronat, A., Martinez, M.C., Reina, M., Puigdomenech, P., and Palau, J. (1986). Isolation and sequencing of a 28kD glutenin-2 gene from maize: Common elements in the 5′ flanking regions among zein and glutenin genes. Plant Sci. **47,** 95–102.

Boston, R.S., Fontes, E.B.P., Shank, B.B., and Wrobel, R.L. (1991). Increased expression of the maize immunoglobulin binding protein homolog b-70 in three zein regulatory mutants. Plant Cell **3,** 497–505.

Bulleid, N.J., and Freedman, R.B. (1988). Defective co-translational formation of disulphide bonds in protein disulphide isomerase deficient microsomes. Nature **335,** 649–651.

Casey, R., Domoney, C., and Ellis, N. (1986). Legume storage proteins and their genes. In Oxford Surveys of Plant Molecular and Cell Biology, Vol. 3, B.J. Miflin, ed (Oxford: Oxford University Press), pp. 1–95.

Casey, R., Domoney, C., and Smith, A.M. (1993). Biochemistry and molecular biology of seed products. In Peas: Genetics, Molecular Biology and Biotechnology, R. Casey and D.R. Davies, eds (Wallingford, UK: CAB International), pp. 121–164.

Ceriotti, A., Pedrazzini, E., Fabrini, M., Zoppe, M., Bollini, R., and Vitale, A. (1991). Expression of the wild-type and mutated vacuolar storage protein phaseolin in *Xenopus* oocytes reveals relationships between assembly and intracellular transport. Eur. J. Biochem. **202,** 959–968.

Chesnut, R.S., Shotwell, M.A., Boyer, S.K., and Larkins, B.A. (1989). Analysis of avenin proteins and the expression of their mRNAs in developing oat seeds. Plant Cell **1,** 913–924.

Chrispeels, M.J. (1991). Sorting of proteins in the secretory system. Annu. Rev. Plant Physiol. Plant Mol. Biol. **42,** 21–53.

Chrispeels, M.J., and Raikhel, N.V. (1992). Short peptide domains target proteins to plant vacuoles. Cell **68,** 613–616.

Crouch, M.L., Tenbarge, K.M., Simon, A.E., and Ferl, R. (1983). cDNA clones for *Brassica napus* seed storage proteins: Evidence from nucleotide sequence analysis that both subunits of napin are cleaved from a precursor polypeptide. J. Mol. Appl. Genet. **2,** 273–283.

D'Amico, L., Valsasina, B., Daminati, M.G., Fabbrini, M.S., Nitti, G., Bollini, R., Ceriotti, A., and Vitale, A. (1992). Bean homologs of the mammalian glucose-regulated proteins: Induction by tunicamycin and interaction with newly synthesized seed storage proteins in the endoplasmic reticulum. Plant J. **2,** 443–455.

De Clercq, A., Vandewiele, M., Van Damme, J., Guerche, P., Van Montagu, M., Vandekerckhove, J., and Krebbers, E. (1990). Stable accumulation of modified 2S albumin seed storage proteins with higher methionine contents in transgenic plants. Plant Physiol. **94,** 970–979.

D'Hondt, K., Bosch, D., Van Damme, J., Goethals, M., Vandekerckhove, J., and Krebbers, E. (1993). An aspartic protease present in seed cleaves *Arabidopsis* 2S albumin precursors *in vitro*. J. Biol. Chem. **268,** 20884–20891.

Dickinson, C.D., Hussein, E.H.A., and Nielsen, N.C. (1989). Role of posttranslational cleavage in glycinin assembly. Plant Cell **1,** 459–469.

Duranti, M., Guerrieri, N., Takajashi, T., and Cerletti, P. (1988). The legumin-like storage proteins of *Lupinus albus* seeds. Phytochemistry **27,** 15–23.

Egorov, T.A. (1988). The amino acid sequence of the 'fast' avenin component (*Avena sativa* L.). J. Cereal Sci. **8,** 289–292.

Entwistle, J., Knudsen, S., Muller, M., and Cameron-Mills, V. (1991). Amber codon suppression: The *in vivo* analysis of two C-hordein genes from barley. Plant Mol. Biol. **17,** 1217–1231.

Ericson, M.L., Rodin, J., Lenman, M., Glimelius, K., Josefsson, L.-G., and Rask, L. (1986). Structure of the rapeseed 1.7 S storage protein, napin, and its precursor. J. Biol. Chem. **261,** 14576–14581.

