Studies on the major seed proteins of some grain crops

Derbyshire, E.

STUDIES ON THE MAJOR SEED PROTEINS OF SOME GRAIN CROPS

A thesis submitted on completion of a course of study under the regulations to enable members of staff to become candidates for the degree of Doctor of Philosophy in the University of Durham.

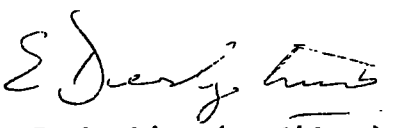
by

E. Derbyshire, B.Sc. (Liverpool), M.Sc. (Wales)

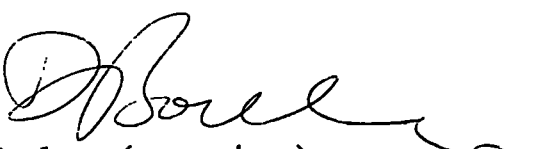
Department of Botany
University of Durham

March 1979

The material presented in this thesis has not previously been submitted by the candidate for a degree in this or any other university.



E. Derbyshire (candidate)



D. Boulter (supervisor)

COLLABORATION

The published work presented in this thesis includes work by collaborating authors. Initial drafts of the section 'Legumin' for the review of Legumin and Vicilin were prepared by D.J. Wright and certain other sections of the two review articles were prepared by Professor D. Boulter. Final drafts of the manuscripts were prepared for publication in collaboration with Professor Boulter who also acted as supervisor. R.D.J. Barker, under the supervision of Dr. A. Yarwood isolated and characterised the major 7S storage protein of *Phaseolus vulgaris*, and J.F. Carasco, under the supervision of Professor D. Boulter performed much of the work on the proteins from *Vigna* species. The comparison of the proteins of *Phaseolus* and *Vigna* spp. was an extension of an undergraduate project by J.N. Yarwood, and E. Neat prepared the subunit profiles of these samples. Dr. N. Harris prepared and interpreted the electron micrographs and R.R.D. Croy performed serological investigations. Professor E.M. Jope supplied the archaeological sample of maize and its provenance. Dr. I. Marta Evans examined the nitrogen and sulphur contents of cowpeas grown at the University of Reading. Apart from these exceptions all the experimental work presented in publications 3-9 were initiated, executed and the results interpreted by the candidate.

E. Derbyshire (candidate)

D. Boulter (supervisor)

ACKNOWLEDGEMENTS

Various aspects of the work presented here have been discussed with other members of the plant protein unit at Durham, and to these, especially Dr. I. Marta Evans, I offer my thanks. I am very grateful also for the collaboration of the several authors whose names have appeared jointly with mine on the publications which form the major part of this thesis. My greatest debt however is to Professor D. Boulter who has guided and encouraged me throughout the work.

I acknowledge also the financial assiatance provided by the Agricultural Research Council and the Ministry of Overseas Development.

ABSTRACT

The general properties, classification and distribution of plant proteins are discussed within an agricultural context and the protein content and composition of several grain crops are tabulated. The major proteins of legume seeds are salt soluble proteins (i.e. globulins) and the methodology for their extraction, separation and characterisation is reviewed. The structure, location and distribution of legume globulins are described and the properties of various purified legume globulins are compared. The major seed proteins of *Pisum sativum* are legumin and vicilin and seed globulins with properties similar to those of legumin and vicilin have been extracted from other legumes. The isolation and partial characterisation of vicilin-like proteins from seeds of *Phaseolus vulgaris* and *Vigna unguiculata* and of legumin-like proteins from these two species and *Phaseolus aureus* is described.

The potential usefulness of the examination of seed proteins by polyacrylamide gel electrophoresis to taxonomic problems in *Phaseolus* and *Vigna* is investigated and the data obtained are discussed in relation to recent re-classifications of the species. Procedures for the extraction and partial characterisation of proteins from an archaeological sample of maize grains are described and the proteins from this sample are compared with the proteins from a modern variety of maize. It is suggested that protein data from archaeological samples could be useful in studies of crop plant origins and in cultural studies, and that the procedures employed may be applicable to archaeological samples of other grains.

Nutritional aspects of plant proteins are discussed briefly. A possible strategy for the improvement of the sulphur amino acid content of legumes with a relatively low legumin content is suggested and it is demonstrated that polyacrylamide gel electrophoresis may provide a basis for a screening procedure for protein quality in legume breeding programmes.

CONTENTS

6. Seed protein of *Phaseolus* and *Vigna*.

7. Seed protein profiles of African representatives of *Vigna* spp.

8. The extraction, composition and intra-cellular distribution of protein in early maize grains from an archaeological site in N.E.Arizona.

9. Proteins of some legumes with reference to environmental factors and nutritional value.

SECTION I

INTRODUCTION

INTRODUCTION

The seed is the first form in which the new spermatophytic plant can exist, *in vivo*, independently of its maternal parent. Its structure is specialised for dispersal and it carries within itself the materials, including proteins, which are required for survival and, under favourable conditions, growth and the development of an adequate photosynthetic system. The bulk of the seed, generally, is carbohydrate and/or oil, however the protein content of seeds can be 10% (cereals), 20% (many legumes) or even 45% (certain varieties of *Glycine max*): for comparison the protein content of many vegetative tissues (as harvested) is less than 5% (1). The seed protein fraction is heterogeneous and it includes many of the types of protein found in the mature plant. Several classifications of plant proteins exist and these have been described (1).

A high proportion of the material synthesised during seed development is laid down as reserve, or storage, material which is utilised subsequently as a source of carbon and nitrogen compounds during germination. More than 80% of the seed protein may be storage protein (1) and this component is therefore of considerable importance both to the seed and to man and his animals as a food source.

In the latter part of the nineteenth century and the early

part of the twentieth century Osborne and his co-workers[+] developed procedures which were of general applicability for the extraction and separation of the major seed proteins. They defined four classes of proteins on the basis of solubility, *viz.*; albumins (water soluble), globulins (salt soluble), prolamins (alcohol soluble), and glutelins (alkali soluble). The percentages of the Osborne solubility classes in various grains have been tabulated (1). The albumin fraction includes the numerous enzymes of the seed and generally contains 10% or less of the seed protein. The major proteins of legumes are globulins, and prolamins and glutelins are either absent or are present in only low concentration. Albumins and globulins occur also in other grains and in some crops, for example hemp and rape, globulins are major proteins. However prolamins and glutelins are the major fractions in cereals and albumins and globulins form only 20% or less of the protein; in rice 80% of the protein is glutelin.

Although the Osborne solubility classes have proved very useful in comparing the grain protein of different crops the percentages referred to above should be regarded only as approximations. Accurate values are difficult to obtain because of artifacts of extraction and fractionation. Some of the problems and precautions which may be taken have been discussed (1, 2); procedures for the isolation of cereal prolamin have been reviewed recently by Miflin and Shewry[++].

During the early investigations by Osborne's group the seed

+ For references see Osborne, T.B. (1924). The Vegetable Proteins, 2nd edn., Longman, Green and Co., London.

++ Miflin, B.J. and Shewry, P.R. (1979). The Biology and Biochemistry of Cereal Seed Prolamins, in Symposium on Seed Protein Improvement, Neuherberg, 1978, I.A.E.A. Vienna, in press.

globulins of *Pisum sativum* were separated into two fractions which were designated legumin and vicilin respectively. Vicilin was more soluble than legumin in dilute salt solution, it precipitated from solution when heated to 95^{o} and it had less sulphur than legumin. The globulins from *Vicia faba, Vicia sativa* and *Lens esculentum* were separated into fractions with properties identical (within the limits of experimentation) to those of legumin and vicilin and these fractions were also termed legumin and vicilin. By contrast the globulins isolated from other legumes, for example *Glycine max* and *Phaseolus vulgaris* differed from legumin and vicilin.

Subsequent to the development of the ultracentrifuge Danielsson[+] prepared legumin and vicilin from *Pisum sativum* by dilution, and by dialysis and isoeletric precipitation, and determined their sedimentation coefficients—legumin 12.6S, vicilin 6.5-8.1S. He further characterised the two fractions by their chemical compositions and by their electrophoretic migration. Danielsson also identified proteins with sedimentation coefficients of approximately 12S and 7S in seed globulin fractions from more than 20 other legumes, which included *Glycine max* and *Phaseolus vulgaris*, and on the sole basis of the sedimentation data he called these proteins legumin and vicilin.

During the last two decades advances in biochemical technique have led to the purification, or partial purification of several legume storage proteins. The methodology for the extraction, separation and characterisation of seed globulins has been discussed and the accumulated

---

+ Danielsson C.E. (1949). Seed globulins of the Gramineae and Leguminosae, Biochem.J. 44: 387-400.

data (to 1976) on their distribution have been reviewed (2). The legumin fraction which Osborne prepared from *Pisum sativum*, although impure, contained mainly a single protein and the term legumin has been retained for this protein. The term vicilin is used here for the 7S storage globulins of *Pisum sativum* and other legumes. A more precise definition of vicilin which would be acceptable in all laboratories is still lacking, and operationally the term is sometimes applied to the globulin fraction which is soluble in dilute salt solution at pH 4.7. However globulins other than the 7S storage proteins may be present in this fraction (for references see 2) and may sometimes be confused with the vicilin. Globulin lectin from *Phaseolus vulgaris* sediments in the ultracentrifuge as a 6.1S species (3) and is not easily detected by this technique when the sample contains high concentrations of 7S globulin. Smaller (2S-4S) seed globulins have been identified in preparations from several legumes; generally they are present in relatively low concentration by comparison with the major 12S and 7S globulins and they have received less attention than the latter.

Legumin-and vicilin-like globulins occur in several (and perhaps in all) genera of the Leguminosae and legumin-like proteins are found also in other dicotyledons. However the relative concentrations, one to another, of legumin and vicilin proteins vary from species to species and even between varieties (1, 2).

At Durham the major proteins of several legumes and some non-legumes have been studied and particular atttention has been given to species of *Phaseolus* and *Vigna*, as well as to members of the Vicieae. The globulins from *Phaseolus aureus* (3), *Phaseolus vulgaris* (3, 4) and *Vigna unguiculata* (5) have been fractionated by use of zonal isoelectric

precipitation. The fractions from the two species of *Phaseolus* were partially characterised by analytical ultracentrifugation and electrophoresis in polyacrylamide gels under non-dissociating and dissociating conditions, and the globulins from *Phaseolus vulgaris* were characterised further by amino acid analysis, N-terminal amino acid determination, carbohydrate determination and by their agglutinating activity against red blood cells. The globulins from *Vigna unguiculata* were partially characterised by analytical ultracentrifugation and gel electrophoresis. The location of the globulins of *Phaseolus vulgaris* in the cotyledons and the separation of the major polypeptides (sub-units) of the globulins of *Vigna unguiculata* and their sequential synthesis during seed development were also described. These investigations give further support to the opinion (2) that legumin- and vicilin-like proteins are distributed widely in the Leguminosae.

Small diferences between equivalent polypeptides in the legumin-like protein from different species, and serological data (see 2) indicate that these proteins are not identical. A similar situation is found with other well defined proteins, for example the haemoglobins, and it is possible, by analogy, that the legumin-like proteins are all forms of legumin. Clearly more critical data, for example amino acid sequences, are required to determine whether or not the legumins are related one to another by common ancestry. Preliminary sequence data has been obtained (Boulter, unpublished) from legumin from members of the Vicieae and a very similar partial sequence has been reported for the 12S globulin from a *Cucurbita spp.* [+]. Recently microheterogeneity within the acidic

---

+ Hara, I., Ohmiya, M. and Matsubara, (1978) Pumpkin (*Cucurbita spp.*) seed globulin III. Plant and Cell Physiology 19, 237-243.

group of legumin subunits from lines of pea has been reported+. This observation and subunit data from pea legumin obtained by use of two dimensional gel electrophoresis (Boulter, unpublished) may be indicative of polymorphism at the sub-specific level, however the possibility that they result from preparative artifacts has not been entirely excluded.

Generally the vicilin fraction of seed globulin is heterogeneous. Only a few vicilin proteins have been purified and characterised and the limited data which is available (2, 3) indicate that the major vicilin protein is not the same in all species. However similarities between the vicilin preparations from different species suggest that only a small number, perhaps five vicilins exist. The ratios of the different vicilins, one to another vary from species to species; with the possible exception of *Phaseolus vulgaris* all the vicilins are not normally observed in a single preparation. Speculation on the possible ancestry of the vicilins is inappropriate because of the paucity of the data.

Proteins are early products of gene action and in general the information contained in their primary structure (amino acid sequence) will be of taxonomic and/or evolutionary significance; other properties of proteins also may vary in a systematic fashion. However a particular protein or a particular property of a protein may not vary systematically always at the same taxonomic level in different taxa. The determination of many properties of proteins, for example amino acid sequence requires the prior purification of the protein and this usually is time consuming. In circumstances where comparison of a large number of samples is required

---

+ Thomson, J.A. and Schroeder, H.E. (1974). Cotyledonary storage proteins in *Pisum sativum, II.* Austral. J. Plant Physiol. 5, 281-294.

protein data may be useful only if they can be obtained by relatively simple and rapid procedures. The potential usefulness of the direct examination of seed proteins and/or their sub-units by polyacrylamide gel electrophoresis to taxonomic problems in *Phaseolus* and *Vigna* has been examined (6, 7). The data obtained from these investigations were consistent with recent revisions of the classification of species of the two genera, and they indicated also that sub-specific variation exists in the seed protein profiles from certain of the species examined. Recently a numerical technique has been employed[+] to process similar data obtained from a larger number of species of *Phaseolus*. The conclusions drawn from the two studies of *Phaseolus* are consistent except in relation to one of the species examined (*Phaseolus aconitifolius*).

In studies of the origins of modern plants most of the botanical data have been obtained from fossils or by comparison of a taxon with contemporary samples of its relatives. Protein data are often of considerable value in establishing relationships between present day taxa however when archaeological plant material has been available usually only morphological data have been recorded and protein data from archaeological samples are sparse (for references see 8). The salt-soluble protein extracted from an archaeological sample of seeds of *Phaseolus vulgaris* differed from that of a modern variety and haemagglutinating activity was absent. Enzyme activity was not detectable in wheat grains recovered from archaeological sites. By contrast prolamin and glutelin proteins similar to modern maize protein have been extracted from an archaeological sample

---

+Sahai, S. and Rana, R.S. (1977). Seed protein homology and elucidation of species relationships in *Phaseolus* and *Vigna* species. New Phytol. 79, 527-534.

of maize grains (8) and the amino acid composition of the early maize has been compared with the composition of a modern variety. It has been concluded that protein data from archaeological samples could be useful in botanical studies of crop plant origins and evolution and in cultural studies, and it has been suggested that the procedures employed with maize may be applicable to archaeological samples of other grains such as wheat, sorghum, teosinte and barley in which similar storage proteins occur.

Nutritionally grain proteins are an important source of amino acids in diets in many regions of the world especially in human diets in the less developed countries. In many localities the quantity and/or the quality of the protein eaten are very probably inadequate especially for young children. Proposals for the improvement of the supply of plant proteins and other nutritional aspects of plant proteins have been discussed briefly, and the importance of adequate biochemical screening methods to breeding programmes for protein quantity and quality has been emphasised (2, 9). Usually the first limiting amino acids in legume grains are the sulphur containing amino acids methionine and cysteine and this is a reflection of the relatively low sulphur amino acid content of the seed globulins. However the sulphur amino acid compositions of different globulins, for example legumin and vicilin, are not identical, one to another (9) and changes in their relative proportions may result therefore in changes in the sulphur amino acid composition of the grain: differences in the relative proportions of the major globulins in different legumes have been demonstrated by the use of polyacrylamide gel electrophoresis (9). The implications of these observations for devising screening procedures which can be applied in breeding programmes for improved protein quality in legumes have been discussed (9).

SECTION II

PUBLICATIONS

A. Surveys of plant proteins

1. The general properties, classification and distribution of plant proteins.

2. Legumin and vicilin, storage proteins of legume seeds.

B. Studies of the isolation and characterisation of seed proteins of *Phaseolus* and *Vigna*.

3. Isolation of legumin-like protein from *Phaseolus aureus* and *Phaseolus vulgaris*.

4. Purification and characterisation of the major storage proteins of *Phaseolus vulgaris* seeds, and their intracellular and cotyledonary distribution.

5. The isolation and characterisation of the major polypeptides of the seed globulin of cowpea (*Vigna unguiculata* L. Walp.) and their sequential synthesis in developing seeds.

C. Applications of seed protein biochemistry to problems in taxonomy, the origin of crop plants, and nutrition.

6. Seed protein of *Phaseolus* and *Vigna*.

7. Seed protein profiles of African representatives of *Vigna* spp.

8. The extraction, composition and intra-cellular distribution of protein in early maize grains from an archaeological site in N.E.Arizona.

9. Protein of some legumes with reference to environmental factors and nutritional value.

New Phytol. (1976) **76**, 283–288.

6

SEED PROTEINS OF *PHASEOLUS* AND *VIGNA*

BY E. DERBYSHIRE, J. N. YARWOOD, E. NEAT AND D. BOULTER

Department of Botany, University of Durham, Durham DH1 3LE

(*Received 4 September 1975*)

SUMMARY

Protein band patterns of globulins extracted from the seeds of a number of different species of *Phaseolus* and *Vigna*, have been obtained by disc electrophoresis and the sub-units of the major seed proteins have been separated on gels by use of continuous SDS electrophoresis. The agglutinin potential of seed extracts with respect to human Group O red blood cells has also been determined. The globulin extracted from *Phaseolus lunatus* was different from that from *P. acutifolius*, *P. coccineus*, *P. dumosus* and *P. vulgaris*, and these globulins were different from the major globulins extracted from other species which differed among themselves. The globulin patterns for *P. atropurpureus* and *P. lathyroides* were not the same as those of the other species examined and saline extracts from the two species lysed red blood cells. Agglutinin activity was detected only in extracts from *P. acutifolius*, *P. coccineus*, *P. dumosus* and *P. vulgaris*. Glycoprotein II, the major seed protein of *P. vulgaris*, was tentatively identified as the major globulin in extracts from *P. acutifolius*, *P. coccineus* and *P. dumosus*, and was not detected in the other species, with the exception possibly of *P. atropurpureus* and *P. lathyroides*. We discuss our data in relation to recent re-classifications of species of *Phaseolus* and *Vigna*.

INTRODUCTION

Approximately 80% of the protein of legume seeds is storage or reserve protein; in those legumes which have been examined it is made up of a small number of different salt-soluble proteins, globulins (for references see Derbyshire, Wright and Boulter, 1975). Usually one or two of these proteins predominate and more than 50% of the total seed protein may occur as a single protein, for example legumin in *Vicia faba* (Wright and Boulter, 1972) and Glycoprotein II in *Phaseolus vulgaris* (Pusztai and Watt, 1970; Racusen and Foote, 1971; Barker *et al.*, 1975).

The composition of the storage protein differs in different legume species (Boulter, Thurman and Derbyshire, 1967; Kloz, 1971; Dudman and Millerd, 1975), but for certain tribes, uniform electrophoretic protein band patterns are obtained when the seed globulins from members of the same tribe are compared (Boulter *et al.*, 1967). In other tribes, however, uniform patterns are not obtained from different members but here it may be possible, as for example in the tribes Phascoleae, Loteae and Glycineae, to distinguish sub-groups, each with a characteristic protein band pattern. It may also prove feasible to use storage protein band patterns to sub-divide the species of a large genus. Thus, the storage protein band patterns of thirty-six species of *Crotalaria* correlate well with subdivisions of the genus based on morphological and cytological data and with geographical distribution (Boulter *et al.*, 1970).

Several revisions of the classification of the Phaseoleae have been proposed in recent years and in these schemes some species of *Phaseolus* have been transferred to other

genera, i.e. *Macroptilium*, *Azukia* and *Vigna* (Hutchinson, 1964; Ohwi, 1965; Ohwi and Ohashi, 1969; Maréchal, 1969; Verdcourt, 1969, 1970). Kloz and co-workers (Kloz, 1971) have examined immunological reactions of the seed proteins of species from the *Phaseolus–Vigna* complex and conclude that seed protein composition correlates with the systematic position of the species from which the protein was obtained. The immuno-electrophoretic techniques employed by Kloz and co-workers (Kloz, 1971) are relatively simple and rapid; however, appropriate antibody samples must be prepared, usually, for each plant group under investigation and, ideally, from pure proteins; these preliminary steps are time-consuming when compared to the direct examination of seed protein extracts or the total seed globulin fraction. The seed globulins and their sub-units from a small sample of species of *Phaseolus* and *Vigna* have been examined by the use of electrophoretic and other techniques in order to assess the potential usefulness of this approach of these genera. Since the major seed globulin of *Phaseolus vulgaris* has been isolated and characterized recently (Pusztai and Watt, 1970; Barker *et al.*, 1975), we are able also to comment on the distribution of this protein in the two genera.

MATERIALS

Seeds of *Phaseolus* species were gifts from J. Smartt and had been collected from plants grown in the Botanical Gardens of the School of Agriculture, Cambridge. Cultivars of *P. vulgaris* were obtained locally and seed of *Vigna unguiculata* was provided by the International Institute of Tropical Agriculture, Ibadan, Nigeria.

METHODS

Seeds were milled in a Glen Creston Mill and each meal was then divided into three portions. Globulin protein was prepared from one portion of each meal by the method of Boulter *et al.* (1967). Total protein was extracted from a second portion by blending with 10 volumes of 2.5% sodium dodecyl sulphate (SDS)—1% 2-mercaptoethanol (2-Me) in 0.1 M sodium phosphate buffer pH 7.0 in a high-speed blender, which was cooled in ice, for 3 min and the resultant slurry was squeezed through two layers of muslin and centrifuged at 23,000 *g* for 10 min. The third portion of meal was stirred with 10 volumes of 1% NaCl in 0.05 M phosphate buffer pH 7.0 and the suspension which passed through two layers of muslin was clarified by centrifugation.

Disc electrophoresis of globulins in 7% acrylamide gels was carried out by the procedure of Ornstein and Davis (1961). Protein extracted in SDS-2-Me was heated for 2 min at 100° C and then cooled in running tap water: electrophoresis of these solutions was carried out by the method of Weber, Pringle and Osborne (1972). Protein bands in gels were detected by staining with Amido-Black 10B and glycoprotein components were identified in duplicate gels by the method of Zacharius *et al.* (1969).

The agglutinating activity of salt extracts was determined by mixing one volume of extract with one volume of a 1% suspension of human Group O red blood cells in 1% NaCl.

RESULTS AND DISCUSSION

The protein band patterns obtained by disc electrophoresis of the seed globulins were of two types (Table 1). In one type a single stained band was dominant and in the other

no single component dominated the pattern, although one stained more heavily than the others (i.e. either R_m 0.21 or 0.31). The single banded pattern was seen in gels of globulins from *Phaseolus acutifolius*, *P. coccineus*, *P. dumosus*, *P. lunatus* and *P. vulgaris*, and the mobility of the major band in each gel, except those of *P. lunatus*, was approximately the same, R_m 0.39; the mobility of the major band of the globulin from *P. lunatus* was R_m 0.33.

Multi-banded patterns were found in gels of the other ten species which were examined and in each of these patterns prominent bands were seen in the regions corresponding to R_m 0.21 and 0.31, except in gels of *P. trilobus*. Other major bands were not seen in gels of *P. aconitifolius*, *P. angularis*, *P. aureus*, *P. mungo* and *Vigna unguiculata*, but were in gels of *Phaseolus calcaratus* (R_m 0.13), *P. filiformis* (R_m 0.26), *P. atropurpureus* (R_m 0.09, 0.24, 0.38) and *P. lathyroides* (R_m 0.37, 0.47). Only the last two species contained a

Table 1. *The prominent electrophoretic components in seed globulins and the prominent subunits of protein extracts*

	Mobilities relative to the mobility of bromo-phenol blue																							
	Globulins												Subunits											
Phaseolus aconitifolius			0.21			0.33									0.20	0.24						0.47	0.51	
P. acutifolius							0.41						0.17	0.20			0.27	0.30	0.41	0.43	0.51	0.54		
P. angularis			0.22			0.31							0.17	0.21	0.24							0.52		
P. atropurpureus	0.09		0.19	0.24	0.30	0.38							0.19		0.25							0.51		
P. aureus			0.23			0.33							0.19	0.21	0.24							0.49	0.51	
P. calcaratus		0.13	0.20			0.28									0.24				0.40	0.44	0.48	0.51		
P. coccineus							0.38							0.22			0.27	0.29	0.41	0.44	0.50	0.54	0.57	
P. dumosus							0.39						0.18	0.21			0.28	0.30	0.42	0.45	0.49	0.52		
P. filiformis			0.22	0.26	0.31									0.20	0.24		0.30	0.42	0.44	0.50	0.54			
P. lathyroides			0.21			0.31	0.37	0.47					0.19		0.24						0.52			
P. lunatus						0.33							0.19	0.22					0.42	0.45	0.50	0.54	0.57	
P. mungo			0.19			0.28								0.22	0.24					0.44	0.48	0.53		
P. trilobus		0.14	0.22	0.24									0.19		0.24						0.50			
P. vulgaris							0.39						0.17	0.21			0.27	0.30	0.42	0.44		0.53		
Vigna unguiculata			0.18			0.31							0.19	0.22	0.24					0.44	0.48			

prominent component which corresponded in mobility on gels to that of the dominant band of the single-banded species. Three bands were prominent in gels of *P. trilobus* and their mobilities were 0.14, 0.22 and 0.24. The storage proteins of many legumes show reversible association/dissociation with changes in ionic strength and pH (for references see Derbyshire *et al.*, 1975). Thus, the differences between multi-banded patterns observed here and those seen in the previous investigation (Boulter *et al.*, 1967) may have resulted from association/dissociation of a single protein.

Electrophoresis of the protein subunits extracted from several cultivars of *P. vulgaris* and from *P. vulgaris* sub-sp. *aborigineus*, gave subunit patterns which were very similar one to another and which were dominated by a pair of bands R_m 0.27 and R_m 0.30 of unequal intensity (Table 1). The band R_m 0.27 was diffuse and intensely stained relative to the band R_m 0.30. More mobile components were prominent in the regions R_m 0.40–0.45, and 0.50–0.57. In certain of the cultivars, for example 'The Prince', a band was observed in the region corresponding to R_m 0.24, but in other cultivars, e.g. 'Masterpiece', this band was not seen. All these subunits were stained as glycoproteins by the PAS method.

A subunit pattern similar to that of *P. vulgaris* and including band R_m 0.24, was obtained from *P. acutifolius*, *P. coccineus* and *P. dumosus*, which suggests that similar proteins exist in relatively high concentrations in each of the four species. A similar conclusion has been arrived at by Klozova and Kloz (1972), who have shown that a

protein immunochemically similar to the major seed protein of *P. vulgaris* occurs in seed extracts from *P. acutifolius*, *P. coccineus* and *P. dumosus*, although the protein from *P. acutifolius* was not identical to that from the other species. In contrast, they showed that extracts from *P. lunatus* did not react immunologically with the antibodies prepared from the protein from *P. vulgaris*. The major seed protein of *P. vulgaris* is a glycoprotein, Glycoprotein II (Pusztai and Watt, 1970; Racusen and Foote, 1971; Barker *et al.*, 1975), which is dissociated by SDS treatment to subunits R_m 0.27 and 0.30, i.e. subunits with mol. wt 50,000 and 47,000 in SDS gels, and the data suggest that this seed protein is a major protein also in *P. acutifolius*, *P. coccineus* and *P. dumosus*. Subunits with mobilities similar to those of Glycoprotein II were observed as very faint bands on gels of *P. lunatus*; the significance of these weak bands, i.e. whether they belong to Glycoprotein II, cannot be assessed. The other components seen on SDS gels of *P. acutifolius*, *P. coccineus*, *P. dumosus* and *P. vulgaris*, were the major components on gels of *P. lunatus*.

The most prominent subunit extracted from each of the other species was also a glycoprotein. Its mobility (R_m 0.24) was the same in each extract and was less than the mobilities of the Glycoprotein II subunits. In gels of *P. atropurpureus* and *P. lathyroides* the relative concentration of subunit R_m 0.24 was less, and that of subunit R_m 0.19 was greater than in the other species of this group. Subunits characteristic of Glycoprotein II were absent from these species except possibly in *P. atropurpureus* and *P. lathyroides*, where bands of R_m 0.27 and R_m 0.30 were seen as minor components. However, Klozova and Kloz (1972) have shown that the major seed protein of *P. vulgaris* was not detected immunologically in extracts from the latter two species or from the other species which we have examined, except *P. acutifolius*, *P. coccineus* and *P. dumosus*.

Two glycoprotein bands were prominent in the region R_m 0.41–0.45 in gels from *P. acutifolius*, *P. calcaratus*, *P. coccineus*, *P. dumosus*, *P. filiformis*, *P. lunatus* and *P. vulgaris*, and human Group O blood cells were agglutinated when mixed with salt extracts from *P. acutifolius*, *P. coccineus*, *P. dumosus* and *P. vulgaris*. The subunits of the agglutinins of *P. vulgaris* are glycoproteins with mol. wt 34,000 and 30,000, which migrate during electrophoresis with R_m 0.41–0.45 (Allan and Crumpton, 1971; Barker *et al.*, 1975), and it is suggested that agglutinins homologous with those of *P. vulgaris* may occur in *P. acutifolius*, *P. coccineus* and *P. dumosus*. Extracts from the samples of *P. calcaratus*, *P. filiformis* and *P. lunatus* which we examined did not agglutinate human Group O red blood cells. Samples of *P. calcaratus* which did not agglutinate this group of erythrocytes have been reported previously (Bird, 1955), and cultivars of *P. lunatus* which contain agglutinins which are active against these cells are rare (Boyd, Waszczenko-Zacharczenko and Goldwasser, 1961), although cultivars which agglutinate human Group A or non-human red blood cells are widespread (for references see Toms and Western, 1971). The agglutinating activity of seed extracts from *P. filiformis* has not been reported previously, as far as we are aware.

A glycoprotein subunit R_m 0.44 was a conspicuous component in gels of *P. mungo* and *Vigna unguiculata* and appeared to be absent from gels of *Phaseolus aconitifolius*, *P. angularis*, *P. atropurpureus*, *P. aureus*, *P. lathyroides* and *P. trilobus*. Human Group O red blood cells were not agglutinated by salt extracts from any of these species; however, they were lysed when mixed with extracts from *P. atropurpureus* and *P. lathyroides*. Boyd *et al.* (1961) also did not detect agglutinin or lytic activity in extracts from *P. angularis*, *P. aureus*, *P. calcaratus* and *Vigna sinensis* (cowpea) when these were tested against human Groups A, B and O erythrocytes (but see Toms and Western, 1971).

Two subunits R_m 0.48 and R_m 0.52, were prominent in the subunit patterns from

Phaseolus aconitifolius, *P. aureus*, *P. calcaratus* and *P. mungo*, and the less mobile of these bands was prominent in gels of *P. trilobus* and *Vigna unguiculata* but not in those of *Phaseolus angularis*, *P. atropurpureus* and *P. lathyroides*; in the latter three species component R_m 0.52 was prominent.

Taken together the protein data discussed above demonstrate that the species of *Phaseolus* can be separated into a number of groups on the basis of their seed protein composition. By this criterion *P. coccineus*, *P. dumosus* and *P. vulgaris* form a closely related group and *P. acutifolius* differs from these only in the slightly different immunological reactivity of its major seed protein. The species in this group differ from *P. lunatus* and other species which do not have Glycoprotein II as their major seed protein, and on the basis of their protein band patterns and the lytic activity of their extracts, *P. atropurpureus* and *P. lathyroides* are different from other species.

Hutchinson (1964) transferred *P. atropurpureus* and *P. lathyroides* to the genus *Macroptilium* and this revision receives some support from the protein data presented here. It has also been proposed that Asiatic species of *Phaseolus*, including *P. angularis*, should be placed in a new genus *Azuika* (Ohwi, 1965), or, more recently, in *Vigna* (Ohwi and Ohashi, 1969). Maréchal (1969) has transferred *Phaseolus aconitifolius* to *Vigna* also. Verdcourt (1970) has extensively revised the genus *Vigna* and has included in it *Phaseolus aconitifolius*, *P. angularis*, *P. aureus*, *P. calcaratus*, and *P. trilobus*, as well as *Vigna unguiculata* (and many species which we have not examined). The protein data are consistent with these revisions and suggest also that species of *Phaseolus* not transferred to other genera can be separated into a group which has the major seed protein Glycoprotein II and others in which this protein is absent.

REFERENCES

Allan, D. & Crumpton, M. J. (1971). Fractionation of the phytohaemagglutinin of *Phaseolus vulgaris*. *Biochem. Biophys. Res. Commun.* **44**, 1143.

Barker, R. D. J., Derbyshire, E., Yarwood, A. & Boulter, D. (1975). Purification and characterization of the major storage proteins of *Phaseolus vulgaris* L. seeds, and their intracellular and cotyledonary distribution. *Phytochemistry* (in press).

Bird, G. W. G. (1955). Hemagglutinins in plants. *A. M. C. Journ.*, **11**, 17.

Boulter, D., Derbyshire, E., Frahm-Leliveld, J. A. & Polhill, R. M. (1970). Cytology and seed proteins of *Crotalaria*. *New Phytol.*, **69**, 117.

Boulter, D., Thurman, D. A. & Derbyshire, E. (1967). A disc electrophoretic study of globulin proteins of legume seeds. *New Phytol.*, **66**, 27.

Boyd, W. C., Waszczenko-Zacharczenko, E. & Goldwasser, S. M. (1961). List of plants tested for hemagglutinating activity. *Transfusion*, **1**, 374.

Derbyshire, E., Wright, D. J. & Boulter, D. (1975). Legumin and vicilin, storage proteins of legume seeds. *Phytochemistry*, **14**.

Dudman, W. F. & Millerd, A. (1975). Immunochemical behaviour of legumin and vicilin from *Vicia faba*; a survey of related proteins in the Leguminosae subfamily Faboideae. *Biochemical Systematics and Ecology*, **2**, 1.

Hutchinson, J. (1964). *The Genera of Flowering Plants*. Vol. 1. Clarendon Press, Oxford.

Kloz, J. (1971). Serology of the Leguminosae. In: *Chemotaxonomy of the Leguminosae* (Ed. by J. B. Harborne, D. Boulter & B. L. Turner), pp. 309–369. Academic Press, London.

Klozova, E. & Kloz, J. (1972). Distribution of the protein "Phaseolin" in some representatives of Viciaceae. *Biol. Plant. (Praha)*, **14**, 379.

Maréchal, R. (1969). Données cytologiques sur les espèces de la sous-tribu des Papilionaceae-Phaseoleae-Phaseolinae. *Bull. Jard. Bot. Nat. Belg.*, **39**, 125.

Ohwi, J. (1965). *Flora of Japan*. Smithsonian Inst., Washington.

Ohwi, J. & Ohashi, H. (1969). Adzuki beans of Asia. *J. Jap. Bot.* **44**, 29.

Ornstein, L. & Davis, B. J. (1961). *Disc Electrophoresis*. Preprint by Distillation Products Industries, Eastman Kodak Co., Rochester, New York.

Pusztai, A. & Watt, W. A. (1970). Glycoprotein II. *Biochim. Biophys. Acta*, **207**, 413.

Racusen, D. & Foote, M. (1971). The major glycoprotein in germinating bean seeds. *Can. J. Bot.*, **49**, 2107.

Toms, G. C. & Western, A. (1971). Phytohaemagglutinins. In: *Chemotaxonomy of the Leguminosae* (Ed. by J. B. Harborne, D. Boulter & B. L. Turner), pp. 367–462. Academic Press, London.

