The Genetics of Legume Storage Proteins

R. Casey and C. Domoney

Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

The genetics of legume storage proteins

By R. Casey and C. Domoney

John Innes Institute, Colney Lane, Norwich, NR4 7UH

The seed storage proteins of *Pisum* (pea) and *Phaseolus vulgaris* (French bean) exhibit genetic variation for polypeptide structure; genetical studies indicate that most of the major storage protein genes exhibit simple, codominant Mendelian inheritance. Biochemical analysis of the storage protein polypeptides and their messenger RNAs shows that the allelic alternatives are probably small families of closely linked structural genes. Two of these genes – those for the major legumin gene family and for convicilin, both from *Pisum sativum* – have been assigned to specific sites on linkage groups. Genes affecting the synthesis of legumin in *Pisum sativum* and of phaseolin in *Phaseolus vulgaris* have been identified.

Introduction

Grain legume seeds characteristically have a high content of protein that, in some cases, can contribute as much as 50 % to the dry weight of the seed. The majority of this protein consists of 'storage protein', the definition of which usually encompasses proteins that (i) constitute a large proportion of the total seed protein; (ii) are synthesized during seed development; (iii) are sequestered in membrane-bound organelles (protein bodies); and (iv) are hydrolysed on germination to provide carbon and nitrogen skeletons for the developing seedling. The storage proteins of legume seeds have been equated with the globulin fraction of the total protein, an equivalence that has gained general, if not universal (Murray 1979), acceptance. Although there are instances of non-protein-body proteins becoming secondarily adapted to a storage role (urease in *Canavalia ensiformis* is an example (Bailey & Boulter 1971)), this article will confine itself to the globulin storage proteins found in the protein bodies of legume seeds. These fall into two main types, the 7 *S* and 11 *S* proteins, each of which consists of a family of closely related molecules.

The two classes of storage protein (11 *S* and 7 *S*) are widely distributed throughout the Leguminosae and have been the subject of comparative studies at the level of peptide mapping Jackson *et al.* 1969), chemistry (Derbyshire *et al.* 1976), serology (Manteuffel 1982), amino acid sequences (Gilroy *et al.* 1979; Moreira *et al.* 1979; Casey *et al.* 1981 *a*, *b*; Hirano *et al.* 1982) and mRNA or genomic sequences (Sun *et al.* 1981; Schuler *et al.* 1982 *a*, *b*; Gatehouse *et al.* 1982, 1983; Lycett *et al.* 1983). Such proteins are not confined to the Leguminosae; 11 *S*-like proteins have, for instance, been detected in the seeds of oats (Brinegar & Peterson 1982), cucurbits (Blagrove & Lilley 1980; Ohmiya *et al.* 1980), rape (Finlayson 1976; Schwenke *et al.* 1979) and sesame (Mori *et al.* 1979). Only in the case of the Leguminosae, however, have any genetic analyses of storage proteins been made and such studies have been confined largely to *Pisum* and *Phaseolus vulgaris*. Studies of the molecular biology and biochemistry of the storage proteins from *Glycine max*, *Phaseolus vulgaris* and *Pisum sativum* (Gilroy *et al.* 1979; Moreira *et al.* 1979; Casey *et al.* 1981 *a*, *b*; Sun *et al.* 1981; Tumer *et al.* 1981; Barton *et al.* 1982;

Hirano *et al.* 1982; Schuler *et al.* 1982a, b; Hall *et al.* 1983; Lycett *et al.* 1983; Slightom *et al.* 1983) have indicated similarities between the storage proteins of these species. This article will consider therefore the genetics of the storage proteins of *Phaseolus vulgaris* and *Pisum* with reference to the molecular biology of the proteins from these and other genera where appropriate. Because homologies across species exist within the two protein classes and the latter are distinct from each other, we have chosen to discuss $11S$ and $7S$ proteins separately.

The structure and genetics of $11S$ (legumin-like) proteins

The $11S$ proteins of legume seeds are broadly defined as oligomeric proteins of molecular mass 3 to 4×10^5 (Badley *et al.* 1975; Casey 1979a; Croy *et al.* 1979; Blagrove & Lilley 1980; Blagrove *et al.* 1980; Khan *et al.* 1980; Matta *et al.* 1981a, b) that contain six acidic (α-) and six basic (β-) subunits disulphide-bonded as $\alpha\beta$ pairs (Derbyshire *et al.* 1976). In a given $11S$ protein preparation, the majority of the α-subunits have molecular masses of approximately 40×10^3 Da and the β-subunits are of molecular mass approximately 20×10^3 Da. The apparent molecular masses of individual acidic and basic subunits can, however, vary considerably; in particular, the acidic subunits have a range of molecular masses from 10×10^3 to 62×10^3 Da (Moreira *et al.* 1979; Khan *et al.* 1980; Matta *et al.* 1981a) and all appear to be capable of participating in $\alpha\beta$ dimers (Khan *et al.* 1980; Staswick *et al.* 1981). The N-terminal amino acid sequence of the small, 10×10^3 molecular mass acidic subunit from glycinin (the $11S$ protein from *Glycine max*) shows it to be homologous to 'normal' acidic subunits; it probably is a truncated form of acid subunit (Moreira *et al.* 1979; Nielsen *et al.* 1981). Such sequence information is still required in other species to clarify the precise relationships between individual polypeptides. While some progress has been made on the structure of $11S$ proteins from *Glycine max*, *Vicia faba* and *Pisum*, only in the case of legumin (the $11S$ protein from *Pisum*) has extensive variation in subunit structure been well documented and used for investigations of genetics.