Esen, A., and Stetler, D.A. (1992). Immunocytochemical location of γ-zein in the protein bodies of maize endosperm. Am. J. Bot. **79,** 243–248.

Field, J.M., Tatham, A.S., and Shewry, P.R. (1987). The structure of a high M_r subunit of durum wheat (*T. durum*) gluten. Biochem. J. **247,** 215–221.

Galau, G.A., Wand, H.Y.-C., and Hughes, D.W. (1992). Cotton *Mat5-A* (C164) gene and *Mat5-D* cDNAs encoding methionine-rich 2S albumin storage proteins. Plant Physiol. **99,** 779–782.

Gatehouse, J.A., Croy, R.R.D., and Boulter, D. (1984). The synthesis and structure of pea storage proteins. CRC Crit. Rev. Plant Sci. **1,** 287–314.

Gayler, K.R., Kolivas, S., MacFarlane, A.J., Lilley, G.G., Baldi, M., Blagrove, R.J., and Johnson, E.D. (1990). Biosynthesis, cDNA and amino acid sequence of a precursor of conglutin δ, a sulphur-rich protein from *Lupinus angustifolium*. Plant Mol. Biol. **15,** 879–893.

Gibbs, P.E.M., Strongin, K.B., and McPherson, A. (1989). Evolution of legume seed storage proteins—A domain common to legumins and vicilins is duplicated in vicilins. Mol. Biol. Evol. **6,** 614–623.

Giorini, S., and Galili, G. (1991). Characterization of HSP-70 cognate proteins from wheat. Theor. Appl. Genet. **82,** 615–620.

Halford, N.G., Forde, J., Anderson, O.D., Greene, F.C., and Shewry, P.R. (1987). The nucleotide and deduced amino acid sequences of an HMW glutenin subunit gene from chromosome 1B of bread wheat (*Triticum aestivum* L.), and comparison with those of genes from chromosomes 1A and 1D. Theor. Appl. Genet. **75,** 117–126.

Hall, T.C., McLeester, R.C., and Bliss, F.A. (1977). Equal expression of the maternal and paternal alleles for polypeptide subunits of the major storage protein of the bean *Phaseolus vulgaris* L. Plant Physiol. **59,** 1122–1124.

Hara-Nishimura, I., Takeuchi, Y., Inoue, K., and Nishimura, M. (1993). Vesicle transport and processing of the precursor to 2S albumin in pumpkin. Plant J. **4,** 793–800.

Harris, N., Henderson, J., Abbot, S.J., Mulcrone, J., and Davies, J.T. (1993). Seed development and structure. In Seed Storage Compounds, Vol. 35, P.R. Shewry and A.K. Stobart, eds (Oxford: Oxford Scientific Publications), pp. 3–21.

Higgins, T.J.V., Chandler, P.M., Randall, P.J., Spencer, D., Beach, L.R., Blagrove, R.J., Kortt, R.J., and Inglis, A.S. (1986). Gene structure, protein structure, and regulation of the synthesis of a sulfur-rich protein in pea seeds. J. Biol. Chem. **261,** 11124–11130.

Hull, G., Halford, N.G., Kreis, M., and Shewry, P.R. (1991). Isolation and characterization of genes encoding rye prolamins containing a highly repetitive sequence motif. Plant Mol. Biol. **17,** 1111–1115.

I'Anson, K.J., Morris, V.J., Shewry, P.R., and Tatham, A.S. (1992). Small angle X-ray scattering studies of the C hordeins of barley. Biochem. J. **287,** 183–185.

Irwin, S.D., Keen, J.N., Findlay, J.B.C., and Lord, J.M. (1990). The *Ricinus communis* 2S albumin precursor: A single preproprotein may be processed into two different heterodimeric storage proteins. Mol. Gen. Genet. **222,** 400–408.

Kasarda, D.D., Autran, J.-C., Lew, E.J.-L., Nimmo, C.C., and Shewry, P.R. (1983). N-terminal amino acid sequences of ω-gliadins and ω-secalins: Implications for the evolution of prolamin genes. Biochim. Biophys. Acta **747,** 138–150.

Kim, W.T., and Okita, T.W. (1988a). Structure, expression, and heterogeneity of the rice seed prolamines. Plant Physiol. **88,** 649–655.

Kim, W.T., and Okita, T.W. (1988b). Nucleotide and primary sequence of a major rice prolamine. FEBS Lett. **231,** 308–310.

Kirihara, J.A., Petri, J.B., and Messing, J. (1988). Isolation and sequence of a gene encoding a methionine-rich 10-kDa zein protein from maize. Gene **71,** 359–370.