VERDCOURT, B. (1969). New combinations in *Vigna savi*. *Kew Bull.*, **23**, 464.
VERDCOURT, B. (1970). Studies in the Leguminosae-Papilionoideae: IV *Kew Bull.*, **24**, 507.
WEBER, K., PRINGLE, J. R. & OSBORN, M. (1972). Measurement of molecular weights by electrophoresis on SDS-acrylamide gel. In: *Methods Enzymol.* 26c (Ed. by C. H. W. Hirs & S. N. Timasheff), p. 3. Academic Press, London.
WRIGHT, D. J. & BOULTER, D. (1972). The characterisation of vicilin during seed development in *Vicia faba* (L.). *Planta (Berl.)* **105**, 60.
ZACHARIUS, R. M., ZELL, T. E., MORRISON, J. H. & WOODLOCK, J. J. (1969). Glycoprotein staining following electrophoresis on acrylamide gels. *Anal. Biochem.*, **30**, 148.

TAXON 26(5/6): 513-516. NOVEMBER 1977

236

7

SEED PROTEIN PROFILES OF AFRICAN REPRESENTATIVES OF VIGNA SPP.

J. F. Carasco, E. Derbyshire** and D. Boulter***

Summary

Representatives of *Vigna* species may be separated into a number of groups on the basis of their seed protein profiles. Since these groups are consistent with contemporary classifications of the species it is suggested that the procedure employed here may be of value in systematic investigations of this genus. The profiles indicate that the different species examined all have the same major seed storage proteins.

Introduction

Vigna a widespread tropical and sub-tropical genus is morphologically very diverse. The systematics of this genus has been subject to prolonged controversy and both the classification of traditionally recognised species of *Vigna* and the composition of the genus have been revised from time to time; for a comprehensive discussion of African representatives of *Vigna* see Verdcourt (1970).

The potential value of electrophoretic patterns of proteins in systematic investigations of plants was discussed by Boulter, Thurman & Turner (1966) and useful data have been obtained for a number of taxa (see for example Boulter, Thurman & Derbyshire, 1967; Vaughan & Denford, 1968; Boulter, Derbyshire, Frahm-Leliveld & Polhill, 1970; Cristofolini, 1971). We now report the seed protein profiles of eleven African representatives of *Vigna*, and suggest that our data indicate that the procedure employed here may be of value in systematic investigations of this genus.

Materials and Methods

Seed samples were obtained from the germplasm collection maintained and catalogued by the International Institute of Tropical Agriculture, Ibadan, and had been harvested from plants grown there under similar conditions of soil and climate. Their accession numbers and the species they represent are shown in Table 1. A seed meal was prepared from each sample by milling several seeds in a Glen Creston Mill equipped with a 0.7 mm screen, and protein was extracted from 10 mg of meal by blending with 1 ml 0.2M sodium phosphate buffer pH 7.0 containing 25 mg sodium dodecyl sulphate (SDS) and 20 µl 2-mercaptoethanol solution. The protein solutions were heated in sealed tubes for 2 min at 100° and then cooled rapidly to room temperature. The density of each solution was increased by the addition of 0.5 ml 80% glycerol and bromo phenol blue dye was added to each to act as a visible marker during electrophoresis. Each solution was then divided into two parts and to one, an equal volume of 0.1% myoglobin in the SDS-phosphate buffer, which had also been heated and cooled, was added as a marker protein. The proteins were subjected to electrophoresis in 10% SDS-acrylamide gels by the procedure of Weber & Osborn (1969) and their locations in the gels were determined by staining with 1% Amido Black in

* Department of Botany, University of Dar Es Salaam, Tanzania.

** Department of Botany, University of Durham, Durham DH1 3LE England: requests for reprints should be sent to this address.

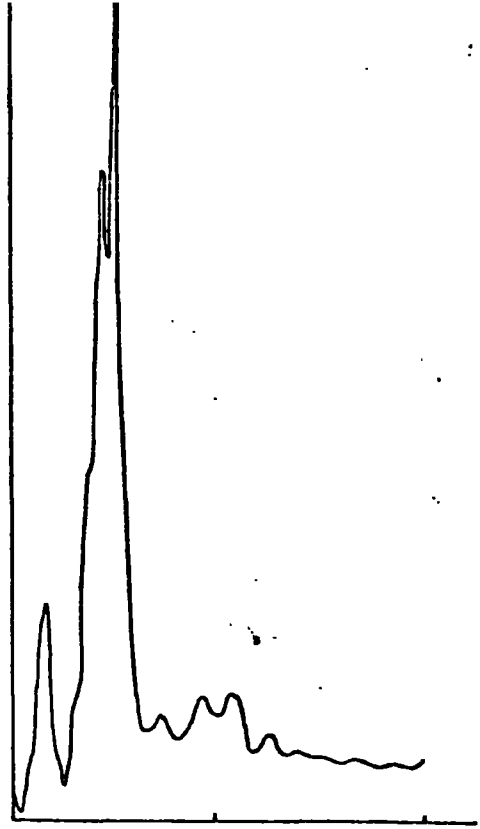


Fig. 1 – Densitometric scan of a 10% SDS-acrylamide gel after electrophoresis of seed protein extracted from *Vigna unguiculata* cv. Prima. *Vertical*: absorbance, *horizontal*: relative mobility (R_m) (scale: 0.0, 0.5, 1.0).

7% acetic acid; excess stain was removed by diffusion in 7% acetic acid and finally the gels were scanned in a Gilford spectrophotometer. The relative rates of migration (R_m) of the individual protein components in the gels were expressed as a decimal fraction of the rate of migration of the bromo phenol blue marker.

Results and Discussion

The seed protein profile of *V. unguiculata* cv. Prima (Fig. 1) was obtained consistently with different samples examined on several occasions; a similar pattern was obtained from cv. Sitao Pole (Table 1). The major components in this profile correspond to the subunits of the major storage proteins of cv. Prima isolated by ion exchange chromatography in 8M urea, (Carasco, 1976) and components with the same R_ms were also the major components in profiles of the other representatives of *Vigna* which we examined; it is probable therefore that the different species of *Vigna* all contain the same major storage proteins.

The profile from *V. unguiculata* included a component, R_m 0.16, which was not prominent in profiles from any of the other samples which we examined and several of the latter contained components with R_m 0.31 and/or R_m 0.48-0.50 which were not prominent in *V. unguiculata*. The profiles of the three samples

TABLE 1. Profiles of seed protein extracted from African representatives of *Vigna spp.* and separated by electrophoresis in 10% SDS-acrylamide gels.

Accession Nos.	Species	R_{ms} of prominent protein components											
Tvu 76	*V. unguiculata* cv. Prima	.09	.16	.21—	.25[1]	.35	.45			.55	.61		
Tvu 22	*V. unguiculata* cv. Sitao Pole	.09	.16	.21—	.25[1]	.35	.45			.55	.61		
Tvnu 36, 40, 41	*V. oblongifolia* var. *parviflora*	.09		.19	.21—	.25[1]	.35	.45		.55	.61		
Tvnu 89	*V. lancifolia*	.09		.19	.21—	.25[1]	.35	.45		.55	.61		
Tvnu 76	*V. vexillata* var. *angustifolia*	.09		.19	.21—	.25[1]	.31	.35	.45	.55	.61		
Tvnu 73, 74	*V. vexillata*	.09		.19	.21—	.25[1]	.31	.35	.45	.48[2]	.55	.61	
Tvnu 65, 75	*V. vexillata*	.09		.19	.21—	.25[1]	.31	.35	.45	.48[2]	.50[2]	.55	.61

(1) Three major components were not completely resolved from one another
(2) Proteins with R_{ms} .48 and .50 were less prominent than the other components listed

of *V. oblongifolia* var. *parviflora* which we examined were identical and the same profile was obtained from accession Tvnu 89 (*V. lancifolia*). This finding is consistent with results obtained by Faris (1965) who showed that *V. oblongifolia, V. parviflora* and *V. lancifolia* were cross fertile, and with the classification of Verdcourt (1970) who has transferred *V. lancifolia* to *V. oblongifolia* and maintains *parviflora* as a separate taxon only at the varietal level. The profiles from the five samples of *V. vexillata* were also substantially similar to each other and they differed from those of the other species examined. However the five profiles from *V. vexillata* were not identical and differences were observed among the minor components with R_ms between 0.48 and 0.50. It must be emphasised that the absence of a component from a profile in Table 1 does not necessarily imply the absence of the corresponding protein from the seed since i) only the more prominent components have been recorded in the table and ii) proteins present in very low concentrations in gels may not have been detected by the staining procedure which we employed.

V. unguiculata, V. oblongifolia and *V. vexillata* belong to different breeding groups (Faris, 1965) and each species is representative of a different sub-genus (Verdcourt, 1970). Thus whilst the data presented here indicate that representatives of *Vigna* may be separated into groups on the basis of their protein profiles it has not demonstrated at which taxonomic level the profiles are characteristic. However, seed protein profiles from several species of *Phaseolus* which have been transferred to *Vigna* (Maréchal, 1969; Verdcourt, 1970) also differed from each other (Derbyshire, Yarwood, Neat & Boulter, 1976) and taken together the two sets of data indicate that species of *Vigna* can be distinguished by use of their seed protein profiles.

References

Boulter, D., Derbyshire, E., Frahm-Leliveld, J. A. & R. M. Polhill 1970 – Cytology and seed proteins of *Crotalaria*. New Phytol. 69: 117-131.

Boulter, D., Thurman, D. A. & E. Derbyshire 1967 – A disc electrophoretic study of globulin proteins of legume seeds. New Phytol. 66: 27-56.

Boulter, D., Thurman, D. A. & B. L. Turner 1966 – The use of disc electrophoresis of plant proteins in systematics. Taxon 15: 135-143.

Carasco, J. F. 1976 – Ph.D. Thesis, University of Durham.

Cristofolini, G. 1971 – Contributo sierodiagnostico alla sistematica di *Euphorbia triflora*. Giornale Botanico Italiano. 105: 145-156.

Derbyshire, E., Yarwood, J. N., Neat, E. & D. Boulter 1976 – Seed proteins of *Phaseolus* and *Vigna*. New Phytol. 76: 283-288.

Faris. D. G. 1965 – The origin and evolution of the cultivated forms of *Vigna sinensis*. Can. J. Genet. Cytol. 7: 433-452.

Maréchal, R. 1969 – Données cytologiques sur les espèces de la sous-tribu des Papilionaceae-Phaseoleae-Phaseolinae. Bull. Jard. Bot. Nat. Belg. 39: 125-165.

Vaughan, J. G. & K. E. Denford 1968 – An acrylamide gel electrophoretic study of the seed proteins of *Brassica* and *Sinapis* species. J. exp. Bot. 19: 724-732.

Verdcourt, B. 1970 – Studies in the Leguminosae-Papilionoideae for the 'Flora of Tropical East Africa': IV. Kew Bulletin 24: 507-569.

Weber, K. & M. Osborn 1969 – The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. J. Biol. Chem. 244: 4406-4412.

New Phytol. (1977) 78, 499–504.

8

THE EXTRACTION, COMPOSITION AND INTRA-CELLULAR DISTRIBUTION OF PROTEIN IN EARLY MAIZE GRAINS FROM AN ARCHAEOLOGICAL SITE IN N.E. ARIZONA

BY E. DERBYSHIRE*, N. HARRIS*, D. BOULTER* and E. M. JOPE†

**Department of Botany, University of Durham, Durham DH1 3LE and*
† Department of Archaeology, Queens University, Belfast BT7 1NN

(*Received 28 June 1976*)

SUMMARY

Maize grains from an archaeological site in N.E. Arizona have been examined by light microscopy and electron microscopy. Their endosperm tissue was intact and was similar to that of modern maize although some of the features seen in other non-viable seeds were present; the embryos, however, were discoloured and damaged. Protein has been extracted from both endosperm and embryos although only a small quantity was extracted from the latter. The protein sub-unit profiles of these extracts and of storage protein fractions prepared by serial extraction of endosperm meal in salt, alcohol and sodium dodecyl sulphate (SDS) solutions have been determined after electrophoresis in SDS-acrylamide gels. The data obtained indicate that zein and glutelins similar to modern maize proteins were the major storage proteins but that their proportions differed to those in modern maizes; these conclusions are supported by the amino acid composition of the wholse seed meal. It is suggested that these procedures may be applicable to other archaeological samples of seeds including those of other species and that the data so obtained may be of use in cultural studies and in botanical studies of crop plant origins and evolution.

INTRODUCTION

The morphology of archaeological plant material has been compared with that of modern varieties in order to trace the probable evolutionary pathway of maize during its domestication (Mangelsdorf, 1974). Morphological structures, however, are a less direct reflection of the genetic information of an organism than the protein profile formed by a direct template mechanism and comparisons of protein profiles have been shown to be of considerable value in establishing relationships of present day taxa (e.g. see Johnson and Hall, 1966; Boulter *et al.*, 1970; Avise, 1974). The investigation reported here demonstrates that proteins can be extracted from archaeological plant material and meaningfully compared with those from modern varieties, indicating that comparative studies of the protein profiles of ancient seeds may be useful both for archaeological investigations and for studies of crop plant origins and evolution.

MATERIALS

Well-preserved grains of 10–12 row maize, average weight 0.3 g, from the excavations of Earl H. Morris (unpublished) on a site in Canyon del Muerto in the N.E. corner of Arizona, near

the more famous Mummy Cave, were kindly provided by Dr J. B. Wheat of the University of Colorado Museum at Boulder. The site has not been precisely dated but this is potentially determinable from tree-ring data (Gladwin, 1957; Schulman, 1949) and is between the Basketmaker stage in the earlier part of the 1st millenium (Morris, 1954) and its time of abandonment sometime between the 11th and 13th centuries (Bushnell, 1957).

METHODS

Starch distribution was examined in fresh sections of imbibed grains stained with iodine in potassium iodide. Protein distribution was examined in 2-μm sections of resin embedded material stained with alcoholic Bromophenol Blue. Material for electron microscopy was fixed in 2.5% glutaraldehyde and 1.5% formaldehyde in 0.1 M phosphate buffer pH 7.0, post-fixed in 1% osmic acid, dehydrated with ethanol and embedded in Spurr's resin. Thin sections were stained with uranyl acetate and alkaline lead citrate.

Meals were prepared from whole seeds and from separated endosperm and embryos by milling in a Glen Creston Mill and were extracted by blending in 2.5% sodium dodecyl sulphate (SDS)–0.01 M 2-Mercaptoethanol (2-Me) –0.1 M sodium phosphate buffer pH 7.0, (100 mg meal/ml) for 3 min. at high speed, then stirred for 15 min. and finally clarified by centrifugation at 23,000 ***g*** for 10 min. Salt soluble protein was extracted from other portions of meals in 0.4 M NaCl-0.1 M phosphate buffer pH 7.0 by a similar procedure. The salt-insoluble residue from the early maize endosperm was serially extracted with 70% ethyl alcohol, 70% ethyl alcohol made 0.01 M with respect to 2-Me, and 2.5% SDS-0.01 M 2Me-0.1 M phosphate buffer; each extraction step, including the saline extraction was repeated twice.

Samples of extracts, in 1% SDS-0.1 M 2-Me-0.1 M phosphate buffer pH 7.0 were heated for 5 min. in closed tubes at 100°C then cooled. Electrophoresis of these samples was carried out in SDS-acrylamide gels by the method of Weber and Osborn (1969). Samples of the saline extracts were examined, in the absence of SDS and 2-Me, in non-dissociating acrylamide gels (Ornstein and Davis, 1961). Gels were stained for total protein with 1% Amido Black in 7% acetic acid and excess stain was removed by diffusion in 7% acetic acid. The esterase, acid phosphatase and leucine amino peptidase activities of saline extracts were determined separately in non-dissociating acrylamide gels (Brewbaker *et al.*, 1968).

Duplicate samples of meals were hydrolysed in 6 M HCl for 20 h in sealed tubes at 105°C and the amino acids were determined with a Locarte automatic loading analyser. Total meal nitrogen was determined by an automated micro-Kjeldahl method (Varley, 1966).

RESULTS AND DISCUSSION

The maize from the Canyon del Muerto did not germinate at 27°C although the imbibed grains had the same moisture content (35%) as viable imbibed grains of a modern maize used for comparison.

Visual inspection and light microscopy showed that the endosperm tissues of the early maize were relatively intact and contained starch and protein whilst those of the embryos were discoloured and damaged, and in some grains destroyed. The general structure of the endosperm in which there was a single layer of aleurone cells was similar to that of modern maizes (Khoo and Wolf, 1970). Electron microscope studies of aleurone cells showed short lengths of endoplasmic reticulum, intact nuclei which were multilobed, and numerous protein bodies and spherosomes (Plate 1, No. 1, and Plate 2, No. 2); the spherosomes

surrounded the nuclei and protein bodies, lined the plasmalemma and filled much of the cells. Other cell organelles were not readily distinguished. Some protein bodies showed clear regions similar to the clear globoids which appear during the initial stages of reserve mobilization in barley (Jones, 1969). Numerous dark dense patches in the chromatin of nuclei (Plate 1, No. 1) and withdrawal of the protoplast from cell walls (Plate 2, No. 2) were observed in the early maize aleurone cells and these are features seen in other non-viable seeds (van Staden, Gilliland and Brown, 1975; Villiers, 1971), however, another feature of non-viable seeds, namely the extensive coalescence of spherosomes (van Staden *et al.*, 1975) was not seen; the majority of the spherosomes were intact and only a small number had coalesced. When examining the inner endosperm we encountered similar practical difficulties to those of mature maize reported by Khoo and Wolf (1970). They distinguished between a sub-aleurone layer and the starch endosperm and commented on differences in protein body staining and size in the two zones. We noted a similar sub-aleurone layer of approximately cubic cells which contrasted in shape with the elongate inner cells and a larger population of small protein bodies in the inner tissues than that in the sub-aleurone layer (Plate 2, No. 3). Larger protein bodies were also present in both the sub-aleurone layer and the five to six layers of inner tissue which we examined; all were membrane bound, and showed the characteristic speckling which had been observed by Khoo and Wolf (1970) only in the sub-aleurone layer. Protein was not confined solely to the protein bodies and formed a matrix in the sub-aleurone layer and in the inner endosperm. The matrix protein was more intensely stained than the protein body protein, however, the total volume of the protein bodies was far in excess of that of the matrix.

The storage protein of maize consists of alcohol-soluble protein (zein) and acid/alkali soluble glutelin which may consist of more than one protein. Zein is laid down in the protein protein bodies of the endosperm and glutelins form a matrix in the endosperm cells (Wolf, Khoo and Seckinger, 1967). Both protein types are extracted by SDS.

The gel protein profiles of SDS-2Me extracts of meals from the whole seed, endosperm and embryo of the early maize show that the endosperm pattern is very similar to that of the whole seed meal indicating that little if any embryo protein is present in the latter; SDS patterns of the embryo confirm this but show that they do contain a small amount of protein. Also included in Fig. 1 is the gel pattern of a modern maize, and comparison with the early maize pattern shows that protein subunits similar to those of modern maize are present in the early maize. The subunit structures of zein and glutelin fractions extracted from the early maize by use of the appropriate solvents have been established by SDS gel electrophoresis Fig. 1, e-g). When these are compared with subunit profiles of similar fractions from modern maize, for example those of Paulis, Bietz and Wall (1975), it can be seen that although the profiles are not identical the subunit structures of zein and most of the glutelin fraction have been preserved in the early material. The relative proportions of the protein types are different, see Fig. 1a, in that early maize contains proportionately less glutelin than modern maize. Further support for this suggestion comes from a consideration of the amino acid composition data of the two meals (Table 1). These are very similar indicating the presence of the same type of proteins but differ in the proportionally lower lysine and methionine and proportionally higher leucine in early maize, probably explicable by the fact that glutelin contains relatively more lysine and methionine and less leucine than zein (Paulis and Wall, 1971; Mertz, 1972). The alcohol-insoluble glutelin (AIG) obtained from early maize in this investigation consisted of a smaller number of subunits than that obtained from AIG of a modern maize by Paulis *et al.*, (1975). also the most prominent

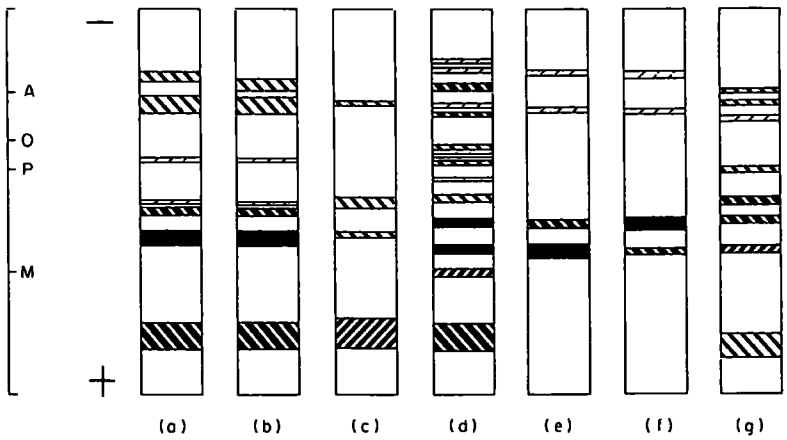


Fig. 1. Sub-unit profiles of maize protein extracts: (a) whole seed, (b) endosperm and (c) embryos of early maize; (d) whole seed of a modern maize; (e) zein, (f) alcohol-soluble glutelin and (g) alcohol insoluble glutelin from early maize. The relative mobilities of bovine serum albumin (A), ovalbumin (O), pepsin (P), and myoglobin (M) in gels are shown for comparison.

Table 1. *The amino acid compositions of early and modern maizes*

	g per 16 g meal N*	
	Early	Modern
Aspartic acid	4.60 ± 0.28	4.92 ± 0.24
Threonine	2.12 ± 0.20	2.20 ± 0.09
Serine	2.76 ± 0.28	2.64 ± 0.35
Glutamic acid	12.53 ± 0.78	13.19 ± 0.53
Proline	5.31 ± 0.42	5.98 ± 0.26
Glycine	2.34 ± 0.16	2.73 ± 0.09
Alanine	4.48 ± 0.30	5.01 ± 0.26
Valine	3.40 ± 0.21	3.52 ± 0.00
Methionine	1.06 ± 0.07	1.67 ± 0.09
Isoleucine	2.55 ± 0.21	2.37 ± 0.02
Leucine	9.27 ± 0.39	8.53 ± 0.31
Tyrosine	1.91 ± 0.26	2.46 ± 0.16
Phenylalanine	4.18 ± 0.28	4.31 ± 0.18
Histidine	1.56 ± 0.16	2.11 ± 0.09
Lysine	1.63 ± 0.14	2.46 ± 0.17
Arginine	2.90 ± 0.21	3.78 ± 0.26

* Means and range of duplicate samples.

subunits in gels of the early maize AIG were more mobile than those of the modern AIG. Since subunits in the latter fraction ranged in molecular weight up to 127,000 daltons, which is an unusually large value for individual polypeptide chains (Smith, 1966), it is possible that the dissociation achieved by Paulis *et al.*, (1975) was incomplete.

The results of this investigation demonstrate the possibility of extracting proteins from archaeological samples for use in cultural studies and in botanical studies of crop plant origins and evolution. Furthermore, it may be possible to use these procedures with archaeo-

logical samples of other seeds such as wheat, sorghum, teosinte and barley in which similar storage proteins occur. Application to dicotyledonous seeds, for example legumes where most of the protein is salt soluble, may be less successful since the non-dissociating gel protein profiles of saline extracts of early maize differed from those of modern maize (Fig. 2). Furthermore, leucine amino peptidase, acid phosphatase and esterase activities were not

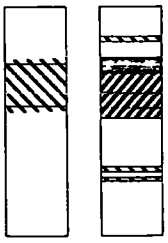


Fig. 2. Protein profiles of saline extracts from (a) early and (b) modern maizes, in non-dissociating acrylamide gels.

detected in gels of early maize when these were stained specifically for the presence of these enzymes. Wheat grains recovered from archaeological sites also did not have detectable enzyme activity (Hallam, 1973). Palozzo and Jaffé (1965) found differences between the salt-soluble proteins of an archaeological sample of legume seeds (*Phaseolus vulgaris*) and those of a modern variety; only weak diffuse bands were obtained following immuno-electrophoresis of the early bean extracts and haemagglutinating activity was absent.

ACKNOWLEDGMENTS

We wish to thank Messrs R. Swinhoe, B. Joicey and L. Al-Shakhs for technical assistance.

REFERENCES

AVISE, J. C. (1974). Systematic Value of Electrophoretic data. *Systematic Zoology*, 23, 465.

BOULTER, D., DERBYSHIRE, E., FRAHM-LELIVELD, J. A. & POLHILL, R. M. (1970). Cytology and seed proteins of *Crotalaria*. *New Phytol.*, 69, 117.

BREWBAKER, J. L., UPADHYA, M.D., MAKINEN, Y. & MACDONALD, T. (1968). Isoenzyme polymorphism in Flowering Plants III. *Physiol. Plant*, 21, 930–940.

BUSHNELL, G. H. (1957). Review of Gladwin, H. S. (1957). *Antiq. J.* 37, 235.

GLADWIN, H. S. (1957). *A History of the Ancient Southwest*. The Bond Wheelwright Co., Portland, Maine.

HALLAM, N. D. (1973). Fine structure of viable and non-viable Rye and other embryos. In *Seed Ecology*, (Ed. by W. Heydecker), p. 99. Butterworths, London.

JOHNSON, B. L. & HALL, O. (1966). Electrophoretic studies of species relationships in *Triticum*. *Acta. Agric. Scand. Suppl.* 16, 222.

JONES, R. L. (1969). Gibberellic acid and fine structure of Barley Aleurone Cells. *Planta (Berl.)*, 88, 73.

KHOO, U. & WOLF, M. J. (1970). Origin and Development of Protein Granules in maize endosperm. *Amer. J. Bot.*, 57, 1042.

MANGELSDORF, P. C. (1974). *Corn; its Origin, Evolution and Improvement*. Harvard University Press, Cambridge, Mass.

MERTZ, E. T. (1972). Recent improvements in corn proteins. In: *Seed Proteins* (Ed. by G. E. Inglett), p. 136 Avi Publishing Co. Westport, Connecticut.

MORRIS, E. H. (1954). *Basketmaker II sites near Durango, Colorado.* Publication 604 Carnegie Institution of Washington, D.C.

ORNSTEIN, L. & DAVIS, B. J. (1961). *Disc Electrophoresis.* Preprint by Distillation Products Industries (Eastman Kodak Co.) Rochester, New York.

PALOZZO, A. & JAFFÉ, W. G. (1965). Reacciones immunológicas de extractos de semillas de habichuelas. *Bolet. Soc. Quim. Peru.*, **31**, 1.

PAULIS, J. W., BIETZ, J. A. & WALL, J. S. (1975). Corn Protein subunits. *J. agric. Food Chem.*, **23**, 197.

PAULIS, J. W. & WALL, J. S. (1971). Fractionation and properties of alkylated-reduced corn glutelin proteins. *Biochim. Biophys. Acta*, **251**, 57.

SCHULMAN, E. (1949). An extension of the Durango Chronology *Tree-Ring Bull.* **16**, No. 2.

SMITH, M. H. (1966). The amino acid composition of proteins. *J. theoret. Biol.* **13**, 261.

VAN STADEN, J., GILLILAND, M. G. & BROWN, N. A. C. (1975). Ultrastructure of dry viable and non-viable *Protea compacta* embryos. *Z. Pflanzenphysiol.* **76**, 28.

VARLEY, J. A. (1966). Automatic methods for the determination of nitrogen, phosphorus and potassium in plant material. *Analyst*, **91**, 119.

VILLIERS, T. A. (1971). Cytological Studies in Dormancy I. *New Phytol.*, **70**, 751.

WEBER, K. & OSBORN, M. (1969). The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. *J. biol. Chem.* **244**, 4406.

WOLF, M. J., KHOO, V. & SECKINGER, H. L. (1967). Subcellular structure of endosperm protein in high lysine and normal corn. *Science*, **157**, 556.

EXPLANATION OF PLATES

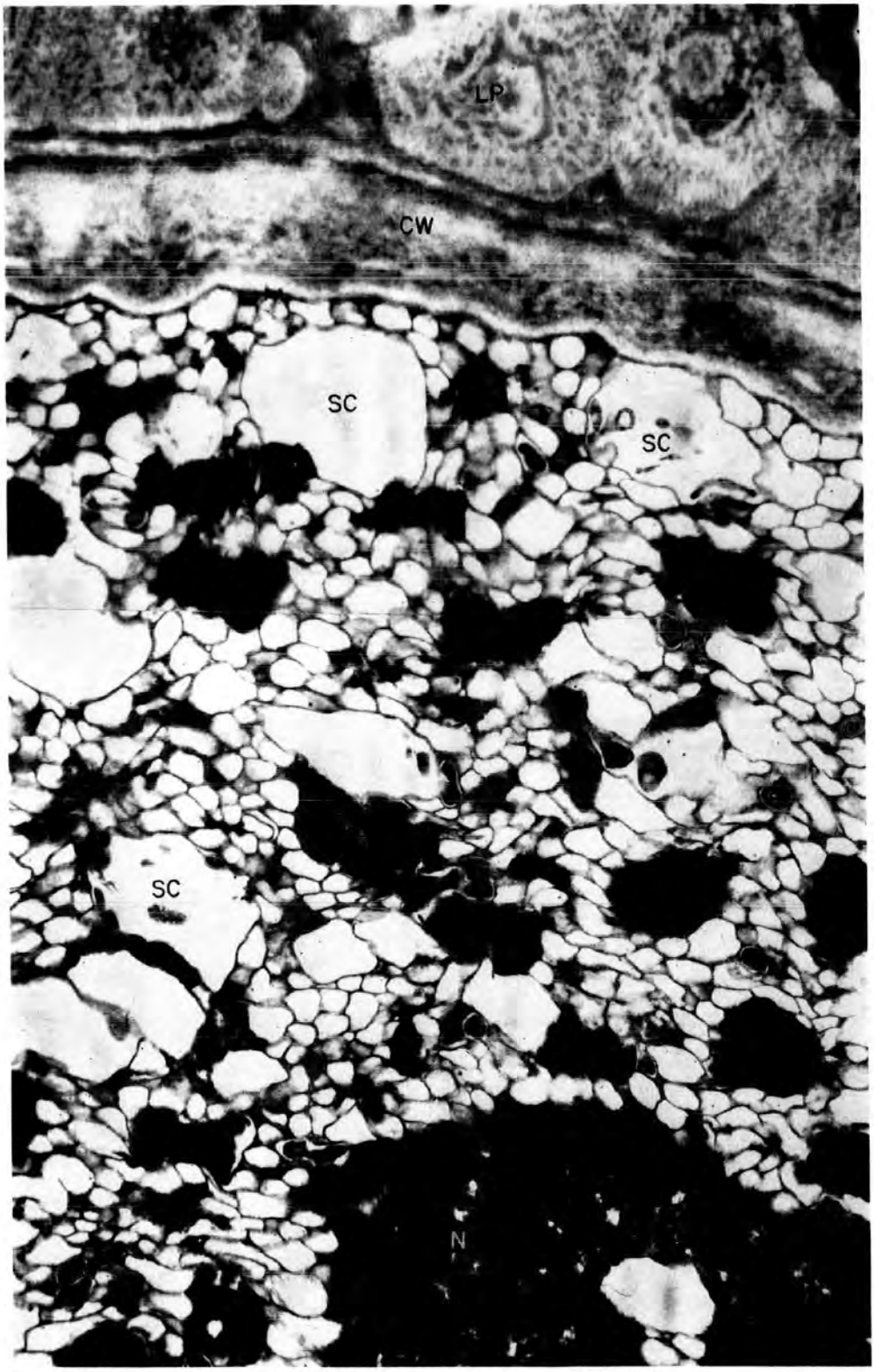
PLATE 1

No. 1. Electron micrograph of endosperm tissues: part of an aleurone cell and adjacent cell in sub-aleurone layer showing nucleus (N) with necrotic features, some spherosome coalescence (SC) in aleurone cell, large protein bodies (LP) of sub-aleurone layer, and cell wall (CW): magnification × 16,000.

PLATE 2

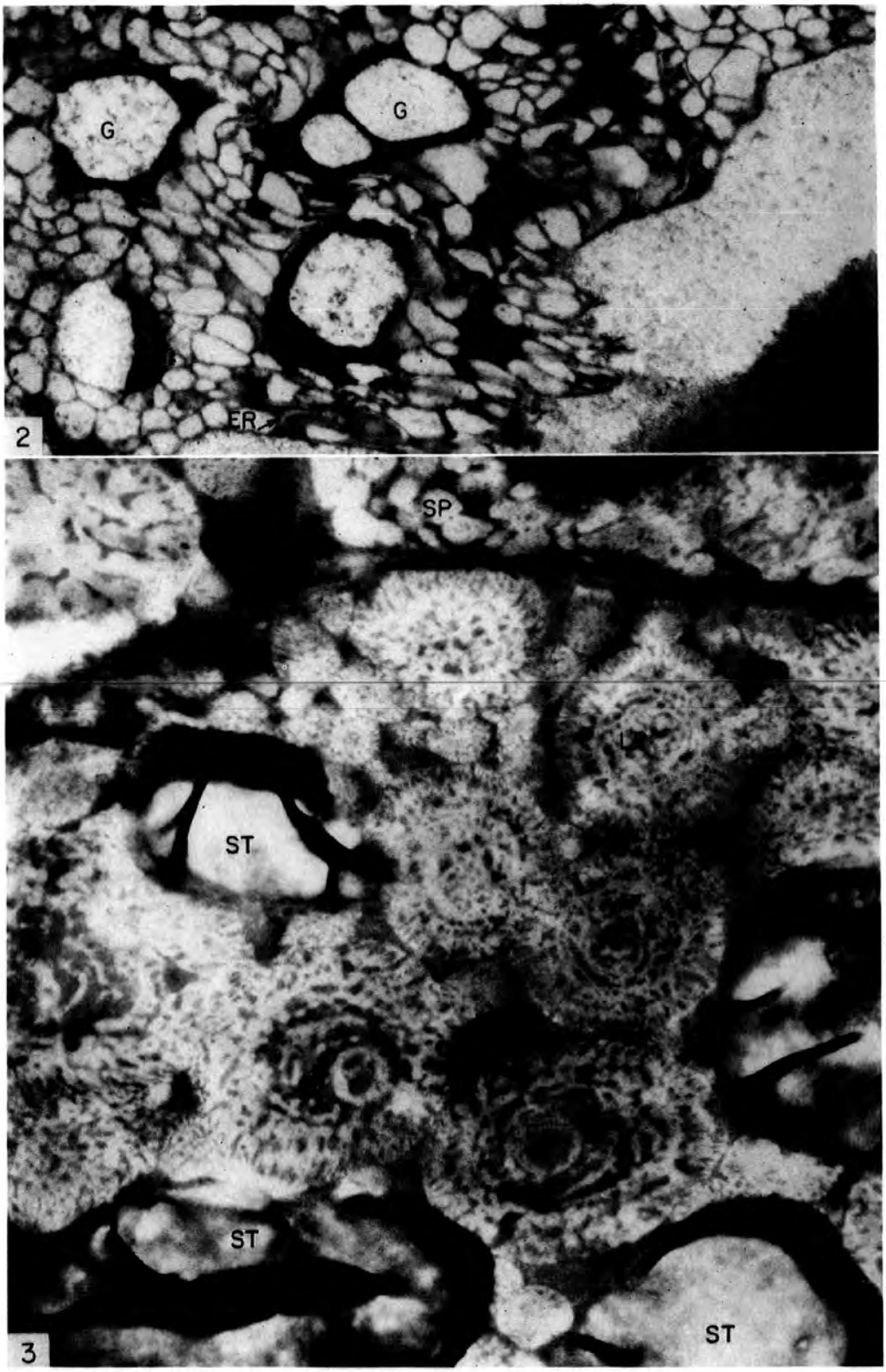
No. 2. Electron micrograph of endosperm tissues: part of aleurone cell showing withdrawal of protoplast from cell wall, clear globoid regions (G) within protein bodies and some endoplasmic reticulum (ER). × 20,000.

