

Pea Legumin Overexpressed in Wheat Endosperm Assembles into an Ordered Paracrystalline Matrix¹

Eva Stöger, Mary Parker, Paul Christou, and Rod Casey*

Molecular Biotechnology Unit (E.S., P.C.) and Department of Applied Genetics (R.C.), John Innes Centre, Norwich Research Park, Colney Lane, Norwich NR4 7UH, United Kingdom; and Institute of Food Research, Norwich Research Park, Colney Lane, Norwich NR4 7UA, United Kingdom (M.P.)

Legumin, a major component of pea seed storage vacuoles, is synthesized by a number of paralogous genes. The polypeptides are cleaved posttranslationally and can form mixed hexamers. This heterogeneity hampers structural studies, based on the production of hexamer crystals *in vitro*. To study a single type of homogenous legumin we produced pea legumin A in transgenic wheat (*Triticum aestivum*) endosperm where prolamins are predominant and only small amounts of globulins accumulate in separate inclusions. We demonstrated that the legumin precursor was cleaved posttranslationally and we confirmed assembly into 11S hexamers. Legumin was deposited within specific regions of the inclusion bodies. Angular legumin crystals extended from the inclusion bodies into the vacuole, correlating with the high legumin content. This suggests that the high-level production of a single type of legumin polypeptide resulted in the spontaneous formation of crystals *in vivo*. The use of a heterologous cereal system such as wheat endosperm to produce, isolate, and recrystallize homogenous 11S legume globulins offers exciting possibilities for structural analysis and characterization of these important seed storage proteins.

There are two major types of storage protein in seeds: the alcohol-soluble prolamins found predominantly in cereal grains, and the salt-soluble globulins found largely in dicots, but also present to a lesser degree in cereals (Shewry and Casey, 1999).

The globulins, which have been most intensively studied in legumes, fall into two distinct structural groups with sedimentation coefficients of about 11S (320–400 kD) and 7S (145–190 kD; Casey et al., 1986). 11S globulins are synthesized, assembled, and sequestered into protein bodies in a complex process that includes several post-translational modifications (Jung et al., 1997 and refs. therein). When purified from seeds, mature 11S globulins are hexameric complexes in which each morphological subunit consists of acidic and basic polypeptides linked by a single disulphide bond (Nielsen, 1984). Each polypeptide pair is initially synthesized as a single precursor. This is cleaved posttranslationally at an Asn-Gly bond (Scott et al., 1992; Hara-Nishimura et al., 1995) after the precursors have assumed a higher order structure and have assembled into 9S trimers within the endoplasmic reticulum (Chrispeels et al., 1982). Legumin, the 11S globulin of peas, is produced by a minimum of 11 genes (Casey et al., 1993; Casey and Domoney, 1999). The individual 11S globulin polypeptides have sequence identities between 50% and 95% (Casey

et al., 1986) and are capable of forming mixed hexamers comprising up to six different polypeptides. Analysis of the mechanisms by which the polypeptides and assembled subunit complexes are targeted to the protein body is, therefore, complicated by the existence of multiple precursors. In addition, heterogeneity arising from this multiplicity of polypeptides has made structural determination more difficult than would be the case with a homogenous population of molecules. Although 11S globulin crystals have been obtained from several sources, they have proved unsatisfactory for structural analyses (Lawrence, 1999) and we still have no detailed crystallographic information on these important proteins.

The production of recombinant 11S globulins *in vitro* and *in vivo* may facilitate the elucidation of various steps in their biogenesis, assembly, and deposition. The production of recombinant globulins in *Escherichia coli* and yeast revealed a low-resolution (approximately 6Å) structure for glycinin, the 11S globulin of soybean (Utsumi et al., 1996). However, such structures always represent the unprocessed polypeptides, because yeast and *E. coli* do not contain the appropriate proteolytic processing activity to produce the mature polypeptides. For this reason the structures thus revealed are unlikely to represent the true situation *in vivo*, where correct assembly to hexamers is dependent upon such proteolysis. Several systems have been devised to overcome this limitation, all of them based on the production of 11S globulin precursors by *in vitro* transcription/translation and their subsequent *in vitro* proteoly-

¹ This work was supported by The John Innes Centre and by the Institute of Food Research, which are supported in part by a grant-in-aid from the Biotechnology and Biological Science Research Council.

* Corresponding author; e-mail rod.casey@bbsrc.ac.uk; fax 44-1603-456844.

sis. These systems are successful in the sense that they produce processed polypeptides of the correct size, which can assemble into hexamers (Yang et al., 1990; Jung et al., 1997), but they offer no information on in vivo assembly and packing and it is difficult to make sufficient material for structural analysis.

An attractive alternative strategy is the production of recombinant 11S globulins, containing a single type of subunit, in transgenic plants. The expression of recombinant legume 11S globulins in species that already make large amounts of such proteins could result in the formation of hybrid globulins, which will impose difficulties with respect to structural analyses and the analysis of assembly and packaging. We, therefore, sought a species that carries out the appropriate processing and assembly, but accumulates endogenous 11S globulin subunits at low levels. Some cereals, including rice, oats, and wheat (*Triticum aestivum*), contain 11S globulin homologs. In rice, the 11S globulin glutelin constitutes about 60% of the total seed protein (Li and Okita, 1993) and is deposited in specialized crystalloid protein bodies. Oat globulin represents a similar proportion of seed protein (Robert et al., 1985). As a consequence, although pea legumin has been produced in rice endosperm to a level of between 1% and 4% of the total seed protein (Sindhu et al., 1997), this is a small amount relative to the endogenous glutelin and the possibility of the extensive formation of hybrid legumin-glutelin molecules still exists. Wheat, in contrast, contains an 11S globulin homolog, triticin, that represents only 5% of the total seed protein (Singh et al., 1991).

This paper describes the high-level production of pea legumin containing only a single type of subunit (LegA; Lycett et al., 1984) in transgenic wheat, and its successful processing and assembly into hexamers in vivo. The recombinant legumin was produced in sufficiently large amounts to form extensive crystals, associated with inclusion bodies, in the four outermost cell layers of the endosperm. The dimensions of the individual structural elements within the crystal were entirely consistent with them containing tightly packed, recombinant pea legumin molecules. Although crystalloid protein regions within protein bodies of seeds are not uncommon (for review, see Lott, 1980), this is the first description of the in vivo crystallization of a recombinant protein in a transgenic cereal seed. Such material will help to clarify a number of issues in relation to 11S globulin assembly in vivo, through directed mutagenesis, and will aid the further study of the determinants of transport and packaging of 11S globulins (Müntz, 1998; Robinson and Hinz, 1999). In addition, an 11S globulin composed of a single type of subunit may be more amenable to crystallization for the analysis of three-dimensional structure.