Both the acidic and basic subunits of *Pisum* legumin exhibit heterogeneity on electrophoresis in polyacrylamide gels containing sodium dodecyl sulphate (SDS gels) (Thomson *et al.* 1978; Casey 1979a, b; Matta *et al.* 1981b) and on two-dimensional isofocusing-SDS gels (Casey 1979a; Krishna *et al.* 1979; Gatehouse *et al.* 1980). N-terminal sequence analysis of the basic (β-) subunits (Casey *et al.* 1981a) indicates the existence of a number of homologous subunits and suggests the existence of a small gene family. Similar results have been observed for the basic subunits of the $11S$ proteins from *Vicia faba* and *Glycine max* (Gilroy *et al.* 1979; Moreira *et al.* 1979).

Two-dimensional gel electrophoresis shows clear multiplicity of the acidic subunits of legumin in *Pisum* and also provides evidence for both major (α^M) and minor (α^m) classes of acidic subunit (Casey 1979b), in agreement with the results of Thomson *et al.* (1978). The structural relationship between the α^M- and α^m-subunits is not known, nor is the relationship between individual α^M-subunits, but it seems likely that they will show sequence homology, as do the several acidic subunits of glycinin (Moreira *et al.* 1979).

Examination of the subunits of legumin from a range of *Pisum* genotypes has shown the existence of extensive genetic variation (Thomson & Schroeder 1978; Casey 1979a, b; Przybylska *et al.* 1979, 1981; Matta & Gatehouse 1982). Each of the three groups to have

examined the genetics of *Pisum* legumin has used a different system of nomenclature (Thomson & Schroeder 1978; Casey 1979*b*; Matta & Gatehouse 1982). The way in which these systems are related has been described elsewhere (Casey 1982) and does not require repetition here.

The α^M-subunits behave as the products of a single Mendelian gene (Thomson & Schroeder 1978; Casey 1979*b*) with at least five different possible alleles (Matta & Gatehouse 1982). Genetical evidence suggests the existence of a small number of legumin genes in a closely linked array (small gene family) which maps near to the r_a locus on chromosome (linkage group) 7 (Davies 1980; Matta & Gatehouse 1982). Other legumin genes must also exist to produce the α^m-subunits; while the inheritance patterns of minor legumin subunits are not clear, variants do exist and present evidence supports their production from a locus other than that for the α^M-subunits (Matta & Gatehouse 1982).

It is clear from biosynthetic studies (Spencer & Higgins 1980; Croy *et al.* 1980*a*) and from DNA sequencing (Croy *et al.* 1982) that legumin polypeptides are synthesized as precursor molecules containing both α- and β-sequences that are later 'nicked' to produce the characteristic αβ disulphide-linked dimers. It is likely, then, that a legumin structural gene codes for both an α- and β-subunit and it follows that the genetic behaviour of the β-subunits should be similar to that of α-subunits; so far, however, detailed study of the inheritance of β-subunits has been hindered by a lack of suitable variants.

When cell-free translation products of poly A-containing RNA, purified from developing pea seeds, are immunoprecipitated with antibody to legumin, several precursors of molecular mass approximately 60×10^3 Da are specifically precipitated (Spencer & Higgins 1980; C. Domoney and R. Casey, unpublished), although some genotypes appear not to show this type of heterogeneity (Croy *et al.* 1980*a*). Hybrid-release translation experiments, in which plasmids containing cDNA inserts homologous to legumin are used, indicate that a number of different mRNA species are selected by such plasmids (Chandler *et al.* 1983; C. Domoney and R. Casey, unpublished) and strongly suggest the existence of several legumin mRNA species in any given preparation of poly A-containing RNA. Direct measurement of the gene copy number for legumin (Croy *et al.* 1982) suggests the existence of four major legumin single-copy genes in *Pisum*, although it must be remembered that this value is obtained in one genotype with one clone.

In *Glycine max* evidence for a small gene family for the 11*S* protein, glycinin, comes from the amino sequence studies of Moreira *et al.* (1979, 1981) and is supported by the size heterogeneity of the glycinin precursors immunoprecipitated from the cell-free translation products of poly A-containing RNA (Tumer *et al.* 1981; Barton *et al.* 1982). Direct measurement of 11 *S* gene numbers suggests approximately three genes per haploid genome for the particular glycinin probes used (Fischer & Goldberg 1982).

Other types of legumin that do not conform to the conventional model have been observed in *Pisum sativum*, in addition to the minor legumin species containing α^m-subunits (see above). Two dimensional non-reducing/reducing gel electrophoresis has pointed to the existence of disulphide-bonded pairs of molecular mass around 40×10^3 Da, in which the acidic subunit is apparently of a similar size to the basic subunit (Domoney 1981; Matta *et al.* 1981*b*). Genetic studies (Matta & Gatehouse 1982) indicate that these 'small' acidic and basic subunits segregate independently of each other.

Legumin subunit pairs which are larger than the normal αβ dimer of about 60×10^3 Da molecular mass have been reported to be present in small amounts in *Vicia faba* and *Vigna*

unguiculata, where the molecular masses of the disulphide-linked dimer species can be as high as 79×10^3 Da (Khan *et al.* 1980; Matta *et al.* 1981*a*). Such species have not been reported in mature *Pisum sativum* seeds, but a legumin-related polypeptide of molecular mass of about 80×10^3 Da has been observed in extracts of immature *Pisum sativum* cotyledons that have been pulse-labelled with radioactive amino acids (Domoney 1981; Chrispeels *et al.* 1982*a*, *b*); this labelled polypeptide specifically binds to a column of antilegumin. A complementary DNA clone, constructed from poly A-containing messenger RNA (mRNA) from immature *Pisum sativum* seeds, has been identified and contains an insert homologous to a mRNA coding for an 80×10^3 Da molecular mass polypeptide. The latter is immunoprecipitable with antibody to legumin, and peptide mapping experiments suggest a structural relationship between it and the conventional (60×10^3 Da molecular mass) legumin precursor (C. Domoney and R. Casey, unpublished). Studies of the accumulation of *Glycine max* proteins during development have also indicated the presence of a short-lived polypeptide of molecular mass 83×10^3 Da that disappears at about the same time as the major acidic and basic glycinin subunits appear (Spielmann *et al.* 1982).