No. 3. Electron micrograph of endosperm tissues: inner endosperm showing starch grains (ST), large protein bodies (LP) and small protein bodies (SP) embedded in protein matrix: × 17,500.



E. DERBYSHIRE *ET AL.—PROTEIN IN EARLY MAIZE GRAINS*

(*facing p.* 504)



E. DERBYSHIRE *ET AL.—PROTEIN IN EARLY MAIZE GRAINS*

PROTEINS OF SOME LEGUMES WITH REFERENCE TO ENVIRONMENTAL FACTORS AND NUTRITIONAL VALUE*

9

DONALD BOULTER, I. MARTA EVANS & ERIC DERBYSHIRE

(Durham)

ABSTRACT

Storage globulins contribute 80–90% of the seed protein of many legumes and consist principally of two proteins called vicilin and legumin in peas. The relative proportions of the two proteins has been shown to differ in different legumes by gel-electrophoresis. The essential amino acid composition of vicilin and legumin of *Pisum* and *Vicia* differ and the possible significance of these findings to screening strategies for improved legume programmes was discussed.

Cowpeas were grown in growth cabinets under standard environmental conditions. Seeds were shown to have the same percentage nitrogen and sulphur on a dry-weight basis, irrespective of the position of the pod on the plant, indicating that legume seeds may have a conservative biochemistry.

Legume seeds may contain up to 50% protein on a dry-weight basis. Most of this protein, of the order of 80–90%, is contributed by the seed storage globulins (Osborne, 1924), which have been operationally divided into two fractions, called legumin and vicilin in peas (Osborne, 1924; Danielsson, 1949). Proteins with similar properties occur in a wide range of legumes (Danielsson, 1949; Altschul, Yatsu, Ory & Englemann, 1966; Boulter & Derbyshire, 1971). Because of their physiological role, it is likely that storage proteins are complex (see Bailey & Boulter, 1971), and, in fact, those of several legume species, e.g. *Pisum sativum, Vicia faba, Cicer arietinum, Glycine max* and *Arachis hypogea,* have been shown to be made up of sub-units (see Boulter & Derbyshire, 1971).

Amino acid analyses of, and rat-feeding experiments with a variety of legume meals, indicate that the sulphur-amino acids, cysteine and methionine, are usually the first-limiting amino acids nutritionally (see Aykroyd & Doughty, 1964), and this is a reflection of the fact that the storage globulins, the major proteins of the meal, are low in sulphur-amino acids.

This paper discusses the proteins of developing legume seeds, in relation-

*Paper presented at the conference of the International Association for Quality Research on food Plants (CIQ) held in common with the Deutsche Gesellschaft für Qualitätsforschung (Pflanzliche Nahrungsmittel) e.V. (DGQ) in Berlin on 6th October 1972

Qual. Plant. – Pl. Fds. hum. Nutr. XXIII, 1/3 : 239–250, 1973

ship to possible strategies for screening in breeding programmes for improved grains legumes. Several different legumes are used as examples.

RESULTS

Table I gives the amino acid compositions of vicilin and legumin of *Pisum* and *Vicia*. It can be seen that with regard to the essential amino acids, legumin of both species is higher than vicilin in the sulphur-amino acids and tryptophan, but lower in isoleucine, lysine and phenylalanine. The remaining essential amino acids are in, more or less, the same proportion in the two fractions.

The results given in Fig. 1 show that vicilin and legumin do not occur in the same relative proportions in different legumes, and that sometimes one and sometimes the other predominates.

We are presently attempting to use total sulphur as a measure of the sulphur-amino acid content of different varieties of cowpeas (Boulter, Evans, Thompson & Yarwood, 1972). In collaboration with Drs. Hughes and Summerfield of the Plant Environment Laboratory of the University of

Table I *Amino acid composition of vicilin and legumin of Pisum and Vicia (g amino acid/16g N)*

	Vicilin		Legumin	
	Pisum sativum	*Vicia faba*	*Pisum sativum*	*Vicia faba*
Arg	7.3	7.8	10.5	11.3
His	2.1	2.4	2.8	3.0
Lys	7.9	8.1	4.9	5.3
Tyr	3.0	3.8	3.3	3.9
Try	0.09	0.08[1]	1.06	1.2[2]
Phe	6.2	6.8	4.9	4.8
Cys	0.35	0.28	0.71	0.74
Met	0.22	0.36	0.65	0.7
Ser	5.8	5.1	4.5	5.1
Thr	3.4	2.9	2.9	3.9
Leu	9.2	9.3	8.1	8.0
Isol	5.1	5.2	4.0	4.0
Val	4.6	4.3	4.6	4.4
Glu	19.3	17.6	21.1	19.2
Asp	12.0	11.9	12.5	11.0
Ala	3.0	3.1	3.7	3.9
Pro	3.5		4.3	
Gly	3.1	2.5	3.4	3.8

Modified from Boulter & Derbyshire (1971).
[1] Bailey & Boulter, 1972. [2] Bailey & Boulter, 1970.

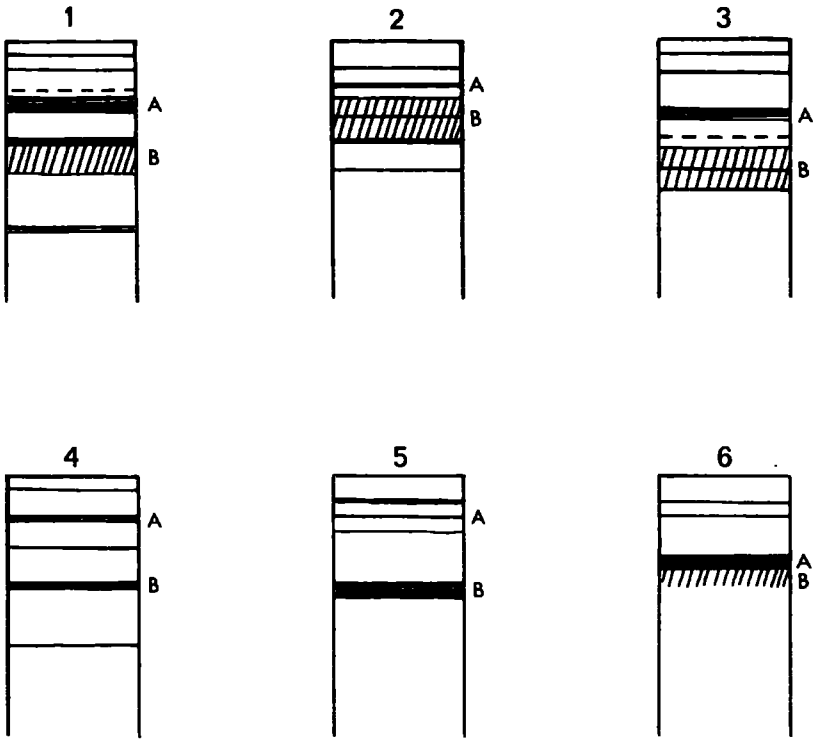


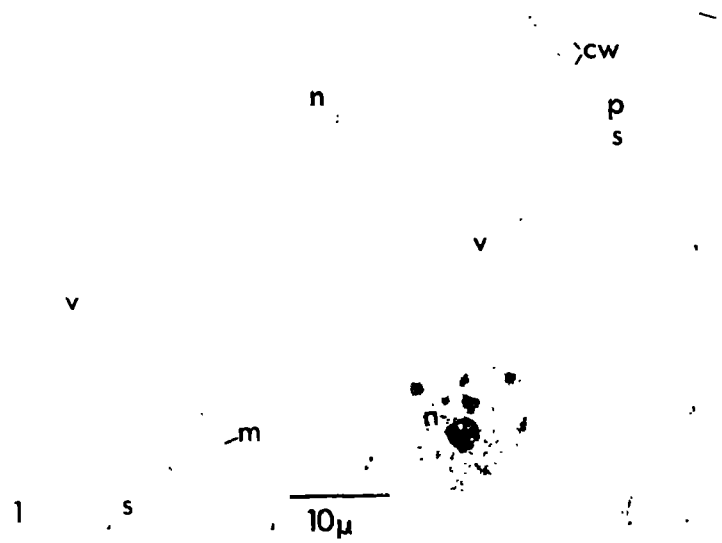
Fig. 1 Title: Electrophoretograms of storage proteins of various legumes separated by anion gel electrophoresis.
Legend: legumin (A); vicilin (B); 1. *Vigna aureus;* 2. *V. mungo;* 3. *Cajanus cajan;* 4. *Phaseolus lunatus;* 5. *P. vulgaris;* 6. *Vicia faba.* Globulins were prepared and electrophorised as described by Bailey & Boulter (1970).

Reading, we have analysed the percentage nitrogen and percentage sulphur of the seeds of pods at various nodes of cowpeas grown, by them, under standard conditions in growth cabinets (daylength 12 h. 30, day temperature 30°, night temperature 21.5°). These results show that the percentage nitrogen and percentage sulphur of seeds in pods at different nodes on the same plant and in pods on different plants, have, within 10% limits, the same percentage nitrogen and percentage sulphur on a dry-weight basis. These results indicate that legume seeds are biochemically conservative and that seeds are produced which conform to the average, otherwise abortion takes place. In our analyses we have ignored aborted seeds which represent less than 0.1% of the total dry weight of the seeds of a pod.

DISCUSSION

Opaque 2, a high lysine maize with an improved nutritional status, has an increased ratio of glutelin to prolamin as compared to normal strains, i.e. a change has taken place in the relative proportions of the storage proteins. By analogy, nutritionally improved grain legumes would be obtained if in a variety of a legume such as *Pisum* or *Vicia*, the ratio of legumin to vicilin were increased in the mature seed, since, as can be seen from the results in Table 1, legumin contains more sulphur-amino acids than vicilin. How far the sulphur-amino acid patterns of vicilin and legumin of *Pisum* and *Vicia* can be extrapolated to these same protein fractions in other legumes, is uncertain, due to difficulties in separating these two fractions cleanly. Since different proportions of these two proteins occur in different legume species (see Fig. 1), the potential for grain improvement by this route also differs.

Vicilin and legumin are synthesised in the protein bodies of the developing seeds of *Vicia faba* (Briarty, Coult & Boulter, 1969; Bailey, Cobb & Boulter, 1970), also see Plates 1–4, and Graham and Gunning (1970) have shown, using serological methods, that both proteins are laid down in the



PLATES 1–4 Are a series of electron micrographs of cotyledon cells from *Vicia faba* (var. Triple White) at various ages after fertilization. They have all been fixed in Gluteraldehyde (4% v/v) post-fixed in Osmic Acid (1% w/v) and after sectioning, stained in Uranyl Acetate followed by Lead citrate.

Plate 1 Cells 25 days after fertilization. They appear highly vacuolate, the cytoplasm containing many free ribosomes; very little endoplasmic reticulum is present.

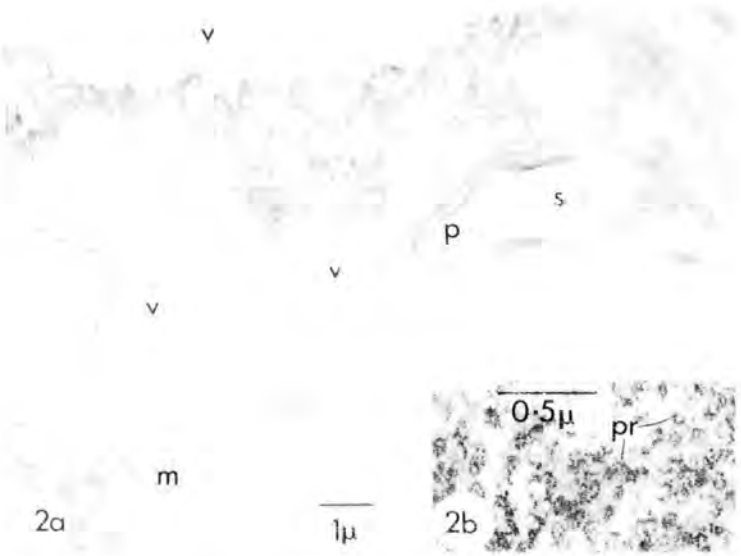


Plate 2 A section through a strand of cytoplasm in a cell 30 days after fertilization. A number of small vacuoles have formed within the cytoplasm, some endoplasmic reticulum has formed, and most of the ribosomes are now in clusters (polyribosomes), see 2b inset.

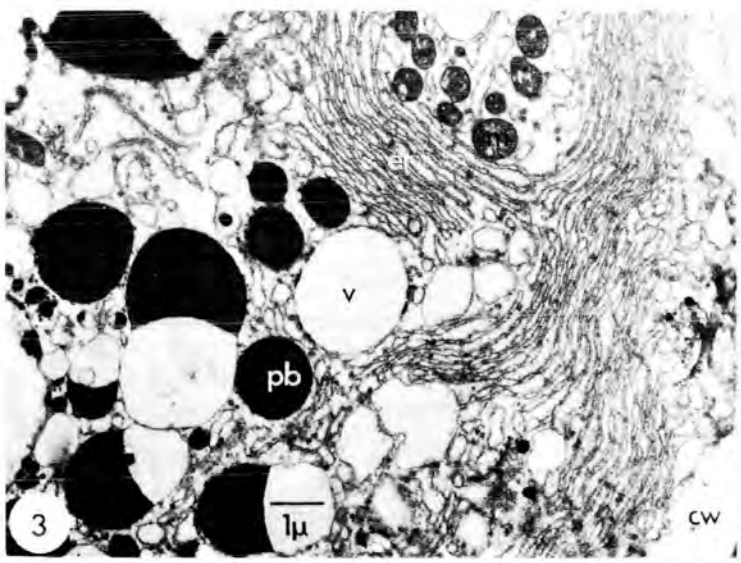


Plate 3 Part of the cytoplasm of a cell 55 days after fertilization. Most of the polyribosomes are now bound to the endoplasmic reticulum and storage protein is being laid down in the vacuoles forming protein bodies.

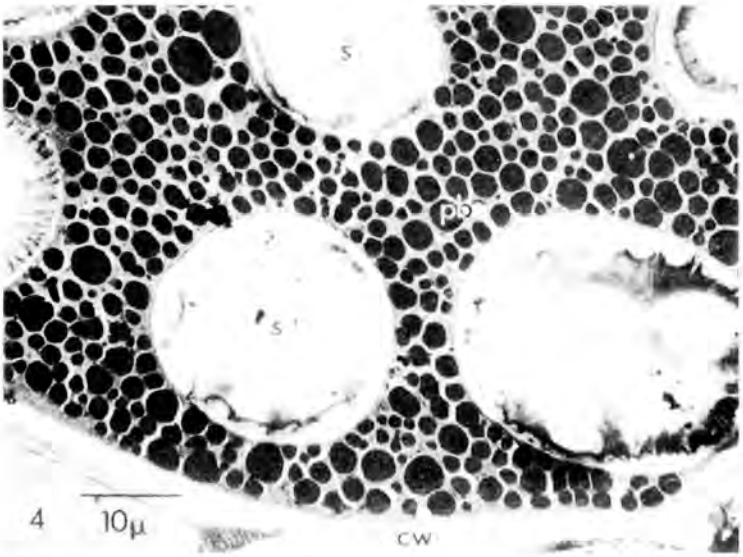


Plate 4 An almost mature cell (85 days after fertilization). It appears almost entirely packed with protein bodies and starch grains.

Abbreviations

cw — cell wall.
er — endoplasmic reticulum.
m — mitochondrion.
n — nucleus.
p — plastid.
pb — protein body.
pr — polyribosome.
s — starch grain.
v — vacuole.

Electron micrographs by courtesy of A. R. Cobb, Department of Botany, University of Durham, England.

same protein bodies. This situation is not directly comparable therefore, with that of maize, where prolamin is laid down in the protein body but glutelin appears as a matrix protein. Thus, the chances of obtaining legume lines with changed proportions of storage proteins, may prove more difficult.

Vicilin is soluble at pH 4.7 (Danielsson, 1949), whereas legumin is not. Unfortunately, complete separation of vicilin and legumin cannot be achieved by this method since not all the vicilin dissolves at pH 4.7, nor does all the legumin remain insoluble (unpublished work on *Vicia, Phaseolus* and *Vigna*). However, comprehensive variations in ionic strength and type of ion have not been fully investigated and, if pursued, might lead to the

elucidation of conditions where there was a virtual complete extraction of vicilin at pH 4.7. If this were so, a screening procedure which determined the total nitrogen of meal after extraction at pH 4.7, would, in fact, be a screening procedure for legumin content, and meals which gave high nitrogen values would be high legumin-containing varieties and hence, with some legumes, be varieties with a high sulphur-amino acid content.

Since legumes are not eaten alone, improved grain legume programmes must be considered in the context of the whole diet. Lines of *Vicia* and *Pisum* which result from an increased proportion of legumin to vicilin will have relatively less of the essential amino acids, lysine, isoleucine and phenylalanine, and, if incorporated into a diet, might in some circumstances lead to an overall deterioration in its nutritional value, e.g. cereal-based diets with limiting lysine. If, therefore, it is accepted that lines with changed proportions of vicilin and legumin are likely to be selected in screening programmes, it is essential to consider the overall amino acid composition of the diet in which these are to be incorporated and not to assume that increased sulphur-amino acid content will necessarily be beneficial.

Danielsson (1952) using peas, and Wright & Boulter (1972) *Vicia faba*, have shown that the onset of the synthesis of the two storage proteins does not occur simultaneously in the developing seeds; for example, in *Vicia faba*, vicilin was formed prior to legumin, although the rate of synthesis of the latter was faster, so that in the mature seed the ratio of legumin to vicilin was about 4 : 1 by weight. These results indicate that these proteins are under separate genetic control and by analogy with Opaque 2, support the suggestion that lines with changed proportions of storage protein may well exist or be generated in breeding programmes. The question, however, still remains to be answered as to how much change can take place in the pattern of seed proteins without loss of viability etc. Probably in the case of the storage proteins which supply nitrogen for the developing seed, the scope for change without loss of function is much greater than for an enzymic protein. The latter, for full activity, has to conform to a strict three-dimensional shape, which depends on the amino acid sequence. Whereas, provided changes in the amino acid composition of storage proteins do not make them indigestible, and providing the resulting pool of amino acids can be used by the germinating seed enzyme systems to provide the nitrogen building block for synthesis, these changes would be tolerated. The results in Fig. 1 show, in fact, that different legume genera do have different proportions of these two proteins, and it is likely this situation could occur in different varieties of the same species.

The changing pattern of protein synthesis in the developing seed underlines the importance of analysing seeds for their amino acid composition at the same physiological development stage, if comparisons between different varieties are to be meaningful.

The other major protein fraction of legume seeds is the albumin fraction (see Boulter & Derbyshire, 1971). Although albumin proteins have not been fully separated and characterised, it is generally accepted that the various enzymes of the seed occur with this fraction. Separations of the albumins of various legume seeds on polyacrylamide gels normally give electrophoretograms with 20–30 protein staining bands (Fox, Thurman & Boulter, 1964). In the case of *Pisum* about ten of these are major bands, and in *Vicia*, about 8 (unpublished observations). This figure contrasts strongly with the probable total number of albumin proteins, which is of the order of several thousand, if we include all of the enzymic proteins of seeds. It is not clear therefore, whether each band on an electrophoretogram consists of one or many proteins. However, making the assumption that there are, in the case of *Pisum*, 10 major proteins representing some 50% of the albumin present, in our view an upper estimate, this would mean that any one albumin protein could only represent of the order of less than one percent of the total protein of the seed. Lines containing this protein with increased sulphur-containing amino acids, therefore, would not play a very significant part in increasing the overall sulphur-amino acid content of the meal. Thus, varieties containing albumin proteins with an improved amino acid profile, which is reflected in the amino acid profile of the whole meal, will not be detected very readily. It is for these reasons that we feel priority should be given to screening strategies based upon analyses of the whole meal, for example, total sulphur or to those based on theoretical considerations of the globulin fraction, such as those outlined above. However, proteins do occur in legumes with high sulphur-amino acid content, e.g. the blue protein of *Phaseolus vulgaris* (Pusztai, 1966). Further, urease in Jackbean (which has a sulphur-amino acid content similar to Smith's (1966) standard protein (Bailey & Boulter, 1971), may contribute as much as 1% of the dry weight of seed.

It is often said that methionine is usually the first limiting amino acid of legumes. However, it is important, in our view, to consider the two sulphur amino acids together, as cysteine can spare methionine. In cowpeas the methionine contents of vicilin and legumin are about the same, whereas that of cysteine in vicilin is much lower (unpublished data), and much of the supplemented methionine used in rat-feeding experiments with cowpea meals, is used to supply cysteine (Boulter, Evans, Thompson & Yarwood, 1972). Probably, in this case, screening for cysteine content may prove more useful than screening for methionine content as a first priority, although screening for both would be the best course of action.

The results obtained with cabinet-grown cowpeas shows that sampling may not be a major problem and that gene/environment interaction may not be a grossly distorting factor in screening programmes with cowpeas.

SUMMARY

The storage proteins, vicilin and legumin, of *Vigna aureus, V. mungo, Cajanus cajan, Phaseolus lunatus, P. vulgaris* and *Vicia faba*, were separated on polyacrylamide gels by electrophoresis and the relative proportions of these two proteins shown to vary in different species. Consideration of the amino acid compositions of vicilin and legumin of *Pisum* and *Vicia*, shows that the essential amino acids, cysteine, methionine and tryptophan, are higher in legumin than in vicilin, whereas isoleucine, lysine and phenylalanine, are lower. By analogy with high lysine corn, it was suggested that lines with relatively high sulphur amino acid patterns could contain an increased proportion of legumin to vicilin. The implications of this for devising screening procedures was discussed, and it was emphasized that the overall essential amino acid composition of the diet as a whole, has to be considered before a screening strategy can be decided. Work with *Vicia faba* has shown that vicilin and legumin are synthesised at different rates in the developing seed, suggesting that the proteins are under separate genetic control.

Vigna unguiculata seeds were grown in growth cabinets under standard environmental conditions. Seeds were shown to have the same percentage nitrogen and sulphur on a dry-weight basis, irrespective of the position of the pod on the plant. These results indicate the possiblity that legume seeds may have a conservative biochemistry, and that under normal conditions, an average type seed is produced. Under other conditions, some average seeds will be produced, but the plant will 'compensate' by abortion of some pods and seeds.

ZUSAMMENFASSUNG

Die Speicherproteine, Vicilin und Legumin, von *Vigna aureus, V. mungo, Cajanus cajan, Phaseolus lunatus, P. vulgaris* und *Vicia faba,* wurden auf Polyacrylamid-Gel durch Elektrophorese getrennt. Das relative Verhältnis dieser beiden Proteine variiert in den verschiedenen Arten. Bei näherer Betrachtung der Aminosäurenzusammensetzung des Vicilins und Legumins von *Pisum* und *Vicia* kann festgestellt werden, daß die essentiellen Aminosäurn Cystein, Methionin und Tryptophan im Legumin stärker vertreten sind als im Vicilin, während Isoleucin, Lysin und Phenylalanin in Vicilin stärker vertreten sind. Analog zu Mais mit hohen Lysingehalten liegt der Schluß nahe, daß Stämme mit relativ viel schwefelhaltigen Aminosäuren einen höheren Anteil von Legumin gegenüber Vicilin enthalten. Die Folgerungen im Hinblick auf Selektionsverfahren wurden diskutiert und es wurde betont, daß die essentiellen Aminosäuren der Nahrung insgesamt berücksichtigt werden müßten, bevor man sich für ein Auswahlverfahren entscheidet. Die Untersuchungen an *Vicia faba* haben gezeigt, daß Vicilin und Legumin in den sich entwickelnden Samen mit unterschiedlicher Geschwindigkeit synthetisiert werden, was auf eine getrennte genetische Kontrolle der Proteine hinweist.

Samen von *Vigna unguiculata* wurden in Klimakammern unter normalen Umweltbedingungen kultiviert. Es zeigte sich, daß die Samen, unabhängig von der Position der Hülse an der Pflanze, bezogen auf die Trockensubstanz, den gleichen Stickstoff- und Schwefelgehalt aufwiesen. Diese Ergebnisse lassen vermuten, daß die Leguminosensamen eine konservative Biochemie aufweisen und daß unter normalen Bedingungen Samen durchschnittlicher Art erzeugt werden. Unter ungünstigen Bedingungen kompensiert die Pflanze dadurch, daß sie nur einige Samen entwickelt, die dem Durchschnitt entsprechen, die anderen werden frühzeitig abgestoßen.

RÉSUMÉ

Les protéines d'accumulation, la viciline et la légumine, de *Vigna aureus, V. mungo, Cajanus cajan, Phaeseolus lunatus, P. vulgaris* et *Vicia faba*, ont été séparés sur des gèles polyacrylamides par electrophorèse; on trouva une variation dans les proportions relatives de ces deux protéines dans des éspèces différentes. Une examination de la composition des aminoacides de viciline et de légumine dans *Pisum* et *Vicia* montra que les niveaux des aminoacides essentiels, cysteine, methionine et tryptophane, sont plus élevés dans la légumine que dans la viciline, tandis que ceux de la lysine et de la phenylalanine sont plus bas. Par analogie avec le maïs à haut teneur de lysine, on proposa que des lignes avec des arrangements d'aminoacides sulfuriques relativement élevés pourraient contenir une proportion augmentée de légumine en relation à la viciline. On discuta les implications de ceci pour l'établissement de procedés de sélection et on souligna que la composition d'ensemble des aminoacides essentiles du régime total doit être considérée avant de décider la stratégie de sélection. Des études avec *Vicia faba* ont montrées que la viciline et la légumine sont syntetisées à des degrés de vitesse différents dans les graines qui se développent, ce qui suggère que les protéines subissent une commande génétique séparée.

Des graines de *Vigna unguiculata* ont été cultivé dans des cabinets de croissance dans des conditions normales du milieu. On trouva que les graines avaient le même pourcentage de soufre et d'azote (comparaison de poids sec) n'importe la position de la cosse sur la plante. Ces résultats indiquent la possibilité que les graines légumineuses ayent une biochimie conservative, et que dans des conditions normales, elles résultent dans une graine de type moyen. Dans des conditions différentes, quelques graines moyennes seront produites, mais la plante 'compensera' par l'avortement de quelques cosses et graines.

ACKNOWLEDGEMENTS

We thank Messrs. R. Swinhoe and T. Allen for technical assistance, and the Overseas Development Administration for financial support.

REFERENCES

Altschul, A. M., Yatsu, L. Y., Ory, R. L. & Engleman, E. M. (1966). Seed proteins. *Ann. Rev. Pl. Physiol.* 17: *113–136.*

Aykroyd, W. R. & Doughty, J. (1964). Legumes in human nutrition. *FAO Nutritional Studies* No. 19, *p. 73.*

Bailey, C. J. & Boulter, D. (1970). The structure of legumin, the storage protein of broad bean *(Vicia faba)* seed. *Eur. J. Biochem.* 17: *460–466.*

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

Bailey, C. J. & Boulter, D. (1972). The structure of vicilin of *Vicia faba. Phytochem.* 11: *59–64.*

Bailey, C. J., Cobb, A. & Boulter, D. (1970). A cotyledon slice system for the electron autoradiographic study of the synthesis and intracellular transport of the seed storage protein of *Vicia faba. Planta (Berl.)* 95: *103–118.*

Boulter, D. & Derbyshire, E. (1971). Taxonomic aspects of the structure of legume proteins. In: Chemotaxonomy of the Leguminosae, ed. Harborne, J. B., Boulter, D. & Turner, B. L. London & New York: Academic Press, pp. 285–308.

Boulter, D., Evans, I. M., Thompson, A. & Yarwood, A. (1973). The amino acid composition of *Vigna unguiculata* (Cowpea) meal in relationship to nutrition. *Proceedings of the PAG Symposium*, Rome.

Briarty, L. G., Coult, D. A. & Boulter, D. (1969). Protein bodies of developing seeds of *Vicia faba*. *J. exp. Bot.* 20: *358–372.*

Danielsson, C. E. (1949). Seed globulins of the Gramineae and Leguminosae. *Biochem. J.* 44: *387–400.*

Danielsson, C. E. (1952). A contribution to the study of the synthesis of the reserve proteins in ripening pea seeds. *Acta chem. scand.* 6: *149–159.*

Fox, D. J., Thurman, D. A. & Boulter, D. (1964). Studies on the proteins of seeds of the Leguminosae. I. Albumins. *Phytochem.* 3: *417–419.*

Graham, T. A. & Gunning, B. E. S. (1970). Localization of legumin and vicilin in bean cotyledon cells using fluorescent antibodies. *Nature (Lond.)* 228: *81–82.*

Osborne, T. B. (1924). The Vegetable Proteins. New York: Longmans-Green.

Pusztai, A. (1966). The isolation of two proteins, Glycoprotein I and a trypsin inhibitor, from the seeds of Kidney Bean *(Phaseolus vulgaris). Biochem. J.* 101: *379–384.*

Smith, M. H. (1966). The amino acid composition of proteins. *J. theoret. Biol.* 13: *261–282.*

Wright, D. J. & Boulter, D. (1972). The characterisation of vicilin during seed development in *Vicia faba* (L.). *Planta (Berl.)* 105: *60–65.*

DISCUSSION

Professor Persson (Ås/Norwegen)

I like to see your paper from a plant breeding point of view. From the varietal difference in *Vicia faba*, can it be suggested that great improvement can be expected in terms of food value?

Secondly as you have pointed out that the synthetic systems of legumin and vicilin are different, is there any information as to the genetic nature of genetic control of the synthesis of globulins, as this will be of interest for plant breeders.

Professor Boulter (Durham/G.B.)

The question as to whether there already exists in the world collections of different legumes, sufficient variation for exploitation via breeding programmes, to give crops with greatly improved food values, has not been fully investigated. However, present indications are that this is likely to be so and this includes our own experiences with *Vicia faba*.

With regard to the second question, at the moment we do not know the genetic nature of the control systems for the synthesis of vicilin and legumin.

McSheehy (Haughley/G.B.)

You have shown very clearly that the production of legumin lags behind that of vicilin. In view of that could there be gene environment interaction. In other words has

one a situation where the production of legumin cannot commence until the intracellular environment contains a predetermined level of vicilin? Have we a situation where a 'repressor gene' controlling the operon for legumin synthesis is 'switched-off' by vicilin? If this is possible then there may be problems associated with trying to alter the final relative concentrations of the two proteins.

Professor Boulter (Durham)

All that can be said at this stage, is that the syntheses of vicilin and legumin are switched on at different times. The particular mechanism of the control is not known. If the mechanism is as suggested by Mr. McSheehy, then an alteration in the final relative concentration of these proteins would have to be effected by their rates of synthesis, which are different, rather than changing the times at which their synthesis is switched on. However, there are very many other possible control mechanisms.

Professor Reinnann-Philipp (Ahrensburg/BRD)

Lassen sich bereits Aussagen über die Situation in Süßlupinen – besonders *L. albus* – bezüglich des Verhältnisses von Vicilin: Legumin machen?

Würden in dieser Hinsicht Süßlupinen mehr dem *Vicia faba*-Typ oder mehr dem *Phaseolus vulgaris*-Typ ähneln?

Professor Boulter (Durham)

I am afraid I am unable to answer your question, apart from saying that lupin, including the species *Lupinus albus*, contains vicilin and legumin-type proteins, but we have not investigated their proportions or times of synthesis.

Professor Schüphan (Geisenheim)

You pointed out that the contents of cystine in *Vicia faba* have been found tenfold in legumin in comparison with vicilin.

As far as I have taken from your table this relation is reverse in methionine. Am I right?

Professor Boulter (Durham)

No. The values for the sulphur amino acids are given in Table 1 of my paper, and, in fact, legumin contains about twice as much methionine as does vicilin.

Authors' address:

Dept. of Botany
University of Durham
Durham
Great Britain

Phytochemistry, 1976, Vol 15, pp. 411-414 Pergamon Press. Printed in England.

ISOLATION OF LEGUMIN-LIKE PROTEIN FROM *PHASEOLUS AUREUS* AND *PHASEOLUS VULGARIS*

E. DERBYSHIRE and D. BOULTER

Department of Botany, University of Durham, Durham City, England

(*Revised received* 6 *October* 1975)

Key Word Index—*Phaseolus aureus*; *P. vulgaris*; Leguminosae; seed storage protein; seed globulins; protein bodies; legumin.

Abstract—An 11S seed globulin has been isolated from *Phaseolus aureus* and *P. vulgaris* by zonal isoelectric precipitation and the MWs of the constituent subunits determined. The protein of *P. vulgaris* occurs in the protein body fraction and its chemical composition, including the *N*-terminal amino acids and amino acid composition has been determined. The similarity between the 11S globulin of the two *Phaseolus* spp. and legumin from other legumes is discussed.

INTRODUCTION

Legumin, an 11S globulin first isolated as an impure fraction from seeds of *Pisum sativum* [1,2] has been purified from *Vicia faba* seeds where it represents approximately 75% of the storage protein [3–5]. Legumin-like globulins are also major seed storage proteins of other legume species for example *Vicia sativa* [6] and *Glycine max* [7–9] and they probably occur widely among flowering plants [10]. Globulins which sediment as 11S components have been identified in several species of *Phaseolus* and were equated with legumin by Danielsson solely on the basis of their sedimentation coefficients [2]. The only species of *Phaseolus* from which an 11S globulin has been isolated previously is *P. aureus* [11] and the subunit composition reported for this globulin differs from that of legumin-like globulins.

We have now isolated a legumin-like globulin from *Phaseolus aureus* and *P. vulgaris* and have characterized that of the latter species.

RESULTS

The crude globulin preparations from *Phaseolus vulgaris* and *P. aureus* were each separated into two protein fractions by zonal isoelectric precipitation [5,12]. One of these fractions consisted of protein which was soluble at pH 4·7 and eluted from the column in a volume of 50 ml immediately after the void volume (130 ml), while the other consisted of protein insoluble at pH 4·7 and which was substantially retarded (V_e approximately equal to 2·5 V_0) and eluted in a volume of 90 ml. The retarded fraction from each species accounted for approximately 15% of the protein recovered (250/300 mg) and was homogeneous in the ultracentrifuge (s_{20} approx. 11S) except for traces (approx. 5%) of a larger component (s_{20} approx. 16S). By comparison the unretarded protein was heterogeneous at pH 7 in the ultracentrifuge and for both species the Schlieren diagrams showed a large 7S peak which was not completely resolved from a small 2S-3S peak: a 6S shoulder was also present in the 7S peak of *P. vulgaris*. When the unretarded fraction of *P. vulgaris* was examined in the ultracentrifuge at pH 6·2 and $I = 0{\cdot}15$ more than 50% of the 7S component observed at pH 7 was replaced by a peak which corresponded to 18S protein and under these conditions the sedimentation coefficient of the 6S component of *P. vulgaris* was 6·1S i.e. it had a calculated MW of approximately 130000. The unretarded protein of *P. aureus* did not associate to 18S at pH 6·2 and $I = 0{\cdot}15$. Agglutinating activity was detected only in the unretarded fractions: agglutination by that of *P. vulgaris* was more rapid than by that of *P. aureus*.

The protein of the retarded fraction from each species migrated during disc electrophoresis as a major band with mobility 0·15 relative to that of the bromophenol blue marker band and two slower and one faster minor bands. The proteins were dissociated by treatment with sodium dodecyl sulphate (SDS) and 2-mercaptoethanol (2Me) into major subunits with apparent MWs 62000 and 20000; only trace amounts of stain were retained by these subunits when half gels were stained for glycoprotein by the PAS procedure.