RESULTS

Generation and Characterization of Transgenic Lines Expressing Legumin

We generated transgenic wheat plants containing a cDNA corresponding to a pea legumin gene (*legA*) and the selectable marker *bar* using two separate plasmids (pLEG and pAHC20). Twenty independent transgenic lines expressing phosphinothricin-acetyltransferase were selected and studied in detail.

Southern-blot analysis of primary transformants confirmed the integration of the legumin gene in 15 lines, resulting in a cotransformation frequency of approximately 75%. The transgene integration patterns were clearly unique for each line and, as expected for plants generated by direct gene transfer, the complexity of integration ranged from one to more than 12 transgene copies per haploid genome. Figure 1 shows a Southern blot for lines cobombarded with pLEG and pAHC20. The enzyme used for digestion cuts only once in pLEG, generating a unique banding pattern for each line and confirming that each plant originates from an independent transformation event. DNA samples in lanes 6 and 11 hybridized with the *bar* probe (data not shown), but did not contain the *legA* gene.

Fifteen immature seeds were collected from each transgenic line that contained the legumin gene to obtain three samples per plant, each comprising a pool of five seeds. Each sample was screened by ELISA and legumin A was detected in nine independent lines. Another 25 mature seeds from each of these nine plants were tested individually to determine segregation of the transgene. At the same time, the embryos were isolated from the seeds and were

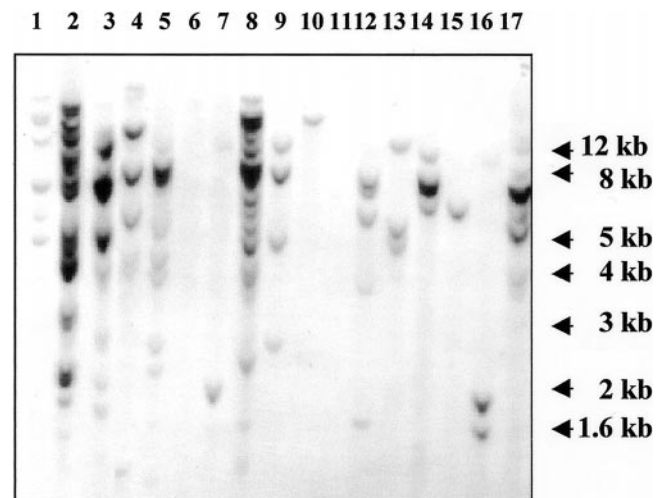


Figure 1. Southern blot of primary transformants. DNA from independent transgenic lines was digested with *Sac*I, which cuts only at one site in the vector pLEG. The blot was probed with the coding sequence of the *legA* gene. Lanes 1–17, Transgenic lines O20, O18, S8, O4, O2, O1, S49, O10, S41, S34, S33, S31, S30, S21, S15, S14, and S11.

germinated under phosphinothricin (PPT) selection. This analysis showed that the *bar* and *legA* genes were linked in all lines and cosegregated as a single locus (Table I), suggesting that all transgenes had integrated at a single site in the genome.

The Low M_r (LMW) Glutenin-Ubiquitin Intron Construct Is Highly Active in the Subaleurone Cells

The legumin gene was driven by the LMW glutenin promoter. We have shown previously that this promoter confers endosperm-specific expression upon a reporter gene (Stöger et al., 1999). In the present study we inserted intron 1 of the maize ubiquitin gene between the LMW promoter and the legumin cDNA. We have used an equivalent construct containing the *gusA* reporter gene for particle bombardment transformation experiments. Analysis of 10 independent transgenic lines containing LMW-I-*gusA* indicated that the reporter gene expression was in most cases confined to the endosperm. In two lines only, GUS (β -glucuronidase) activity was also detected in pollen and stigmata, and in one of these lines GUS was also found in the scutellum. No GUS activity could be detected in leaves of any of the lines. We used the two lines with the highest GUS activity to determine the amounts of GUS at different stages of seed development (Fig. 2). The average GUS activity in mature seeds of the 10 independent lines (19 nmol 4-methylumbelliferone $\text{min}^{-1} \text{mg}^{-1}$ protein; range: 3–35 nmol 4-methylumbelliferone $\text{min}^{-1} \text{mg}^{-1}$) was higher than that obtained previously with the ubiquitin promoter or with the LMW promoter alone (Stöger et al., 1999). Within the endosperm, the highest GUS activity was found in the subaleurone cells.

Legumin A Is Expressed in an Endosperm-Specific Manner

Legumin was detectable from 5 d after pollination (dap) throughout maturation. We observed considerable variation in legumin levels among individual mature seeds from the same line, which is not unex-

Table I. Legumin content and T1 segregation

Line	Legumin Content		Segregation (Legumin Content)	Segregation (Herbicide Resistance)
	Maximum	Minimum		
	$\mu\text{g g}^{-1}$ dry seed wt			+:-
S8	570	310	3:22	3:22
S11	590	100	19:6	19:6
S21	550	150	18:7	18:7
S30	480	100	18:7	18:7
S31	450	<100	19:6	19:6
S41	300	<100	17:8	17:8
O2	320	<100	20:5	20:5
O4	210	<100	12:13	12:13
O10	470	<100	20:5	20:5

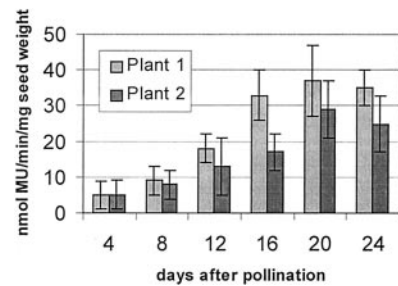


Figure 2. GUS activity (nanomoles per minute per milligram of TSP) at different stages of seed development in two independent plants transformed with LMW-I-*gusA*.

pected since triploid endosperm may contain one, two, or three transgene loci. Similar dosage effects have been shown previously, e.g. for the phaseolin gene in rice (Zheng et al., 1995). The maximum and minimum legumin contents of positive seeds from the sample of 25 seeds per line are shown in Table I. Figure 3 shows the mean legumin content for each line. The highest level was close to $500 \mu\text{g g}^{-1}$ dry weight, equivalent to 1.5% total soluble protein (TSP). We consistently detected no legumin in leaves and embryos of positive lines, which contained legumin A in the endosperm.

We also confirmed that most of the legumin could be extracted with buffer containing 0.14 M NaCl at room temperature (Fig. 4A). The endogenous 11S globulin of wheat, triticin, is not soluble under these conditions (Singh et al., 1991).