THE STRUCTURE AND GENETICS OF 7S (VICILIN-LIKE) PROTEINS

The 7S proteins of legume seeds have generally been considered to be more complex than the 11S proteins, because of their substantial physical and serological heterogeneity, a tendency towards association/dissociation behaviour in response to changes in pH and ionic strength and, in some cases, simply because of complexity in apparent subunit structure. Some of this confusion has arisen from early studies that (necessarily) considered whole proteins. Recent work on the sequences of 7S protein subunits, their mRNAs and genes (Sun *et al.* 1981; Gatehouse *et al.* 1982; Hirano *et al.* 1982; Schuler *et al.* 1982*a*, *b*; Hall *et al.* 1983; Lycett *et al.* 1983; Slightom *et al.* 1983) and on the biosynthesis of these proteins (Gatehouse *et al.* 1981; Chrispeels *et al.* 1982*b*) has considerably clarified matters. From studies of *Pisum* vicilin, for example, it has become clear that the plethora of subunits found in vicilin from mature seeds is derived from two types of precursor subunit, one of which is relatively short-lived (Gatehouse *et al.* 1981, 1982; Chrispeels *et al.* 1982*b*). The 7S proteins which have been studied in any depth are *Pisum* vicilin and convicilin, *Glycine max* conglycinin and *Phaseolus vulgaris* phaseolus (Glycoprotein II); of these, all except conglycinin have been analysed genetically.

Pisum vicilin

Vicilin from *Pisum* is a heterogeneous 7S protein that contains a number of different polypeptides, that vary in molecular mass from 12×10^3 to 50×10^3 Da (Thomson & Schroeder 1978; Gatehouse *et al.* 1981). The native protein has a molecular mass of about 1.5×10^5 Da and probably has a trimeric structure with subunits of molecular mass around 50×10^3 Da. All native vicilin preparations seem to contain polypeptides of molecular mass 33×10^3, 19×10^3 and 12.5×10^3 Da in addition to the major 50×10^3 Da molecular mass species, together with small amounts of polypeptides of molecular mass 35×10^3, 16×10^3 and 13.5×10^3 Da. Sequencing studies of proteins and DNA have shown clearly that all the smaller polypeptides are derived by proteolytic processing of a precursor polypeptide of molecular mass approximately 50×10^3 Da; the 16×10^3 Da molecular mass polypeptide is probably a glycosylated form of the 12.5×10^3 Da molecular mass polypeptide (Gatehouse *et al.* 1982).

Hence, as suggested by Thomson *et al.* (1980), the fundamental subunit structure of native vicilins is simpler than their apparent polypeptide composition suggests. Thus the structural genes for vicilin are for precursor polypeptides of molecular mass approximately 50×10^3 Da and one would expect genetical data for polypeptides of molecular mass less than 50×10^3 Da to reflect those for their precursors.

Only one extensive study of variability in the polypetides of vicilin from mature *Pisum* seed has been described (Thomson & Schroeder 1978; Thomson *et al.* 1980). Variation in the molecular masses of the 50×10^3 Da and of the approximately 30×10^3 Da polypeptides, and in the charge of the 12 and 14×10^3 Da molecular mass polypeptides has been observed and the inheritance of variant forms studied. Such observations show all three types of polypeptide to behave as the products of single genes that exhibit simple Mendelian inheritance patterns. None of the vicilin genes has been assigned to a linkage group; it may be assumed that the 33×10^3 Da molecular mass, the 12 and 14×10^3 Da molecular mass and the precursor polypeptides from which these are derived are all the products of the same locus. The only unequivocal evidence for multiple genes for *Pisum* 7S proteins comes from the studies of Gatehouse *et al.* (1982) and Lycett *et al.* (1983), that clearly show the existence of a number of copies of each of two homologous sequences for polypeptides of molecular mass approximately 50×10^3 Da; one of these (classes of) polypeptides is processed into smaller fragments while the other is not. Nothing is known of the genetical relationship of these two types of polypeptides.

Pisum convicilin

Convicilin (Casey & Sanger 1980; Croy *et al.* 1980 *b*) is normally a minor constituent of *Pisum* protein but can become a major one in certain genotypes (Casey & Sanger 1980; Casey *et al.* 1982). It is a 7 S protein of molecular mass $2.2–2.8 \times 10^5$ Da that has a single size class of subunit with a molecular mass of around 70×10^3 Da; it is therefore either a trimer or a tetramer. It is serologically closely related to vicilin, but differs from it in subunit size, amino acid composition and carbohydrate content (Casey & Sanger 1980; Croy *et al.* 1980 *b*). Convicilin exhibits heterogeneity in both the size and charge of its subunits and variation in the molecular mass of convicilin subunits has been reported (Thomson & Schroeder 1978; Casey & Sanger 1980; Matta & Gatehouse 1982). Genetical analyses of molecular mass variants of convicilin subunits show the convicilin locus to behave as a single codominant gene (Casey 1982; Matta & Gatehouse 1982) that is not linked to the locus for legumin α^M-subunits (Casey & Sanger 1980; Matta & Gatehouse 1982). The convicilin structural locus, which is probably identical to that designated *VcA* by Thomson & Schroeder (1978), has been assigned to linkage group 2 and resides between genes *s* and *k*, 3 ± 1 map units from the *k* locus (Matta & Gatehouse 1982).