Ko, T.-P., Ng, J.D., and McPherson, A. (1993). The three-dimensional structure of canavalin from jack bean (*Canavalia ensiformis*). Plant Physiol. **101,** 729–744.

Kortt, A.A., and Caldwell, J.B. (1990). Low molecular weight albumins from sunflower seed: Identification of a methionine-rich albumin. Phytochemistry **29,** 2805–2810.

Kortt, A.A., Caldwell, J.B., Lilley, G.G., and Higgins, T.J.V. (1991). Amino acid and cDNA sequences of a methionine-rich 2S protein from sunflower seed (*Helianthus annuus* L.). Eur. J. Biochem. **195,** 329–334.

Krebbers, E., Da Silva Conceicao, A., Denis, M., D'Hondt, K., and Vandekerckhove, J. (1993). Modification of plant seed storage proteins. In Seed Storage Compounds: Biosynthesis, Interactions and Manipulation, P.R. Shewry and A.K. Stobart, eds (Oxford: Oxford University Press), pp. 317–324.

Kreis, M., and Shewry, P.R. (1989). Unusual features of seed protein structure and evolution. BioEssays **10,** 201–207.

Kreis, M., Forde, B.G., Rahman, S., Miflin, B.J., and Shewry, P.R. (1985a). Molecular evolution of the seed storage proteins of barley, rye and wheat. J. Mol. Biol. **183,** 499–502.

Kreis, M., Shewry, P.R., Forde, B.G., Forde, J., and Miflin, B.J. (1985b). Structure and evolution of seed storage proteins and their genes, with particular reference to those of wheat, barley and rye. In Oxford Surveys of Plant Cell and Molecular Biology, Vol. 2, B.J. Miflin, ed (Oxford: Oxford University Press), pp. 253–317.

Krishnan, H.B., Franceschi, V.R., and Okita, T.W. (1986). Immunochemical studies on the role of the Golgi complex in protein-body formation in rice seeds. Planta **169,** 471–480.

Larkins, B.A., and Hurkman, W.J. (1978). Synthesis and deposition of zein in protein bodies of maize endosperm. Plant Physiol. **62,** 256–263.

Lawrence, M.C., Suzuki, E., Varghese, J.N., Davis, P.C., Van Donkelaar, A., Tulloch, P.A., and Colman, P.M. (1990). The three dimensional structure of the seed storage protein phaseolin at 3 Å resolution. EMBO J. **9,** 9–15.

Lawrence, M.C., Izard, T., Beuchat, M., Blagrove, R.J., and Colman, P.M. (1994). Structure of phaseolin at 2.2Å resolution. Implications for a common vicilin/legumin structure and the genetic engineering of seed storage proteins. J. Mol. Biol. **238,** 748–776.

Lending, C.R., Chesnut, R.S., Shaw, K.L., and Larkins, B.A. (1989). Immunolocalization of avenin and globulin storage proteins in developing endosperm of *Avena sativa* L. Planta **178,** 315–324.

Lending, C.R., Chesnut, R.S., Shaw, K.L., and Larkins, B.A. (1992). Synthesis of zeins and their potential for amino acid modification. In Plant Protein Engineering, P.R. Shewry and S. Gutteridge, eds (Cambridge: Cambridge University Press), pp. 209–218.

Levanony, H., Rubin, R., Altschuler, Y., and Galili, G. (1992). Evidence for a novel route of wheat storage proteins to vacuoles. J. Cell Biol. **119,** 1117–1128.

Li, X., Franceschi, V.R., and Okita, T.W. (1993a). Segregation of storage protein mRNAs on the rough endoplasmic reticulum membranes of rice endosperm cells. Cell **72,** 869–879.

Li, X., Wu, Y., Zhang, D.-Z., Gillikin, J.W., Boston, R.S., Franceschi, V.R., and Okita, T.W. (1993b). Rice prolamine protein body biogenesis: A BiP-mediated process. Science **262,** 1054–1056.

Lilley, G.G., and Inglis, A.S. (1986). Amino acid sequence of conglutins, a sulfur-rich seed protein of *Lupinus angustifolium* L. FEBS Lett. **195,** 235–241.

Marks, M.D., and Larkins, B.A. (1982). Analysis of sequence microheterogeneity among zein messenger RNAs. J. Biol. Chem. **257,** 9976–9983.

Marks, M.D., Lindell, J.S., and Larkins, B.A. (1985). Nucleotide sequence analysis of zein mRNAs from maize endosperm. J. Biol. Chem. **260,** 16451–16459.