Further characterization was then carried out on the protein recovered from the retarded fraction of both species. The protein was first precipitated by the addition of $(NH_4)_2SO_4$ to 70% saturation and then redissolved in a small volume of 0·01 M 2-Me. This method of recovery was used rather than precipitation by dialysis since protein precipitated by dialysis against deionized water did not redissolve in 0·4 M NaCl–0·01 2 Me–0·05 M NaH_2PO_4, pH7. The recovered protein of the retarded fractions sedimented in the ultracentrifuge with $s_{20w} = 11{\cdot}6$S (*P. vulgaris*) and 11·4S (*P. aureus*) at pH 7·2 and $I = 0{\cdot}4$ except for a minor component with

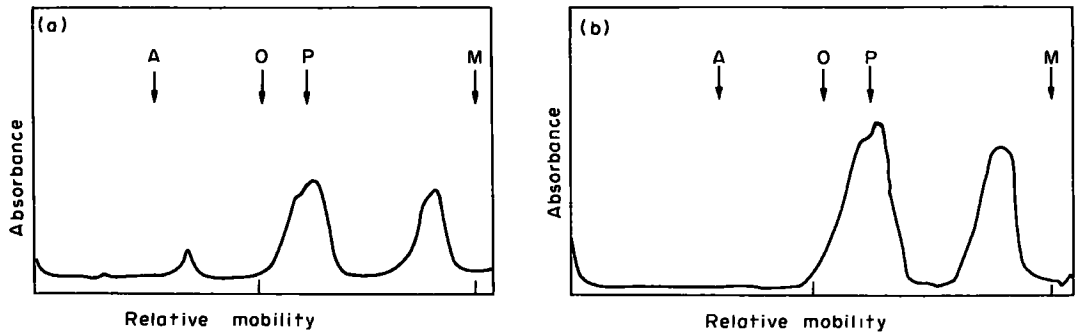


Fig. 1. SDS polyacrylamide gel electrophoresis of reduced dissociated 11S globulins from (a) *Phaseolus aureus*. and (b) *Phaseolus vulgaris*. Proteins (80 μg) were run on 10% gels and stained in Amido Black. The mobilities of serum albumin (A), ovalbumin (O) and pepsin (P) relative to that of myoglobin (M) are shown for comparison.

s_{20w} = 16S. The approximate MWs of the 11S components are 340000 and 330000 respectively. The gel patterns obtained by disc electrophoresis of the native proteins and continuous electrophoresis of the subunits of protein dissociated with SDS and 2 Me were the same as those obtained directly with the retarded fractions before the precipitation and recovery step. However, SDS-2Me treatment preceded by incubation of the proteins in 6 M urea (1 hr at 45°) dissociated them both to subunits with MWs 37000, 34000 and 20000 (Fig. 1).

The nitrogen and sulphur contents of the 11S globulin of *P. vulgaris* were 17·8 (w/w) and 0·44% (w/w) respectively, and 0·5% (w/w) carbohydrate was present (Table 1). Glycine, leucine, threonine and methionine were determined as *N*-terminal amino acids, and the amino acid composition of the protein is presented in Table 1.

Table 1. The chemical composition of 11S globulin of *Phaseolus vulgaris*

Amino acid composition (g/100 g protein)*	
Asp	8·6
Thr	3·9
Ser	5·0
Glu	12·8
Pro	3.8
Gly	3·6
Ala	3·8
Val	5·5
Met	1.4
Ile	4.5
Leu	7·7
Tyr	3·7
Phe	5·2
Lys	7·9
His	3.2
Arg	6.0
½ Cys	0·5 ± 0·05†
Trp	1·0 ± 0·1†
Kjeldahl N	17·8 ± ·2% (w/w)†
Total S	0·41 ± ·1% (w/w)†
Carbohydrate	0·5 ± ·07% (w/w)†
N-terminal amino acids	glycine, leucine, threonine, methionine

* Best values of means of duplicate samples hydrolysed for 20 and 70 hr, except values for Thr and Ser which were obtained by extrapolation to zero time, ½ Cys (which was determined as cysteic acid after 20 h hydrolysis only), and Trp which was determined colorimetrically [41]. † Means and errors of duplicate samples.

Approximately 10% of the protein of the fraction (G1A) which was precipitated from an ascorbate extract of *P. vulgaris* sedimented in the ultracentrifuge as an 11S component at pH 7·2 and I = 0·4. although it was not completely resolved from the major 7S peak of this fraction. Its probable identity with the legumin-like globulin referred to above is indicated by the presence of subunits with MWs of 62000 and 20000 as minor components when the G1A protein was dissociated by treatment with SDS and 2Me: other minor bands. MWs 55000 and 33000, were present in the gels and the MWs of major components were 50000, 47000 and 24000. The G1A fraction agglutinated red blood cells.

An 11S molecular component was detected also when the protein of the protein body fraction of *P. vulgaris* was examined in the ultracentrifuge at a protein concentration of 10 mg/ml at pH 7·2 and I = 0·4; it was a minor component only and was not completely resolved from the major 7S peak, which had a 6S shoulder. Subunits with MW of 62000 and 20000 were minor components of protein body protein dissociated without the use of urea, and subunit with MW 50000, 47000, 34000, 31000 and 24000 were the major components of this fraction.

DISCUSSION

Zonal isoelectric precipitation has been employed previously to isolate pure legumin from protein precipitated by ammonium sulphate precipitation from initial extracts of *Vicia faba* [5] and *V. sativa* [12,13]. However we have found it necessary, when isolating the 11S globulins from *Phaseolus* spp. by this procedure to first remove small molecules from the initial extracts by molecular sieving and also to inhibit oxidative processes. The 11S globulins isolated from *Phaseolus aureus* and *P. vulgaris* are insoluble at pH 4·7 at I = 0·3 and they each have sedimentation coefficients, MWs and subunit sizes which are similar to those of legumin of *Vicia faba* [5] and *V. sativa* [13,14]. The 11S globulin of *P. vulgaris* and legumin of *V. faba* are similar in amino acid composition, *N*-terminal amino acids and in their low content of carbohydrate.

The *N*-terminal amino acids of legumin of members of the Viciae are glycine. leucine and threonine [3.5,13,15,16] and those of *Glycine max* are glycine, leucine and phenylalanine [8]: unlike the situation in *Phaseolus vulgaris* therefore, *N*-terminal methionine

has not been reported in these legumins. Several explanations for the apparent finding of *N*-terminal methionine can be proposed. It may be there as a true *N*-terminal of the protein or it could be a contaminant if the protein were not completely purified. A further possibility arises since the synthesis of polypeptide chains in the cytoplasm of eukaryotes is initiated by *N*-terminal methionyl residues which are subsequently cleaved from the completed chains [17,18]. Thus the *N*-terminal methionyl residue detected may be the initiating methionyl residue.

The presence of contaminating material could explain the low levels of carbohydrate found in several 11S globulin preparations [10]; however it could also indicate that legumin contains a glycoprotein subunit. The carbohydrate content of subunits isolated from 11S globulins of legumes have not been reported but a 2·7S subunit of the 12S globulin of *Brassica napus* contains 4–5% carbohydrate [19].

The three banded pattern of subunits with MWs 37000, 34000 and 20000 found in SDS gels of dissociated 11S globulin from *Phaseolus aureus* and *P. vulgaris* is also found in dissociated legumin from *Vicia faba* [5] and *V. sativa* [14]. The subunit with MW 62000 seen in some of the *Phaseolus* preparations would appear to be an association product and a component of similar MW is also seen in dissociated legumin preparations from *V. faba* and *V. sativa* under special conditions. In dissociated 11S globulin from *Glycine max* only two sizes of subunits, with MWs 37000 and 22000 have been separated so far [9]. Dissociated 11S globulin from *Lupinus angustifolius* has subunits with MWs 63000, 53000, 40000 and 20000 which differ somewhat from the above [20]. However the complete dissociation and separation of the 11S globulin subunits are very dependent on the conditions employed and a critical comparison can be made between subunit compositions of various globulins only when the latter have been dissociated under identical conditions and their subunits examined by a variety of techniques.

The high nitrogen and dicarboxylic acid/amide contents of the 11S globulin of *Phaseolus vulgaris* relative to those of a standard protein [21], its occurrence in the protein body fraction and its similarity to legumin suggest strongly that this protein is, like other legumins, a storage protein. Whereas the 11S globulins of *Vicia faba* and *Glycine max* contribute substantially to the nutritional value of these beans, the major globulin of *Phaseolus vulgaris* is Glycoprotein II [22,23] and legumin represents a much smaller proportion of the protein of these beans. However, the composition of the storage proteins is different in different cultivars of the same species [24–27] and other cultivars of *Phaseolus aureus* and *P. vulgaris* may have higher legumin contents and be of greater nutritional value than those used in the present investigation. Procedures based on those employed here may play a part in screening programmes for the selection of lines with improved nutritional status.

The unretarded fractions will not be discussed in detail here but the presence in *Phaseolus aureus* of high concentrations of 7S protein which does not associate to an 18S molecular species suggests, contrary to the suggestion of Ericson and Chrispeels [11], that this protein is not identical to the 7S Glycoprotein II of *P. vulgaris* [22]. Agglutinins with sedimentation coefficients of approximately 6S have also been isolated from *P. vulgaris* [28.29] and it is probable that these were the proteins of the unretarded fraction which sedimented as 6.1S species.

The GI fraction of ascorbate extracts prepared from *Phaseolus vulgaris* by Sun *et al.* [30.31] is regarded by them as Glycoprotein II and an 11S protein was not detected, although they had reported previously that the sedimentation coefficient of GI (32) was 11S. The subunits identified in dissociated GI protein were also major components of a dissociated GI fraction prepared by isoelectric precipitation [31] and they equate Glycoprotein II procedurally with legumin [31]. By contrast the GIA fraction which we isolated was heterogeneous and contained 11S protein, 7S protein (Glycoprotein II), and possibly a globulin agglutinin. The 11S globulins which have been isolated from various species as (relatively) pure proteins are very similar to legumin of *Vicia faba* when characterized by modern biochemical techniques [10]. They have similar sedimentation coefficients, MWs, subunit sizes, amino acid composition, N-terminal amino acids, low contents of carbohydrate and low solubilities at pH 4·7 at $I = 0{\cdot}3$, and we regard them all as legumin by analogy with other well characterized proteins, for example the haemoglobins, although only when their amino acid sequences have been determined will it be clear whether they are functionally similar proteins or whether they are a group of proteins related by a common ancestry.

EXPERIMENTAL

Isolation of 11S globulins. Seed meals of *Phaseolus aureus* (green gram) and *P. vulgaris* cv. "Streamline", were dispersed with 10 g insoluble polyvinyl-pyrrolidone in 0·4 M NaCl–0·01 M 2 Me-0·1 M NaH_2PO_4 pH 7 (10 g meal/100 ml), and were blended for 3 min in a high-speed blender which was cooled in ice, and then stirred for 15 min. Each extract was filtered through 2 layers of muslin and clarified by centrifugation, then 60 ml clarified extract, containing approximately 1 g protein was applied to a column of Sephadex G-25 (6·6 cm × 20 cm), which had been equilibrated with 0·2 M NaCl-0·035 M NaH_2PO_4 pH 7–0·01 M 2 Me. The column was developed with the latter buffer. The chromatographic fractions which contained protein excluded from the Sephadex G-25 were bulked and adjusted to 70% saturation with respect to $(NH_4)_2SO_4$ and the protein which precipitated was quickly collected by centrifugation and redispersed in 10 ml 0·2 M NaCl–0·05 M NaH_2PO_4, pH 8–0·01 M 2 Me (pH 8 buffer soln). Ammonium sulphate was removed from the protein soln by passing it through a column (3·4 × 20 cm) of Sephadex G-25, which had been equilibrated with the pH 8 buffer solution. After five-fold concentration of the protein solution by use of an Amicon ultrafiltration cell 8 ml, containing approximately 300 mg protein, was applied to a column (3·4 × 35 cm) of Sephadex G-50 which had been equilibrated with 0·2 M NaCl–0·05 M citric acid pH 4·7–0·01 M 2 Me and the column was then developed with the pH 8 buffer solution (zonal iso-electric precipitation (12)) at a flow rate of 25 ml/hr. Buffer solns which were employed had been deaerated and flushed with N_2 before the addition of 2 Me and were stored under N. The Sephadex used for chromatography had been deaerated before equilibration. Column eluates were monitored continuously at 280 nm by use of an LKB Uvicord II analyser.

Ascorbate extract. An ascorbate extract of meal of *Phaseolus vulgaris* was prepared by the method of McLeester *et al.* [32] and clarified by centrifugation. The clarified extract was diluted with two volumes of dionized water at 4° and the protein which precipitated was collected by centrifugation and designated the GIA fraction.

Protein bodies. These were isolated by centrifugation through glycerol, a method suggested by M. J. Chrispeels.

Seed meal (500 mg) of *Phaseolus vulgaris* was suspended in 5 ml 80% glycerol in 0·05 M NaH_2PO_4 pH 8 and centrifuged at 3000 *g* for 20 min and then at 23000 *g* for 20 min. The supernatant was layered over 10 ml 90% glycerol in the same buffer and centrifuged at 23000 *g* for 20 min. The pellet was resuspended in 5 ml 80% glycerol in buffer, large aggregates of material were removed by low speed centrifugation and the protein bodies were pelleted at 23000 *g*.

Analytical centrifugation. This was carried out at 40000 rpm and 20° in an Omega II 70000 ultracentrifuge. Proteins were centrifuged at concentrations of 5 mg/ml and 1 mg/ml and sedimentation coefficients, s_{20} and s_{20w} values were determined as described by Svedberg and Pedersen [33]. The Schlieren peaks obtained at the lower protein concentration were too small for accurate measurement and values quoted in the text were obtained at concentrations of 5 mg/ml, except where stated otherwise. Approximate MW's were calculated by the method of Halsall [34].

Dissociation of protein. Equal volumes of protein soln (2 mg/ml) and 2·5% SDS–0·1 M 2 Me–0·05 M Pi buffer pH 7 were mixed and heated for 3 min in a boiling water bath and then cooled in running tap water. Myoglobin was employed as a marker protein and after heating in SDS–2Me–Pi buffer pH 7, it was added to a portion of each dissociated protein sample to a final concentration of 0·7 mg/ml.

Polyacrylamide-gel electrophoresis. Undissociated proteins were examined in 7·5% gels by the procedure of Ornstein and Davis [35] but without the use of spacer gel. Dissociated proteins, 20 μl and 80 μl vols, were loaded under buffer onto 10% gels of 60 mm length in a continuous 0·1 M Pi, pH 7–10 mM sodium thioglycolate system (11,31): 15 mA/gel was maintained for 3 hr. Gels were stained for protein in 1% (w/v) Amido Black in 7% HOAc and destained by diffusion in 7% HOAc. Certains gels, on which the subunits of 11S globulin had been separated, were split longitudinally and one half of each gel was stained for protein while the other was stained for glycoprotein by the periodic acid–Schiffs base procedure [37]. The electrophoretic mobilities of the subunits relative to that of the marker protein were determined after staining and the apparent MW's of the subunits were estimated by reference to a calibration curve prepared using cytochrome *c*, myoglogin, pepsin, ovalbumin and serum albumin.

Agglutinin activity. Protein fractions were adjusted to pH 7 and a protein concentration of 1 mg/ml and then mixed with an equal vol of a 2% suspension of human group O red blood cells in 1% NaCl at RT.

Chemical composition. Total nitrogen was determined by an automated micro-Kjeldahl method [38], and total sulphur was estimated by wet digestion with perchloric-nitric acid mixture, and subsequent turbidimetric analysis of $BaSO_4$ [39]. Carbohydrate was determined by the method of Dubois *et al.* [40]. The amino acid composition of acid hydrolysed (20 and 70 hr) protein was analysed on a Locarte automatic loading amino acid analyser. Tryptophan was determined colorimetrically [41] and cysteine plus cystine were determined by performic acid oxidation [42] and subsequent analysis of cysteic acid on the Locarte analyser. *N*-Terminal amino acids were determined by the procedure of Gros and Labouesse [43]; after hydrolysis each sample was loaded on one side of a polyamide sheet, a mixture of dansylated amino acids was loaded onto the reverse side and the *N*-terminal amino acids of the sample were identified after chromatographic separation by reference to the pattern obtained from the amino acid mixture.

Acknowledgements—We wish to thank R. Swinhoe for determination of nitrogen and sulphur, and the Agricultural Research Council for financial support.

REFERENCES

1. Osborne, T. B., (1924) *The Vegetable Proteins*, 2nd edn., Longmans, Green, London.
2. Danielsson, C. E. (1949) *Biochem. J.* **44**, 387.
3. Bailey, C. J. and Boulter, D. (1970) *European J. Biochem.* **17**, 460.
4. Wright, D. J. and Boulter, D. (1972) *Planta (Berlin)* **105**, 60.
5. Wright, D. J. and Boulter, D. (1974) *Biochem. J.* **141**, 143.
6. Vaintraub, I. A. and Shutov, A. D. (1966) *Biokhimiya* **31**, 726.
7. Wolf, W. J. (1972) *Food Technol.* **26**, 44.
8. Catsimpoolas, N., Rogers, D. A., Circle, S. J. and Meyer, E. W. (1967) *Cereal Chem.* **44**, 631.
9. Catsimpoolas, N., Kenney, J. A., Meyer, E. W. and Szuhaj, B. F. (1971) *J. Sci. Fd. Agric.* **22**, 448.
10. Derbyshire, E., Wright, D. J. and Boulter, D. (1976) *Phytochemistry* **15**, 3.
11. Ericson, M. C. and Chrispeels, M. J. (1973) *Plant Physiol.* **52**, 98.
12. Shutov, A. D. and Vaintraub, I. A. (1965) *Ukr. Biokhim. Zh.* **37**, 177.
13. Vaintraub, I. A. and Nguyen Thanh Thien (1968) *Dokl. Akad. Nauk SSSR* **180**, 1239.
14. Vaintraub, I. A. and Nguyen Thanh Thien (1971) *Mol. Biol.* **5**, 59.
15. Grant, D. R. and Lawrence, J. M. (1964) *Arch.. Biochem. Biophys.* **108**, 552.
16. Jackson, P., Boulter, D. and Thurman, D. A. (1969) *New Phytol.* **68**, 25.
17. Boulter, D., Ellis, R. J. and Yarwood, A. (1972) *Biol. Rev.* **47**, 113.
18. Zalik, S. and Jones, B. L. (1973) *Ann. Rev. Plant Physiol.* **24**, 47.
19. Goding, L. A., Bhatty, R. S. and Finlayson, A. J. (1970) *Can. J. Biochem.* **48**, 1096.
20. Blagrove, R. J. and Gillespie, J. M. (1975) *Aust. J. Plant Physiol.* **2**, 13.
21. Smith, M. H. (1966) *J. Theoret. Biol.* **13**, 261.
22. Pusztai, A. and Watt, W. A. (1970) *Biochim. Biophys. Acta* **207**, 413.
23. Racusen, D. and Foote, M. (1971) *Can. J. Botany* **49**, 2107.
24. Tombs, M. P. (1964) *Nature, London* **200**, 1321.
25. Adriaanse, A., Klop, W. and Robbens, J. E. (1969)) *J. Sci. Fd Agric.* **20**, 647.
26. Gillespie, J. M. and Blagrove, R. J. (1975) *Aust. J. Plant Physiol.* **2**, 29.
27. Davis, D. R. (1973) *Nature (New Biol.)* **245**, 30.
28. Pusztai, A. and Watt, W. A. (1974) *Biochim. Biophys. Acta* **365**, 57.
29. Barker, R. D. J., Derbyshire, E. and Boulter, D. (1976) *Phytochemistry* **15**, (in press).
30. Sun, S. M., McLeester, R. C., Bliss, F. B. and Hall, T. C. (1974) *J. Biol. Chem.* **249**, 2118.
31. Sun, S. M. and Hall, T. C. (1975) *J. Agric. Food Chem.* **23**, 184.
32. McLeester, R. C., Hall, T. C., Sun, S. M. and Bliss, F. A. (1973) *Phytochemistry* **12**, 85.
33. Svedberg, T. and Pedersen, K. O. (1940) *The Ultracentrifuge*, Oxford University Press, London.
34. Halsall, H. B. (1967) *Nature, London* **215**, 880.
35. Ornstein, L. and Davis, B. J. (1961) *Disc Electrophoresis*, Preprint by Distillation Products Inductries (Eastman Kodak Co.), Rochester, New York.
36. Weber, K., Pringle, J. R. and Osborn, M. (1972) *Methods Enzymol.* **26c**, 3.
37. Zacharius, R. M., Zell, T. E., Morrison, J. H. and Woodlock, J. J. (1969) *Anal. Biochem.* **30**, 148.
38. Varley, J. A. (1966) *Analyst* **91**, 119.
39. Mottershead, B. E. (1971) *Laboratory Practice* **20**, 483.
40. Dubois, M., Gillies, K. A., Hamilton, J. K., Rebers, P. A. and Smith, F. (1956) *Anal. Chem.* **28**, 350.
41. Bencze, W. L. and Schmidt, K. (1957) *Anal. Chem.* **29**, 1193.
42. Moore, S. (1963) *J. Biol. Chem.* **238**, 235.
43. Gros, C. and Labouesse, B. (1969) *European J. Biochem.* **7**, 463.

Phytochemistry, 1976, Vol. 15, pp. 751 757. Pergamon Press. Printed in England.

PURIFICATION AND CHARACTERIZATION OF THE MAJOR STORAGE PROTEINS OF *PHASEOLUS VULGARIS* SEEDS, AND THEIR INTRACELLULAR AND COTYLEDONARY DISTRIBUTION

RICHARD D. J. BARKER*, ERIC DERBYSHIRE, ALAN YARWOOD and DONALD BOULTER
Department of Botany, University of Durham, Durham DH1 3LE, England

(*Received* 3 *November* 1975)

Key Word Index—*Phaseolus vulgaris*; Leguminosae; amino acid composition; protein bodies; storage proteins; *N*-terminus.

Abstract—Several extraction and fractionation procedures have been employed to isolate the major storage proteins of mature seeds of *Phaseolus vulgaris* cv. "Seafarer"; three proteins which were soluble at pH 4.7, and one that was insoluble at that pH were identified. The characteristic subunits of the three pH 4.7 soluble proteins had MW's 50000 and 47000, 32000, and 23000 respectively; those of the pH 4.7 insoluble fractions had MW 60000 and 20000. Amino acid compositions, *N*-terminal amino acid residues and the presence of carbohydrate in these proteins have been determined. All these proteins occurred in the protein body fraction and their relative amounts were different in the outer and central parts of the cotyledons.

INTRODUCTION

Protein laid down during seed development and utilized as a nitrogen and carbon source during germination is termed storage protein; in legumes 80% of the seed protein may be storage protein. Pusztai has purified and characterized several seed proteins from *Phaseolus vulgaris* [1–4] and Racusen and Foote [5] and Hall and his associates [6–8] have also investigated the seed storage proteins of this species; an 11S storage protein from *P. vulgaris* has recently been described [9]. However, when interpreting previously described results, problems arise since the variety of methods of isolation and purification used make comparison difficult. Also, the relationship of a purified fraction to the total protein complement is often unclear, and the analytical methods used, e.g. ultracentrifugation, often do not unequivocally identify the protein.

In this investigation of the storage proteins of *Phaseolus vulgaris*, sodium dodecyl sulphate (SDS)-polyacrylamide gel electrophoresis has been used to identify the component subunits and to assess the purity of the protein fractions obtained. The intracellular localization of the storage proteins and their distribution within the cotyledons has also been determined.

RESULTS

Extraction and fractionation of storage proteins

Alkaline salt extracts. The protein pattern obtained by disc electrophoresis of an alkaline salt extract of cultivar "Seafarer" showed several components which were not always completely resolved. There was a diffuse major band, R_m 0.35–0.4, and in addition there were prominent components, R_m 0.19, 0.1, and several other minor components. A densitometric trace of the protein band pattern obtained by SDS-gel electrophoresis of this extract is shown in Fig. 1. There are two main subunits, MW 50000 and 47000, but four other prominent components, MW 60000, 32000, 23000 and 20000 are also present. SDS extraction and electrophoretic analysis of meal gave a similar electrophoretic profile. SDS extractions of meal of a different cultivar, "Canadian Wonder", gave a similar subunit pattern, except that in this cultivar the equivalent subunits to the 50000 and 47000 MW components were of lower MW, 49000 and 45000 respectively, and an additional major band of MW 53000 was present.

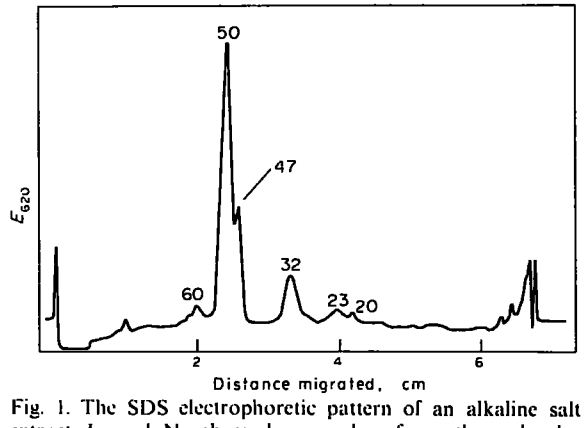


Fig. 1. The SDS electrophoretic pattern of an alkaline salt extract. *Legend.* Numbers above peaks refer to the molecular weights ($\times 10^{-3}$) of the polypeptide chains. A similar band pattern and molecular weights were obtained when a cetyltrimethylammonium-bromide (CTAB)-containing system at pH 5.7 was used [40].

*Present address: Department of Biochemistry, The University, Sheffield 10.

The precipitate, obtained from alkaline salt extracts by addition of ammonium sulphate to 70% saturation, was enriched in the 60000, 32000, 23000 and 20000 MW subunits. Two fractions were obtained by zonal isoelectric precipitation of the redissolved precipitate: one eluted immediately after the void volume, in citrate buffer pH 4.7, and was pH 4.7 soluble; the second fraction which accounted for approximately 20% of the protein applied to the column, was insoluble at pH 4.7, and was retarded during chromatography. When these fractions were analysed by disc electrophoresis, the pH 4.7 soluble fraction migrated as a major, poorly defined band, R_m 0.35–0.4, with a subsidiary band R_m 0.19; the pH 4.7 insoluble fraction, which was poorly soluble in the electrophoresis sample buffer, migrated as a single band R_m 0.1. SDS gel electrophoresis showed that the major components of the initial extract, i.e. the MW 50000 and 47000 components, together with the lower MW subunits of 32000 and 23000 MW were confined to the pH 4.7 soluble fraction, and the pH 4.7 insoluble fraction consisted almost entirely of subunits of MW about 80000, 60000 and 20000. This partial purification of components was confirmed when the main band cut from the gel after disc electrophoresis of an alkaline extract was shown, in the SDS gel system, to be mainly composed of 50000 and 47000 MW subunits, with small amounts of the 32000 and 23000 MW subunits. It is of interest to note that when the subsidiary band (R_m 0.19), found on disc gel electrophoresis of the pH 4.7 soluble fraction was similarly cut out and analysed, it also consisted of the 50000, 47000 and 32000 MW subunits. In the absence of 2-mercaptoethanol, the pH 4.7 insoluble fraction gave a single band of MW about 80000 on SDS gel electrophoresis; treatment with 1% (v/v) 2-mercaptoethanol resulted in loss of this component and the concomitant appearance of subunits of MW 60000 and 20000, but a small amount of the 80000 MW band always remained. The pH 4.7 soluble fraction gave the same band pattern in the presence or absence of 2-mercaptoethanol. Carboxymethylation caused no change in the band pattern of any fraction.

When alkaline salt extracts were chromatographed on DEAE-cellulose, the protein was separated into two fractions. Unadsorbed protein was eluted as a small peak with the starting buffer and when analysed by disc electrophoresis it gave a diffuse band R_m 0.27–0.39, together with a band migrating with the bromo-phenol blue marker, R_m 1.0; SDS gel electrophoresis showed this fraction to be mainly composed of the 32000 MW subunit. Protein adsorbed to the column was eluted as a large peak with 0.17 M NaCl; while the leading edge of the elution profile was very sharp, a shoulder occurred on the trailing edge. Disc electrophoresis of the leading edge of this peak resulted in an intense band R_m 0.37, with a weak band R_m 0.21; the protein of the trailing edge contained a slow-moving band R_m 0.1, in addition to these components. The leading edge of the main peak was shown by SDS gel electrophoresis to contain the subunits of 50000 and 47000 MW, together with a trace of the 23000 MW subunit. The middle of the peak contained subunits of MW 80000, 60000, 50000, 47000 and 20000, i.e. the 23000 MW subunit was absent from this region of the peak. The trailing edge contained all the above subunits with enrichment of a 23000 MW subunit. The 32000 MW subunit was not observed in this peak.

Rechromatography of fractions from the leading edge of the main peak failed to remove completely the trace amount of the 23000 MW subunit present. When samples of the trailing edge of the peak were rechromatographed, fractions enriched in the 60000 and 20000 subunits, relative to the 50000 and 47000 MW subunits, were obtained, but the latter subunits were still the major components on SDS gels. Alteration of the pH at which chromatography was carried out, the dimensions of the column, the amount of protein applied, the flow rate, and the steepness of the sodium chloride elution gradient, all failed to give increased resolution of the proteins of the main peak.

When samples of alkaline salt extracts were centrifuged in sucrose density gradients a single broad peak was obtained, although some protein remained at the top of the gradient. SDS gel electrophoresis of fractions from the gradient showed that (a) the pH 4.7 insoluble fraction subunits, MW 60000 and 20000, were confined to the lower edge of the peak; (b) as judged by SDS-gel electrophoresis, most of the 50000, 47000 and 32000 MW subunits were found in the broad peak, although some protein composed of these subunits remained at the top of the gradient; (c) while the protein which remained at the top of the gradient was enriched in subunits of MW 23000, traces of subunits of this size were also seen on the SDS gels of fractions from the main peak.

Water extracts. When water extracts of meal were cooled a substantial cryoprecipitate formed. When analysed by disc electrophoresis this was found to be composed of a main component which migrated as a wide band, R_m 0.35–0.4, together with a subsidiary band R_m 0.2. The supernatant fluid after removal of the cryoprecipitate contained approximately equal amounts of components with R_ms 0.37 and 0.1.

Subunit analysis by SDS gel electrophoresis showed that the cryoprotein consisted of the 50000 and 47000 MW subunits, with traces of the 32000 and 23000 MW subunits; when the supernatant fluid after removal of this protein was analysed, it contained an increased proportion of the 80000, 60000, 32000, 23000 and 20000 MW subunits and about 30% of the total amount of the 50000 and 47000 MW subunits. When the cryoprotein was analysed on disc gels and the main band (R_m 0.35–0.4) cut from the gel and the protein analysed in the SDS system, it was shown to consist of the 50000 and 47000 MW subunits, with a very small amount (*ca* 3% of the total protein) of the 23000 MW subunit; the 32000 MW subunit was not observed.

When the cryoprecipitate from a water extract was subjected to gel isoelectric-focusing in 7 M urea over the pH range 3–6, at least eighteen bands were obtained in the pH region 5.0–5.75. However, similarly complex patterns were also obtained from the pH 4.7 soluble fraction on isoelectric focusing in 7 M urea gels.

Acidic extracts. Two acidic extraction methods were used (see Experimental) and both gave the same results when the extracts were examined by gel electrophoresis. When acidic extracts were subjected to disc electrophoresis, the patterns obtained were similar to those obtained with alkaline salt extracts, but owing to the poorer resolution obtained as compared to SDS gels, detailed comparison was not profitable. SDS gel electrophoretic analysis showed that the acidic extracts were also similar in subunit content to alkaline salt extracts.

Table 1. The amino acid compositions and *N*-terminal amino acid residues of protein fractions. All results are expressed as mol%. The amino acid composition of Glycoprotein II [3], recalculated to mol%, is also included

Amino acid	pH 4.7 insoluble fraction	Non-adsorbed ion-exchange fraction	Leading edge of main ion exchange peak	Glycoprotein II [3]
Asp	11.9	14.8	14.9	13.4
Thr	4.1	8.2	3.8	4.2
Ser	6.6	8.3	7.1	9.6
Glu	18.3	7.3	16.2	14.6
Pro	6.5	4.6	3.6	3.7
Gly	4.7	8.0	6.1	5.8
Ala	5.5	7.2	5.2	5.2
½-Cys	N.D.[a]	0.8	0.4	0.2
Val	6.9	7.2	5.9	6.5
Met	0.8	0.3	0.8	0.7
Ile	4.1	5.5	5.4	6.2
Leu	8.2	10.6	9.5	10.0
Tyr	2.6	2.7	2.6	2.7
Phe	2.9	5.5	5.4	5.6
Lys	8.3	5.3	6.4	5.4
His	3.7	1.4	2.5	2.4
Arg	4.9	3.0	4.8	4.0
N-terminal amino acids	Gly.Thr.	N.D.*	Thr. Ser. Leu trace Glu	—

* Not determined.

On fractionation of the acidic ascorbate extract, subunits of 80000, 60000 and 20000 MW were confined to the F-I fraction (see Experimental for terminology of fractions) but those of MW 50000 and 47000 were found in all three fractions. The 23000 MW subunit was enriched in the F-I fraction, while the 32000 MW subunit was confined to the F-II fractions.

When acidic extracts were prepared in the absence of ascorbate, the F-I fraction contained the same subunits as the F-I fraction from the ascorbate-containing extraction, but there was markedly less of the 50000 and 47000 MW subunits, i.e. the 80000, 60000 and 20000 subunits were enriched. Whilst the presence or absence of ascorbate had less effect on the composition of the F-II fraction, nevertheless, in the absence of ascorbate the F-IIa fraction consisted almost entirely of the 50000 and 47000 MW subunits, and the F-IIb fraction of the 32000 MW subunit. Using either extractant, the F-II fractions agglutinated human group O erythrocytes.

In contrast to the situation with acidic extracts, protein was not precipitated by two-fold dilution of alkaline salt extracts. When the latter were dialysed overnight against running tap water, a fraction with subunits similar to those of the F-I fraction from the ascorbate-containing extract was precipitated.

Characterisation of protein fractions

Amino acid analysis. The amino acid compositions of the pH 4.7 insoluble fraction (60000 and 20000 MW subunits), the non-absorbed fraction from DEAE-cellulose ion-exchange chromatography (32000 MW subunit), and the leading edge of the main peak eluted during ion-exchange chromatography (50000 and 47000 MW subunits), are presented in Table 1, together with the *N*-terminal amino acid residues found for each fraction. The amino acid composition of Glycoprotein II (3), recalculated to mole%, is included in Table 1 for comparison.

Carbohydrate content. When SDS gels of the pH 4.7 soluble and insoluble fractions were stained for the presence of carbohydrate, only the 50000, 36000 and 32000 MW subunits took up stain. The 32000 MW subunit stained relatively strongly; the 36000 MW subunit, which was an insignificant component on the gels when stained for protein, showed up as a prominent component on periodic acid-Schiff (PAS) staining. When gels of SDS extracts of *P. vulgaris* cv. "Canadian Wonder", were similarly stained, the 53000, 49000 and 32000 MW subunits took up stain.

Protein composition of protein bodies

SDS gel electrophoresis showed that all the main storage protein subunits (see Fig. 1) were found in the protein body fraction. Protein bodies sedimented during sucrose density gradient centrifugation as an opaque band centred at approximately 80% (w/v) sucrose. When fractionation of the gradient was monitored at 280 nm, the protein body peak was usually symmetrical with respect to 280 nm absorption. Occasionally shoulders of apparently more or less dense particles were observed, but no differences in subunit pattern were detected when these regions were analysed. Only about 10% of the protein applied to the gradients was recovered in the protein body fraction. All the subunits present in the protein body fraction were also found at the top of the gradients.