Phaseolus vulgaris phaseolin

A large percentage of the storage protein of *Phaseolus vulgaris* consists of a 7 S protein variously known as phaseolin (Sun *et al.* 1981), G1 fraction (Sun *et al.* 1974) and Glycoprotein II (Pusztai & Watt 1970). The native protein undergoes pH-dependent reversible association/dissociation between a number of molecular forms, the monomeric form having a molecular mass at neutral pH of $1.4–1.6 \times 10^5$ Da (Pusztai & Watt 1970; Sun *et al.* 1974). The native molecule dissociates under denaturing conditions to give polypeptides in the range 43×10^3–53×10^3 Da (McLeester *et al.* 1973; Pusztai & Stewart 1980). These polypeptides have been classified as

α, β and γ on the basis of their electrophoretic mobility in SDS-gels. Microheterogeneity of the native protein has been demonstrated by a number of techniques (Pusztai & Stewart 1980) while two dimensional gel electrophoresis under denaturing conditions has shown the existence of charge microheterogeneity within the polypeptide size classes (Brown *et al.* 1980); peptide mapping (Ma *et al.* 1980) indicates that the α, β and γ polypeptides are structurally related. Genetic variation in phaseolin polypeptides has been detected using one and two dimensional gel electrophoresis (Romero *et al.* 1975; Hall *et al.* 1977; Brown *et al.* 1981 *b*; Bollini & Vitale 1981). Such variation has been used for analyses of the inheritance of phaseolin polypeptides. The structural genes coding for these proteins are linked, codominant and inherited in a block as a single Mendelian gene (Brown *et al.* 1981 *a, b*). The loci for phaseolin and for those proteins defined as G2/albumin by McLeester *et al.* (1973) are not linked to each other, but their chromosomal location has yet to be identified.

The sequencing of cDNAs from phaseolin mRNAs has confirmed the similarity between the mRNAs for various phaseolin polypeptides (Hall *et al.* 1983) while direct measurement of the copy number of phaseolin structural genes suggests the presence of at least four copies of the phaseolin gene per haploid genome (Brown *et al.* 1982; Hall *et al.* 1983). Similar studies of cDNA sequences for the α and α' subunits of conglycinin (Schuler *et al.* 1982 *a, b*) are also consistent with the production of these homologous *Glycine max* 7S protein subunits from small gene families and with the estimate of approximately five α/α' subunit genes per haploid genome (Goldberg *et al.* 1981).

Variation in the storage proteins of other legumes

Variation in the storage proteins of *Arachis hypogaea, Lupinus angustifolius, Psophocarpus tetragonolobus, Vicia faba* and *Glycine max* has been detected by a number of electrophoretic techniques (Gillespie & Blagrove 1975; Blagrove & Gillespie 1978; Basha 1979; Utsumi *et al.* 1980; Mori *et al.* 1981). In none of these cases, however, has this variation been used to establish the genetics of any storage protein subunit.

Genes which have an effect on storage protein synthesis

Our current knowledge of the control of storage protein synthesis in legumes is at an early stage and present efforts are directed towards a definition of the molecular events associated with the onset of synthesis and accumulation of storage proteins; such studies have been confined to *Pisum sativum, Glycine max* and *Phaseolus vulgaris*. Analyses of this kind are often facilitated by the use of mutants that in some way perturb the synthesis of specific storage proteins and three such genes have been reported to date. In lupins an incompletely dominant gene, the so-called beta subunit controller (Bsc) has been shown to have an influence on the synthesis of the major conglutin β subunit, although it appears that other loci with smaller effects also influence the synthesis of this subunit (Oram *et al.* 1981). Genes have been identified in *Phaseolus vulgaris* that either enhance or diminish the expression of phaseolin (Bliss & Brown 1982; Brown *et al.* 1982) and in at least one case major repression of phaseolin synthesis has been correlated with the presence of a single gene (Bliss & Brown 1982). The r_a locus in *Pisum sativum* is closely linked to the major legumin structural locus and has a clear effect on legumin expression (Davies 1980). In R_aR_a genotypes legumin proportions tend to be considerably

higher than in corresponding near-isogenic $r_a r_a$ genotypes. The r_a locus, however, influences many metabolic pathways including those for the synthesis of sugars, lipids and starch (Coxon & Davies 1982; Kooistra 1962). Thus the effect on legumin synthesis may be secondary to one or more of the pleiotropic effects of the r_a locus.

THE EVOLUTION OF LEGUME SEED STORAGE PROTEINS

It is becoming clear from nucleic acid and protein sequencing studies that storage proteins in a given legume may have homologous counterparts in other legume species. Sequences corresponding to *Pisum sativum* vicilin, for example, are to be found in *Phaseolus vulgaris* (phaseolin) and *Glycine max* (conglycinin), while the subunit sequences of *Pisum sativum* legumin and *Glycine max* glycinin are extremely similar (for references see above). In addition, there is considerable sequence homology between the lectins from *Phaseolus vulgaris*, *Pisum sativum*, *Vicia faba*, *Lens culinaris* and *Canavalia ensiformis* (Cunningham *et al.* 1975; Foriers *et al.* 1977, 1981; Hoffmann *et al.* 1982; Hopp *et al.* 1982). It is reasonable to suggest that such homologous sequences evolved from a common ancestor and have been conserved during evolution because they are in some way important to the synthesis, structure or stability of the storage proteins. Schuler *et al.* (1982*a*, *b*), for instance, have postulated that certain conserved conglycinin sequences may be necessary for mRNA stability, because similar sequences are also found in the 3′ *non*coding regions of mRNAs for other *Glycine max* polypeptides. Lycett *et al.* (1983) suggest that such sequences, which also have homologous counterparts in *Pisum sativum* vicilin, might equally well have been conserved because they are important to the structure or function of the protein.