Maschke, O. (1859). Ueber den Bau und die Bestandtheile der Kleberbläschen in Bertholletia, deren Entwickelung in Ricinus, nebst einigen Bemerkungen über Amylonbläschen. Botanische Zeitung **17,** 409–447.

Masumura, T., Shibata, D., Hibino, T., Kato, T., Kawabe, K., Takeba, G., Tanaka, K., and Fugii, S. (1989). cDNA cloning of an mRNA encoding a sulfur-rich 10 kDa prolamin polypeptide in rice seeds. Plant Mol. Biol. **12,** 123–130.

Masumura, T., Hibino, T., Kidzu, K., Mitsukawa, N., Tanaka, K., and Fujii, S. (1990). Cloning and characterization of a cDNA encoding a rice 13 kDa prolamin. Mol. Gen. Genet. **221,** 1–7.

Menéndez-Arias, L., Monsalve, R.I., Gavilanes, J.G., and Rodriguez, R. (1987). Molecular and spectroscopic characterization of a low molecular weight seed storage protein from yellow mustard (*Sinapis alba* L.). Int. J. Biochem. **19,** 899–907.

Miflin, B.J., Field, J.M., and Shewry, P.R. (1983). Cereal storage proteins and their effects on technological properties. In Seed Proteins, J. Daussant, J. Mosse, and J. Vaughan, eds (London: Academic Press), pp. 255–319.

Miles, M.J., Carr, H.J., McMaster, T., I'Anson, K.J., Belton, P.S., Morris, V.J., Field, J.M., Shewry, P.R., and Tatham, A.S. (1991). Scanning tunnelling microscopy of a wheat gluten protein reveals

details of an unusual supersecondary structure. Proc. Natl. Acad. Sci. USA **88,** 68–71.

Müntz, K. (1989). Intracellular protein sorting and the formation of protein reserves in storage tissue cells of plant seeds. Biochem. Physiol. Pflanzen **185,** 315–335.

Müntz, K., Jung, R., and Saalbach, G. (1993). Synthesis, processing and targeting of legume seed proteins. In Seed Storage Compounds: Biosynthesis, Interactions, and Manipulation, P.R. Shewry and A.K. Stobart, eds (Oxford: Oxford University Press), pp. 128–146.

Osborne, T.B. (1924). The Vegetable Proteins. (London: Longmans, Green).

Patel, S., Cudney, R., and McPherson, A. (1994). Crystallographic characterization and molecular symmetry of edestin, a legumin from hemp. J. Mol. Biol. **235,** 361–363.

Payne, P.I. (1987). Genetics of wheat storage proteins and the effect of allelic variation on breadmaking quality. Annu. Rev. Plant Physiol. **38,** 141–153.

Pedersen, K., Devereux, J., Wilson, D.R., Sheldon, E., and Larkins, B.A. (1982). Cloning and sequence analysis reveal structural variation among related zein genes in maize. Cell **29,** 1015–1026.

Pedersen, K., Argos, P., Naravana, S.V.L., and Larkins, B.A. (1986). Sequence analysis and characterization of a maize gene encoding a high-sulfur zein protein of M_r 15,000. J. Biol. Chem. **261,** 6279–6284.

Pedrazzini, E., Giovinazzo, G., Bollini, R., Ceriotti, A., and Vitale, A. (1994). Binding of BiP to an assembly-defective protein in plant cells. Plant J. **5,** 103–110.

Pelham, H.R.B. (1994). About turn for COPs? Cell **79,** 1125–1127.

Rechinger, K.B., Simpson, D.J., Svendson, I., and Cameron-Mills, V. (1993). A role for γ-3 hordein in the transport and targeting of prolamin peptides to the vacuole of developing barley endosperm. Plant J. **4,** 841–853.

Roden, L.T., Miflin, B.J., and Freedman, R.B. (1982). Protein-disulphide isomerase is located in the endoplasmic reticulum of developing wheat endosperm. FEBS Lett. **138,** 121–124.

Rodin, J., and Rask, L. (1990). Characterization of matteuccin, the 2.2S storage protein of the ostrich fern. Evolutionary relationship to angiosperm seed storage proteins. Eur. J. Biochem. **192,** 101–107.

Rubin, R., Levanony, H., and Galili, G. (1992). Evidence for the presence of two different types of protein bodies in wheat endosperm. Plant Physiol. **99,** 718–724.

Saalbach, G., Jung, R., Kunze, G., Saalbach, I., Adler, K., and Müntz, K. (1991). Different legumin protein domains act as vacuolar targeting signals. Plant Cell **3,** 695–708.