Isolation of protein bodies with glycerol at pH 5.0 resulted in a protein body fraction which accounted for approximately 40% of the total protein extracted from meal. The protein subunit composition of these protein body preparations resembled that of those isolated by sucrose density gradient centrifugation but in contrast to the sucrose method none of the pH 4.7 insoluble com-

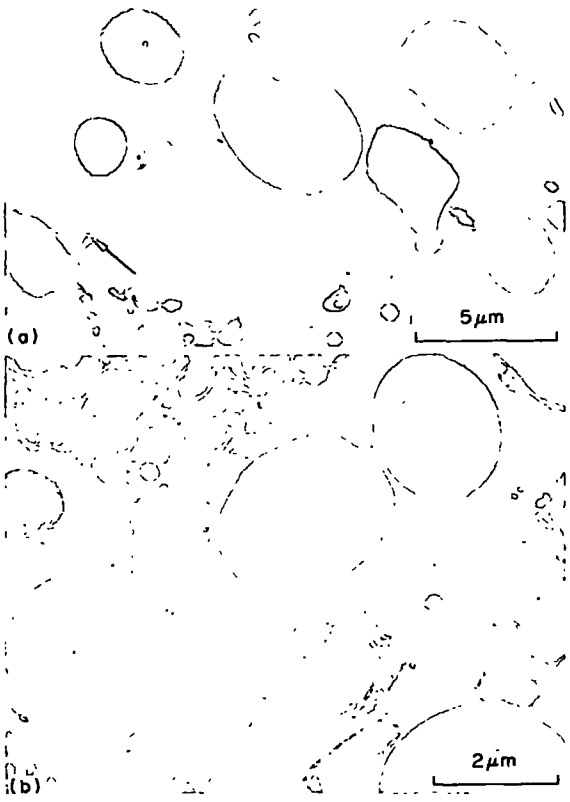


Fig. 2. Electron micrographs of protein body preparations. *Legend*. (a) prepared by sucrose density gradient centrifugation: the arrow indicates one of the protuberances noted in the text. The bar in the lower right-hand corner represents 5 μm. (b) prepared by the glycerol method: the bar represents 2 μm.

ponents were found in the supernatant. When the glycerol procedure was carried out at pH 7.5, some of the pH 4.7 insoluble fraction protein subunits were found in the supernatant, but at this pH the yield of protein bodies was lower (about 25% of the protein present in the pH 5.0 protein body fraction) and the protein body fraction contained an even higher proportion of the pH 4.7 insoluble fraction subunits relative to the total.

Figure 2a is an electron micrograph of protein bodies prepared by sucrose gradient centrifugation. The protein bodies are membrane-bound, although the membranes of several have been ruptured and leakage of their contents is apparent. In some protein bodies, distinct and often sharp protuberances (arrowed) are found.

Protein bodies prepared by the glycerol method at pH 5.0 were more intact (Fig. 2b), and their membranes were more clearly defined. However, small fragments of cell wall were present and protein bodies were frequently present in clusters, surrounded by cytoplasmic material.

Protein distribution within the cotyledon

Salt extracts from abaxial sections of cotyledons were richer in protein than those from the centre and trypsin inhibitor activity was also greater in the abaxial sections. Not only is there relatively more protein in the abaxial portion of the cotyledon, but the proportion of the subunits there differs from that of the central region. Thus, the abaxial region is enriched relative to the central portions in 60000 and 20000 MW subunits and impoverished in 32000 and 23000 MW subunits. By comparison with the whole seed the abaxial region has less 47000 MW subunit relative to the 50000 MW subunit. In contrast, the adaxial surface contained relatively large amounts of 32000 and 23000 MW subunits when compared with other parts of the seed.

The distribution of protein subunits established by the direct analysis of SDS extracts of different regions of the cotyledon was confirmed when protein bodies were isolated, using the glycerol procedure at pH 5.0, and their constituent protein was analysed by SDS gel electrophoresis.

DISCUSSION

The major storage protein of *P. vulgaris* cv. "Seafarer" accounts for about 60% of the total seed protein and is composed of two subunits with MW 50000 and 47000, in the approximate ratio 2·3:1. This protein is pH 4.7 soluble and, as described by Derbyshire and Boulter [9], sediments in the ultracentrifuge at pH 7.0 as a 7S molecular species which partially associates to an 18S species at pH 6.2. Pusztai and Watt (3) have isolated a glycoprotein, Glycoprotein II, from *P. vulgaris* and Racusen and Foote [5] have equated this protein with the major salt soluble protein of *P. vulgaris* on the evidence of disc electrophoresis and amino acid composition. Although the subunit composition of Glycoprotein II has not been reported, the similarity between the major purified protein of this investigation and Glycoprotein II in sedimentation coefficient, association behaviour and amino acid compositions clearly indicate that they are the same protein. Furthermore, although the mobility in 5% acrylamide gels obtained by us (not given here) was slightly greater than that reported by Racusen and Foote [5], there was a general similarity in the appearance of the band on disc electrophoresis in the two investigations.

This major protein behaved as a cryoprotein and precipitated when a water extract was cooled to 0·4°. A vicilin (i.e. a 7S, pH 4.7 soluble protein) from *Pisum sativum*, which was cryoprecipitated from a water extract of the meal, has also been described [10]. In contrast, a cryoprotein obtained from water extracts of *Glycine max* [11] has been shown to be composed mainly of an 11S protein (legumin-like) [12]. Cryoprecipitation has been used as an initial step in the purification of this protein [13] but the effectiveness of water as an extractant for seed proteins varies from legume to legume [14].

A second pH 4.7 soluble fraction, corresponding to the 6S fraction observed in the ultracentrifuge by Derbyshire and Boulter [9], was composed of glycoprotein subunits with MW approximately 32000. The amino acid composition of this fraction is in reasonable agreement with that of the lectin (subunit MW 30000–32000) isolated by Andrews [15] from navy bean; a notable common feature is a low glutamic acid content as compared to the other storage proteins of this species.

Recently, Pusztai and Watt [4] have separated and partially characterized a number of agglutinins from *P. vulgaris*, which differ in their isoelectric points but are also characterized by low glutamic acid contents. It would appear, therefore, that the protein fraction described in this investigation is related to the preparations

of Pusztai and Watt, but their precise relationship is not clear.

The pH 4.7 insoluble fraction obtained here by zonal isoelectric precipitation is similar to that described in greater detail by Derbyshire and Boulter [9], i.e. it is a legumin-like protein. It is the only one of the storage proteins described here that appears to contain inter-subunit disulphide bonds.

The presence of at least a fourth protein is suggested by the occurrence of subunits of MW 23000 which cannot be assigned to any of the three previously described proteins. Subunits of this size were found in two different fractions during ion-exchange chromatography and sucrose density gradient centrifugation. One fraction eluted and sedimented with Glycoprotein II, while the other sedimented more slowly and was eluted at a higher salt concentration during ion-exchange chromatography.

The use of an acidic extractant either with or without ascorbate to prevent oxidation, has been advocated by Hall and his associates [6–8]. These same procedures were used here but our interpretation of the results is different to that of Hall and his colleagues. We consider Hall's G-I fraction to be equivalent to our FI and FIIa fractions (prepared in the absence of ascorbate) combined, i.e. to an impure preparation of Glycoprotein II. Hall's G-II fraction, containing a single major subunit of MW 30000 appears similar to our F-IIb which we equate with the agglutinins described from *Phaseolus* by Pusztai and Watt [4]. The use of an acid extraction method with *Vicia faba* led to the formation of lower MW polypeptides [16] and the method cannot be recommended for general use with legumes.

Isoelectric focusing in urea gels of several globulin preparations has led to the demonstration of more protein components than expected from the results obtained with other methods [17–19] and this is also true for the data on Glycoprotein II presented here. It is not possible to say whether some components are artefacts caused by carbamylation [20] and/or deamidation [21] or if, in fact, all the components are native subunits, the complexity of the pattern resulting from microheterogeneity, possibly in the carbohydrate moiety [22].

The major storage proteins of *P. vulgaris* described above have been shown to be located in the protein body preparations by the presence of their subunits and the protein bodies were morphologically similar to those previously described *in situ* by Opik [23]. The glycerol method for protein body preparation is rapid and gives a higher yield than from sucrose density gradient centrifugation, but glycerol preparations appeared to be less pure.

Differences in the amount of protein in different parts of legume and cereal seeds have been reported [24, 25]. Vogel and Wood [26] demonstrated that the outer regions of the cotyledons of *Phaseolus vulgaris* were richer in protein-bound sulphydryl groups than the inner regions; Wood and Cole [27] subsequently showed that individual disc electrophoretic components were found in different ratios in the outer and central regions of the cotyledons. We have confirmed the gradient of protein and trypsin inhibitor activity described by Zimmermann *et al.* [25] and have shown differences in the distribution of storage protein subunits in different parts of the cotyledon. Although the sulphur content of different regions has not been analysed, the presence of greater amounts of legumin subunits in the abaxial region of the cotyledon is compatible with a higher protein bound sulphydryl concentration there.

A striking result was the difference in the ratio of the 50000 and 47000 MW subunits observed between the abaxial surface and the central regions of the cotyledon. The fact that this difference was also seen when the protein subunits of protein bodies derived from these regions were analysed by SDS gel electrophoresis suggests that the difference is not simply an extraction artefact. Further work will be necessary to determine whether this difference is upheld in purified preparations of this protein isolated from different regions of the cotyledons.

EXPERIMENTAL

Preparation of seed meals. Seeds or cotyledons of *Phaseolus vulgaris* cvs. "Seafarer" and "Canadian Wonder" were finely ground in a Janke and Kunkel water-cooled mill.

Extraction of proteins. (a) *Alkaline salt extracts.* 6 g of meal was stirred with 60 ml 0.5 M NaCl, 0.05 M NaPi buffer pH 7.5 for 30 min at 4°. Slurry was clarified by centrifugation at 38000 *g* for 90 min at 4°.

(b) *Water extracts.* 6 g meal was stirred with 60 ml distilled water for 30 min at room temp (18–20°). After clarification at 38000 *g* for 90 min (at 20°) the supernatant fluid was cooled in an ice bath for 90 min. The ppt. which formed during cooling was collected by centrifugation at 23000 *g* for 30 min at 4°.

(c) *Acidic extraction and fractionation.* All procedures were at 4°. Extraction and fractionation were carried out as described by McLeester *et al.* [6] using 0.5 M NaCl, 0.25 M ascorbic acid (measured pH 2.4); a similar procedure was followed in the absence of ascorbate, using 0.5 M NaCl, adjusted to pH 2.0 with dil HCl, as the extractant. The solvent to meal ratio was 10:1 (v/w). The fraction precipitated by addition of 2 vols of distilled water to the initial supernatant was collected by centrifugation, and is referred to as the F-I fraction. Overnight dialysis of the supernatant, after removal of the F-I fraction, against running H_2O resulted in the formation of a ppt: this was collected by centrifugation and is referred to as the F-IIa fraction. A further ppt. was obtained by dialysis of the remaining supernatant fluid against several changes of dist. H_2O: this was also collected and is referred to as the F-IIb fraction.

(d) *SDS extraction.* 20 mg samples of meal were stirred with 4 ml of 0.01 M NaPi buffer pH 7.0, 1.0% (w/v) SDS and 1.0% (v/v) 2-mercaptoethanol for 60 min at 37°; they were then heated to 100°C for 3 min and allowed to cool at room temp; extracts were clarified by centrifugation and analysed by SDS gel electrophoresis.

Preparation of protein bodies. (a) *By sucrose density gradient centrifugation.* All operations were carried out at 0–4°. Seeds were soaked overnight in running H_2O and their testas and embryo axes removed. 3 g cotyledons were finely sliced and extracted with 10 ml of 35% (w/v) sucrose in 0.025 M NaPi, 0.025 M citric acid, pH 7.5, by gentle grinding for 1 min in a chilled mortar. Extracts were stirred for 10 min and then centrifuged at 165 *g* for 5 min. Approximately 1.5–2.0 ml of the supernatant fluid was carefully layered onto linear 50–90% (w/v) sucrose gradients in 0.025 M NaPi, 0.025 M citric acid, pH 7.5, prepared in 23 ml polycarbonate tubes. Gradients were centrifuged for 3 hr in a 3 × 23 aluminium swing-out rotor operating at 4° and 30000 rpm (R av. 94000 *g*). Gradients were fractionated by inserting a thin capillary tube through the gradient to the bottom of the tube and pumping the contents out. The gradient was monitored at 280 nm. Sucrose concentration was measured using a Bellingham and Stanley sugar refractometer. Samples were taken for electron microscopy, SDS gel electrophoresis and protein estimation.

(b) *By centrifugation in glycerol solutions.* The method used was communicated by Chrispeels [28]. 3 g meal was gently blended with 30 ml 80% glycerol in 0.05 M NaPi, 0.05 M citric acid, pH 5.0, containing 0.1% 2-mercaptoethanol at room

temp. The suspension was centrifuged for 10 min at 2000 *g*. When the supernatant was layered onto a cushion of 8 ml 90% (v/v) glycerol in water in a 50 ml centrifuge tube and centrifuged at 38000 *g* for 60 min at 15°, protein bodies were obtained as a pellet. Samples of some pellets were taken for electron microscopy and for SDS-gel electrophoresis; pellets were resuspended in 0.5 M NaCl, 0.05 M NaPi buffer, pH 7.5, for protein estimation. The procedure was also carried out in the glycerol Pi–citrate buffer adjusted to pH 7.5 with NaOH.

Column chromatography. (a) *Zonal isoelectric precipitation.* Supernatant fluid from an alkaline salt extract was adjusted to 70% saturation with respect to ammonium sulphate; precipitated protein was collected by centrifugation, dispersed in 0.2 M NaCl, 0.05 M NaPi, pH 8.0, made 0.1% in 2-mercaptoethanol, and dialysed against this buffer. It was then subjected to zonal isoelectric precipitation (29,30) using buffer of the same composition for elution.

(b) *DEAE-cellulose ion-exchange chromatography.* Ion-exchange chromatography was carried out on columns of Whatman DE 52 cellulose equilibrated in 0.025 M NaPi buffer, pH 7.2. Samples for chromatography were dialysed against this buffer prior to application. Elution was commenced with buffer of the same composition, subsequently with a linear NaCl gradient (routinely 0–0.5 M) in the same buffer.

Sucrose density gradient centrifugation of proteins. Samples of clarified alkaline salt extracts were dialysed against excess 0.5 M NaCl, 0.05 M NaPi pH 7.6. Approximately 1.5 ml was then carefully layered on to linear 5–20% (w/v) sucrose gradients in the same buffer and the gradients were centrifuged for 18 hr at 4° and 30000 rpm (R av. 94000 *g*) in a 3 × 23 ml swing-out rotor. Gradients were fractionated as described for the preparation of protein bodies by sucrose density gradient centrifugation. Fractions of 1.0 ml were collected.

Estimation of protein. Protein concentration was determined by the method of Lowry *et al.* [31] using bovine serum albumin as standard. Crude globulin preparations, prepared by lyophilisation of the precipitate formed by overnight dialysis against running tap water of an alkaline salt extract of *Phaseolus vulgaris*, gave a similar calibration curve up to a protein concentration of 200 μg/ml.

Polyacrylamide gel electrophoresis. Undissociated proteins were examined in 7.0% (w/v) acrylamide gels by the disc electrophoretic procedure of Ornstein and Davis [32] except that the spacer gel was omitted. The method of Weber and Osborn [33] was used to determine the apparent MWs of the subunits of protein dissociated with SDS (1%, w/v) and 2-mercaptoethanol (0.2%, v/v). MWs were routinely determined on 7.0% and 10.0% (w/v) acrylamide gels. To obtain maximum dissociation of subunits (i.e. the lowest MW's) it was found necessary to incubate samples in a boiling water bath for 3 min. Gels were stained for protein in 0.2% (w/v) Amido Black in 7.0% (v/v) HOAc and destained by diffusion in 7.0% (v/v) HOAc; densitometric traces were obtained by scanning the gels in transmission at 620 nm using a Joyce–Loebl Chromoscan. Gels were stained for glycoprotein by the PAS method of Zacharius *et al.* [34]. The location of bands after disc electrophoresis is described by their mobility relative to that of the bromophenol blue marker used (relative mobility, R_m). Molecular weights were determined in the SDS system by comparison of the mobility of subunits with that of standard proteins of known MW's.

S-Carboxymethylation. This was carried out by the method of Crestfield *et al.* [35].

Direct analysis of bands in disc gels. After electrophoresis of undissociated proteins in disc gels, the region corresponding to a particular band was cut from several gels in which the same sample had been electrophoresed; the pooled gel sections were ground and extracted with 0.025 M NaPi buffer pH 7.0 containing 2% (w/v) SDS and 2% (v/v) 2-mercaptoethanol.

Gel isoelectric focusing in 7 M urea. The method used was that of Wrigley [36].

Amino acid analysis. Duplicate samples of proteins in 6 N-HCl were hydrolysed *in vacuo* at 105° for 24 and 72 hr; their amino acid compositions were determined on a Locarte automatic-loading amino acid analyser. The cysteine-cystine content was determined as cysteic acid following performic acid oxidation [37].

N-terminal amino acid analysis. *N*-terminal amino acids were determined by the dansyl-method of Gros and Labouesse [38].

Electron microscopy. Samples of protein bodies were fixed in 2.5% (v/v) glutaraldehyde, 1% (w/v) osmic acid, 0.05 M sodium cacodylate buffer pH 7.0 in 80% sucrose (for samples from sucrose gradients) or in 80% glycerol (for samples from the glycerol procedure) for 3 hr at 4°. They were dehydrated in an alcohol series and embedded in Spurr's resin. Thin sections were post-stained with uranyl acetate and alkaline lead citrate and examined in an A.E.I. EM 66 electron microscope.

Agglutination test. Protein samples were prepared in 0.9% (w/v) sodium chloride pH 7.0, and mixed with an equal vol of 2% (v/v) suspension of Human group O erythrocytes in the same solution, at room temp.

Trypsin inhibitor activity. This was determined by the effect of alkaline salt extracts on the trypsin-catalysed hydrolysis of benzoyl-arginine-*p*-nitroanilide (BAPA); the second method described by Erlanger *et al.* [39] was used except that the assay was carried out in 0.05 M NaPi, 0.05 M citric acid buffer pH 7.5, at 35°. Assays were carried out such that the total inhibition did not exceed 60% of the control trypsin activity.

Acknowledgements—We wish to thank Dr. N. Harris for the electron microscopic analysis of protein bodies and the preparation of micrographs; Mr. R. Swinhoe for the performic acid oxidation of proteins, and the Ministry of Overseas Development for financial support.

REFERENCES

1. Pusztai, A. (1966) *Biochem. J.* **101**, 379.
2. Pusztai, A. and Duncan, I. (1971) *Biochim. Biophys. Acta* **229**, 785.
3. Pusztai, A. and Watt, W. B. (1970) *Biochim. Biophys. Acta* **207**, 413.
4. Pusztai, A. and Watt, W. B. (1974) *Biochim. Biophys. Acta* **365**, 57.
5. Racusen, D. and Foote, M. (1971) *Can. J. Botany* **49**, 2107.
6. McLeester, R. C., Hall, T. C., Sun, S. M. and Bliss, F. A. (1973) *Phytochemistry* **12**, 85.
7. Sun, S. M., McLeester, R. C., Bliss, F. A. and Hall, T. C. (1974) *J. Biol. Chem.* **249**, 2118.
8. Sun, S. M. and Hall, T. C. (1975) *J. Agr. Food Chem.* **23**, 184.
9. Derbyshire, E. and Boulter, D. (1976) *Phytochemistry* **15**, 411.
10. Buzila, L. (1967) *Rev. Roum. Biochim.* **4**, 103.
11. Briggs, D. R. and Mann, R. L. (1950) *Cereal Chem.* **27**, 243.
12. Briggs, D. R. and Wolf, W. J. (1957) *Arch. Biochem. Biophys.* **72**, 127.
13. Eldridge, A. C. and Wolf, W. J. (1967) *Cereal Chem.* **44**, 645.
14. Smith, C. R. Jr., Earle, F. R., Wolff, I. A. and Jones, Q. (1959) *Agr. Food Chem.* **7**, 133.
15. Andrews, A. T. (1974) *Biochem. J.* **139**, 421.
16. Wright, D. J. and Boulter, D. (1973) *Phytochemistry* **12**, 79.
17. Catsimpoolas, N. (1969) *FEBS Letters* **4**, 259.
18. Catsimpoolas, N. and Wang, J. (1971) *Analyt. Biochem.* **44**, 436.
19. Wright, D. J. (1973) Ph.D. Thesis, University of Durham.
20. Stark, G. R., Stein, W. H. and Moore, S. (1960) *J. Biol. Chem.* **235**, 3177.
21. McKerrow, J. H. and Robinson, A. B. (1971) *Analyt. Biochem.* **42**, 565.

22. Schmid, K. (1968) In: *Biochemistry of Glycoproteins and Related Substances* (Rossi, E., Stoll, E., eds.), p. 4. Karger, New York.
23. Opik, H. (1968) *J. Exp. Botany* **19**, 64.
24. Normand, F. L., Hogan, J. T. and Deobald, H. J. (1965) *Cereal Chem.* **42**, 359.
25. Zimmermann, G., Weissmann, S. and Yannai, S. (1967) *J. Food Sci.* **32**, 129.
26. Vogel, K. P. and Wood, D. R. (1971) *Crop Sci.* **11**, 249.
27. Wood, D. R. and Cole, C. V. (1973) In: *Nutritional Improvement of Food Legumes by Breeding*, p. 325. Protein Advisory Group of the United Nations, New York.
28. Chrispeels, M. J. (1974) Personal communication.
29. Shutov, A. D. and Vaintraub, I. A. (1965) *Ukr. Biokhim. Zh.* **37**, 177.
30. Wright, D. J. and Boulter, D. (1974) *Biochem. J.* **141**, 413.
31. Lowry, O. H., Rosebrough, N. J., Farr, A. L. and Randall, R. J. (1951) *J. Biol. Chem.* **193**, 265.
32. Ornstein, L. and Davis, B. J. (1961) *Disc Electrophoresis*. Preprint by Distillation Products Industries (Eastman Kodak Co.), Rochester, New York.
33. Weber, K. and Osborn, M. (1969) *J. Biol. Chem.* **244**, 4406.
34. Zacharius, R. M., Zell, T. E., Morrison, J. H. and Woodlock, J. J. (1969) *Analyt. Biochem.* **30**, 148.
35. Crestfield, A. M., Moore, S. and Stein, W. H. (1963) *J. Biol. Chem.* **238**, 622.
36. Wrigley, C. W. (1971) In: *Methods in Enzymol.* (Jakoby, W. B., ed.) Vol. XXII, p. 559. Academic Press, New York.
37. Moore, S. (1963) *J. Biol. Chem.* **238**, 235.
38. Gros, C. and Labouesse, B. (1969) *European J. Biochem.* **7**, 463.
39. Erlanger, B. F., Kokowsky, N. and Cohen, W. (1961) *Arch. Biochem. Biophys.* **95**, 271.
40. Barker, R. D. J. (1975) Ph.D. Thesis, University of Durham.

Journal of Experimental Botany, Vol. 29, No. 109, pp. 309–323, April 1978

The Isolation and Characterization of the Major Polypeptides of the Seed Globulin of Cowpea (*Vigna unguiculata* L. Walp) and their Sequential Synthesis in Developing Seeds

J. F. CARASCO,[1] R. CROY, E. DERBYSHIRE, AND D. BOULTER

Department of Botany, University of Durham, Durham DH1 3LE, England

Received 14 June 1977

ABSTRACT

Globulin protein formed the major protein fraction of mature seeds of cowpea; it was found to be heterogeneous when examined by using chromatography and zonal isoelectric precipitation. Both 7S and 11S globulin were present and the fraction was dissociated by SDS treatment into three major subunits with apparent molecular weights 56 000, 54 000, and 52 000 as determined in SDS-acrylamide gels. The individual major subunits were purified by using ion-exchange chromatography in 8 M urea and further characterized by chemical analysis. Two of the major subunits had low but different contents of S-containing amino acid residues and were probably subunits of 7S glycoproteins.

The pattern of seed development was investigated and four phases were identified. The protein profile and the amino acid composition of the seeds changed during development; most of the seed globulin was synthesized in the third and fourth phases and the net synthesis of the three major subunits of the globulin differed. The profile of the amino acid composition of the pod walls did not change during the period of synthesis of the seed globulin fraction. The essential amino acid content of the meal is determined by that of the globulin fraction except for cysteine/cystine which is mostly supplied by the albumin fraction, the latter probably containing some proteins rich in S-containing amino acids.

INTRODUCTION

The cowpea is cultivated in many tropical and subtropical regions and, in West Africa, it is the most important pulse crop (Rachie, 1973). It is a main source of dietary protein in Nigeria (Oyenuga, 1967) in spite of the fact that cowpeas are nutritionally poor in S-containing amino acids (Boulter, Evans, Thompson, and Yarwood, 1973) and that the overall seed yields are low. An extensive breeding programme directed towards the improvement of cowpea has been established at the International Institute of Tropical Agriculture, Ibadan.

Work with other legumes, for example *Vicia faba*, *Pisum sativum* (Wright and Boulter, 1972; Boulter, Evans, and Derbyshire, 1973), and soya bean (*Glycine*

[1] Present address: Department of Botany, University of Dar Es Salaam, Tanzania.

Abbreviations: SDS, sodium dodecyl sulphate; ME, 2-mercaptoethanol; P_i = inorganic phosphate.

max) (Wolf, 1970), has shown that the different proteins of seeds are of differing relative importance nutritionally and that globulin protein is usually the most important protein fraction. The aims of the present work were to characterize cowpea globulin with respect to the number of globulin components present and their amino acid composition and to follow the synthesis of individual polypeptides during seed development. Harris and Boulter, (1976) have recently published a developmental study of the fine structure of cowpea cotyledon cells.

MATERIALS AND METHODS

Materials

Mature air-dried seeds of *Vigna unguiculata* cv. Prima were supplied by I.I.T.A. Ibadan, Nigeria and were grown under glass in pots of John Innes No. 1 compost inoculated after 10 d with a *Rhizobium* strain specific for cowpea. The first flowers appeared 8 weeks after planting; usually 8–14 seeds were formed in each pod and seeds underdeveloped by comparison with the other seeds in a pod were discarded.

Methods

Meals, prepared by milling mature seeds in a water-cooled mill, were sieved through a 40 mesh sieve and defatted with hexane before use.

Protein extraction. Defatted meal was dispersed in 0·4 M NaCl–0·1 M P_i buffer, pH 8·0, (2 g meal/50 ml) and blended for 2 min in a high speed blender, which was cooled in ice. The brei was stirred for 15 min, filtered through two layers of muslin, and clarified by centrifugation. Protein fractions were then prepared by prolonged dialysis of extracts against 0·033 M sodium acetate buffer, pH 5·0. The protein which precipitated was designated 'globulin' and the pH 5·0-soluble protein was designated albumin. Alkali-soluble protein was extracted from defatted meal by blending with 1 N NaOH at a meal weight: solvent volume ratio of 1 : 5.

Chromatography. Ion-exchange chromatography in 8 M urea was performed with Whatman DE52-cellulose in a 1·25 cm × 25 cm column, in 0·076 M Tris-citrate buffer, pH 8·7, containing urea at 8 M ME at 0·01 M. Chromatography was carried out at room temperature using a 0–0·35 M NaCl elution gradient. Zonal isoelectric precipitation was performed using the procedure described by Derbyshire and Boulter (1976).

Ultracentrifugation. Protein solutions were examined at 40 000 rev. min^{-1} and 20 °C in a Christ Omega II 70 000 ultracentrifuge. Sedimentation coefficients and $s_{20,w}$ values were determined as described by Svedberg and Pedersen (1940).

Dissociation of protein. Equal volumes of protein solution (2 mg ml^{-1}) and a solution containing 2·5% (w/v) SDS, 0·1 M ME, and 0·05 M P_i buffer, pH 7, were mixed and heated for 3 min in a boiling water bath and then cooled in running tap water. Myoglobin was employed as a marker protein and, after heating in SDS-ME-P_i buffer (pH 7) solution, it was added to a portion of each dissociated protein sample to give a final concentration of 0·7 mg ml^{-1}.

Polyacrylamide gel electrophoresis. The method of Weber and Osborn (1969) was employed to determine the subunit composition of dissociated proteins. The apparent molecular weights of subunits were determined on 10% acrylamide-SDS gels by reference to the relative mobilities of serum albumin, ovalbumin, γ-globulin, pepsin, myoglobin, lysozyme, and cytochrome *c*. Gels were stained for protein in either 1% (w/v) Amido Black in 7% (v/v) acetic acid or 0·1% (w/v) Coomassie brilliant blue-R250 in methanol-acetic acid-water (90 : 7 : 43, by vol.) and for glycoprotein by the procedure of Zacharius, Zell, Morrison, and Woodcock (1969).

Serological methods. Polyvalent antiserum against the total proteins in a buffered salt extract from mature seeds was obtained from rabbits, which had received a series of 20–24 intravenous injections over a period of 7–8 weeks. Five days after the final injections the animals were killed and their blood collected. Immunoelectrophoresis using the Shandon micro-immunoelectrophoresis apparatus, was performed essentially as described by Dudman and Millerd (1975). 0·75% (w/v) Agarose in Tris-EDTA-borate buffer, pH 9·0 (0·25 M Tris, 0·8 mM EDTA, 0·375 M borate, adjusted to pH 9·0 with 6 M HCl) was spread on precoated microscope slides to give a gel 1–2 mm thick, and allowed to set at 4 °C. Seed protein solutions containing bromophenol blue were placed in precut wells and electrophoresed at 200 V at room temperature until the marker dye had

migrated 30–40 mm from the origin. Antiserum (0·2 ml) was then placed in precut troughs between the samples and allowed to diffuse for 2 d at 4 °C, in a humidified chamber. After washing in buffered saline, precipitin arcs were stained with 0·1% (w/v) Coomassie blue in methanol-acetic acid-water and excess stain was removed by washing in methanol-acetic acid-water.

Chemical composition. Dry weights were determined by drying at 105 °C to constant weight. Total N was determined by an automated micro-Kjeldahl method (Varley, 1966) and total S was determined by wet digestion with perchloric-nitric acid mixture and subsequent turbidimetric analysis of $BaSO_4$ (Mottershead, 1971). Protein was determined by the method of Lowry, Rosebrough, Farr, and Randall (1951) and the amino acid compositions of proteins were analysed, after hydrolysis in 6 N HCl *in vacuo* at 105 °C for 22 h and 72 h, on a Locarte automatic loading amino acid analyser; cysteine plus cystine were determined as cysteic acid after performic acid oxidation. Tryptophan was determined colorimetrically (Bencze and Schmidt, 1957). The free amino acid contents of ethanol extracts of developing seeds and pod walls were determined by a ninhydrin method, using L-leucine as a standard (Moore and Stein, 1948, 1954). N-terminal amino acids were determined by the procedure of Gros and Labouesse (1969). Neutral sugars were determined by the method of Dubois, Giles, Hamilton, Rebers, and Smith, (1956), and total hexosamines were determined by use of the Morgan–Elson reaction (Morgan and Elson, 1934) as recommended by Good and Bessman (1964) after prior hydrolysis in 4 N HCl for 6 h (Spiro, 1966); Glucosamine-HCl was employed as a standard.

RESULTS

Seed storage protein

The N content of defatted air-dried meals of mature cowpea seeds was 4·33% of which 72% was extracted by use of buffered salt solutions. Approximately 15% of this extracted N diffused through the membrane during prolonged dialysis of the extract against the salt buffer and 85% of the N which remained in the sac precipitated during dialysis against distilled water, i.e. was globulin nitrogen (Osborne, 1924).

The protein in the extract after dialysis against salt buffer sedimented in the ultracentrifuge to give a broad Schlieren peak with a small shoulder on its trailing edge. This protein fraction was composed mainly of globulin, a large proportion of which was 7S protein (the rate of displacement of the maximum of the Schlieren peak corresponded to $s_{20,w} = 7{\cdot}2$).

The major subunits of the undialysed buffered salt extract (Fig. 1) had molecular weights of 56 000, 54 000, and 52 000, the 56 000 subunit being less prominent than the other two; carbohydrate was detected in the three subunits by the periodic acid–Schiff's reagent (PAS) procedure.

Protein with subunits of molecular weights 56 000, 40 000, and 20 000 was retarded when proteins extracted in salt buffer were examined by zonal isoelectric precipitation at pH 4·7. This protein fraction sedimented in the ultracentrifuge as a mixture of 11S and 7S proteins.

The individual major subunits of the globulin fraction were purified by elution from DEAE-cellulose in the presence of 8 M urea (Fig. 2). In addition to three pure fractions several minor fractions which contained mixtures of other subunits were eluted: the molecular weights of the subunits and the S contents of these minor fractions are tabulated in Table 1. The chemical compositions of the purified major subunits are tabulated in Table 2 and confirm the presence of carbohydrate in the subunits with molecular weights of 54 000 and 52 000. Table 2 also shows the amino acid compositions of the total cowpea globulin, the albumin (water-soluble)

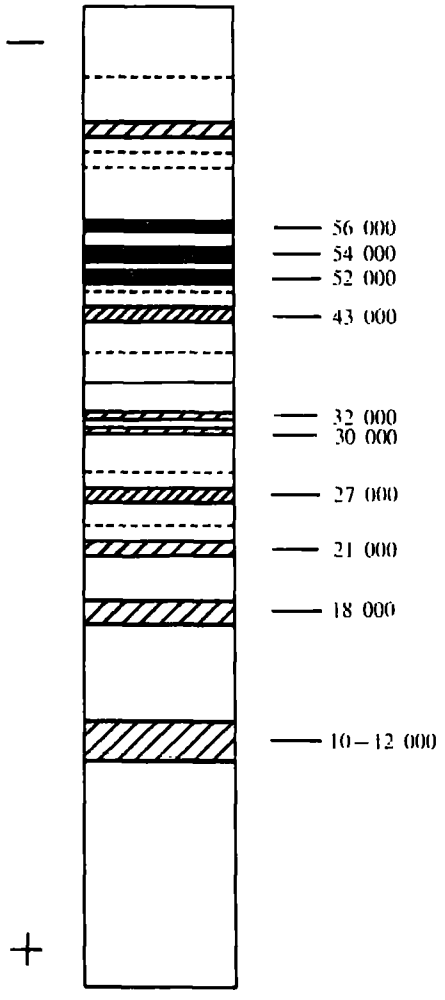


FIG. 1. The subunit profile of cowpea seed proteins extracted in 0·4 M NaCl–0·1 M P_i buffer (pH 8)–0·01 M ME (1 g/25 ml). Dissociated protein was examined by electrophoresis on 10% SDS-acrylamide gels in a continuous 0·1% SDS–0·1 M phosphate, pH 7·0, system (Weber and Osborn, 1969). The molecular weights of prominent subunits are shown.

protein, and the seed meal. The amino acid compositions of the purified subunits were similar to that of the total globulin fraction, i.e. they contained relatively high concentrations of aspartic acid and glutamic acid residues and were also rich in

TABLE 1. *The molecular weights of the subunits in minor protein peaks eluted from DEAE-cellulose in the presence of 8 M urea, and their cysteine/cystine content*

Fraction	Mol. wts of subunits				Cysteine (g/100 g protein)
6–8	27 000;	13 000			0·71
10–12	36 000;	16 000	13 000		0·54
18–21	36 000:	31 000	27 000	18 000	0·70
22–25	36 000:	30 000	18 000		0·30

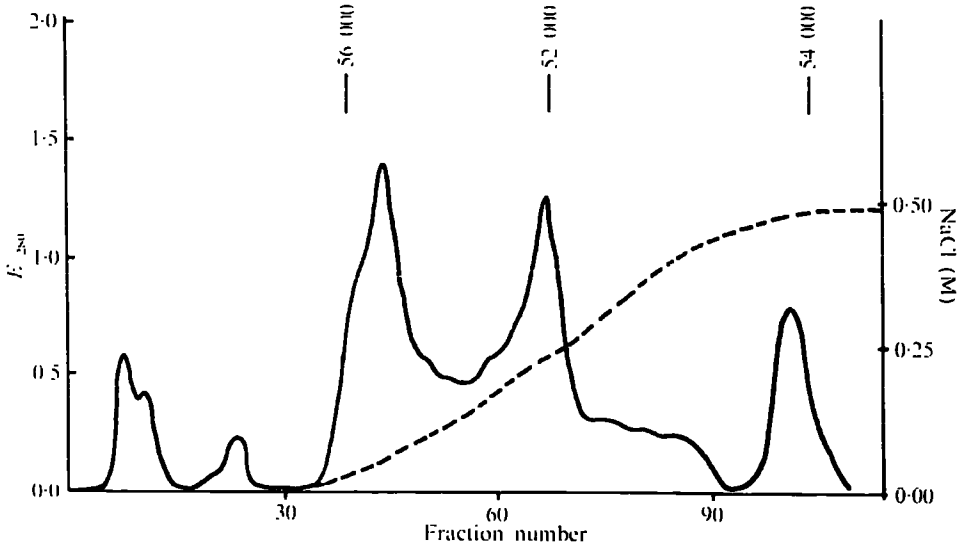


FIG. 2. The elution profile of cowpea seed globulins separated by ion-exchange chromatography with a NaCl gradient on Whatman DE52-cellulose in 0·076 M Tris-citrate buffer (pH 8·7)–8 M urea. The approximate positions at which the major subunits (mol. wt. 56 000, 54 000, and 52 000) were eluted are shown.