If one ascribes to a storage protein the 'functions' of being transported, sequestered, stored in a dry state and recognized by proteolytic enzymes on seed germination, it might be assumed that there have been limitations to change in these proteins during their evolution. This is not irreconcilable with the considerable degree of variation observed in, for example, the major legumins from *Pisum*, for in no case is this variation so extreme as to alter the three-dimensional structure of the protein. The fact that legumin-like proteins are found in families as diverse as cucurbits and cereals (see above) suggests a need to conserve at least certain elements of the genes for 11 *S* proteins; alternatively, it is possible that convergent evolution to an 11 *S*-like protein has taken place in response to the requirements of the 'functions' outlined above.

It has been suggested that the ancestors of storage proteins may have had some clear function, such as enzymic activity; the observed oxaloacetate decarboxylase activity of cucurbitin (Hara *et al.* 1976) may give a clue to the origins of 11 *S* storage proteins.

The discovery of small regions of internal homology in the amino acid sequence of glycinin acidic subunits (Moreira *et al.* 1981) raises the possibility that these repeated domains could be the products of a series of gene duplications within the eventual structural loci; the glycinin family of proteins may have evolved as a consequence of a complicated series of gene duplications, to give regions of homology both within and between glycinin polypeptides. Considerably more sequence analysis of the storage proteins and their putative progenitors is required, however, before detailed consideration of the paths of their evolution becomes feasible.

356 R. CASEY AND C. DOMONEY

References

Badley, R. E., Atkinson, D., Hauser, H., Oldani, D., Green, J. P. & Stubbs, J. M. 1975 The structure, physical and chemical properties of the soybean protein glycinin. *Biochim. biophys. Acta* **412**, 214–228.

Bailey, C. J. & Boulter, D. 1971 Urease, a typical seed protein of the Leguminosae. In *Chemotaxonomy of the Leguminosae* (ed. J. B. Harborne, D. Boulter & B. L. Turner), pp. 485–502. London and New York: Academic Press.

Barton, K. A., Thompson, J. F., Madison, J. T., Rosenthal, R., Jarvis, N. P. & Beachy, R. N. 1982 The biosynthesis and processing of high molecular weight precursors of soybean glycinin subunits. *J. biol. Chem.* **257**, 6089–6095.

Basha, S. M. M. 1979 Identification of cultivar differences in seed polypeptide composition of peanuts (*Arachis hypogaea* L.) by two-dimensional polyacrylamide gel electrophoresis. *Pl. Physiol.* **63**, 301–306.

Blagrove, R. J. & Gillespie, J. M. 1978 Variability of the seed globulins of winged bean *Psophocarpus tetragonolobus* (L.) DC. *Aust. J. Pl. Physiol.* **5**, 371–375.

Blagrove, R. J. & Lilley, G. G. 1980 Characterisation of cucurbitin from various species of the Cucurbitaceae. *Eur. J. Biochem.* **103**, 577–584.

Blagrove, R. J., Lilley, G. G. & Davey, R. 1980 Molecular weight of legumin from *Pisum sativum*. *Aust. J. Pl. Physiol.* **7**, 221–225.

Bliss, F. A. & Brown, J. W. S. 1982 Genetic control of phaseolin protein expression in seeds of common bean, *Phaseolus vulgaris* L. *Qualitas Pl. Pl. Fds Hum. Nutr.* **31**, 269–279.

Bollini, R. & Vitale, A. 1981 Genetic variability in charge microheterogeneity and polypeptide composition of phaseolin, the major storage protein of *Phaseolus vulgaris*; and peptide maps of its three major subunits. *Physiologia Pl.* **52**, 96–100.

Brinegar, A. C. & Peterson, D. M. 1982 Synthesis of oat globulin precursors. *Pl. Physiol.* **70**, 1767–1769.

Brown, J. W. S., Bliss, F. A. & Hall, T. C. 1980 Microheterogeneity of globulin-1 storage protein from french bean with isoelectrofocusing. *Pl. Physiol.* **66**, 838–840.

Brown, J. W. S., Bliss, F. A. & Hall, T. C. 1981*a* Linkage relationship between genes controlling seed proteins in french bean. *Theor. appl. Genet.* **60**, 251–259.

Brown, J. W. S., Ersland, D. R. & Hall, T. C. 1982 Molecular aspects of storage protein synthesis during seed development. In *The physiology and biochemistry of seed development, dormancy and germination* (ed. A. A. Khan), pp. 3–42. New York, Amsterdam: Elsevier Biomedical Press.

Brown, J. W. S., Ma, Y., Bliss, F. A. & Hall, T. C. 1981*b* Genetic variation in the subunits of globulin-1 storage protein of french bean. *Theor. appl. Genet.* **59**, 83–88.

Casey, R. 1979*a* Immunoaffinity chromatography as a means of purifying legumin from *Pisum* (pea) seeds. *Biochem. J.* **177**, 509–520.