Pea Legumin A Is Processed and Assembled in Transgenic Wheat

Seed extracts of plants producing the highest legumin levels (S8, S11, S21, S30, O2, and O10) were further analyzed by western blotting, using a monospecific polyclonal rabbit anti-legumin A IgG (IgG). Under non-reducing conditions a band was detected in the transformed seeds, which comigrated with the authentic pea legumin (molecular mass = 58 kD; Fig. 4A). Under reducing conditions three polypeptides

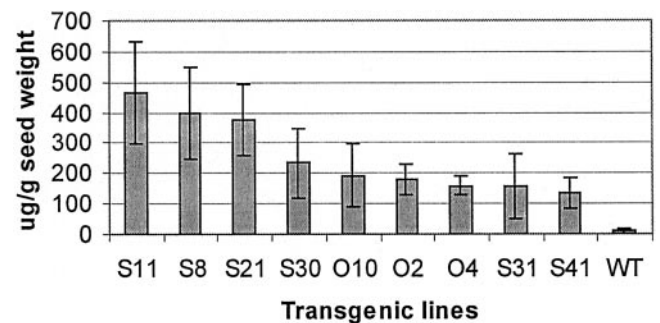


Figure 3. Legumin content in different transgenic wheat lines. Twenty-five seeds from each primary transformant were tested individually and mean values were calculated from the positive measurements.

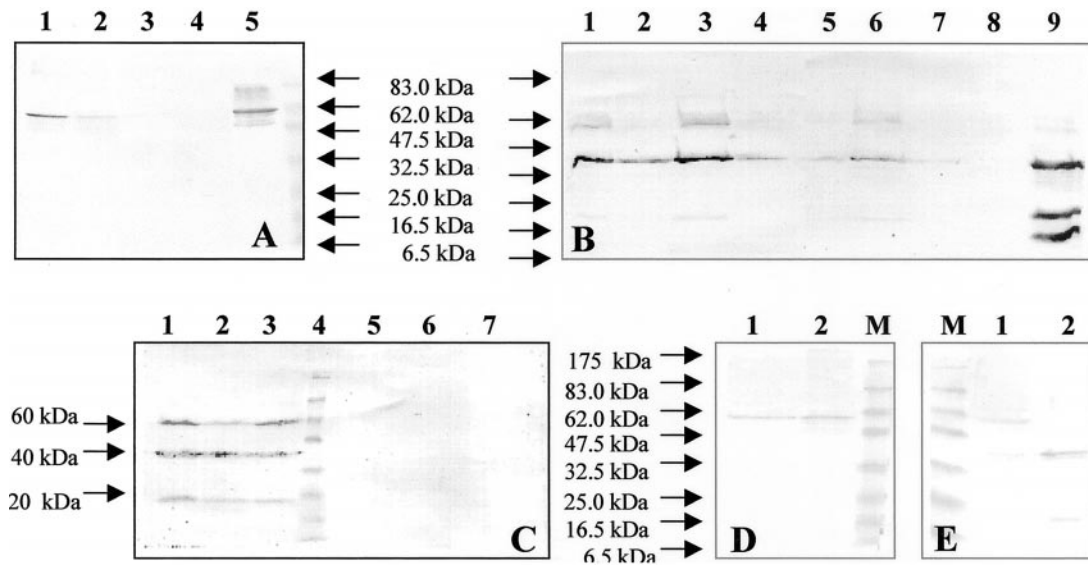


Figure 4. Immunoblot analysis of proteins fractionated by SDS-PAGE and probed with rabbit antibody raised against pea legumin. Reducing agent (10% [w/v] 2-mercaptoethanol) was added to samples in B, C, and E. A, Ground seeds from line S11 were sequentially extracted with 0.14 M NaCl (lane 1) and 1 M NaCl (lane 2) at room temperature and then with 1 M NaCl at 60°C (lane 3). Proteins were separated under non-reducing conditions. Lane 4, Untransformed seed. Lane 5, Pea legumin standard. B, Seed extracts from the highest legumin-producing plants (reducing conditions). Lanes 1 through 7, Seed extracts from lines S8, S30, S11, O2, O10, S21, and O4. Lane 8, Untransformed wheat seed. Lane 9, Pea legumin standard. C, Lanes 1 through 3, Seed extracts from line S21 at different stages of seed development (7, 14, and 21 dap). Lane 4, Size marker. Lanes 5 and 6, Leaf and embryo extracts from the same plant (S21). Lane 7, Untransformed seeds (14 dap). D and E, Fractions obtained after density gradient centrifugation were separated under non-reducing (D) and reducing conditions (E). Lane M, Size marker. Proteins comigrating with the 7S fraction (lane 1) or with the 11S fraction (lane 2).

of 60, 40, and 20 kD were observed (Fig. 4B). The 60-kD protein corresponds to the expected size of the legumin precursor molecule (Lycett et al., 1984). The 40- and 20-kD bands correspond to the acidic and basic legumin polypeptides, respectively. These could be detected at several developmental stages (Fig. 4C). This indicates that in wheat endosperm, the legumin precursor molecule is endoproteolytically cleaved into polypeptides of a similar size to those found in the native pea cotyledons. However, considerable amounts of the unprocessed precursor molecule were also present in wheat extracts, indicating that the efficiency of processing may be lower in the heterologous system. No immunoreactive proteins were detected in non-transgenic seed extracts or in extracts from leaves or embryos of transgenic plants (Fig. 4C, lanes 5–7).

To investigate the oligomeric nature of the processed polypeptides and unprocessed precursors of pea legumin in transgenic wheat seeds we analyzed concentrated extracts by density gradient centrifugation. For this experiment we used seeds from T1 plants derived from S11, the line containing the highest level of legumin. Gradient fractions were separated by 10% (w/v) SDS-PAGE and were analyzed by western blotting (Fig. 4, D and E).

The results confirmed that the processed polypeptides comigrated with the 12S fraction and were probably assembled into hexamers. The unprocessed precursor molecules comigrated with the 7S vicilin,

consistent with it being a trimer (Casey and Doherty, 1999).