Sayanova, O., Mekhedov, S., Zhelnin, L., Khoklova, T., and Ananiev, E. (1993). Nucleotide sequence of a barley C hordein gene. Genetika **29,** 1070–1079.

Scott, M.P., Jung, R., Müntz, K., and Nielsen, N.C. (1992). A protease responsible for post-translational cleavage of a conserved Asn-Gly linkage in glycinin, the major seed storage protein of soybean. Proc. Natl. Acad. Sci. USA **89,** 658–662.

Sharief, F.S., and Li, S.-L. (1982). Amino acid sequence of small and large subunits of seed storage protein from *Ricinus communis.* J. Biol. Chem. **257,** 14753–14759.

Shewry, P.R. (1995). Plant storage proteins. Biol. Rev. **70,** 375–426.

Shewry, P.R., Halford, N.G., and Tatham, A.S. (1989). The high molecular weight subunits of wheat, barley and rye: Genetics, molecular biology, chemistry and role of wheat gluten structure and functionality. In Oxford Surveys of Plant Molecular and Cell Biology, Vol. 6, B.J. Miflin, ed (Oxford: Oxford University Press), pp. 163–219.

Shewry, P.R., Halford, N.G., and Tatham, A.S. (1992). The high molecular weight subunits of wheat glutenin. J. Cereal Sci. **15,** 105–120.

Shewry, P.R., Miles, M.J., and Tatham, A.S. (1993). The prolamin storage proteins of wheat and related species. Prog. Biophys. Mol. Biol. **61,** 37–59.

Stamnes, M.A., Rutherford, S.L., and Zuker, C.S. (1992). Cyclophilins: A new family of proteins involved in intracellular folding. Trends Cell Biol. **2,** 272–276.

Tabe, L.M., Higgins, C.M., McNabb, W.C., and Higgins, T.J.V. (1993). Genetic engineering of grain and pasture legumes for improved nutritive value. Genetika **90,** 181–200.

Tamas, L., Greenfield, J., Halford, N.G., Tatham, A.S., and Shewry, P.R. (1994). A β-turn rich barley seed protein is correctly folded in *E. coli.* Protein Expression Purif. **5,** 357–363.

Tatham, A.S., Shewry, P.R, and Belton, P.S. (1985). ^{13}C NMR study of C hordein. Biochem. J. **232,** 617–620.

Tatham, A.S., Masson, P., and Popineau, Y. (1990). Conformational studies of peptides derived from the enzymic hydrolysis of a γ-gliadin. J. Cereal Sci. **11,** 1–13.

Tatham, A.S., Field, J.M., Morris, V.J., l'Anson, K.J., Cardle, L., Dufton, M.J., and Shewry, P.R. (1993). Solution conformational analysis of the α-zein proteins of maize. J. Biol. Chem. **268,** 26253–26259.

Templeman, T.S., Demaggio, A.E., and Stetler, D.A. (1987). Biochemistry of fern spore germination: Globulin storage proteins in *Matteuccia struthiopteris* L. Plant Physiol. **85,** 343–349.

Thanh, V.H., and Shibasaki, K. (1979). Major proteins of soybean seeds. Reversible and irreversible dissociation of β-conglycinin. J. Agric. Food Chem. **27,** 805–811.

Vandekerckhove, J., Van Damme, J., Van Lijsebettens, M., Botterman, J., De Block, M., Vandewiele, M., De Clercq, A., Leemans, J., Van Montagu, M., and Krebbers, E. (1989). Enkephalins produced in transgenic plants using modified 2S seed storage proteins. Bio/Technology **7,** 929–932.

Vitale, A., Ceriotti, A., and Denecke, J. (1993). The role of the endoplasmic reticulum in protein synthesis, modification and intracellular transport. J. Exp. Bot. **44,** 1417–1444.

Wallace, J.C., Galili, G., Kawata, E.E., Cuellar, R.E., Shotwell, M.A., and Larkins, B.A. (1988). Aggregation of lysine-containing zeins into protein bodies in *Xenopus* oocytes. Science **240,** 662–664.

Wickner, W.T. (1994). How ATP drives proteins across membranes. Science **266,** 1197–1198.

Youle, R.J., and Huang, A.H.C. (1981). Occurrence of low molecular weight and high cysteine containing albumin storage proteins in oilseeds of diverse species. Am. J. Bot. **68,** 44–48.

Zhang, F., and Boston, R.S. (1992). Increases in binding protein (BiP) accompany changes in protein body morphology in three high-lysine mutants of maize. Protoplasma **171,** 142–152.