Casey, R. 1979*b* Genetic variability in the structure of the α-subunits of legumin from *Pisum* – a two-dimensional gel electrophoresis study. *Heredity, Lond.* **43**, 265–272.

Casey, R. 1982 The genetics of pea seed storage proteins. *Qualitas Pl. Pl. Fds Hum. Nutr.* **31**, 281–295.

Casey, R., March, J. F. & Sanger, E. 1981*a* N-terminal amino acid sequence of β-subunits of legumin from *Pisum sativum*. *Phytochemistry* **20**, 161–163.

Casey, R., March, J. F., Sharman, J. E. & Short, M. N. 1981*b* The purification, N-terminal amino acid sequence and some other properties of an α^M subunit of legumin from the pea (*Pisum sativum* L.). *Biochim. biophys. Acta* **670**, 428–432.

Casey, R. & Sanger, E. 1980 Purification and some properties of a 7 S seed storage protein from *Pisum* (pea). *Biochem. Soc. Trans.* **8**, 658.

Casey, R., Sharman, J. E., Wright, D. J., Bacon, J. R. & Guldager, P. 1982 Quantitative variability in *Pisum* seed globulins: its assessment and significance. *Qualitas Pl. Pl. Fds Hum. Nutr.* **31**, 333–346.

Chandler, P. M., Higgins, T. J. V., Randall, P. J. & Spencer, D. 1983 Regulation of legumin levels in developing pea seeds under conditions of sulfur deficiency. *Pl. Physiol.* **71**, 47–54.

Chrispeels, M. J., Higgins, T. J. V., Craig, S. & Spencer, D. 1982*a* Role of the endoplasmic reticulum in the synthesis of reserve proteins and the kinetics of their transport to protein bodies in developing pea cotyledons. *J. Cell Biol.* **93**, 5–14.

Chrispeels, M. J., Higgins, T. J. V. & Spencer, D. 1982*b* 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.

Coxon, D. T. & Davies, D. R. 1982 The effect of the r_a and r_b loci on the lipid content of the seed of *Pisum sativum*. *Theor. appl. Genet.* **64**, 47–50.

Croy, R. R. D., Derbyshire, E., Krishna, T. G. & Boulter, D. 1979 Legumin of *Pisum sativum* and *Vicia faba*. *New Phytol.* **83**, 29–35.

Croy, R. R. D., Gatehouse, J. A., Evans, I. M. & Boulter, D. 1980*a* Characterisation of the storage protein subunits synthesized *in vitro* by polyribosomes and RNA from developing pea (*Pisum sativum* L.) I. Legumin. *Planta* **148**, 49–56.

Croy, R. R. D., Gatehouse, J. A., Tyler, M. & Boulter, D. 1980*b* The purification and characterization of a third storage protein (convicilin) from the seeds of pea (*Pisum sativum* L.) *Biochem. J.* **191**, 509–516.

Croy, R. R. D., Lycett, G. W., Gatehouse, J. A., Yarwood, J. N. & Boulter, D. 1982 Cloning and analysis of cDNAs encoding plant storage protein precursors. *Nature, Lond.* **295**, 76–79.

Cunningham, B. A., Wang, J. L., Waxdal, M. J. & Edelman, G. M. 1975 Covalent and three-dimensional structure of concanavalin A. *J. biol. Chem.* **250**, 1503–1512.

Davies, D. R. 1980 The r_a locus and legumin synthesis in *Pisum sativum. Biochem. Genet.* **18**, 1207–1219.

Derbyshire, E., Wright, D. J. & Boulter, D. 1976 Legumin and vicilin storage proteins of legume seeds. *Phytochemistry* **15**, 3–24.

Domoney, C. A. M. 1981 Legumin synthesis and accumulation in *Pisum sativum* L. Ph.D. Thesis, University of East Anglia.

Finlayson, A. J. 1976 The seed protein contents of some Cruciferae. In *The biology and chemistry of the Cruciferae* (ed. J. G. Vaughan, A. J. McLeod & B. M. G. Jones), pp. 279–306. New York: Academic Press.

Fischer, R. L. & Goldberg, R. B. 1982 Structure and flanking regions of soybean seed protein genes. *Cell* **29**, 651–660.

Foriers, A., Lebrun, E., van Rapensbusch, R., de Neve, R. & Strosberg, A. D. 1981 The structure of the lentil (*Lens culinaris*) lectin. *J. biol. Chem.* **256**, 5550–5560.

Foriers, A., Wuilmart, C., Sharon, N. & Strosberg, A. D. 1977 Extensive sequence homologies among lectins from leguminous plants. *Biochem. biophys. Res. Commun.* **75**, 980–986.

Gatehouse, J. A., Croy, R. R. D. & Boulter, D. 1980 Isoelectric-focusing properties and carbohydrate content of pea (*Pisum sativum*) legumin. *Biochem. J.* **185**, 497–503.

Gatehouse, J. A., Croy, R. R. D., Morton, H., Tyler, M. & Boulter D. 1981 Characterisation and subunit structures of the vicilin storage proteins of pea (*Pisum sativum* L.). *Eur. J. Biochem.* **118**, 627–633.

Gatehouse, J. A., Lycett, G. W., Croy, R. R. D. & Boulter, D. 1982 The post-translational proteolysis of the subunits of vicilin from pea (*Pisum sativum* L.). *Biochem. J.* **207**, 629–632.

Gatehouse, J. A., Lycett, G. W., Delauney, A. J., Croy, R. R. D. & Boulter, D. 1983 Sequence specificity of the post-translational proteolytic cleavage of vicilin, a seed storage protein of pea (*Pisum sativum* L.) *Biochem. J.* **212**, 427–432.