Legumin Is Confined to the Outer Four Layers of Wheat Endosperm Cells and Forms Crystals

After removing the embryo, single immature seeds (typically 16 dap) or mature seeds from two of the highest expressing lines (S11 and S21) were cut in one-half and were tested for legumin by ELISA. Legumin contents were rated and the remaining halves of the seeds were fixed and embedded for immunolocalization. We confirmed by immunolocalization that the deposition of legumin was confined to the subaleurone of developing wheat endosperm, consistent with the histochemical detection of GUS activity (Fig. 5A). Using a polyclonal antibody we demonstrated that legumin was associated with the storage protein bodies (Fig. 5B, arrows) in the outer endosperm cells. At the light microscope level, the presence of angular crystals (Fig. 5C, arrows) at the periphery of the protein bodies in these cells was closely correlated with a high legumin content in each seed as measured by ELISA (Table II). We also reconfirmed this correlation between legumin content and crystal formation using seeds from T1 plants derived from S11. At 16 dap, gluten storage protein was found in small protein bodies in the cytoplasm and as partially fused, clustered protein bodies in large central vacuoles (Figs. 5C and 6A). Examination

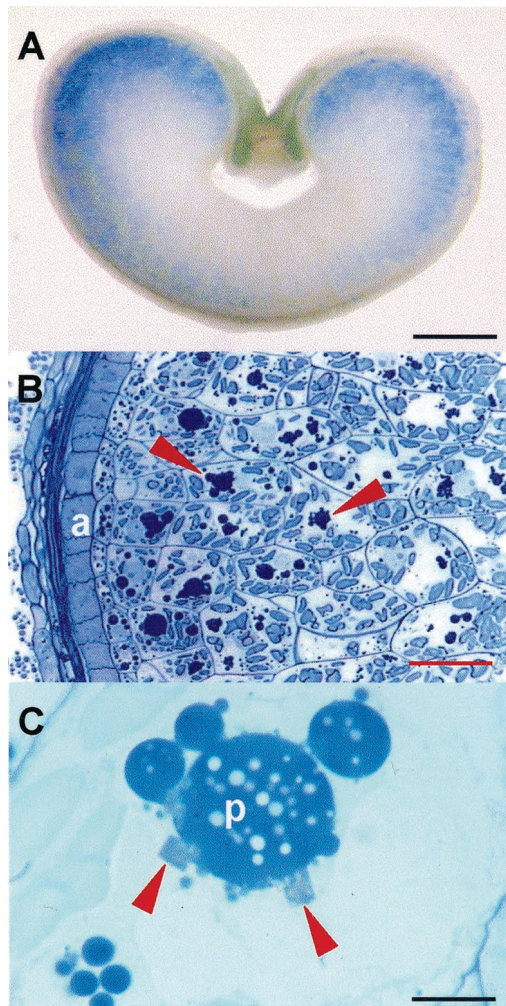


Figure 5. Light microscopy of developing transformed wheat grains. A, Distribution of GUS activity in the outer cells of the endosperm. Bar = 1 mm. B, Outer layers of transformed wheat. The legumin is associated with storage protein bodies (arrows) in endosperm cells adjacent to the aleurone layer (a). Bar = 60 μm . C, Angular crystals (arrows) are located on the periphery of the storage protein bodies (p). Bar = 10 μm .

by transmission electron microscopy showed that at the periphery of these protein masses, and sometimes trapped between them, were the dark-staining spherical inclusion bodies containing tritacin (Fig. 6A). The crystals seen by light microscopy appeared as dark-staining angular bodies at the periphery of the storage protein (Fig. 6A, arrows). Crystals up to 4 μm in length presented triangular (Fig. 6B), trapezoid (Fig. 6C), and rectangular (Fig. 6D) profiles. They were often associated with the inclusion bodies and extended from them into the vacuole as illustrated in Figure 6D. Within the inclusion bodies of the outer endosperm were dark-stained areas, often spherical in outline (Fig. 6D), and the crystals appeared to grow from these. The crystals may have regions of imperfection and can be associated with more than one inclusion body. Higher magnification of the area

marked in Figure 6D showed the lattice pattern present in the crystal and the dark-stained areas within the inclusion bodies (Fig. 6E). Sometimes the crystals and dark-stained areas of the inclusion bodies were amorphous and appeared to lack a lattice structure, but on tilting the section in the electron microscope, the lattice structure was resolved.

In high-expressing seeds, immunogold labeling was observed in two distinct locations in the outer endosperm cells. Label was found over the dark-stained zones (arrows) in some of the inclusion bodies (Fig. 7, A and inset) and over the crystals (Figs. 7B). Several different lattice patterns could be observed when sections were tilted; one of the major crystal planes is illustrated in Figure 7C. This lattice is square and the spacing, relative to the overlying 15-nm colloidal gold marker, is approximately 9 nm. The density of gold label is consistent with a reported labeling efficiency of about 0.5% for tissue embedded in epoxy resin (Barber et al., 1998).

In low-expressing seeds, which lacked crystals, the immunogold label was only found in the dark-stained areas of the inclusion bodies of the outer endosperm (Table II). In nonexpressing (out-segregating) seeds (Fig. 7D), labeling intensity was similar to the very light background level seen over the inclusion bodies of the inner endosperm, the inclusion bodies of untransformed wheat, and in control transformed material in which the primary antibody had been omitted. No labeling was observed on the protein bodies of the aleurone layer.

DISCUSSION

Unlike dicots, cereals synthesize different classes of storage proteins with different solubilities. The prolamins are soluble in dilute alcohol mixtures, whereas the globulins are salt soluble. Globulins and prolamins in cereals accumulate in separate phases. In rice they are deposited in completely separate protein bodies (Krishnan et al., 1986). In oat, globulins are the predominant storage proteins and the relatively small quantity of prolamins occupies a discrete phase within the globulin protein bodies (Lending et al., 1989). In contrast, wheat synthesizes only

Table II. Correlation between legumin content and crystal formation

Grain	ELISA	Crystals	Label in Inclusion Bodies
	μg^{-1}		
S11-1	270	++	Yes
S11-2	460	+++	Yes
S11-3	<100	None	Yes
S11-4	<150	+	Yes
S11-5	310	+++	Yes
S11-6	Not detected	None	No
S11-7	490	+++	Yes
S11-8	430	+++	Yes

small amounts (less than 10% TSP) of globulin, and this is stored in dark-stained spherical inclusion bodies embedded in the periphery of the gluten (prolamin) protein bodies (Bechtel et al., 1991).

To investigate the behavior of globulin-type legume storage proteins in a cereal background we generated transgenic wheat plants overexpressing pea legumin A under the control of a wheat endosperm-specific promoter. Our main objective was to investigate how the over-production of an 11S globulin in transgenic wheat affected its synthesis, processing, assembly, transport, and deposition. The use of a heterologous system allowed us to study a single form of legumin in isolation, whereas in pea, several different legumins are co-expressed in the seed to form a complex molecular milieu. Although