TABLE 2. *The chemical composition of* Vigna unguiculata *var. 'Prima'*

				Subunits		
	Meal (g/100 g meal)	Albumin (g/100 g protein)	Globulin (g/100 g protein)	Mol. wt. 56 000 (g/100 g protein)	Mol. wt. 54 000 (g/100 g protein)	Mol. wt. 52 000 (g/100 g protein)
Aspartic acid	9·65	12·26	11·29	11·5	8·42	12·00
Threonine	3·47	4·62	2·99	2·48	1·65	2·62
Serine	4·89	4·56	4·26	5·70	3·44	4·98
Glutamic acid	15·60	12·00	15·75	13·6	10·74	13·22
Proline	3·28	3·34	3·78	3·14	3·08	3·76
Glycine	4·01	3·76	2·72	2·96	1·97	2·21
Alanine	3·81	4·28	2·83	3·14	1·93	2·99
Cysteine	1·01[a]	2·69[a]	0·55[a]	0·6[a]	0·37[a]	0·23[a]
Valine	5·09	3·75	4·96	3·40	2·52	4·30
Methionine	1·40	1·06	0·98	trace	0·72	0·81
Isoleucine	4·46	2·43	3·86	3·06	2·13	4·15
Leucine	7·47	2·96	6·36	6·51	4·54	9·42
Tyrosine	2·86	2·65	2·23	2·37	1·68	3·12
Phenylalanine	5·05	2·78	6·66	4·64	2·53	7·05
Histidine	3·06	2·64	3·27	2·12	2·72	2·67
Lysine	6·27	5·55	6·75	4·03	4·25	5·56
Arginine	6·36	4·32	6·00	4·51	5·53	9·42
Tryptophan	0·85	n.d.	n.d.	n.d.	n.d.	n.d.
N-terminal amino acids				n.d.	Glycine	Leucine/ isoleucine
Neutral sugars				n.d.	2·2%	1·9%

[a] Determined as cysteic acid after performic oxidation.
n.d. = not determined.

leucine, phenylalanine, and arginine but low in S-containing amino acid residues. The cystine/cysteine contents of the two subunits for which these residues were determined were not identical and subunit mol. wt. 54 000 was richer in these residues than subunit mol. wt. 52 000. For most of the amino acids, the seed meal composition was closer to that of the globulin than to that of the albumin fraction, reflecting the higher concentration of the former in the seed; however the cystine/cysteine content of the seeds was determined mainly by the albumin fraction, which was rich in these residues compared to that of the globulin fraction.

Seed pod development

The development of mature pods of cowpeas grown under glass at Durham occupied approximately 28 d from anthesis to maturity compared to the 18–21 d required in West Africa. Except during the early stages, development was not well synchronized between seeds of the same chronological age, for example the fresh weights of individual seeds 14 d after anthesis ranged from 51 to 180 mg and their N contents differed also (20–80 mg). In contrast, except for seeds which were drying out, seeds of equal fresh weight were similar to one another in seed coat colour, dry weight, and N and S contents even though their ages varied by up to ± 5 d. (Seeds which were drying out were easily distinguished from younger seeds of equal fresh weight since seed desiccation coincided with drying out of pods.) The various parameters which have been investigated have been determined therefore in relation to seed fresh weight rather than to chronological age and this data is presented in Fig. 3 (pod wall) and Fig. 4 (seeds). However, for comparison with the work of others, the approximate time course of development is also indicated.

As in other legumes, the pod walls developed more rapidly than the seeds during the early part of development and after approximately 7 d from anthesis they were one third the length of mature pods and had three times as much N as the seeds contained within them. During the next few days the pod wall N increased temporarily and then declined to a level which remained relatively constant throughout the second half of the development period. About this time there was an increase in free amino acids in the wall and an acceleration in the N and dry weight accumulation by the seed. The route by which N and other material was transported from the wall to the seeds was not identified, however electron micrograph studies (Harris, unpublished) show that, at this stage, the seed wall is in close contact with the pod wall and has vascular tissue near its periphery. Pod wall protein was probably not a source of N for the seeds since the 'total' protein fraction (alkali-soluble) of the pod wall remained relatively constant whilst these early changes were occurring. When dehydration of the pods commenced the soluble protein fraction increased, although the amino acid composition profile of the pod walls (Table 3) was not changed.

From the data presented in Fig. 4, four phases of seed development have been identified. During phase I, which is a phase of active cell division, the seeds remained small and were rich in free amino acids. In phase II the growth rate accelerated and the seed dry weight, N, and S contents doubled whilst the proportion of free amino acids present declined. Globulin protein was first detected

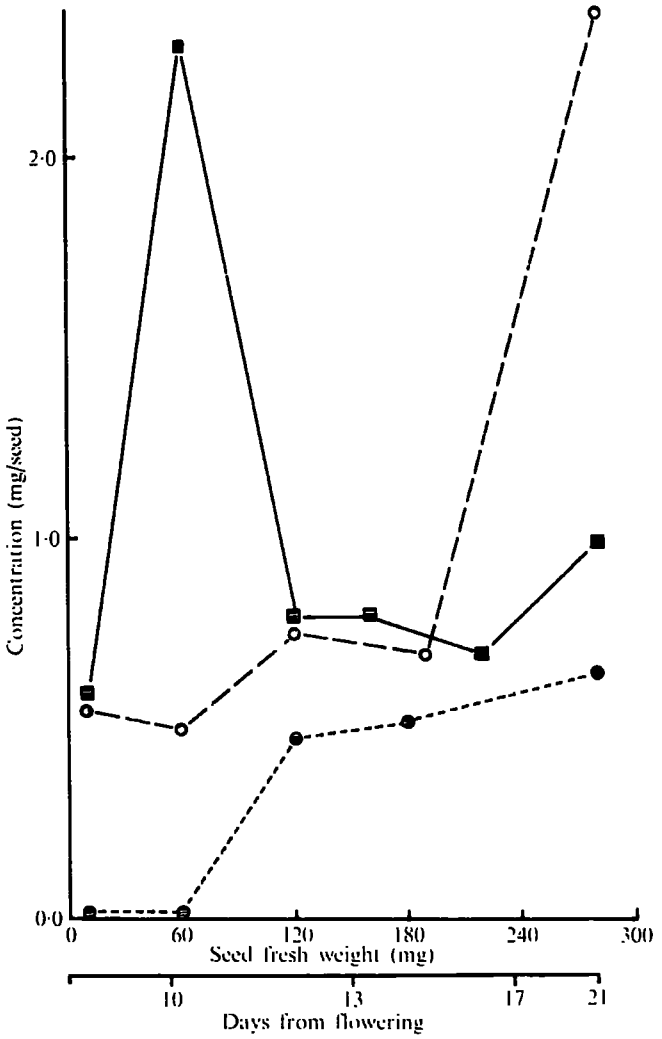


FIG. 3. Changes in the pod wall during the development of cowpea seeds. The data were calculated on a per seed basis, and the approximate time course of development is indicated. Nitrogen, ■—■; alkali-soluble protein, O—O; free amino acids, ●—●.

in this phase during which the cell vacuoles in the cells of the cotyledons divide prior to protein body formation (Harris and Boulter, 1976). Phase III corresponds with the period of maximum growth and with the enlargement of cell vesicles (protein bodies) in the cotyledons (Harris and Boulter, 1976). Nitrogen increased to a level approximately 80% of that found in mature seeds and S and protein, including globulin protein, also increased. Development during this phase was not uniform and the rates of increase in dry weight and N were temporarily retarded at approximately 150 mg fresh weight. After this temporary decline the rate of S accumulation decreased and by the end of phase III desiccation had commenced.

In the final phase (IV), during which desiccation continued and was completed, the rates of dry weight, N, and S accumulation declined; net synthesis of globulin continued.

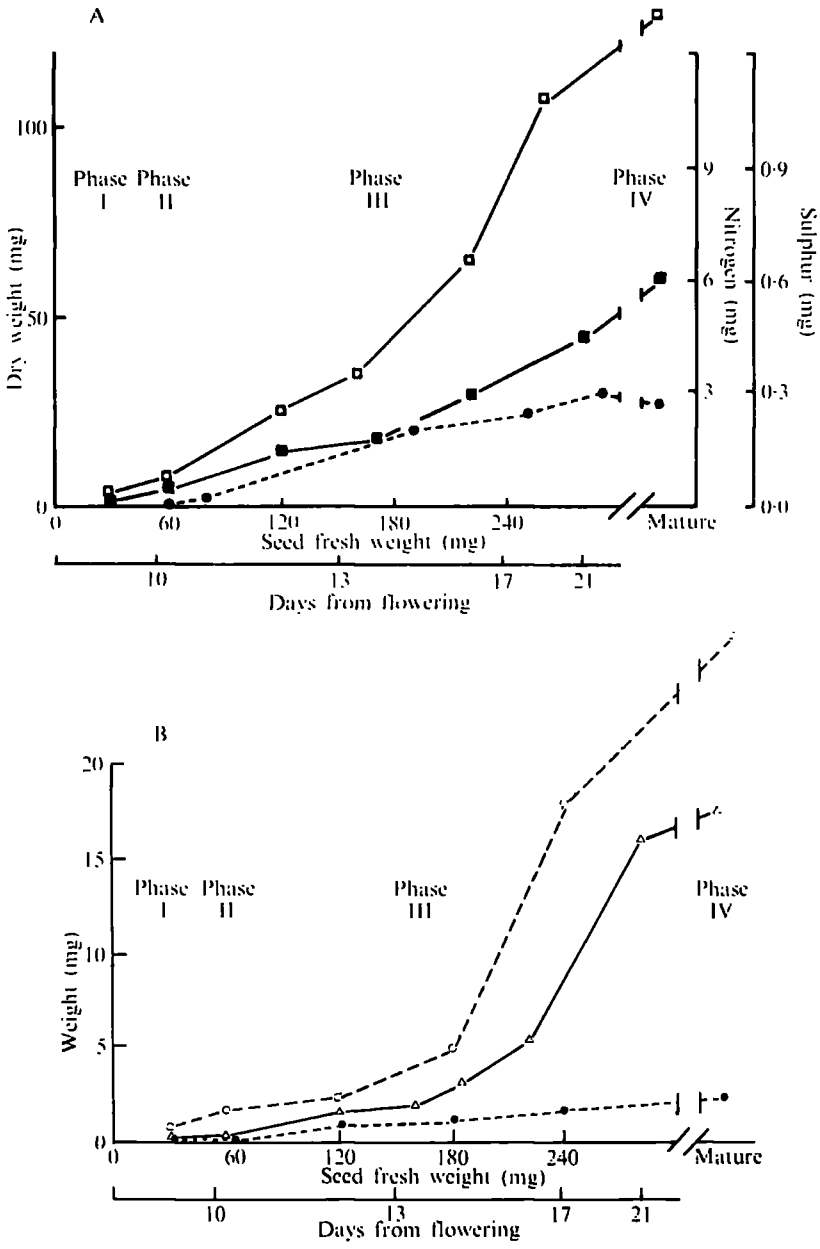


FIG. 4. Changes in cowpea seeds during their development. The data were calculated on a per seed basis, and the approximate time course of development is indicated. A. Dry weight, □—□; N, ■—■; S, ●—●. B. Alkali-soluble protein, ○—○; globulin, △—△; free amino acids, ●—● (leucine equivalents).

The seed globulin fraction changed qualitatively as well as quantitatively during development and this probably explains the change in the profile of the seed amino acid composition (Table 3). The number of serologically different proteins detected by immunodiffusion experiments increased from three at the commencement of phase III to five by the end of phase IV (Table 4; Fig. 5) and the number of major

TABLE 3. *The amino acid compositions of pod walls and seeds of cowpeas at two stages during their development; (i) seed fresh weight 70 mg, (ii) seed fresh weight 240 mg*

Data are expressed as g amino acid residues/100 g freeze dried sample. Cystine plus cysteine was determined on separate samples after performic acid oxidation.

Amino acid	Pod		Seeds	
	(i)	(ii)	(i)	(ii)
Aspartic acid	1·26	0·96	2·06	3·20
Threonine	0·50	0·31	1·13	0·84
Serine	0·50	0·34	1·04	1·03
Glutamic acid	1·13	0·84	2·16	2·96
Proline	0·44	0·29	0·96	0·87
Glycine	0·43	0·28	0·86	0·79
Alanine	0·53	0·33	1·07	0·91
Cysteine	0·17	0·12	0·37	0·31
Valine	0·61	0·37	1·27	1·18
Methionine	0·15	0·09	0·43	0·30
Isoleucine	0·50	0·31	1·27	1·03
Leucine	0·85	0·55	1·56	1·72
Tyrosine	0·35	0·21	0·96	0·69
Phenylalanine	0·56	0·33	1·20	1·27
Histidine	0·29	0·16	1·31	0·91
Lysine	0·69	0·48	1·46	1·53
Arginine	0·59	0·33	1·82	1·77

subunits seen in SDS-acrylamide gels (Fig. 6) changed from the two seen early in phase III to three at later stages (Table 4). The two major subunits which appeared early in phase III had molecular weights of 54 000 and 52 000 and were the most prominent subunits seen in later stages of development also. Their concentration in seed extracts increased progressively, as judged by their staining intensity in gels, and their relative concentrations changed, since subunit mol. wt. 52 000 retained less stain in gels early in phase III compared to subunit mol. wt. 54 000 but was the major component in phase IV. The third major subunit (mol. wt. 56 000) was detected only after the temporary decline in growth rate which occurred during

TABLE 4. *The antigenic globulins and the molecular weights of the subunits of globulins in extracts from developing cowpeas*

The molecular weights of the major subunits at each stage of development are underlined.

Stage of development (mg. seed fresh weight)	Antigens	Subunits of dissociated globulins (mol. wt. × 10^{-3})
8	None	
62	None	
110	a, b, c	54; 52; 30; 27; 18; 10/12
164	a, b, c	56; 54; 52; 30; 27; 21; 18; 10/12
220	a, b, c, d	56; 54; 52; 43; 30; 27; 21; 18; 10/12
270	a, b, c, d	56; 54; 52; 43; 30; 27; 21; 18; 10/12
Mature	a, b, c, d, e	56; 54; 52; 43; 32; 30; 27; 21; 18; 10/12

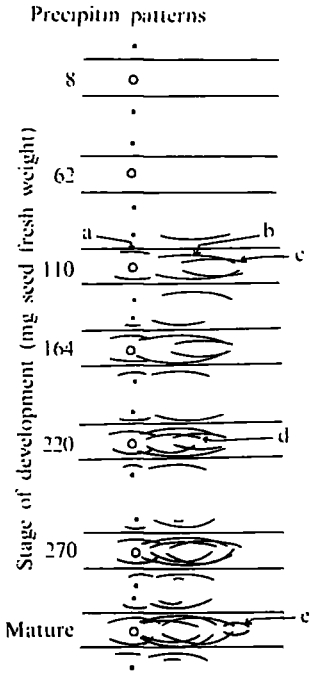


FIG. 5. Precipitin patterns given by protein extracts from seeds at different stages of development. Precipitin arcs which were not detected in earlier stages are labelled a–e.

phase III and it was less prominent than the other two major subunits at all stages of development. New minor subunits also were detected in gels during phases III and IV (Table 4). The correspondence, if any, between the gel patterns and serological profiles was not established.

DISCUSSION

The major seed proteins of legumes are globulins. They include 7S and 11S globulins and the ratio of these is different in different species, for example it is 1 : 4 in *Vicia faba* (Wright and Boulter, 1972), and 9 : 1 in *Phaseolus vulgaris* (Derbyshire and Boulter, 1976). The small number of 11S globulins which have been adequately characterized are similar to each other (Derbyshire, Wright, and Boulter, 1976) and legumin from *Vicia faba* for example is typical of the group. In contrast, the major 7S globulins in extracts from soya bean, pea, and *Phaseolus vulgaris* are not identical, although all three are glycoproteins, and furthermore, the 7S globulin fractions from these species are heterogeneous. The major protein fraction in cowpea is 7S globulin and the data obtained here by ultracentrifugation together with electrophoretic data (Derbyshire, Yarwood, Neat, and Boulter, 1976) indicates that it is heterogeneous; although cowpea has been placed close to *Phaseolus vulgaris* in traditional classifications, the 7S globulin of cowpea is different from that of *Phaseolus vulgaris* (Joubert, 1957; Derbyshire *et al.*, 1976; Barker, Derbyshire, Yarwood, and Boulter, 1976). The 11S globulin from cowpea has not been extensively characterized but the limited data available suggest that it is a typical legumin-like 11S globulin.

Subunits of 7S seed globulin have been isolated and characterized previously only from soya bean (Masaki and Soejima, 1972; Thanh and Shibasaki, 1977). Both soya bean preparations dissociated into subunits of more than one molecular weight class; however, the values obtained were not identical but included subunits with molecular weights similar to one or other of the subunits isolated from cowpea. From the limited amino acid data presented (Masaki and Soejima, 1972) the soya bean subunits would appear to differ significantly in their cysteine/cystine content.

The accumulation of N in the pod wall and its subsequent decrease with the onset of rapid synthesis of seed protein suggests that the wall fractions act as a N reservoir, at least in the early stages of development. The pathway by which N may be transported to the seeds at this stage was not determined but the presence of vascular tissue in seed walls, which are in close contact with the pod wall (Harris, unpublished), suggests that these tissues may participate in the process. If nitrate reductase occurs in the testas of developing cowpea seeds as it does in *Phaseolus vulgaris* and *Vicia faba* (Schleiser and Muntz, 1974) nitrate may be one of the forms in which N is transported from the pods to young seeds.

The growth rate of ripening cowpea seeds declined temporaily when their fresh weight was approximately 150 mg and the protein profile of the globulin fraction after this lag phase was different from that of younger seeds. A similar lag phase followed on the initial increase in seed protein synthesis in *Pisum sativum* (Burrows, 1967; Flinn and Pate, 1968), in *Vicia faba* (Fig. 1A in Briarty, Coult, and Boulter, 1969), and in maize (Ingle, Beitz, and Hageman, 1965); in *P. sativum* it coincided with the disappearance of the endosperm. If the nutrients transported from the endosperm to the embryo differ from those transported to the seed from maternal tissues (Paramonova, 1975) the disappearance of endosperm would result in a change in the composition of the nutrients entering the embryo.

The individual globulins of cowpea, including the minor components, are probably under separate genetic control during synthesis, as made evident by the changing serological and electrophoretic patterns. A similar pattern of sequential protein synthesis has been observed previously with *Vicia faba* (Wright and Boulter, 1972), *Pisum* (Millerd, 1975), and soya bean (Hill and Breidenbach, 1974) which suggests that it may be of general occurrence.

The different cowpea proteins and subunits differ in their S-containing amino acid contents and they also vary in relative concentration during development. Since it would appear that at least in some cases they are under separate genetic control it may be possible to manipulate their proportions to give rise to seeds with an increased proportion of S-containing amino acids. Table 1 shows that the albumin fraction as a whole has a nutritionally better amino acid profile than the globulin fraction and this is reflected in the seed composition. However, it may be much more difficult to manipulate the albumin fraction, qualitatively or quantitatively, than to manipulate the globulins, since the basic metabolism of the seed is more likely to be disturbed by changes in the albumins than by changes in the globulins, which probably function only as storage material.

Globulin synthesis continues in cowpea during drying out of the seeds and less

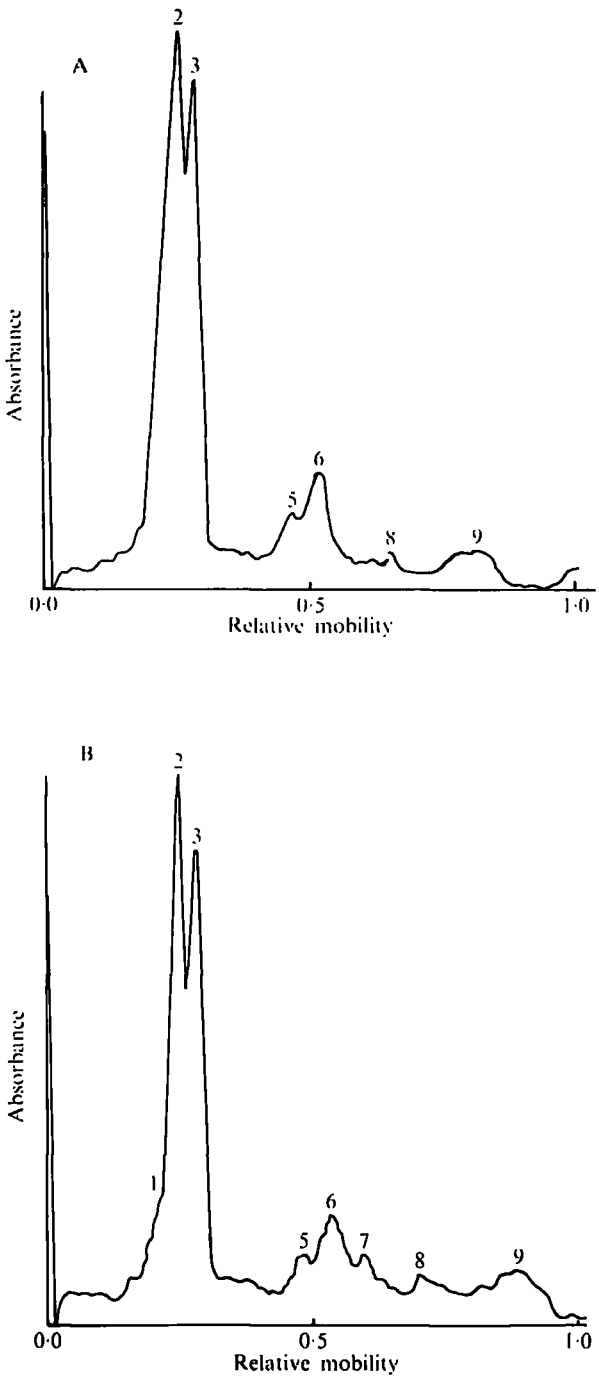
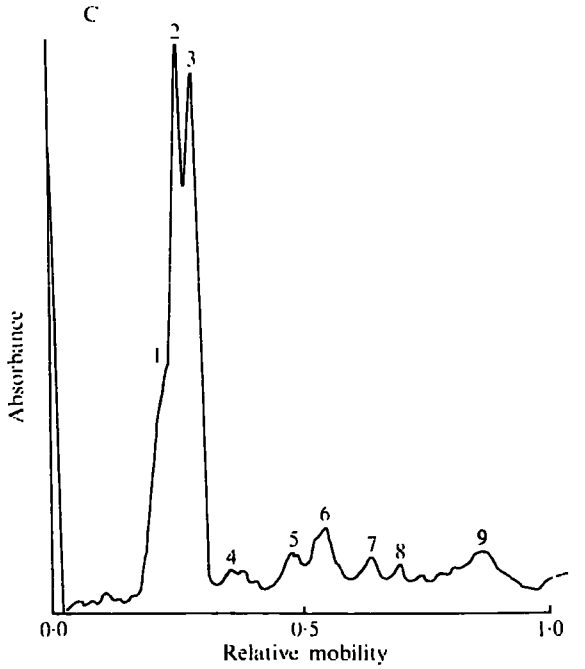


FIG. 6. The subunit profiles of cowpea seed proteins during development. Proteins were extracted in 0·4 M NaCl–0·1 M P_i buffer (pH 8)–0·01 M ME. Dissociated protein was examined by electrophoresis in 10% SDS—acrylamide gels in a continuous 0·1% SDS—0·1 M P_i (pH 7·0) system (Weber and Osborn, 1969). The molecular weights corresponding to absorbance peaks 1–9 are: peak 1, 56 000; peak 2, 54 000; peak 3, 52 000; peak 4, 43 000; peak 5, 31 000; peak 6, 27 000; peak 7, 21 000; peak 8, 18 000; peak 9, 10/12 000. Seed fresh weights: A, 110 mg; B, 164 mg; C, 220 mg; D, mature seed.



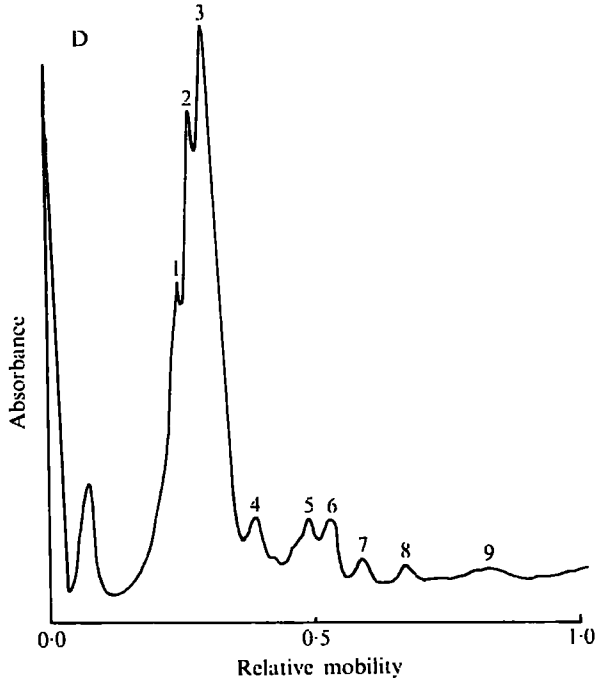


FIG. 6—(*continued*)

11S globulin than 7S globulin is synthesized. These results contrast with those found in *P. sativum*, *Vicia faba*, and soya bean, in which globulin synthesis is almost completed before dehydration commences and 11S globulin is synthesized in equal or greater amounts than the 7S fraction, (Danielsson, 1952; Flinn and Pate, 1968; Wright and Boulter, 1972; Hill and Breidenbach, 1974). This apparent correlation between the physiological stage at which seed storage protein is synthesized and the ratio of 11S globulin to 7S globulin in the mature seed suggests that it may prove possible to increase this ratio in cowpea seeds by postponing the onset of dehydration.

LITERATURE CITED

BARKER, R. D. J., DERBYSHIRE, E., YARWOOD, A., and BOULTER, D., 1976. *Phytochemistry*, **15**, 751–7.

BENCZE, W. L., and SCHMIDT, K., 1957. *Analyt. Chem.* **29**, 1193–6.

BOULTER, D., EVANS, I. M., and DERBYSHIRE, E., 1973. *Qualitas Plant.-Pl. Fds. Hum. Nutr.* **23**, 239–50.

—— —— THOMPSON, A., and YARWOOD, A., 1973. In *Nutritional improvement of food legumes by breeding* (PAG, U.N., New York). Pp. 205–15.

BRIARTY, L. G., COULT, D. A., and BOULTER, D., 1969. *J. exp. Bot.* **20**, 358–72.

BURROWS, W. J., 1967. Ph.D. thesis. Queens University, Belfast.

DANIELSSON, C. E., 1952. *Acta chem. scand.* **6**, 149–59.

DERBYSHIRE, E., and BOULTER, D., 1976. *Phytochemistry*, **15**, 411–14.

—— WRIGHT, D. J., and BOULTER, D., 1976. Ibid **15**, 3–24.

—— YARWOOD, J. N., NEAT, E., and BOULTER, D., 1976. *New Phytol.* **76**, 283–8.

DUBOIS, M., GILES, K. A., HAMILTON, J. K., REBERS, P. A., and SMITH, F., 1956. *Analyt. Chem.* **28**, 350–52.

DUDMAN, W. F., and MILLERD, A., 1975. *Biochem. Syst. Eco.* **3**, 25–33.

FLINN, A. M., and PATE, J. S., 1968. *Ann. Bot.* **32**, 479–95.

GOOD, T. A., and BESSMAN, S. P., 1964. *Analyt. Biochem.* **9**, 253–62.

GROS, C., and LABOUESSE, B., 1969. *Eur. J. Biochem.* **7**, 463–70.

HARRIS, N., and BOULTER, D., 1976. *Ann. Bot.* **40**, 739–44.

HILL, J. E., and BREIDENBACH, R. W., 1974. *Pl. Physiol., Lancaster*, **53**, 742–6.

INGLE, J., BIETZ, D., and HAGEMAN, R. H., 1965. Ibid. **40**, 835–9.

JOUBERT, F. J., 1957. *J. S. Afr. Chem. Inst.* **10**, 16–20.

LOWRY, O. H., ROSEBROUGH, N. J., FARR, A. L., and RANDALL, R. J., 1951. *J. biol. Chem.* **193**, 265–75.

MASAKI, M., and SOEJIMA, M., 1972. *Sci. Rep. Fac. Agric. Ibaraki Univ.* **20**, 35–40.

MOORE, S., and STEIN, W. H., 1948. *J. biol. Chem.* **176**, 367–88.

—— —— 1954. Ibid. **211**, 907–13.

MORGAN, W. T. J., and ELSON, L. A., 1934. *Biochem. J.* **28**, 988–95.

MILLERD, A., 1975. *A. Rev. Pl. Physiol.* **26**, 53–72.

MOTTERSHEAD, B. E., 1971. *Lab. Pract.* **20**, 483–4.

OSBORNE, T. B., 1924. *The vegetable proteins* (2nd edn.). Longmans Green, London.

OYENUGA, V. A., 1967. *Agriculture in Nigeria*, F.A.O., Rome.

PARAMONOVA, N. V., 1975. *Fiziologiya Rast.* **22**, 324–32.

RACHIE, K. O., 1973. In *Nutritional improvement of food legumes by breeding*. PAG, U.N., New York. Pp. 83–92.

SCHLEISER, G., and MUENTZ, K., 1974. *Biochem. Physiol. Pfl.* **166**, 87–93.

SPIRO, R. G., 1966. In *Methods in Enzymology* Ed. S. P. Colowick and N. O. Kaplan. Academic Press, New York. Vol. 8, pp. 3–26.

SVEDBERG, T., and PEDERSEN, K. D., 1940. *The ultracentrifuge*. Oxford Univesity Press, London.

THANH, V. H., and SHIBASAKI, K., 1977. *Biochim. biophys. Acta*, **490**, 370–84.

VARLEY, J. A., 1966. *Analyst (Lond.)*, **91**, 119–26.

WEBER, K., and OSBORN, M., 1969. *J. biol. Chem.* **244,** 4406–12.
WOLF, W. J., 1970. *J. Am. Oil Chem. Soc.* **47,** 107–10.
WRIGHT, D. J., and BOULTER, D., 1972. *Planta,* **105,** 60–65.
ZACHARIUS, R. M., ZELL, T. E., MORRISON, J. H., and WOODCOCK, J. J., 1969. *Analyt. Biochem.* **30,** 148–52.

Plant Proteins (ed. G.Norton). Butterworths, London. pp 3-24, (1978).

1

THE GENERAL PROPERTIES, CLASSIFICATION AND DISTRIBUTION OF PLANT PROTEINS

D. BOULTER
E. DERBYSHIRE
Department of Botany, University of Durham

Introduction

The objectives of this paper are to give a general description of plant proteins within the context of this meeting, and to set the scene for many of the topics dealt with in detail in later papers. The structure of proteins will not be given in any detail as this has been adequately covered elsewhere (Dickerson and Geis, 1969).

There are very many different plant proteins, at least several thousands, which have many different functions. As yet, there are few data relating their detailed structure to these functions, but the use of the methods of physical chemistry, especially X-ray crystallography, for determining the structure of proteins is rapidly supplying information in this field, and the mechanisms of action of a few enzymes are now known.

General Properties

The properties of biological systems are inscribed by those of the proteins and nucleic acids which they synthesize. Diffusion tends to destroy biological systems, and large molecules, therefore, confer stability since the speed of diffusion is reduced with increase in mass. Proteins consist of one or more polypeptide chains and because of their large size, they occur as free macromolecules in colloidal solution and as associated colloids in a supramolecular condition. The colloidal state is correlated with low thermal conductivity, low diffusion coefficients and high viscosity, properties which buffer the system from the effects of a changing environment. Proteins are polyelectrolytes and in supramolecular structures allow energy transduction to take place, including the transduction of light into chemical energy under physiological conditions. Similarly, the many chemical processes which take place in biological systems could only proceed at the necessary rate through the agency of enzymes, all of which are proteins. Furthermore, many specific activities, such as cell interaction, cell motility, gene repression, mitogenesis, compartmentation and intracellular transport are mediated by proteins.

Biological systems in addition to having great stability also have the capacity to reproduce their structures. The information for specifying the structures and processes of the system from birth to death is stored and transferred to offspring by

the activity of nucleic acids; this information must also be transcribed and translated to give rise to the proteins of the system.

Proteins consist of single or several polypeptide chains with unique sequence(s) of amino acids giving rise to unique three-dimensional structures which determine their highly specific activity. Isoenzymes, that is proteins with the same enzymic activity but different structures, are, however, quite common in plants; Scandalios (1974) lists 45 examples. Although relatively little is known in detail it can be safely assumed from our general knowledge of proteins that small differences will be found in the three-dimensional structures of different isomers. Furthermore, kinetic differences between isoenzymes have sometimes been demonstrated, suggesting that the different forms of an enzyme may have physiological significance, especially as the pattern of isoenzymes may change during development and differentiation. Usually, isoenzymes consist of two or more different polypeptide chains. In the simplest case of a dimeric molecule specified by two different alleles (a,b) at one locus, three forms of the protein may exist in the heterozygote, namely aa, bb and ab. If there are more than two chains, more than two alleles or more than one locus, the possible number of different protein molecules increases, although not all of these may occur, e.g. haemoglobin with the theoretically possible tetramer of four identical α-chains is not found in nature (Williamson, 1969).

As a first approximation it can be taken that all molecules of a particular protein species are identical. Thus, unique sequences have been deduced for several proteins e.g. insulin (Sanger, 1956), suggesting that within the sensitivity of the method, i.e. approximately 5 per cent, molecules of these proteins are identical. Amino acid incorporation studies also lead to the same conclusion and show that whereas *in vitro* a small number of mistakes can occur in protein synthesis, such mistakes are extremely rare *in vivo*. Thus, haemoglobin contains no isoleucine and when ^{14}C isoleucine is fed to rabbit reticulocyte cells, they synthesize haemoglobin *in vivo* which contains no radioactivity; haemoglobin synthesized *in vitro* contains a small amount of isoleucine. There is evidence, however, that mistakes do occur rarely in the translation of the haemoglobin message *in vivo* (Carbon, Berg and Yanofsky, 1966) and physical techniques, such as diffusion and ultracentrifugation, suggest that some microheterogeneity may exist (Cole, 1964).

Since the biological activity of proteins is determined by their amino acid sequence, they cannot be synthesized as are the other compounds of the cell by an enzymic method; this would necessitate a large number of specific enzymes which in turn would require others *ad infinitum*. Probably all proteins are synthesized by a template mechanism which involves interaction of mRNA with ribosomes. Even though the process requires the co-ordinated activity of about 100 enzymes the rate of synthesis is very fast, approximately 15 amino acids per second. The main site of synthesis is in the general cytoplasm but mitochondria and chloroplasts also synthesize some proteins, although most of the proteins of these organelles are synthesized elsewhere in the cytoplasm prior to their transport to and incorporation into the organelles. Protein bodies (*see later*) probably do not synthesize proteins, in spite of earlier claims to the contrary. However, recently, Burr and Burr (1976) have suggested that zein, the major storage protein of maize, is synthesized on rough endoplasmic reticulum attached to the outside of protein bodies.