Gillespie, J. M. & Blagrove, R. J. 1975 Variability in the proportion and type of subunits in lupin storage globulins. *Aust. J. Pl. Physiol.* **2**, 29–39.

Gilroy, J., Wright, D. J. & Boulter, D. 1979 Homology of basic subunits of legumin from *Glycine max* and *Vicia faba. Phytochemistry* **18**, 315–316.

Goldberg, R. B., Hoschek, G., Ditta, G. S. & Breidenbach, R. W. 1981 Developmental regulation of cloned superabundant embryo mRNAs in soybean. *Dev. Biol.* **83**, 218–231.

Hall, T. C., McLeester, R. C. & 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. *Pl. Physiol.* **59**, 1122–1124.

Hall, T. C., Slightom, J. L., Ersland, D. R., Murray, M. G., Hoffman, L. M., Adang, M. J., Brown, J. W. S., Ma, Y., Matthews, J. A., Cramer, J. H., Barker, R. F., Sutton, D. W. & Kemp, J. D. 1983 Phaseolin: nucleotide sequence explains molecular weight and charge heterogeneity of a small multigene family and also assists vector construction for gene expression in alien tissue. In *Structure and function of plant genomes* (ed. O. Ciferri & L. Dure). New York: Plenum Press (In the press).

Hara, I., Wada, K., Wakabayashi, S. & Matsubara, J. 1976 Pumpkin (*Cucurbita* sp.) seed globulin I. Purification, characterisation and subunit structure. *Pl. Cell Physiol.* **17**, 799–814.

Hirano, H., Gatehouse, J. A. & Boulter, D. 1982 The complete amino acid sequence of a subunit of the vicilin seed storage protein of pea (*Pisum sativum* L.) *FEBS Lett.* **145**, 99–102.

Hoffman, L. M., Ma, Y. & Barker, R. F. 1982 Molecular cloning of Phaseolus vulgaris lectin mRNA and use of cDNA as a probe to estimate lectin transcript levels in various tissues. *Nucl. Acids Res.* **10**, 7819–7828.

Hopp, T. P., Hemperly, J. J. & Cunningham, B. A. 1982 Amino acid sequence and variant forms of favin, a lectin from *Vicia faba. J. biol. Chem.* **257**, 4473–4483.

Jackson, P., Boulter, D. & Thurman, D. A. 1969 A comparison of some properties of vicilin and legumin isolated from seeds of *Pisum sativum, Vicia faba* and *Cicer arietinum. New Phytol.* **68**, 25–33.

Khan, M. R. I., Gatehouse, J. A. & Boulter, D. 1980 The seed proteins of cowpea (*Vigna unguiculata* L. Walp). *J. exp. Bot.* **31**, 1599–1611.

Kooistra, E. 1962 On the differences between smooth and three types of wrinkled peas. *Euphytica* **11**, 357–373.

Krishna, T. G., Croy, R. R. D. & Boulter, D. 1979 Heterogeneity in subunit composition of the legumin of *Pisum sativum. Phytochemistry* **18**, 1879–1880.

Lycett, G. W., Delauney, A. J., Gatehouse, J. A., Gilroy, J., Croy, R. R. D. & Boulter, D. 1983 The vicilin gene family of pea (*Pisum sativum* L.): a complete cDNA coding sequence for preprovicilin. *Nucl. Acids Res.* **11**, 2367–2380.

Ma, Y., Bliss, F. A. & Hall, T. C. 1980 Peptide mapping reveals considerable sequence homology among the three polypeptide subunits of G1 storage protein from french bean seed. *Pl. Physiol.* **66**, 897–902.

Manteuffel, R. 1982 Immunology. In *Nucleic acids and proteins in plants I. Structure, biochemistry and physiology of proteins* (ed. D. Boulter & B. Parthier), pp. 457–502. Berlin, Heidelberg and New York: Springer-Verlag.

Matta, N. K. & Gatehouse, J. A. 1982 Inheritance and mapping of storage protein genes in *Pisum sativum* L. *Heredity, Lond.* **48**, 383–392.

Matta, N., Gatehouse, J. A. & Boulter, D. 1981a The structure of legumin of *Vicia faba* L. – a reappraisal. *J. exp. Bot.* **32**, 183–197.

Matta, N. K., Gatehouse, J. A. & Boulter, D. 1981 *b* Molecular and subunit heterogeneity of legumin of *Pisum sativum* L. (garden pea) – a multi-dimensional gel electrophoretic study. *J. exp. Bot.* **32**, 1295–1307.

McLeester, R. C., Hall, T. C., Sun, S. M. & Bliss, F. A. 1973 Comparison of globulin proteins from *Phaseolus vulgaris* with those from *Vicia faba. Phytochemistry* **2**, 85–93.

Moreira, M. A., Hermodson, M. A., Larkins, B. A. & Nielsen, N. C. 1979 Partial characterisation of the acidic and basic polypeptides of glycinin. *J. biol. Chem..* **254**, 9921–9926.

Moreira, M. A., Hermodson, M. A., Larkins, B. A. & Nielsen, N. C. 1981 Comparison of the primary structure of the acidic polypeptides of glycinin. *Archs Biochem. Biophys.* **210**, 633–642.

Mori, T., Utsumi, S. & Inaba, H. 1979 Interaction involving disulfide bridges between subunits of soybean seed globulin and between subunits of soybean and sesame seed globulins. *Agric. biol. Chem.* **43**, 2317–2322.