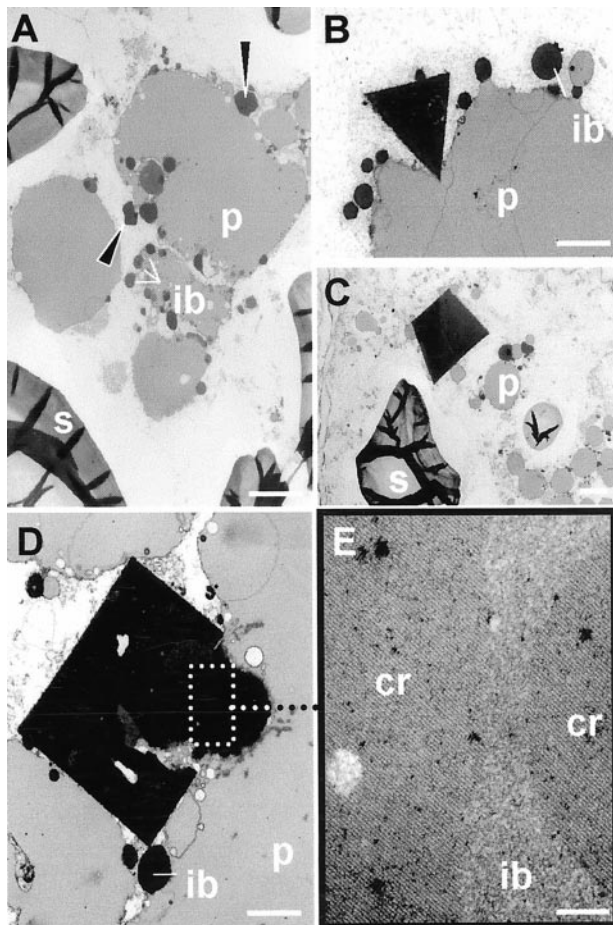


Figure 6. Transmission electron microscopy of developing outer endosperm cells of transformed wheat. Cr, Crystalline lattice; ib, inclusion body; p, protein body; s, starch. A, Protein bodies cluster and fuse in the central vacuole with dark-stained inclusion bodies and crystals (arrows) at the periphery. Bar = 2 μm . B, Triangular crystal at periphery of protein body. Bar = 2 μm . C, Trapezoid crystal associated with small protein bodies. Bar = 2 μm . D, Detail of inclusion body with electron-opaque regions and associated rectangular crystal with imperfections. Bar = 1 μm . E, Enlargement of region outlined in D showing crystalline lattice in the inclusion body and crystal. Bar = 100 nm.

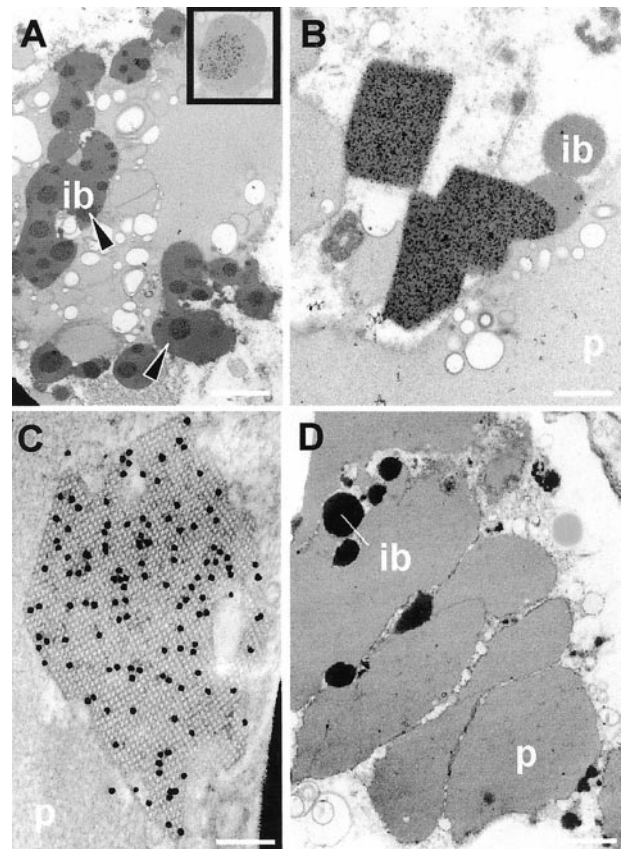


Figure 7. Immunogold localization of legumin in endosperm of transformed wheat 16 dap. Ib, Inclusion body; p, protein body. A, Discrete, dark-stained regions (arrows) of the inclusion bodies immunolabel heavily (inset shows gold particles) for legumin. Bar = 1 μm . B, Angular legumin-containing crystals are heavily immunolabeled and project from the inclusion bodies into the vacuole. Bar = 500 nm. C, Crystal with 9-nm lattice structure decorated with 15-nm immunogold label for legumin. Bar = 100 nm. D, Transformed wheat not expressing legumin. Background immunogold labeling over protein bodies and inclusion bodies is low and similar to that in control untransformed wheat. Bar = 2 μm .

this does not provide a complete picture, this transgenic approach removes some of the complexities involved in studying these processes in a homologous system.

In pea, the endoproteolytic cleavage of the nascent legumin polypeptide is thought to occur within 2 to 3 h of synthesis, on the way from the endoplasmic reticulum to the protein body or within the protein body itself (Chrispeels et al., 1982). In transgenic wheat we found that the legumin precursor was post-translationally processed, but this processing was incomplete. A similar observation was reported when pea legumin was expressed in transgenic rice (Sindhu et al., 1997), whereas in transgenic tobacco complete processing was reported (Ellis et al., 1988). We confirmed that in wheat endosperm the processed polypeptides were assembled into 11S hexamers, whereas the unprocessed precursor polypeptides had a lower sedimentation rate, suggesting they were

present as 8S trimers. This demonstrated that processing and assembly is carried out properly and supports earlier reports in which processing was shown to be necessary for assembly and deposition (Dickinson et al., 1989; Jung et al., 1997, 1998; Nielsen and Nam, 1999). Triticin contains the highly conserved post-translational proteolytic cleavage site between the large, acidic and small, basic polypeptides (between Asn and Gly residues) and two conserved Cys residues, which are thought to link the two polypeptides by a disulphide bond before the post-translational proteolytic cleavage (Jung et al., 1998). Therefore, correct cleavage of legumin is expected in wheat. In contrast to legumin, triticin contains a high proportion of Cys residues, which may form intermolecular sulfide bridges and support the assembly of tetramers and larger polymers (Singh et al., 1991). Legumin, however, is assembled properly into hexamers in wheat and may therefore form separate deposits rather than interspersing with triticin.

There have been few previous studies concerning the localization of recombinant seed storage proteins in heterologous plants. In transgenic tobacco seeds recombinant pea legumin was deposited in vacuolar protein bodies of the endosperm and embryo (Croy et al., 1988). In rice no immunolocalization of recombinant legumin has been reported, but a 7S globulin from bean (β -phaseolin) was expressed in seeds and shown to be localized predominantly in the type II protein bodies, which normally contain glutelin, the highly abundant globulin-type storage protein endogenous to rice (Zheng et al., 1995). Since specialized protein bodies exist for globulins in rice, this distribution was not unexpected. In wheat, however, only a small amount of endogenous globulin (triticin) is produced, and this is stored in dense inclusions on the surface of the protein bodies. It was of interest to establish how the wheat seed would accommodate the increased globulin synthesis consequent to the expression of LegA. In other systems the sudden accumulation of an "unusual" storage protein has induced the formation of novel protein bodies, suggesting perhaps that the machinery is present, although dormant, in many diverse plants. For example, when the prolamin-type maize storage protein γ -zein was expressed in tobacco (a dicot plant in which protein bodies are normally derived from the vacuoles) it accumulated in leaves and seeds in novel endoplasmic reticulum-derived protein bodies (Bagga et al., 1995).