Normally, a cell produces several molecules of each specific protein-type but in cases where enzymes may be under regulatory synthetic control, one or a few proteins may be produced in great excess. Clarke (1974) in considering microbial enzymes, has divided them on this criterion into three groups:

1. The central metabolic pathway enzymes, which are needed whatever the nature of the primary growth substance, e.g. enzymes of the glucose → pyruvate pathway, which are normally constitutive.
2. The biosynthetic enzymes which determine the synthesis of cell constituents from the intermediates of the metabolic pools, e.g. the enzymes of the aspartate → arginine pathway, which are normally de-repressed except in the presence of exogenous cell constituents.
3. The peripheral enzymes which convert the primary food substances into compounds which can enter the central metabolic pathways, e.g. β-galactosidase, which hydrolyses lactose to glucose. These enzymes are catabolic and initiate the attack on substrates. Normally, they are inducible and almost always subject to catabolite repression by other carbon compounds, i.e. those more directly metabolized.

However, the extent and importance of induction and repression of enzyme synthesis in plants has not been established. Seed storage proteins are an example from plants where proteins may be produced in great excess (*see later*), although the underlying mechanism is not known in this case.

Classification

Classification is an artificial process reflecting the use of purpose of the classifier (Davis and Heywood, 1963). Present methods do not give exclusive classes and a particular protein can be classified in several different ways; for example, histones may be considered as regulatory or as structural proteins, permeases as membrane or transporting proteins, etc. Possibly when sufficient three-dimensional data is available, the ideal classification may be based on the mechanism of protein action. Alternatively, the omega classification (Davis and Heywood, 1963) may be based on evolutionary relationships since it is possible that on this basis there are fewer than 500 different types of protein. Rossmann, Moras and Olsen (1974), for example, have suggested that the NAD-nucleotide binding enzymes, including various dehydrogenases, kinases and flavodoxins, are homologously related, since they possess a common nucleotide binding domain.

Present methods of classification relate to chemical structure, mechanism of action, biological function, location and the separatory procedures employed in protein purification.

CHEMICAL STRUCTURE

Proteins may be classified as acidic, basic or neutral; alternatively, they may be classed as conjugated or unconjugated. Unconjugated proteins are composed entirely of amino acids, whilst conjugated proteins have other chemical groups bound to the polypeptide chain(s) in stoichiometric proportions. A variety of chemical groups may be involved, including carbohydrates, lipids, metals, porphyrins, open-chain tetrapyrroles and flavins. Emphasis on stoichiometry is important, since there are many examples where a protein has been classed as a conjugated

protein and later found not to be so because the non-amino group was not stoichiometrically bound, e.g. Concanavalin A from jack bean, first thought to be a glycoprotein (Olson and Liener, 1967; Agrawal and Goldstein, 1968).

MECHANISM

An example of a mechanistic classification is the grouping together of the mammalian proteolytic enzymes, trypsin, chymotrypsin, elastase and thrombin plasmin, on the basis of the similarity of their mechanism of action (Hartley, 1974); they all have the same serine catalytic centre and charge transfer groups. Interestingly, subtilisin from *Bacillus subtilis* has a similar catalytic mechanism to these enzymes but an entirely different amino acid sequence otherwise; mechanistically, therefore, it belongs to the same group but evolutionarily is distinct. As more three-dimensional structures of proteins are solved, and as a consequence their mechanism of action understood, this method of classification will become more widely used.

FUNCTION

Proteins are classified on the basis of their function into enzymic, enzyme inhibiting, recognition, regulatory, transporting, storage and structural proteins.

Enzymic

Enzymes form the largest groups in any classification scheme of proteins. Referenc should be made to Dixon and Webb (1964) or to the reports of the IUPAC-IU of Biochemistry (IUB, 1973) for details of their official nomenclature. In these schen enzymes are classified according to the nature of their reaction towards specific sub strates, often determined *in vitro* though these may not always reflect completely their activities *in vivo.*

Enzyme inhibitors

Proteinaceous inhibitors identified in plants include those of trypsin, chymotrypsin, other mammalian proteinases, insect and bacterial proteinases, amylase, invertase and lipase (Liener, 1969).

The role *in vivo* of proteinase inhibitors has often been regarded as protective (Ryan, 1973). Alternatively, a storage role has been assigned to them since, in for example soya bean, barley and potatoes, they are present in relatively high concentrations (Ryan, 1973). Most proteinase inhibitors which have been studied do not inhibit endogenous proteolytic activity from the same source (Ryan, 1973); however, proteolytic activity in lettuce and barley seeds increases during germination and this increase is correlated with a disappearance of the endogenous inhibitor(s) (Ryan, 1973), suggesting that the latter may regulate proteolytic activity during germination. In cow peas, Gennis and Cantor (1976) have shown that the

proteinase inhibitor is complexed with an endoproteinase, suggesting that its role may be to stabilise its activity.

Recognition

Recognition proteins may be loosely defined as those which determine the compatibility or otherwise of incoming materials by specific interactions. The classic examples of recognition proteins are those involved in the immune response in animals. Little is known about recognition proteins in plants but their importance is rapidly becoming acknowledged as our information about them increases. They include membrane receptor proteins (not yet clearly identified in plants), algal and fungal 'sexual' recognition proteins (Weise, 1969; Crandall, Lawrence and Saunders, 1974), pollen proteins which determine the compatibility/incompatibility reaction on receptive or non-receptive stigma, respectively (Heslop-Harrison, 1975), and those proteins which may control the specificity of roots towards rhizobial strains (Bohlool and Schmidt, 1974).

Many proteins from plants have been shown to agglutinate red blood cells and are, therefore, called phytohaemagglutinins; since many of them show selective agglutination effects they are also called lectins, a term sometimes used synonymously with phytohaemagglutinins. These proteins occur especially in the seeds of the Leguminosae and Euphorbiaceae but are not exclusive to this organ or to these families (Toms and Western, 1971). Eventually it may be found that many recognition proteins are also phytohaemagglutinins, since both recognition proteins and phytohaemagglutinins are usually glycoproteins. This is not invariably so however. Concanavalin A, for example, the best known plant lectin, is not a glycoprotein. In addition to agglutination, phytohaemagglutinin preparations often show other effects on animal cells, e.g. toxicity or mitogenicity (Liener, 1969). However, in most instances the preparations are not of sufficient purity to decide whether one or more of the proteins present are responsible for the different functions.

Regulatory

The transcription and replication of DNA is probably regulated in part by histones, either by inhibition of DNA-dependent RNA polymerases or by interaction with the DNA template. Histones can be fractionated into five groups, at least one of which, Histone IV, is probably homologous between species (Delange and Smith, 1971); further heterogeneity of histones may result from side-chain modifications of the amino acids after the synthesis of the primary sequence. Lack of specificity of histones, however, precludes them from being responsible for specific genetic regulation during differentiation and development and it is now generally accepted that they are non-specific repressors which can interact with most regions of DNA and so reduce the template activity of the latter.

In contrast to the non-specific histones, highly specific regulatory proteins have been shown to control the synthesis of some proteins in micro-organisms (Jacob

and Monod, 1961a and 1961b); specific regulatory proteins of this type have not been isolated from higher plants but can be assumed to be present.

Other regulatory plant proteins include the photosynthetic receptors, chlorophyll *a* and accessory pigments such as chlorophyll *b*, the algal biliproteins and morphogenic proteins. Chlorophyll and the other photosynthetic pigments are photoreceptors which thereby regulate photosynthesis and hence, indirectly, the biochemical processes which are dependent on its products. Photochlorophyllide holochrome and phytochrome, on the other hand, are morphogenic photoreceptors. Conformational changes in etioplast structure, synthesis of chlorophyll and synthesis of some specific proteins, all follow the photoreduction of protochlorophyllide holochrome to chlorophyllide holochrome and precede the development of photosynthetic competence (Zucker, 1972). Phytochrome consists of a blue chromophore covalently bound to a protein; it exists in two forms P_R and P_{FR}, which are interconvertible and the ratio P_R:P_{FR} is controlled by light quality and intensity or by the length of duration of darkness. One or other form of phytochrome initiates or modifies developmental processes such as germination, etiolation and flowering by mechanisms as yet unknown; there is little direct evidence, for example, to support the suggestion by Mohr (1966) that phytochrome regulation involves gene activation.

Other morphogenic proteins are associated with microtubules, which are involved in chromosome movements during mitosis and microfilament orientation in plant cell-wall formation; still others are associated with microfilaments involved in motility phenomena such as protoplasmic streaming (Hepler and Palevitz, 1974).

The tubulin protein of plant microtubules which differs from actin, its animal counterpart, has been extracted from several plants and shown to be similar in all cases (Hepler and Palevitz, 1974). Proteins isolated from phloem of higher plants, P proteins, are also suspected to function morphogenically. They are heterogeneous in form, but are probably all made up of subunits of mol. wt approximately 60 000, which in 0.1 mol KCl polymerize into filaments (Weber and Kleinig, 1971; Kleinig, Dörr and Kollmann, 1971; Kleinig *et al.*, 1971).

A major class of regulatory proteins and polypeptides found in animals are hormones: comparable proteins have not been identified in higher plants, although a protein which induces sexual morphogenesis has been isolated from a culture medium in which *Volvox carteri* was grown (Starr and Jaenicke, 1974).

Transporting proteins

Proteins located in the membranes of animal and bacterial cells mediate the transport of several types of molecules, including amino acids and sugars. Similar proteins have not been isolated or characterised from higher plants but presumptive evidence for them exists. Thus, the organisation and structure of plant, animal and bacterial membranes are very similar as are the kinetics of amino acid uptake in these different organisms. The transport of oxygen in legume nodules is controlled by leghaemoglobin, which is genetically coded for by the host plant (Dilworth, 1969); two leghaemoglobins have been isolated from soya bean nodules, one of mol. wt approximately 15 600 and the other with a mol. wt about 15 900.

Storage proteins

Proteins which are laid down in one phase of development but function subsequently at a later metabolic phase to supply intermediary nitrogen compounds for biosynthesis are called storage proteins. The most important of these are the seed storage proteins which will be dealt with in some detail later. Apart from the storage organs of plants, e.g. seeds and tubers, storage proteins also occur in the growing meristems at some stages in the life cycle (Newcomb, 1967).

Structural proteins

Structural proteins have no enzymic activity and a purely structural function although some of the enzymes associated with membranes and ribosomes, etc. may also play an important structural role. There have been considerable technical problems to overcome in the isolation and characterisation of structural proteins and often denatured proteins have been confused with them. However, during the last few years considerable progress has been made in understanding the structural proteins of the cell membranes (Guidotti, 1972) and also those of the bacterial 70S ribosome. An important structural protein of plants is the cell-wall glycoprotein, extensin; it occurs closely bound to the α-cellulose fraction of the wall and its function has been discussed by Lamport (1970).

Location

Subcellular organelles have multi-enzyme complexes associated with them in which the component proteins occur in a fixed spatial arrangement upon which the proper functioning of the system depends. Some of these proteins are only found in a particular organelle, e.g. cytochrome-oxidase in mitochondria, Fraction I protein in chloroplasts, and hence proteins are sometimes classified according to their subcellular location, i.e. chloroplastic, mitochondrial, nuclear, etc.

Some proteins occur exclusively in a particular organ and this affords another method of locational classification, e.g. leaf protein, seed protein (Altschul *et al.* 1964). Because of their importance, seed proteins have been the subject of several recent reviews (Danielsson, 1956; Stahmann, 1963; Altschul *et al.*, 1966; Inglett, 1972; Millerd, 1975).

Methodological

Osborne's (1924) classification, based mainly on solubility criteria, is still the main classification of this type. He grouped proteins into albumins, soluble in water; globulins, soluble in salt; prolamins, soluble in 70–80 per cent aqueous ethanol; and glutelins, soluble in dilute acids and alkalis. In spite of its many shortcomings, this operational classification, or modifications of it, is still widely used (*see later*).

Proteins are normally purified by taking advantage of differences in their size, shape or charge, using a variety of chromatographic, electrophoretic, molecular sieving and centrifugal techniques. Some classifications have been based on these

separatory techniques. Serological cross-reactivity has also been used (Grabor, Benhamou and Daussant, 1962).

Distribution of Plant Proteins

INTRODUCTION

Cells do not synthesize all the proteins for which they have the genetic information. The enzymes of the basic metabolic processes such as respiration and protein synthesis will be present in all 'living' cells but differentiation leads to the formation of many specialised cells and in these the complement of additional proteins will vary. Thus, cells specialised to produce storage protein in the seed will contain proteins which are not found in cells of the leaf specialised for photosynthesis, etc. Furthermore, cell specialisation may lead to the loss of some of the basic enzymes.

During the course of evolution the genetic material coding for a particular protein will have changed so that the amino acid sequence of this protein will vary in different organisms; differences in homologous proteins from plants, e.g. cytochrome *c*, have been used in an attempt to establish plant phylogenetic relationships (Boulter, 1973). Non-homologous proteins evolve at different rates, for example, plant plastocyanin has evolved about one and a half times as fast as plant cytochrome *c* (Boulter *et al.*, 1977).

It is not intended here to give a comprehensive account of the distribution of plant proteins but only to mention some aspects of the proteins of the major food crops.

GRAIN CROPS

Seeds are normally high in protein relative to other parts of the plant as harvested, although on a global basis the latter may synthesize a greater total quantity of protein. The firm foundation of our present knowledge of the seed proteins was laid more than half a century ago during 40 years of study by Osborne and his colleagues (Osborne, 1924) and their classification of the seed proteins has been retained on the grounds of its usefulness in comparing the results obtained for different crops by various workers. Broadly speaking, the albumins correspond to metabolic proteins whilst the other solubility fractions from seeds consist mainly of storage proteins. These fractions are not single proteins, although characterisation of the individual proteins of which they consist has not advanced very far, and care must be exercised in interpreting comparative data, since it is probable that an homologous protein from two different species may not always appear in the same solubility class. Even for a single species, the same protein may appear in different solubility classes if extracted at different developmental stages of the seed, partly due to protein–protein interactions and partly to changes in extractability caused by the presence of other major deposits. The properties of seed proteins may also change during storage or with the method of seed-drying (Nash, Kwolek and Wolf, 1971; Munck, 1972). The order in which the solvents

are used during extraction can also affect the solubility of a protein, due primarily, it is thought, to denaturation (Preaux and Lontie, 1975).

Osborne's methodology has been modified and extended by the substitution of isopropanol for ethanol and by the addition of dissociating and reducing agents (for example, Paulis, Bietz and Wall, 1975; Preaux and Lontie, 1975); these modifications usually result in a glutelin fraction which differs in its composition from that described by Osborne; a new fraction, prolamin II or alcohol soluble glutelin is obtained for example from barley and maize, when an alcohol insoluble residue is extracted with alcohol in the presence of 2-mercaptoethanol. In practice it may be particularly difficult to separate quantitatively albumins from globulins due to the phase characteristics of colloids. Furthermore, confusion can arise if water extracts are referred to as albumins since some globulins will be soluble under these conditions due to the presence of endogenous salts; in the absence of experimental details an unusually high ratio of albumins to globulins is suspect. Usually some denatured protein belonging to the other solubility classes is extracted with the glutelin fraction.

CEREALS

Table 1.1 gives typical values for the protein contents of the major cereals. These are of the order of 10 per cent but show that generally oats have a higher protein content than sorghum, barley and wheat, whilst maize and rice have the lowest protein contents. Individual varieties, however, often deviate from the mean by two or more percentage points and, for example, a high protein variety of rice, BP1-76-1, has a protein content of 14 per cent (Juliano, 1972), which is higher than the mean for barley and wheat. Usually more than 80 per cent of the total protein is endospermic; however, the concentration of protein is higher in the embryo since this is only one-tenth or less of the seed by weight (Wall, 1964).

Also given in *Table 1.1* are the percentages of the various solubility classes as defined in the modified Osborne classification. In cereals, two of the four solubility classes, prolamins and glutelins, predominate. In rice these occur in the protein bodies and can, therefore, be called storage proteins. In barley and maize, whilst both of these proteins co-dominate, evidence suggests that only prolamin occurs in the typical protein bodies (Inglett, 1972; Munck, 1972). Even so, both prolamins and glutelins behave as storage proteins during germination insofar as they are broken down and supply nitrogen building blocks (Folkes and Yemm, 1956). Although prolamins and glutelins usually predominate, their ratio differs in different cereals; for example, in barley and wheat they are present in almost equal proportions whilst in rice nearly all of the storage protein is glutelin. The oft-quoted statement that oats contain a high proportion of globulin stems from an incorrect value quoted in a major review (*Table 2* in Brohult and Sandegren, 1954); reference to the original (Lüers and Siegert, 1924) and more recent publications (Ewart, 1968; Wu *et al.*, 1972) gives values which range between 6 and 13 per cent. Usually the fractions are heterogeneous, for example glutelins in maize (Paulis, Bietz and Wall, 1975), prolamins in oats (Waldschmidt-Lietz and Zwisler, 1963) and globulins in rice and other cereals (Danielsson, 1952; Morita and Yoshida, 1968); however, the prolamin of maize may consist of a single size of polypeptide chain (Paulis, Bietz and Wall, 1975).

Table 1.1 Protein content of cereal meals

	Total[1] *protein* (g/100 g meal)	*Albumins*[2] (g/100 g protein)	*Globulins*[2] (g/100 g protein)	*Prolamins*[2] (g/100 g protein)	*Glutelins*[2] *and residue* (g/100 g protein)
Barley	12	5	15	40	40
Wheat	12	5	10	45	40
Maize	9	5	5	50	40
Rice	8	5	10	5	80
Oats	13	1	13	18	68
Sorghum	10	8	8	52	32

[1] FAO (1970)
[2] Data for protein classes from: Lüers and Siegert (1924); Frey (1951); Brohult and Sandegren (1954); Waldschmidt-Leitz and Zwisler (1963); Wall (1964); Juliano (1972); Wu *et al.* (1972); Whitehouse (1973)

In addition to the major storage proteins, other proteins, for example phytohaemagglutinins in wheat, may occur in concentrations significantly greater than those in which the basic metabolic enzymes are found.

Since the different protein fractions have overall different amino acid compositions (Munck, 1972; *see also* Chapter 13 this volume), cereal meals containing different proportions of the protein classes reflect this in their amino acid compositions (*see Table 2* of Whitehouse, 1973).

LEGUMES

Legumes are often called protein crops because their seeds have a high proportion of protein, usually 20–30 per cent of the dry weight, and even higher in the case of soya bean. Six legumes, chick pea, cow pea, dry bean, peanut, pigeon pea and soya bean, are extensively cultivated as grain crops (Roberts, 1970) and *Table 1.2*

Table 1.2 Protein content of legume meals[1]

	g/100 g meal
Chick pea	20.1
Cow pea	23.4
Dry bean	22.1
Peanut	25.6
Pigeon pea	20.9
Soya bean	38.0
Broad bean	23.4
Pea	22.5

[1] Data from FAO (1970)

gives their seed protein contents, together with those of broad bean and pea which are major crops in Europe; extraction of oil, from oilseeds for example, leaves a residue which is even richer in protein. By contrast with the situation in cereals there is usually little or no endosperm present in mature legume seeds and the storage tissues are almost entirely cotyledonary. Complete balance sheets of the proportions of the Osborne solubility classes are comparatively rare in the literature but indirect evidence suggests that most crop legumes contain about 70 per cent globulin, 10–20 per cent albumin and 10–20 per cent glutelin with very little alcohol soluble protein (*Table 1.3*). There are several different storage globulins as characterised by sedimentation value and amino acid composition and they are extracted in the protein body fraction (Derbyshire, Wright and Boulter, 1976). Proteins which sediment as 11S or 7S molecular species are widely distributed;

Table 1.3 The proportions of solubility classes of protein in legumes[1]

	Albumins (g/100 g protein)	*Globulins* (g/100 g protein)	*Glutelins* (g/100 g protein)
Mung bean	4	67	29
Broad bean	20	60	15
Pea	21	66	12
Peanut	15	70	10
Soya bean	10	90	0

[1] Data calculated from: Fox (1964); Blagoveschensky (1967); Dawson (1971); Wolf (1972)

Table 1.4 The essential amino acid content of legumes[1]

Essential amino acids (g/16 g N)	*Mung bean*	*Broad bean*	*Soya bean*	*Yam bean*	*WHO*[2] *requirement pattern* (g/100 g protein)
Cystine	0.7	0.8	1.3	1.5	} 3.5
Methionine	0.5	0.7	1.3	1.9	
Lysine	8.1	6.5	6.4	6.8	5.5
Isoleucine	3.6	4.0	4.5	4.4	4.0
Leucine	7.0	7.1	7.8	7.7	7.0
Phenylalanine	4.9	4.3	4.9	5.3	} 6.0
Tyrosine	2.5	3.2	3.1	4.3	
Threonine	3.3	3.4	3.9	4.1	4.0
Tryptophan	–	–	1.3	–	1.0
Valine	4.1	4.4	4.8	5.6	5.0

[1] Data from FAO (1970)
[2] World Health Organisation (1973)

however, the exact number of globulins and the extent to which counterparts occur in different legumes, are still not known except in specific cases. Different globulin storage proteins have different amino acid compositions; for example, the 11S legumin of *Phaseolus vulgaris* contains higher amounts of sulphur-amino acids than the 7S glycoprotein II, the major storage protein present (Derbyshire, Wright and Boulter, 1976). Since different legumes contain different proportions of the individual storage globulins (Derbyshire, Wright and Boulter, 1976) this is reflected in the amino acid compositions of their meals (*Table 1.4*). However, as is the case with cereals, other proteins, for example urease in jack bean, phytohaemagglutinins and trypsin inhibitors in *Phaseolus vulgaris*, which occur in significant amounts, may be an important influence on the overall amino acid composition.

OTHER GRAIN CROPS

These include several of the major oil seeds, for example cotton, rape and sunflower and others in more localised cultivation. They represent a broader taxonomic collection than either the legumes or the cereals and this is reflected in their seed structure. Different species have different ratios of endosperm to embryo and in a few species a substantial quantity of nucellar tissue may persist as perisperm in the mature seeds, for example in sugarbeet (Esau, 1953). The seed protein contents of several crops on a fresh weight basis are given in *Table 1.5*, which shows that they vary from species to species and in some, for example rape and cotton, may be as high as the protein contents of many legumes, even before extraction of oil; the amino acid composition of the protein also varies considerably (*Table 1.5*). Data on the protein composition of this group of seeds are much more limited than those on cereals and legumes; however, the major proteins are usually globulins (Brohult and Sandegren, 1954). Globulins which sediment in the ultracentrifuge as 11–12S, 7S or 2S proteins have been identified in many species (Derbyshire, Wright and Boulter, 1976), and it is possible, therefore, that these may be counterparts of legume storage globulins.

Table 1.5 The total protein and essential amino acids of other grain crops[1]

	Cotton	*Rape*	*Sesame*	*Sunflower*
Total protein (g/100 g meal)	20.2	24.0[2]	18.1	12.6
Essential amino acids (g/16 g N)				
Cystine	1.6	2.4	1.8	1.5
Methionine	1.3	1.8	2.8	1.9
Lysine	4.4	5.8	2.7	3.6
Isoleucine	3.3	3.6	3.6	4.3
Leucine	5.9	6.3	6.7	6.4
Phenylalanine	5.2	3.5	4.4	4.5
Tyrosine	2.9	2.6	3.1	1.9
Threonine	3.3	3.8	3.6	3.7
Tryptophan	1.2	1.2	–	1.4
Valine	4.6	4.8	4.6	5.1

[1] Data from FAO (1970)
[2] Bhatty, McKenzie and Finlayson (1968)

Individual protein fractions which have been investigated extensively include edestin from hemp (St. Angelo, Yatsu and Altschul, 1968), amadin from almond (Steenkamp and Joubert, 1967) and the globulins from rape (Bhatty, McKenzie and Finlayson, 1968).

OTHER CROPS

Other major sources of food are roots, tubers and fleshy fruits; they include cassava, yam, potato and banana (plantain). These are generally considered to be calorie crops and they out-yield both legumes and cereals in this respect. On a fresh weight (as harvested) basis they contain less protein than legumes and cereals, of the order of 2 per cent for roots and tubers and less for fleshy fruits (*Table 1.6*). However, post-harvest processing may reduce the moisture

Table 1.6 The total protein and essential amino acids of roots, tubers and fleshy fruits

	Potato[1]	*Yam*[1]	*Cassava*[1]	*Banana*[1]	*Plantain*[2]
Total protein (g/100 g vegetable)	2.0	2.4	1.6	1.15	0.43
Essential amino acids (g/16 g N)					
Cystine	0.6	1.2	1.4	2.7	–
Methionine	1.3	1.6	1.3	2.0	0.6
Lysine	4.8	4.1	4.1	4.1	5.6
Isoleucine	3.8	3.7	2.8	2.9	5.1
Leucine	6.0	6.5	4.0	4.7	5.4
Phenylalanine	4.0	4.8	2.5	3.9	4.5
Tyrosine	2.7	3.2	1.6	2.6	–
Threonine	3.8	3.6	2.6	3.4	2.7
Tryptophan	–	1.3	1.2	–	0.7
Valine	4.7	4.7	3.3	4.0	4.4

[1] FAO (1970)
[2] Harvey (1970)

content of the crop; for example removal of toxins from cassava and drying of potatoes for storage removes more than 80 per cent of the moisture and gives a product with a protein content comparable to or greater than that of some grain crops: at the International Potato Centre, Lima, Peru, potato strains with a protein content of 20 per cent dry weight have been developed. The proteins of these storage tissues, except potato protein, have received little attention in the past; however, it is known that much of the protein can be expressed in the sap (McKee, 1962). The most extensively characterised proteins of potato tubers are the protease inhibitors which represent only 10 per cent of the protein (Ryan, 1973); however, approximately 70–80 per cent of potato protein is storage protein and, as in seeds, this is heterogeneous and highly amidated (Stegemann, 1975).

Minor proteins which are of special interest in a dietary context are the sweeteners which have been isolated from several sources, for example miracle fruit and serendipity berries (Inglett, 1975). They are small proteins, mol. wt 10 000–42 000 and the protein from miracle fruit is a glycoprotein. Although

it is known that the sweeteners act directly on the taste buds the mechanism of action is not fully understood.

The shoots and leaves of plants are also used as sources of dietary protein. Two-thirds of the world's agricultural land is given over to herbage which supplies energy and protein to ruminant animals (Cooke, 1976). Generally, the levels of protein in fresh leaves, for example leafy vegetables, are of the order of 3 per cent (*Table 1.7*), and, as with roots and tubers, post-harvest drying may

Table 1.7 The total protein and essential amino acids of leafy vegetables[1]

	Broccoli	*Cabbage*	*Lettuce*	*Spinach*
Total protein (g/100 g vegetable)	4.3	1.6	1.3	2.2
Essential amino acids (g/16 g N)				
Cystine	1.1	1.1	–	1.6
Methionine	1.4	1.0	1.8	2.1
Lysine	5.1	3.1	3.8	7.3
Isoleucine	4.3	3.1	3.8	4.9
Leucine	5.5	5.3	6.3	9.5
Phenylalanine	4.1	3.0	5.1	6.1
Tyrosine	–	1.8	2.7	5.0
Threonine	3.7	3.8	4.1	5.3
Tryptophan	1.1	–	–	–
Valine	4.9	4.2	5.4	6.1

[1] Data from FAO (1970)

increase the protein content to approximately 20 per cent or more. Leaf proteins are often divided into two fractions, 1 and 2, on the basis of their solubility in ammonium sulphate solutions (Wildman and Bonner, 1947). Fraction 1 is now known as ribulose bisphosphate carboxylase and may account for more than 50 per cent of the total leaf protein; fraction 2 is heterogeneous and contains the rest of the proteins of the leaf. This predominance of the same protein in leaves of different species may account for the fact that the amino acid compositions of different species are rather similar, although not identical (Byers, 1971). The total proteins of leaves have been investigated extensively by Pirie and his colleagues (Pirie, 1971; *see also* Chapter 10, this volume) and fraction 1 protein is discussed further by Ellis (*see* Chapter 2, this volume).

Nutritional Aspects of Plant Proteins

PROTEINS IN DIETS

There are three types of protein which are important in this context:

1. Plant proteins used as food staples
2. Anti-metabolic or toxic proteins
3. Proteins which may afford protection against pests and diseases

Proteins are needed in food to supply the essential amino acids which cannot be synthesized by animals and humans, and as a source of the non-essential amino acids or of nitrogen for their synthesis.

There are eight essential amino acids, phenylalanine, methionine, lysine, leucine, threonine, tryptophan, valine and isoleucine; in addition, cysteine and tyrosine are often also included since the former can spare methionine and the latter phenylalanine. Unless the amino acid composition of the protein in the food is balanced, i.e. in approximately the proportions for the requirement of essential amino acids (WHO, 1973), the protein will be utilised only to the level of the limiting amino acid. Usually only the essential amino acids which are present in low concentration are considered in this context; however, when high levels of leucine occur, for example in maize and sorghum, these may interfere with the availability of other amino acids (Harper, 1964).

With diets containing a mixture of foods, the overall amino acid composition of their protein is usually balanced; but in developing countries where often one plant food forms the major part of the diet, this may not be so. By comparing the amino acid composition of the cereal meals (*see Table 2* of Whitehouse, 1973) with the WHO pattern, it can be seen that lysine is limiting in all these crops, although oats, rice and barley have higher levels than the other cereals. Diets made up largely of cereals, therefore, might also be limiting in lysine; diets in which legumes are the major protein source may be limiting in sulphur amino acids. If other components of diets supply adequate amounts of these amino acids, other amino acids might still be inadequate. Thus, sorghum is also low in methionine, phenylalanine, tryptophan and arginine, the latter being an essential amino acid for some animals; maize is low in tryptophan and arginine, and both maize and sorghum have a large excess of leucine. Rice probably has the best spectrum of the cereals and that of oats is almost as satisfactory. After the sulphur amino acids, legumes may be low in tryptophan, threonine or isoleucine.

The oft-quoted statement that plant proteins are nutritionally inferior to animal proteins stems from these considerations. The statement is not correct if one compares the same proteins from an animal and a plant source, e.g. cytochrome *c* or histone IV from animals is not nutritionally superior to that from plants (Dayhoff, 1972 and 1973). However, plant foods often consist predominantly of one or a few storage proteins and in these specialised proteins one or more essential amino acids are present in relative concentrations lower than those recommended for dietary protein by WHO (1973); for example glutelins in rice are relatively low in lysine and threonine and glycoprotein II in seeds of *Phaseolus vulgaris* is relatively low in sulphur amino acids. Food from animal sources, on the other hand, usually contains a predominant protein which has a good amino acid profile, for example ovalbumin in egg protein, or myosin in meat (*Table 1.8*), or they contain a mixture of proteins with an overall balanced amino acid composition. It is often suggested also that plant proteins are less available or less digestible than animal proteins but the evidence for these general statements is not strong and appears to be based on data from protein mixtures. Much more work is needed in this very important area, since, due to increasing world population, plant proteins will be used more extensively in the future.

Inhibitors of a wide range of proteolytic and glycolytic enzymes occur in plants including cereals and legumes, and their presence in uncooked diets seriously reduces utilisation of the food. The balance of evidence, however, from rat-feeding trials is that these compounds do not have an adverse effect

Table 1.8 The essential amino acid content of eggs and meat and their major proteins

	g/100 g protein			
	Egg[1]	*Meat*[2]	*Ovalbumin*[3]	*Myosin*[3]
Cystine	2.4	1.2	1.6	0.9
Methionine	3.1	2.7	4.6	2.8
Lysine	6.4	9.1	5.6	10.9
Isoleucine	6.6	4.5	7.1	4.7
Leucine	8.8	8.2	7.9	8.9
Phenylalanine	5.8	4.2	8.5	5.0
Tyrosine	4.2	3.5	3.2	2.9
Threonine	5.1	4.6	3.5	4.1
Tryptophan	1.6	1.2	3.6	–
Valine	7.3	4.8	6.0	4.2

[1] Kofranyi (1972)
[2] Bigwood (1972)
[3] Haurowitz (1963)

when they are inactivated by the correct cooking procedures prior to being fed (*see* Chapter 7). During cooking, toxic proteins similar to ricin in castor bean will, if present, be inactivated also.

Another important group of proteins, which if present in uncooked diets reduce the food value, are the phytohaemagglutinins. These may not always be completely inactivated by cooking procedures and so will continue to reduce the nutritional value of the food (*see* Chapter 7). Their activity in this situation is not a direct one on the blood cells but possibly involves interaction with other proteins in the gut or with the gut wall (*see* Chapter 8 for another explanation). Many of the agglutinins which have been isolated are glycoproteins. In legumes they are usually present in concentrations which are low relative to those of the major storage glycoproteins. The dietary significance of glycoproteins generally has not been assessed.

IMPROVING THE SUPPLY OF PLANT PROTEIN

There is now considerable evidence that many regional diets in developing countries contain inadequate quantities of protein as defined by the WHO requirements (1973). In most cases the deficiency of protein could be made up by increasing the amount of food consumed (*see* Chapter 14, this volume). However, there are some diets, for example those where the staple is a mainly starchy crop such as cassava or yam, in which the protein content is so low and of poor quality as to preclude meeting protein requirements solely by increased intake of food. In some cases, it may be possible to introduce other food crops containing higher protein contents and this may overcome the deficiency and additionally improve protein quality. However, attempts to introduce new food sources into traditional diets are likely to be impeded by sociopsychological factors which are difficult to overcome.

Many diets could also be improved if the quantity and/or the quality of the protein in utilised crops were enhanced. There are two possible ways of accomplishing this: one is by breeding programmes, either conventional or mutation breeding, and the other by improved management, e.g. fertiliser treatments for those

soils which limit the amount and quality of the protein of the plants grown in them. For these programmes to be successful, satisfactory screening methods for total protein and protein quality must be available. These must be simple, fast and not require expensive equipment or much plant material (Evans and Boulter, 1974). Methods exist for determining total protein either by dye-binding, automated Kjeldahl or by the use of IR reflectance methods (Neotec, 1971; Evans and Boulter, 1974), and for screening for lysine and tryptophan in cereals (Villegas and Mertz, 1971), and for the sulphur amino acids of legumes (Evans and Boulter, 1974; Boulter and Evans, 1975).

Since about 70 per cent of the world's protein comes from cereals, and only about 20 per cent from legumes, at first sight it might appear that improving the yield and/or protein content of cereals would be the best strategy. However, as pointed out by Sinclair and de Wit (1975), 1 unit of glucose gives rise to approximately 0.8 units of polysaccharide but to only 0.4 units of protein, synthesis of the latter having a higher energy requirement. Therefore, for the same quantity of photosynthate transported to the seed there must be a 'trade-off' between seed protein content and seed weight. While this situation might be accepted by nutritionists in regions where the available carbohydrate is more than adequate, it would not be accepted by farmers except in special circumstances. This situation underscores the difficulties involved in increasing the protein content of cereals. However, Johnson and his colleagues have shown that a negative correlation between protein content and total yield is not invariable. They have recently released a new productive, high protein, hard winter wheat variety derived from Atlas 66, with genetic potential for 2 percentage points higher grain protein content (Johnson, 1976).

An alternative strategy is to improve the yield of legumes which, apart from soya beans, have not received as much attention as cereals. Their potential has been pointed out by Jain (1971), Borlaug (1973), Boulter (1976) and others and there are several reports of legume lines which contain much higher percentages of protein than the usual varieties (for example, Luse, 1975). If the problems of instability of total and protein yields and susceptibility to pests and diseases could be solved, or at least in part, then improving legumes could be an excellent strategy.