Mori, T., Utsumi, S., Inaba, H., Kitamura, K. & Harada, K. 1981 Differences in subunit composition of glycinin among soybean cultivars. *J. agric. Fd Chem.* **29**, 20–23.

Murray, D. R. 1979 A storage role for albumins in pea cotyledons. *Pl. Cell Environ.* **2**, 221–226.

Nielsen, N. C., Moreira, M., Staswick, P., Hermodson, M. A., Tumer, N. & Thanh, V. H. 1981 The structure of glycinin from soybeans. *Abh. Akad. Wiss. DDR Abt. Math. Naturwiss. Techn.* **5**, 161–173.

Ohmiya, M., Hara, I. & Matsubara, H. 1980 Pumpkin (*Cucurbita* sp.) seed globulin. IV. Terminal sequences of the acidic and basic peptide chains and identification of a pyroglutamyl peptide chain. *Pl. Cell Physiol.* **21**, 157–167.

Oram, R. N., Gillespie, J. M. & Blagrove, R. J. 1981 Genetic effects on the proportions of conglutin γ and a subunit of conglutin β in the storage proteins of *Lupinus angustifolius* L. seeds. *Aust. J. Pl. Physiol.* **8**, 277–284.

Przybylska, J., Hurich, J. & Blixt, S. 1981 Variation of basic legumin components in the genus *Pisum* revealed by isoelectrofocusing. *Pisum Newsl.* **13**, 44–45.

Przybylska, J., Hurich, J. & Zimniak-Przybylska, Z. 1979 Comparative study of seed proteins in the genus *Pisum*. IV. Electrophoretic patterns of legumin and vicilin components. *Genet. pol.* **20**, 517–528.

Pusztai, A. & Stewart, J. C. 1980 Molecular size, subunit structure and microheterogeneity of glycoprotein II from the seeds of kidney bean (*Phaseolus vulgaris* L.). *Biochim. biophys. Acta* **623**, 418–428.

Pusztai, A. & Watt, W. B. 1970 Glycoproteins II. The isolation and characterisation of a major antigenic and non-haemagglutinating glycoprotein from *Phaseolus vulgaris*. *Biochim. biophys. Acta*. **207**, 413–431.

Romero, J., Sun, S. M., McLeester, R. C., Bliss, F. A. & Hall, T. C. 1975 Heritable variation in a polypeptide subunit of the major storage protein of the bean, *Phaseolus vulgaris* L. *Pl. Physiol.* **56**, 776–779.

Schwenke, K. D., Elkowicz, K. & Kizlowska, H. 1979 Solubility of the 12*S* globulin from rapeseed effected by phytic acid. *Nahrung* **23**, 967–970.

Schuler, M. A., Ladin, B. F., Pollaco, J. C., Freyer, G. & Beachy, R. N. 1982 *a* Structural sequences are conserved in the genes coding for the α, α′ and β-subunits of the soybean 7 *S* seed storage protein. *Nucl. Acids Res.* **10**, 8244–8261.

Schuler, M. A., Schnutt, E. S. & Beachy, R. N. 1982 *b* Closely related families of genes code for the α and α′ subunits of the soybean 7 *S* storage protein complex. *Nucl. Acids Res.* **10**, 8225–8244.

Slightom, J. L., Sun, S. M. & Hall, T. C. 1983 Complete nucleotide sequence of a French bean storage protein gene: phaseolin. *Proc. natn. Acad, Sci. U.S.A.* **80**, 1897–1901.

Spencer, D. & Higgins, T. J. V. 1980 The biosynthesis of legumin in maturing pea seeds. *Biochem. Int.* **1**, 502–509.

Spielmann, A., Schurmann, P. & Stutz, E. 1982 Gel electrophoretic characterisation of protein fractions from soybean during seed development. *Pl. Sci. Lett.* **24**, 137–145.

Staswick, P. E., Hermodson, M. A. & Nielsen, N. C. 1981 Identification of the acidic and basic subunit complexes of glycinin. *J. biol. Chem.* **256**, 8752–8755.

Sun, S. M., McLeester, R. C., Bliss, F. A. & Hall, T. C. 1974 Reversible and irreversible dissociation of globulins from *Phaseolus vulgaris* seed. *J. biol. Chem.* **249**, 2118–2120.

Sun, S. M., Slightom, J. L. & Hall, T. C. 1981 Intervening sequences in a plant gene – comparison of the partial sequence of cDNA and genomic DNA of french bean phaseolin. *Nature, Lond.* **289**, 37–41.

Thomson, J. A. & Schroeder, H. E. 1978 Cotyledonary storage proteins in *Pisum sativum*. II. Hereditary variation in components of the legumin and vicilin fractions. *Aust. J. Pl. Physiol.* **5**, 281–294.

Thomson, J. A., Schroeder, H. E. & Dudman, W. F. 1978 Cotyledonary storage proteins in *Pisum sativum* I. Molecular heterogeneity. *Aust. J. Pl. Physiol.* **5**, 263–279.

Thomson, J. A., Schroeder, H. E. & Tassie, A. M. 1980 Cotyledonary storage proteins in *Pisum sativum*. V. Further studies on molecular heterogeneity in the vicilin series of holoproteins. *Aust. J. Pl. Physiol.* **7**, 271–282.

Tumer, N. E., Thanh, V. H. & Nielsen, N. C. 1981 Purification and characterisation of mRNA from soybean seeds. *J. biol. Chem.* **256**, 8756–8760.

Utsumi, S., Yokoyama, Z. & Mori, T. 1980 Comparative studies of subunit compositions of legumins from various cultivars of *Vicia faba* L. seeds. *Agric. biol. Chem.* **44**, 595–601.