Wheat prolamins are highly hydrophobic and, therefore, would naturally associate and condense into densely-packed spheres in an aqueous environment. Triticin, which is more soluble, aggregates in separate inclusion bodies located on the surface of the gluten protein bodies. We have demonstrated the deposition of legumin in transgenic wheat endosperm within specific regions of the inclusion bodies.

It is perhaps not surprising that proteins with similar solubility characteristics should colocalize, but it is interesting that the legumin is able to condense within the inclusion body and eventually form crystals that extend into the vacuole. The legumin crystals are mostly found in the large central vacuole of the cell where inclusion bodies coalesce, increasing the availability of legumin for crystallization. The crystals have an internal lattice structure, and optical diffraction analysis indicated that the molecular packing was probably square (Fig. 7) with 8.9-nm spacing, consistent with the size of legumin.

The presence of highly ordered legumin as crystals in the endosperm of transgenic wheat is an unexpected and exciting result. In the few pea genotypes that have been examined at the electron microscope level no legumin crystals have been observed. This may be because transgenic wheat contains a single type of legumin, whereas endogenous pea legumin is expressed from at least 11 paralogous genes, allowing the formation of heteromeric hexamers (Casey et al., 1986). The endogenous 11S globulin in wheat, triticin, is only approximately 40% identical (50% similar) in protein sequence to legumin A. Furthermore, upstream of the processing site, a Lys-rich repetitive domain about 65 amino acids is present in triticin (Singh et al., 1993), but absent from legumin, and this may reduce the likelihood of the formation of hybrid triticin-legumin molecules or may be instrumental in sorting the proteins to different locations. The tendency of triticin to form polymers may also prevent hybrids. Thus, pea legumin A may form crystals solely due to its deposition in an unusually homogeneous form. In a similar manner, triticin may not form crystals in hexaploid wheat because it is also a mixture of various gene products and because of intermolecular S-S interactions.

Our initial examination showed that legumin crystals could be detected in immature and mature grains of transgenic wheat. A segregating population of line S11 was chosen for ELISA and parallel immunolocalization of the same grain so that nonexpressing, out-segregating seeds from the same spike could be used as negative controls in addition to untransformed material. Seeds that contained different levels of legumin, but that were grown under identical conditions, could also be compared in this way. The legumin content varied greatly between seeds of the same (heterozygous) plant, as expected since the transgene locus may be present in one, two, or three copies in the triploid endosperm cells. Crystals were detected only in seeds with relatively high levels of legumin ($> 100 \mu\text{g g}^{-1}$), whereas in seeds accumulating lower levels of legumin at the same developmental stage, the protein could be detected only in the dense inclusion bodies, and no crystals were formed. Thus, the formation of crystals appeared to require a homogeneous polypeptide and a minimum level of legumin production. The amount of legumin measured

in individual wheat seeds accounted for up to 2% TSP. One factor favoring high accumulation of recombinant legumin is the exceptional stability of this protein, particularly in hexamer form. It has been shown that hexamers are not attacked by vacuolar processing enzyme, indicating that the trimer-to-hexamer transition renders legumin resistant to the legumin-like Cys endopeptidase. In a similar manner, a preparation of mature legumin could not be degraded by proteinase B, a legumain from germinating vetch seeds (Becker et al., 1995; Müntz, 1998). It has been suggested that legumin could be used as a fusion protein to help stabilize other recombinant proteins (Hall et al., 1990). Processing, stability, and subcellular location are crucial factors for the accumulation of recombinant proteins and an understanding of these mechanisms is of practical importance for the production of valuable proteins in transgenic cereal seeds.

This is the first report to our knowledge concerning *in vivo* crystallization of an 11S globulin in a transgenic seed. Such material could help to clarify a number of issues relating to 11S globulin structure, assembly, transport, and packaging.

Despite their importance no crystal structure for 11S storage globulins is available to date, with the exception of edestin, an 11S globulin from hemp (Patel et al., 1994). Although 11S proteins have been crystallized in abundant quantities from several plants, the crystals have generally been small and disordered and, therefore, not ideal for high-resolution x-ray diffraction analysis. We believe that this is partly due to the natural structural heterogeneity of 11 S storage proteins. Processed pea legumin comprising hexamers of identical protein subunits could be isolated from the transgenic wheat endosperm free of other 11 S proteins. Recrystallization would then allow x-ray diffraction analysis. Such production of recombinant 11S globulins in transgenic wheat endosperm is, therefore, an attractive strategy to generate homogeneous material for three-dimensional protein structure analysis.

MATERIALS AND METHODS

Vectors and Plant Transformation

The LMW promoter, extending from -326 to 30 relative to the transcriptional start site (Colot et al., 1987; kindly provided by M. Hammond-Kosak, John Innes Center), and the nopaline synthase terminator were subcloned into pUC19. The maize ubiquitin-1 intron was excised from pAHC20, a plasmid carrying the selectable PPT-resistance gene *bar* (Christensen and Quail, 1996), with *Bgl*II and *Bam*HI. The resulting fragment was inserted into the pUC-based vector using a *Bam*HI site located between the LMW promoter and the nopaline synthase terminator. The resulting plasmid was again linearized with *Bam*HI to insert the 1,729-bp legumin A cDNA fragment previously excised

from plasmid pRC924 (EMBL accession no. AJ132614) with *Bam*HI. The final plasmid was designated pLEG. In a similar manner, LMW-I-*gusA* was obtained by inserting the *gusA* gene instead of *legA*. Wheat (*Triticum aestivum* cv Bobwhite, genomic complement AABBDD) embryos were cotransformed with pLEG and pAHC20 or LMW-I-*gusA* and pAHC20, respectively. Immature embryos were bombarded and selected on medium containing PPT as described (Altpeter et al., 1996).

Plant Material for Transformation

Wheat plants were grown in the greenhouse or in growth rooms at 15°C/12°C day/night temperature with a 10-h photoperiod during the first 40 d, followed by maintenance at 21°C/18°C day/night temperature with a 16-h photoperiod thereafter.

Southern-Blot Analysis

DNA was prepared from leaf tissue according to Dellaporta et al. (1984). Aliquots of DNA (15 µg) were digested with *Sac*I and were fractionated by 0.9% (w/v) agarose gel electrophoresis. Transfer to nylon membranes and hybridization procedures were carried out according to standard protocols (Sambrook et al., 1989). ³²P-Labeled hybridization probes, comprising the coding region of the *legA* and *bar* genes, respectively, were prepared using the random primer labeling kit (Gibco-BRL, Cleveland).

Progeny Analysis

The segregation of functional *bar* loci was determined as described previously (Stöger et al., 1999).

GUS Assays

Fluorometric GUS assays were carried out using 4-methylumbelliferyl glucuronide as the substrate according to Jefferson (1987). To determine GUS activity during seed development, five independent measurements were carried out using seeds at approximately the same stage. Histochemical staining for GUS activity was performed according to Jefferson (1987), with the exception that the substrate buffer contained 20% (w/v) methanol.

Immunoblot Analysis

Tissue samples were homogenized in PBS (phosphate-buffered saline containing 0.14 or 1 M NaCl, respectively), centrifuged for 15 min, and the supernatant was collected. Aliquots of 10 or 30 µg TSP were denatured and fractionated by 10% (w/v) SDS-PAGE according to standard procedures (Sambrook et al., 1989). Proteins were transferred to nitrocellulose membranes and were probed with polyclonal rabbit anti-legumin IgG. Goat anti-rabbit IgG alkaline phosphatase (AP) conjugate (Promega, Madison, WI)

was used as the secondary antibody. Detection was carried out using AP substrate buffer (Sigma, St. Louis). TSP levels were estimated in tissue extracts by the Bradford protein assay (Bio-Rad, Hercules, CA).

ELISA

Endosperm halves were weighed, ground, and extracted with 300 μL of PBS. For pooled seed samples, 1 mL of buffer was used per 100 mg. After a 30-min incubation on ice with occasional shaking, the homogenate was centrifuged and the supernatant was kept for Bradford and for ELISA assays, using rabbit-anti-legumin primary antibodies and an AP-conjugated goat-anti-rabbit secondary antibody. Purified pea legumin (Casey, 1979) was used as a standard for estimating the legumin content in seeds.

Gradient Centrifugation

Protein extracts were separated at 36,000 rpm using Sorvall rotor type TH641 on a Suc gradient (5%–20% [w/v]) as described (Casey, 1979). Fractions (200 μL) were collected and used for Bradford analysis and ELISA assays. Fractions containing legumin were concentrated and separated by SDS-PAGE.

Antibody Production

Monospecific, polyclonal antibodies were raised against pea legumin in rabbits and were purified by repeated positive and negative immunoaffinity chromatography on columns of legumin and vicilin (Casey, 1979). Affinity purified IgG was stored at 1 mg mL^{-1} in PBS at -20°C .

Microscopy

Developing grains (16 dap) were cut in one-half transversely and the embryo-containing one-half of each grain was assessed for legumin content by ELISA. Small blocks of endosperm with the aleurone layer attached were cut from the upper one-half of the grain and were fixed in 3% (w/v) glutaraldehyde in 0.1 M cacodylate buffer, pH 7.4, for 3 h, and in 1% (w/v) aqueous osmium tetroxide overnight. The tissue was dehydrated in an ethanol series, transferred to acetone, then infiltrated and polymerized in Spurr epoxy resin (Spurr, 1969).

For light microscopy, sections approximately 0.5- μm thick were stained in 1% (w/v) toluidine blue in 1% (w/v) borax, pH 11. For electron microscopy, sections showing silver interference colors were stained sequentially in uranyl acetate (saturated solution in 50% [w/v] ethanol) and Reynold's lead citrate. The sections were examined in a JEOL 1200EX transmission electron microscope.

Immunogold Labeling

Sections mounted on gold grids were treated with 0.56 M sodium metaperiodate (NaIO_4) for 30 min and 0.1 N HCl to

unmask the antigenic determinants (Craig and Goodchild, 1984). All grids were preincubated in 3% (w/v) bovine serum albumin (BSA Fraction V, A 2153, Sigma) in PBS at pH 7.4, and were then incubated in primary antibody (2 $\mu\text{g mL}^{-1}$; affinity purified polyclonal rabbit anti-legumin) in PBS containing 1% (w/v) BSA (Kim et al., 1988). Grids were then immersed in secondary antibody, goat anti-rabbit 15-nm gold (British BioCell International, Cardiff, UK), diluted 1:50 with PBS containing 1% (w/v) BSA. Sections, unstained or stained with uranyl acetate and lead citrate, were examined in the transmission electron microscope. Control material consisted of transformed seeds with negative legumin ELISA levels, untransformed seeds, and omission of the primary antibody.

ACKNOWLEDGMENT

The authors thank Dr. Richard Twyman for critical reading of the manuscript and help with its preparation.

Received October 10, 2000; returned for revision December 3, 2000; accepted December 22, 2000.

LITERATURE CITED

- Altpeter F, Vasil V, Srivastava V, Stöger E, Vasil IK (1996) Accelerated production of transgenic wheat (*Triticum aestivum* L.) plants. *Plant Cell Rep* **16**: 12–17
- Bagga S, Adams H, Kemp JD, Sengupta-Gopalan C (1995) Accumulation of 15-kilodalton zein in novel protein bodies in transgenic tobacco. *Plant Physiol* **107**: 13–23
- Barber DL, Lott JNA, Yang H (1998) Immunocytochemical reactions to intact protein bodies in rice (*Oryza sativa* L.): using antibodies to purified fractions of some rice polypeptides. *J Cereal Sci* **27**: 71–81
- Bechtel DB, Wilson JD, Shewry PR (1991) Immunocytochemical localization of wheat storage protein triticin in developing endosperm tissue. *Cereal Chem* **68**: 573–577
- Becker C, Shutov AD, Nong VH, Senyuk VI, Jung R, Horstmann C, Fischer J, Nielsen NC, Müntz K (1995) Purification, cDNA cloning and characterization of Proteinase-B, an asparagine-specific endopeptidase from germinating vetch (*Vicia sativa*) seeds. *Eur J Biochem* **228**: 456–462
- Casey R (1979) Immunoaffinity chromatography as a means of purifying legumin from *Pisum* (pea) seeds. *Biochem J* **177**: 509–520
- Casey R, Domoney C (1999) Pea globulins. In PR Shewry, R Casey, eds, *Seed Proteins*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 171–208
- Casey R, Domoney C, Ellis N (1986) Legume storage proteins and their genes. *Oxford Surveys Plant Mol Cell Biol* **3**: 1–95
- Casey R, Domoney C, Smith AM (1993) Biochemistry and molecular biology of seed products. In R Casey, DR Davies, eds, *Peas: Genetics, Molecular Biology and Biotechnology*. CAB International, Wallingford, UK, pp 121–163

- Chrispeels MJ, Higgins TJV, Spencer D** (1982) Assembly of storage protein oligomers in the endoplasmic reticulum and processing of the polypeptides in the protein bodies of developing pea cotyledons. *J Cell Biol* **93**: 306–313
- Christensen AH, Quail PH** (1996) Ubiquitin promoter-based vectors for high level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Res* **5**: 213–218
- 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. *EMBO J* **6**: 3559–3564
- Craig S, Goodchild DJ** (1984) Periodate-acid treatment of sections permits on-grid immunogold localization of pea seed vicillin in ER and Golgi. *Protoplasma* **122**: 35–44
- Croy RRD, Evans IM, Yarwood JN, Harris N, Gatehouse JA, Shirsat AH, Kang A, Ellis JR, Thompson A, Boulter D** (1988) Expression of pea legumin sequences in pea, *Nicotiana* and yeast. *Biochem Physiol Pflanzen* **183**: 183–197
- Dellaporta SL, Wood J, Hicks JB** (1984) Maize DNA miniprep. In J Messing, I Sussex, R Malmberg, eds, *Molecular Biology of Plants: A Laboratory Course Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 36–37
- Dickinson CD, Hussein EHA, Nielsen NC** (1989) Role of post-translational cleavage in glycinin assembly. *Plant Cell* **1**: 459–469
- Ellis JR, Shirsat AH, Hepher JN, Gatehouse JA, Croy RRD, Boulter D** (1988) Tissue-specific expression of a pea legumin gene in seeds of *Nicotiana plumbaginifolia*. *Plant Mol Biol* **10**: 203–214
- Hall TC, Bustos MM, Anthony JL, Yang LJ, Domoney C, Casey R** (1990) Opportunities for Bioactive Compounds in Transgenic Plants: Ciba Foundation Symposium 154. Wiley & Sons, Chichester, UK, pp 177–197
- Hara-Nishimura I, Shimada T, Hiraiwa N, Nishimura M** (1995) Vacuolar processing enzyme responsible for maturation of seed proteins. *J Plant Physiol* **145**: 632–640
- Jefferson RA** (1987) Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Mol Biol Rep* **5**: 387–405
- Jung R, Nam Y-W, Saalbach I, Müntz K, Nielsen NC** (1997) Role of the sulfhydryl redox state and disulfide bonds in processing and assembly of 11S seed globulins. *Plant Cell* **9**: 2037–2050
- Jung R, Scott MP, Nam Y-W, Beaman TW, Bassüner R, Saalbach I, Müntz K, Nielsen NC** (1998) The role of proteolysis in the processing and assembly of 11S seed globulins. *Plant Cell* **10**: 343–357
- Kim WT, Franceschi VR, Krishnan HB, Okita TW** (1988) Formation of wheat protein bodies: involvement of the Golgi apparatus in gliadin transport. *Planta* **176**: 173–182
- Krishnan HB, Franceschi VR, Okita TW** (1986) Immunological studies on the role of the Golgi complex in protein-body formation in rice seeds. *Planta* **169**: 471–480
- Lawrence MC** (1999) Structural relationships of 7S and 11S globulins. In PR Shewry, R Casey, eds, *Seed Proteins*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 517–541
- 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, Okita TW** (1993) Accumulation of prolamins and glutelins during rice seed development: a quantitative evaluation. *Plant Cell Physiol* **34**: 385–390
- Lott JNA** (1980) Protein bodies. In NE Tolbert, ed, *The Biochemistry of Plants, Vol 1, The Plant Cell*. Academic Press, New York, pp 589–623
- Lycett GW, Croy RRD, Shirsat AH, Boulter D** (1984) The complete nucleotide sequence of a legumin gene from pea (*Pisum sativum* L.). *Nucleic Acids Res* **12**: 4493–4506
- Müntz K** (1998) Deposition of storage proteins. *Plant Mol Biol* **38**: 77–99
- Nielsen NC** (1984) The chemistry of legume storage proteins. *Phil Trans Royal Soc B* **304**: 287–299
- Nielsen NC, Nam Y-W** (1999) Soybean globulins. In PR Shewry, R Casey, eds, *Seed Proteins*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 285–313
- Patel S, Cudney R, McPherson A** (1994) Crystallographic characterization and molecular symmetry of edestin, a legumin from hemp. *J Mol Biol* **235**: 361–363
- Robert LS, Nozzolillo C, Altosaar I** (1985) Homology between rice glutelin and oat 12S globulin. *Biochim Biophys Acta* **829**: 19–26
- Robinson DG, Hinz G** (1999) Golgi-mediated transport of seed storage proteins. *Seed Science Res* **9**: 267–283
- Sambrook J, Fritsch EF, Maniatis T** (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Scott MP, Jung R, Müntz K, Nielsen NC** (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
- Shewry PR, Casey R** (1999). *Seed Proteins*. Kluwer Academic Publishers, Dordrecht, The Netherlands
- Sindhu AS, Zheng Z, Murai N** (1997) The pea seed storage protein legumin was synthesized, processed, and accumulated stably in transgenic rice endosperm. *Plant Sci* **130**: 189–196
- Singh NK, Donovan GR, Carpenter HC, Skerritt JH, Langridge P** (1993) Isolation and characterization of wheat triticin cDNA revealing a unique lysine-rich repetitive domain. *Plant Mol Biol* **22**: 227–237
- Singh NK, Shepherd KW, Langridge P, Gruen LC** (1991) Biochemical characterization of triticin, a legumin-like protein in wheat endosperm. *J Cereal Sci* **13**: 207–219
- Spurr AR** (1969) A low-viscosity epoxy resin embedding medium for electron microscopy. *J Ultrastruct Res* **26**: 31–43
- Stöger E, Williams S, Keen D, Christou P** (1999) Constitutive versus seed specific expression in transgenic wheat: temporal and spatial control. *Transgenic Res* **8**: 73–82

Utsumi S, Gidamis AB, Takenaka Y, Maruyama N, Adachi M, Mikami B (1996) Crystallization and analysis of normal and modified recombinant soybean proglycinins: three-dimensional structure of normal proglycinin at 6Å resolution. *In* N Parris, A Kato, LK Creamer, J Pearce, eds, *Macromolecular Interactions in Food Technology*. American Chemical Society, Washington, DC, pp 257–270

Yang L, Domoney C, Casey R, Hall TC (1990) Processing *Pisum sativum* seed storage protein precursors *in vitro*. *Cell Res* **2**: 153–162

Zheng Z, Sumi K, Tanaka K, Murai N (1995) The bean seed storage protein β -phaseolin is synthesized, processed, and accumulated in the vacuolar type-II protein bodies of transgenic rice endosperm. *Plant Physiol* **109**: 777–786