Not only protein content but also protein quality has come under intensive investigation by breeders and agronomists. Much of the work has been concentrated on improving amino acid profiles and considerable progress has been made, mostly with cereals (*see* Chapters 12 and 13). In addition however, the extent and nutritional significance of toxic and antimetabolic proteins, and other compounds in plants, e.g. tannins, saponins and polyphenols, will need to be better understood. As more 'exotic' plant materials are incorporated into breeding programmes these compounds may have to be 'bred out'. On the other hand, this could be detrimental if it were established that a compound was a natural protective agent against pests and diseases of the crop (Ryan, 1973), either in the field or in storage. For example, when gossypol, a nutritionally undesirable polyphenol was bred out of cotton seeds, these cotton plant lines were more susceptible to pests. In some cases the compounds in question may be otherwise highly nutritious; thus, some protease inhibitors of legumes have exceptionally high sulphur amino acid contents, for example the blue protein inhibitor in *Phaseolus vulgaris* (Pusztai, 1966). For those compounds which are completely inactivated during normal cooking procedures,

the correct strategy might be to breed for increased rather than decreased amounts.

In the first agricultural revolution which took place in Neolithic times, the major crop plants which supply most of the world's food, except sugar beet, were brought into cultivation. It is not certain upon what factors the choice and successful domestication depended, but an important factor must have been the ability of the technology of that time to minimise the effects of toxic substances, and to store and prepare food. As the world population continues to increase, it would appear that in order to grow sufficient food a second agricultural revolution will be needed before the turn of the century, in which new crops, new technology and new cropping patterns will play a part.

Acknowledgements

We gratefully acknowledge the financial assistance of the Ministry of Overseas Development, and Mrs. M. Creighton for help in the preparation of the manuscript.

References

AGRAWAL, B.B. and GOLDSTEIN, I.J. (1968). *Archs Biochem. Biophys.*, **124**, 218

ALTSCHUL, A.M., DECHARY, J.M., NEUCERE, N.J. and WOODHAM, A.A. (1964). *Nature, Lond.* **203**, 501

ALTSCHUL, A.M., YATSU, L.Y., ORY, R.L. and ENGLEMAN, E.M. (1966). *A. Rev. Pl. Physiol.* **17**, 113

BHATTY, R.S., McKENZIE, S.L. and FINLAYSON, A.J. (1968). *Can. J. Biochem.* **46**, 1192

BIGWOOD, E.J. (1972). In *Protein and Amino Acid Functions*, p.215. Ed. by E.J. Bigwood. Pergamon Press, Oxford

BLAGOVESCHENSKY, A.V. (1967). *Biochemistry of Leguminous Plants*, Akad. Nauk., SSSR

BOHLOOL, B.B. and SCHMIDT, E.L. (1974). *Science* **185**, 269

BORLAUG, N.E. (1973). In *Nutritional Improvement of Food Legumes by Breeding*, p.7. Ed. by M. Milner. Proc. Symp. sponsored by PAG, Rome, 1972. John Wiley and Sons, New York

BOULTER, D. (1973). *Nobel Symposium* **25**, 211

BOULTER, D. (1976). 'Improving protein content and quality in legumes.' *Proceedings of the International Symposium on 'Genetic Control of Diversity in Plants'*, Lahore, Pakistan, March 1976 (in press)

BOULTER, D. and EVANS, I.M. (1975). *Proceedings of the International Institute of Tropical Agriculture Collaborators' Meeting on Grain Legume Improvement*, p.115. Ed. by R.A. Luse and K.O. Rachie

BOULTER, D., HASLETT, B.G., PEACOCK, D., RAMSHAW, J.A.M. and SCAWEN, M.D. (1977). 'The chemistry, function and evolution of plastocyanin.' In *International Review of Biochemistry – Series II. Plant Biochemistry.* Ed. by D.H. Northcote. Medical and Technical Publishing, Lancaster (in press).

BROHULT, S. and SANDEGREN, E. (1954). In *The Proteins* Vol.IIA, p.487. Ed. by H. Neurath and K. Bailey. Academic Press, New York

BURR, R. and BURR, F.A. (1976). *Proc. natn. Acad. Sci. U.S.A.* **73**, 515

BYERS, M. (1971). *J. Sci. Fd Agric.* **22**, 242

CARBON, J., BERG, P. and YANOFSKY, C. (1966). *Cold Spring Harb. Symp. quant. Biol.* **31**, 487

CLARKE, P.H. (1974). In *Evolution in the Microbial World*, p.183. Ed. by M.J. Carlile and J.J. Skehel. *Symp. Soc. gen. Microbiol.* **24**. Cambridge University Press

COLE, R.D. (1964). In *Proteins and their Reactions*, p.3. Ed. by H.W. Schultz and A.F. Anglemier. Avi Publishing, Westport, Connecticut

COOKE, G.W. (1976). In *Fertilizer Use and Protein Production*, p.29. Proc. 11th Coll. International Potash Institute, Ronne-Bornholm, Denmark 1975. IPI, Berne

CRANDALL, M., LAWRENCE, L.M. and SAUNDERS, R.M. (1974). *Proc. natn. Acad. Sci. U.S.A.* **71**, 26

DANIELSSON, C.E. (1952). *Svensk. Kem. Tidskr.* **64**, 43

DANIELSSON, C.E. (1956). *A. Rev. Pl. Physiol.* **7**, 215

DAVIS, P.H. and HEYWOOD, V.H. (1963). *Principles of Angiosperm Taxonomy*. Oliver and Boyd, Edinburgh and London

DAWSON, R. (1971). *Analyt. Biochem.* **41**, 305

DAYHOFF, M.O. (1972). *Atlas of Protein Sequence and Structure*, Vol.5. National Biomedical Research Foundation, Maryland, USA

DAYHOFF, M.O. (1973). *Atlas of Protein Sequence and Structure*, Vol.5, Supplement. National Biomedical Research Foundation, Maryland, USA

DELANGE, R.J. and SMITH, E.L. (1971). *A. Rev. Biochem.* **40**, 279

DERBYSHIRE, E., WRIGHT, D.J. and BOULTER, D. (1976). *Phytochemistry* **15**, 3

DICKERSON, R.E. and GEIS, I. (1969). *The Structure and Action of Proteins*. Harper and Row, New York

DILWORTH, M.J. (1969). *Biochim. biophys. Acta* **184**, 432

DIXON, M. and WEBB, E.C. (1964). *Enzymes*, 2nd edn. Longmans, Green and Co. London

ESAU, K. (1953). *Plant Anatomy*. John Wiley and Sons, New York

EVANS, I.M. and BOULTER, D. (1974). *J. Sci. Fd Agric.* **25**, 311

EWART, J.A.D. (1968). *J. Sci. Fd Agric.* **19**, 241

FAO (1970). *FAO Nutritional Studies*, No.24, Rome

FOLKES, B.F. and YEMM, E.W. (1956). *Biochem. J.* **62**, 4

FOX, D. (1964). PhD. thesis, University of Liverpool

FREY, K.J. (1951). *Cereal Chem.* **28**, 506

GENNIS, L.S. and CANTOR, C.R. (1976). *J. biol. Chem.* **251**, 734

GRABOR, P., BENHAMOU, N. and DAUSSANT, J. (1962). *Archs Biochem. Biophys. Suppl.* **1**, 187

GUIDOTTI, G. (1972). *A. Rev. Biochem.* **41**, 731

HARPER, A.E. (1964). *Mammalian Protein Metabolism*, Vol.2, p.87. Ed. by N.H. Munrow and J.B. Allison. Academic Press, New York

HARTLEY, B.S. (1974). In *Evolution in the Microbial World*, p.151. Ed. by M.J. Carlile and J.J. Skehel. *Symp. Soc. gen. Microbiol.* **24**, Cambridge University Press

HARVEY, D. (1970). *Tables of the Amino Acids in Foods and Feeding Stuffs*, 2nd edn. Commonwealth Agricultural Bureaux, Farnham Royal

HAUROWITZ, F. (1963). *Chemistry and Functions of Proteins.* Academic Press, New York

HEPLER, P.K. and PALEVITZ, B.A. (1974). *A. Rev. Pl. Physiol.* **25**, 309

HESLOP-HARRISON, J. (1975). *A. Rev. Pl. Physiol.* **26**, 403

INGLETT, G. (1972). In *Symposium: Seed Proteins*, p.176. Ed. by G.E. Inglett. Avi Publishing, Westport, Connecticut

INGLETT, G. (1975). In *The Chemistry and Biochemistry of Plant Proteins*, p.265. Ed. by J.B. Harborne and C.F. van Sumere. Academic Press, London

IUB (1973). *Enzyme Nomenclature.* Elsevier, Amsterdam

JACOB, F. and MONOD, J. (1961a). *J. molec. Biol.* **3**, 318

JACOB, F. and MONOD, J. (1961b). *Cold Spring Harb. Symp. quant. Biol.* **26**, 193

JAIN, H.K. (1971). *Indian Fmg.*, November

JOHNSON, V.A. (1976). Wheat proteins. *Proceedings of the International Symposium on 'Genetic Control of Diversity in Plants'*, Lahore, Pakistan, March 1976 (in press)

JULIANO, B.O. (1972). In *Symposium: Seed Proteins*, p.114. Ed. by G.E. Inglett. Avi Publishing, Westport, Connecticut

KLEINIG, H., DÖRR, I. and KOLLMANN, R. (1971). *Protoplasma* **73**, 293

KLEINIG, H., DÖRR, I., WEBER, C. and KOLLMANN, R. (1971). *Nature, New Biology* **229**, 152

KOFRANYI, E. (1972). In *Protein and Amino Acid Functions*, p.1. Ed. by E.J. Bigwood. Pergamon Press, Oxford

LAMPORT, D.T.A. (1970). *A. Rev. Pl. Physiol.* **21**, 235

LIENER, I.E. (Ed.) (1969). *Toxic Constituents of Plant Foodstuffs.* Academic Press, New York and London

LÜERS, H. and SIEGERT, M. (1924). *Biochem. Z.* **144**, 467

LUSE, R.A. (1975). *Proceedings of the International Institute of Tropical Agriculture Collaborators' Meeting on Grain Legume Improvement* p.110. Ed. by R.A. Luse and K.O. Rachie

McKEE, H. (1962). *Nitrogen metabolism in Plants.* Clarendon Press, Oxford

MILLERD, A. (1975). *A. Rev. Pl. Physiol.* **26**, 53

MOHR, H. (1966). *Photochem. Photobiol.* **5**, 469

MORITA, Y. and YOSHIDA, C. (1968). *Agr. Biol. Chem. (Jap.)* **32**, 664

MUNCK, L. (1972). *Hereditas* **72**, 1

NASH, A.M., KWOLEK, W.F. and WOLF, W.J. (1971). *Cereal Chem.* **44**, 183

NEOTEC INSTRUMENTS INC. (1971). Grain Quality Analyzer. Technical Bulletin. Rockville, Maryland, USA

NEWCOMB, E.H. (1967). *J. Cell Biol.* **33**, 143

OLSON, M.O.J. and LIENER, I.E. (1967). *Biochemistry* **6**, 105

OSBORNE, T.B. (1924). *The Vegetable Proteins*, 2nd edn. Longmans, Green and Co., London

PAULIS, J.W., BIETZ, J.A. and WALL, J.S. (1975). *J. Agric. Fd Chem.*, **23**, 197

PIRIE, N.W. (1971). *Leaf Proteins.* Blackwell, Oxford

PREAUX, G. and LONTIE, R. (1975). In *The Chemistry and Biochemistry of Plant Proteins*, p.89. Ed. by J.B. Harborne and C.F. van Sumere. Academic Press, London

PUSZTAI, A. (1966). *Biochem. J.* **101**, 379

ROBERTS, L.M. (1970). *The Food Legumes: recommendation for expansion and acceleration of research.* Rockefeller Foundation Report

ROSSMANN, M.G., MORAS, D. and OLSEN, K.W. (1974). *Nature, Lond.* **250**, 194

RYAN, C.A. (1973). *A. Rev. Pl. Physiol.* **24**, 173

SANGER, F. (1956). In *Currents in Biochemical Research*. Ed. by D.E. Green. Interscience, New York

SCANDALIOS, J.G. (1974). *A. Rev. Pl. Physiol.* **25**, 225

SINCLAIR, T.R. and DE WIT, C.T. (1975). *Science* **189**, 565

ST. ANGELO, A.J., YATSU, L.Y. and ALTSCHUL, A.M. (1968). *Archs Biochem. Biophys.* **124**, 199

STAHMANN, M.A. (1963). *A. Rev. Pl. Physiol.* **14**, 137

STARR, R.C. and JAENICKE, L. (1974). *Proc. natn. Acad. Sci. U.S.A.* **71**, 1050

STEENKAMP, D.J. and JOUBERT, F.J. (1967). *Jl. S. Afr. chem. Inst.* **20**, 81

STEGEMANN, H. (1975). In *The Chemistry and Biochemistry of Plant Proteins*, p.71. Ed. by J.B. Harborne and C.F. van Sumere. Academic Press, London

TOMS, G.C. and WESTERN, A. (1971). In *Chemotaxonomy of the Leguminosae*, p.367. Ed. by J.B. Harborne, D. Boulter and B.L. Turner. Academic Press, London

VILLEGAS, E. and MERTZ, E.T. (1971). *Chemical screening methods for maize protein quality at CIMMYT.* International Maize and Wheat Improvement Center, Mexico. *Res. Bull.* No.20

WALDSCHMIDT-LEITZ, E. and ZWISLER, O. (1963). *Hoppe-Seyler's Z. physiol. Chem.* **332**, 216

WALL, J.S. (1964). In *Proteins and their Reactions*, p.315. Ed. by H.W. Schultz and A.F. Anglemier. Avi Publishing, Westport, Connecticut

WEBER, C. and KLEINIG, H. (1971). *Planta (Berl.)*, **99**, 179

WEISE, L. (1969). In *Fertilisation*, Vol.2. Academic Press, New York

WHITEHOUSE, R.N.H. (1973). In *The Biological Efficiency of Protein Production*, p.83. Ed. by J.G.W. Jones. Cambridge University Press

WILDMAN, S.G. and BONNER, J. (1947). *Archs Biochem. Biophys.* **14**, 381

WILLIAMSON, A.R. (1969). In *Essays in Biochemistry*, Vol.5, p.139. Ed. by P.N. Campbell and G.D. Greville. Academic Press, London

WOLF, W.J. (1972). *Fd Technol., Champaign* **26**, 44

WORLD HEALTH ORGANIZATION (1973). Energy and Protein Requirements. *WHO Tech. Report Series 522*

WU, Y.V., SEXSON, K.R., CAVINS, J.F. and INGLETT, G.E. (1972). *J. agric. Fd Chem.*, **20**, 757

ZUCKER, M. (1972). *A. Rev. Pl. Physiol.* **23**, 133

Phytochemistry, 1976, Vol 15, pp 3–24. Pergamon Press. Printed in England.

REVIEW

LEGUMIN AND VICILIN, STORAGE PROTEINS OF LEGUME SEEDS

E. DERBYSHIRE, D. J. WRIGHT* and D. BOULTER

Department of Botany, University of Durham, Durham DH1 3LE, England

(*Received* 2 *June* 1975)

Key Word Index—Leguminosae; legumes; protein bodies; seed storage protein; seed globulins; legumin; vicilin.

Abstract—The structure, location in the seed and distribution of the storage protein of legume seeds are described. Methods which have been employed for the extraction, purification and characterisation of seed globulins are reviewed in relation to modern biochemical practice. The physical, chemical, and immunological characteristics of the classical legumin and vicilin preparations from *Pisum sativum* are summarised and the distributions of proteins with sedimentation coefficients and/or immunological determinants similar to those of legumin and vicilin, are tabulated. The structure and composition of various purified legumin and vicilin-type proteins from a variety of legumes, are compared.

INTRODUCTION

The biological role of the seed

Prior to the development of the seed habit, plant dispersal was the function of the spore. Once the megaspore was retained and fertilisation took place on the parent plant, the developing embryo could be nourished by the parent tissues, which was a considerable biological advantage, but the role of dispersal by the spore was lost; this function was taken over by the seed. The developing seed builds up a substantial store of reserve material and is shed as an independent propagule. There may be a period of dormancy but eventually, if conditions are favourable, germination will take place and the reserve material is used as a source of nitrogen and carbon compounds for the developing seedling; the nitrogen is mainly in the form of proteins, the carbon in the form of starch or oil, or both.

On average, the percentage of protein in cereal grains is 10–15% of dry matter, and in legume seeds 20–25%. In contrast, a typical vegetative organ, such as the leaf, has only 3–5% of its dry matter as protein. Seeds, therefore, and particularly legume seeds, are a high protein food source for man and his animals either directly or, more recently, as 'textured vegetable protein foods' for man.

A definition of storage protein

During the course of seed development from the fertilised ovule to maturity, protein is laid down at a variable rate. Usually the rate of protein deposition increases dramatically about one-third of the way through the development cycle [1–7]. Concomitantly, membrane-bound vesicles, protein bodies, appear which become filled with protein; this can be demonstrated by their staining properties in the light microscope [8] and by immunofluorescent techniques using the electron microscope [9]. Furthermore, it has been shown with a variety of plants that the protein of the protein bodies is degraded on germination [10–13], and serves as the source of nitrogen for the various new nitrogen compounds synthesized by the developing seedling [14]. This protein is, therefore, called storage protein, since it is laid down at one stage of the life cycle for future use at a metabolically more active stage.

Many proteins occur in seeds and the problem is to distinguish the storage proteins from those which have metabolic or structural roles etc. Since it has rarely been shown that a protein isolated from seeds occurred *in vivo* in the protein bodies, arbitrarily, proteins extracted from seeds, which constitute 5% or more of the total protein, may be suspected to be storage proteins. This review describes the structure, and distribution of the major storage proteins of legumes.

The 'classical' vicilin and legumin fractions of Osborne [15] *and Danielsson* [16]

Osborne and Campbell [17] showed that much of the protein of legume seeds was salt soluble globulin, and they were able to separate this fraction from *Pisum sativum* into 2 major fractions, legumin and vicilin, using repeated precipitation by dilution or heat treatment. The methods made use of the fact that legumin was less soluble than vicilin in dilute salt solutions, and also that legumin solutions were not heat-coagulable, whereas those of vicilin coagulated at 95°. Chemical analyses of legumin and vicilin showed that their nitrogen and sulphur contents were 18·04% N and 0·42% S and 17·4% N and 0·18% S respectively.

Osborne demonstrated that similar protein fractions could be extracted from other legume seeds, e.g. *Phaseolus vulgaris* [18–20] and *Glycine max* [21], but that their

* Present address: Unilever Ltd., Colworth/Welwyn Laboratory, Colworth House, Sharnbrook, Bedford MK44 1LQ.

chemical compositions differed. In view of this, Osborne [15] refuted Ritthausen's idea that only a comparatively small number of vegetable proteins occurred in nature.

The innovation of ultracentrifugation [22] and electrophoresis [23] as analytical tools for high molecular weight substances, led to a closer investigation of the seed globulins. Ultracentrifugation analysis of seed proteins was performed initially by Svedberg [24] and later and more extensively by Danielsson [16]. Danielsson [16, 25] separated globulins from *Pisum sativum* by the use of ammonium sulphate precipitation and isoelectric precipitation at pH 4·7, and on the basis of their solubility, heat stability and % nitrogen, he equated them with the legumin and vicilin preparations described by Osborne and Campbell [17]. He demonstrated [16, 26–28] that vicilin and legumin: (a) sedimented in the ultracentrifuge as single components at different protein concentrations with $S_{20,w}$ = 6·5–8·1 and 12·64 respectively, and had molecular weights of 186000 and 331000 respectively: (b) migrated as single components during free flow electrophoresis in presence of 0·2 M NaCl; as judged by zero mobility, the isoelectric point of legumin was pH 4·8 and vicilin pH 5·5: (c) were degraded during germination, i.e. functioned as storage proteins: and (d) that vicilin contained more lysine, and less tryptophan and glutamic acid than legumin. Tryptophan is easily destroyed during acid hydrolysis and as this method is usually employed, values for tryptophan are often missing from published amino acid compositions. Danielsson, however, emphasised the difference between the ratios of tyrosine to tryptophan in vicilin (10:1) and legumin (4:1).

Danielsson [16] also examined the globulin fraction of 34 species of legumes from 8 tribes of the Leguminosae, (Fabaceae [29]), and found, with a few exceptions, that they all contained two globulins with sedimentation coefficients of approximately 7S and 11S which, on the sole basis of their sedimentation values, he called vicilin and legumin. For many proteins, sedimentation velocity is dependent on the concentration, and the range of values between 6·6 and 8·3 recorded for the 7S component by Danielsson [16] from different sources, may have been due, in part, to differences in concentration of an equivalent protein in the preparations. However, it may also reflect some degree of heterogeneity in the protein contributing to the 7S peaks. Notable among the exceptions referred to above were *Acacia longifolia*, *A. penninervis*, *A. verticillata* and *Trifolium repens*, all of which contained a 7S component, but not an 11S, although the sensitivity of the technique is insufficient to demonstrate the presence of very small amounts of the 11S component.

Another method of comparing these globulins from different legumes has been to follow the immunochemical cross-reactivity of legumin and vicilin prepared by methods similar to those of Danielsson [27]. Earlier serological investigations particularly those of Kloz and coworkers (see e.g. Kloz [30]) using immunoelectrophoresis, suggested that vicilin might be more widely distributed than legumin. More recently, Dudman and Millerd [31] have investigated genera from eleven tribes of the Leguminosae [32] and, in agreement with Kloz, have shown proteins, immunologically related to vicilin and legumin, to be widely distributed in the Fabeae (Vicieae) and Trifolieae. However, by using several different immunological techniques they conclude, contrary to Kloz and co-workers, that legumin-like proteins are more widely distributed than vicilin-like proteins in the other tribes investigated; the Phaseoleae gave no cross-reactions with either protein from *Vicia faba*.

Detailed characterisation of the storage proteins requires their isolation in a form which satisfies stringent homogeneity criteria. The advent of column chromatography and gel electrophoresis demonstrated the heterogeneous nature of globulins prepared by the earlier techniques, and the traditional methods of protein purification, e.g. crystallisation and salt fractionation have now been either entirely replaced or supplemented by techniques which utilise more directly the fundamental differences between the proteins to be separated, viz. molecular weight, shape and charge. The following description and critique of the various methods used to purify globulins is necessary, in order to understand the status of the results found in the literature; the physico-chemical properties assigned to some of these proteins must be viewed with suitable caution.

THE STORAGE PROTEINS OF LEGUMES

Extraction and purification

The complex and diverse chemical nature of seeds is such that the methodology used for extracting proteins may not be universally applicable. However, certain precautions outlined below must be taken unless it is demonstrated for a particular seed source, that they are unimportant.

Freshly harvested, not stored, air-dried seeds should be used as starting material. If controlled growth environment facilities are not available to give adequate supplies of freshly harvested seeds, then stored seeds are used, but the results obtained should be compared with those from freshly harvested seeds, e.g. storage of seeds or of ground meal results in a decrease in the quantity of protein which can be extracted by water and salt [33, 34].

Testas should be removed prior to extraction since they often contain substances, e.g. pigments, phenolics, etc. which may later interfere with the isolation process. If freshly harvested seed is being used this presents no problem; if air-dried seed is being used it is often not possible to remove the testa readily, without soaking the seed and almost certainly some protein hydrolysis will occur during this process. Imbibition can be expedited by scarifying the seed and proteolysis can be reduced by carrying out the imbibition at low temperature.

Air-dried seed is usually milled and defatted prior to extraction, although it has been suggested that this step is not necessary in some cases even with seeds of high oil content [7]. It is sometimes necessary to remove pigment prior to protein extraction; Joubert [35–37] with *Lupinus* achieved this by washing with ethanol and water.

Initial extraction of the protein of freshly harvested seed is normally accomplished by homogenisation in a high-speed blender. Meal can be extracted in an ultrasonicator or by homogenisation. Protein–phytate interactions are enhanced by the presence of cations and hence the ionic strength of the extractant should be kept as low as possible. Precipitation of protein-phytate complexes takes place at acidic pH values, between pH 4·5 and pH 6·6 [38], and the pH should be maintained above

this range. On the other hand, association/dissociation reactions of globulins are often promoted by high pH values and non-enzymic deamidation is also enhanced at pH 10 [39]. Deamidation proceeds more rapidly, at pH 7·4, in the presence of phosphate ions as compared to the rate in borate and tris–glycine buffers.

Sulphydryl reagents such as 2-mercaptoethanol and dithiothreitol at low concentration, inhibit the formation of disulphide bridges between proteins with free sulphydryl groups and hence reduce polymerisation and subsequent insolubility; thus, it has been reported that 2-mercaptoethanol in the extraction medium increases the amount of protein extracted [40]. Aklylation of SH groups with *N*-ethylmaleimide will also prevent disulphide bridge formation, but little is known about the effects of these reagents on the internal sulphydryl and disulphide groups of the protein.

EDTA, a neutral chelating agent, has proved useful in enzyme isolation and it has also given increased extraction of the protein from seeds. Inclusion in the extraction medium of polyvinylpyrrolidone or other similar reagents which complex with and so remove phenolic plant constituents, have been shown to be necessary for the isolation of active enzymes and mitochondria from some plants [41–43], although these agents have not been used generally when extracting seed globulins.

One subject which perhaps deserves more attention than it is receiving at the present time, is that of proteolytic activity in seed extracts, since this could modify substantially the proteins under investigation. Proteinase inhibitors such as diisopropyl fluorophosphate and phenylmethylsulphonyl fluoride, have been shown to be effective in overcoming this problem in the isolation of other proteins [44–48], and it is probable that their use will be extended to the isolation of the seed globulins, if only to determine whether the observed heterogeneity in some protein preparations is inherent or due to proteolytic activity. Proteolytic degradation could be minimised by conducting extraction and purification procedures with small amounts of material at low temperatures, but the cryoprecipitation of some globulins [16, 49, 50] may make this procedure impractical. Proteolysis can be minimised also by the use of phenol-containing solvents [51]. The value of phenol as an efficient extractant of plant protein has been known for a considerable time [52], and its particular value in the preparation of seed protein free from loosely bound carbohydrate has been emphasised recently by Pusztai [51]. Interactions between proteins and other polyelectrolytes are decreased in phenol containing solvents and proteins can be recovered in their native state.

Proteins which agglutinate animal erythrocytes are found in many legume seeds and often several agglutinins with different specificities exist in the same species [53]. Some of these proteins have also been shown to interact with other animal cells via the glycoproteins of the membranes. Since glycoproteins occur extensively in seeds, interaction between these and the seed agglutinins, although not investigated so far, could well take place in seed extracts. This is especially important in the present context, since some storage globulins are glycoproteins.

Even when extracts have been prepared under these conditions, it is advisable as Danielsson [54] recommended, to remove other non-protein materials from the extract as soon as possible: this may be effected by molecular sieving, by cryoprecipitation [49], or by salting out the proteins by the addition of ammonium sulphate [55]. The use of ammonium sulphate allows the simultaneous fractionation of the proteins and results in precipitates which can be redissolved directly or stored in a relatively stable condition. However, fractionation is usually incomplete and it is necessary eventually, to remove the salt either by chromatography or dialysis; the aggregation of protein molecules which occurs during precipitation, may, itself, enhance protein–protein interaction despite the presence of the salt.

Further purification of the crude globulin preparation can be obtained by several procedures, but purification of individual proteins should be carried out as quickly as possible to reduce interactions: if required, samples between purification stages can be stored frozen, freeze-dried or under a saturated solution of ammonium sulphate. The most successful method of purifying legumin (11S protein) has been its isoelectric precipitation leaving the vicilin (7S protein) in solution [16, 56]: the exact pH used varies from one protein source to another, but is in the region of 4·7. For complete separation this procedure must be repeated several times [16, 57], but its greatest disadvantage is that some of the precipitated protein will not redissolve when the pH is readjusted to the original value [55, 58–60]. The formation of disulphide-bridged polymers has been connected with this phenomenon [34], but another explanation for it could be the production during isoelectric precipitation of local areas of high acidity which leads to the dissociation of the globulins.

A modification of this method, termed zonal isoelectric precipitation [61–64], seems to overcome these difficulties. Because of the multiple precipitations involved and also since the 11S protein is not subjected to pH values below the isoelectric point, this technique affords relatively pure, readily solubilised legumin preparations. The vicilin fraction obtained by this method is contaminated with legumin and is also impure [65].

Even with those sources where another protein is present with a similar isoelectric point to that of legumin, e.g. the 7S protein of soyabean [66], some separation may still be possible by the selection of a suitable ionic strength. Since phytate has been reported as a major contaminant of globulins prepared by isoelectric precipitation [38], it is advisable to remove as much of it as possible from the extract by dialysis or by treatment with Dowex ion-exchange resin [38] before attempting isoelectric precipitation.

Gel filtration can be utilised to remove low molecular weight proteins and also non-protein contaminants during the purification of globulins. This technique has proved far less successful in the separation of vicilin and legumin however, than would at first appear from the difference in their sedimentation coefficients (7S and 11S) and molecular weights (186000 and 330000). Hasegawa *et al.* [67] separated two 7S proteins and a 6S protein from the 11S protein on Sephadex G-200, using extracts of *Glycine max*, but a third 7S protein co-eluted with the 11S protein. Some 11S globulins have been purified using this method, e.g. the 11S protein from seeds of *Helianthus annuus* was separated from its 7S dissociation product by using repeated chromatography on Sephadex G-200 columns [68].

Koshiyama [69], using a globulin preparation from *Glycine max* was unable even with the use of 200 cm-long

columns, to completely separate the 11S and 7S components; he did, however, obtain a pure preparation of the 11S protein, glycinin, by taking only those fractions corresponding to the leading edge of the eluted peak. The failure to achieve reasonable separation of these two components on dextran gels is probably due to the fact that they both have approximately the same Stokes' radius [69]. Another complication is illustrated by the results obtained with the 7S globulins of *Pisum sativum*. These were retarded relative to standard proteins during thin-layer chromatography on cross-linked dextran [70] and this anomalous behaviour led to an underestimate of their molecular weights [71].

Interaction between proteins and cross-linked dextran may lead to the retention of proteins during molecular sieving and to selective losses. The degree of cross-linkage of the dextran has a differential effect, e.g. Con A is not retained by more highly cross-linked dextrans [72].

A more effective method of separating the 7S and 11S proteins is by chromatography on hydroxylapatite (calcium phosphate) gels [73, 74]. The degree of resolution afforded by this method is enhanced significantly by operating at low elution rates [75].

A more recent innovation has been the use of affinity chromatography with either monospecific antibodies or agglutinating proteins such as Concanavalin A, attached to solid supports. In the case of globulins of *Glycine max* the 7S protein contains more carbohydrate than the 11S [76], and is adsorbed during passage through a column of Con A Sepharose [77]; however, the carbohydrate content of seed globulins varies considerably and the future usefulness of the technique in the separation of 7S and 11S globulins from other species, is difficult to assess. Two 7S globulins, β conglycinin and γ conglycinin, from *Glycine max* have been separated, one from another, by use of monospecific antibodies bound to Sepharose 6B (Koshiyama, I., personal communication). The specificity of this type of affinity chromatography suggests that it will be employed extensively in the future, even though the production of antibodies is time-consuming.

Ion-exchange chromatography on DEAE-cellulose and DEAE-Sephadex has been used in the preparation of globulins from a wide range of seeds including *Vicia sativa* [78], *Pisum sativum* [79], *Phaseolus aureus* [80], *Glycine max* [81, 82], *Brassica napus* [83], and *Prunus amygdalus* [84]. Because of its sensitivity, this technique affords better results when employed in the final stages of purification. It has also been used as a criterion of homogeneity [85] and as a form of fingerprinting technique to compare preparations of the same protein obtained by different procedures [86].

Sucrose density gradient centrifugation has proved useful for the preparation of small but relatively pure quantities of the 11S globulins of *Glycine max* [7], *Phaseolus aureus* [80] and *Vicia faba* [2]. Electrophoretic procedures have proved useful in the separation of particular globulins from other proteins, e.g. Glycoprotein II of *Phaseolus vulgaris* was partially purified by use of free flow electrophoresis [87] and legumin and vicilin of *Pisum* have been isolated from gels after electrophoretic separation (Thomson, J., personal communication).

The progress of the purification is monitored by procedures such as gel electrophoresis. Once purification is thought to be complete, a variety of methods is used to establish the homogeneity of the preparation, although none of these, either singly or in combination, is necessarily definitive proof of the presence of only one protein. The ultracentrifuge still remains a powerful analytical tool for this purpose, since it yields information on both the purity of the preparation and on molecular size. However, the powers of resolution of the technique are inferior to those of electrophoresis. Free flow electrophoresis has been employed to characterise globulins in a number of laboratories [27, 74, 88–90], but is time-consuming and today the more convenient procedure of electrophoresis using a stationary support is employed. Solid supports such as paper, cellulose acetate membranes, starch and polyacrylamide gels have been used [79, 90–92] and of these polyacrylamide gels usually give the greatest resolution; they can also be used on a preparative scale. The pH of electrophoretic buffers can be selected to modify the net charge on the proteins and 0·5 M NaCl [69] and 2-mercaptoethanol [77] can be included in the gel system to prevent the globulin undergoing dissociation or aggregation during electrophoresis. Isoelectric focusing in solvent stabilised by the incorporation of sucrose [93] or on polyacrylamide gel [94], has also been used and affords a very high degree of resolution. Its usefulness with regard to the seed globulins is limited, since these proteins tend to precipitate at their isoelectric points and also since the technique only operates effectively at very low ionic strengths. Under those conditions it is necessary to solubilise the seed globulins with reagents such as urea, and these affect the quaternary structure of the protein. Nevertheless, the procedure has found application with the more soluble globulins, e.g. glycinin [69].

Immunoelectrophoresis can be used to check the purity of preparations since it is an excellent method for separating mixtures of legumin and vicilin [31, 95, 96]. Extracts of *Vicia faba*, for example, give clearly separated bands of vicilin and legumin as well as a third protein of intermediate electrophoretic mobility. However, the positions of immunoelectrophoretic bands is influenced by the concentration of the protein antigens, and it is not always possible, therefore, to be sure when an analysis gives just a single band, whether it represents vicilin or legumin. The identity of bands may be confirmed, however, by the use of the Osserman technique [97] which combines both immunodiffusion and immunoelectrophoresis, although this method is relatively insensitive. Various artefacts are possible in the immunological analysis of unfractionated seed extracts and these are discussed by Kloz [30] and by Dudman and Millerd [31].

Often a single *N*-terminal amino acid is used as a measure of the homogeneity of a protein preparation. Since the seed globulins consist of more than one different polypeptide chain in this case the test of purity is that the number of different *N*-terminal amino acids found should be consistent with the number of different polypeptide chains in the molecule. Three different labels, dinitrophenyl-(DNP-), phenylisothiocyanate and dansyl-(DNS-), have been employed to determine the *N*-terminal amino acids of storage globulins and the first of these has been used most frequently. It is anticipated however, that the greater sensitivity (× 100) of the DNS-method and the greater stability of the labelled derivatives during protein hydrolysis, will lead eventually to the adoption of the latter as the preferred method of

N-terminal analysis, especially when this is applied to the small quantities of protein eluted from gels.

Methods of characterisation

In addition to characteristics such as size (ultracentrifuge), isoelectric point (electrophoresis) and number of subunits (electrophoresis in sodium dodecyl sulphate) which may be determined as by-products of homogeneity checks, further characterisation of the globulins involves the determination of their chemical composition (e.g. amino acid composition, carbohydrate content, etc.) and the separation, isolation and characterisation of their constituent subunits. Often outdated methods of chemical analysis have been used and these should now be replaced by modern methods such as described in Methods in Enzymology [98], to which reference should be made.

Dissociation of globulins can be brought about by alteration of pH [33, 74, 99, 100] and by exposure to dissociating agents, e.g. urea, formamide, guanidine hydrochloride, detergents, β-mercaptoethanol [101] and dithiothreitol [102]. The latter 2 reagents are employed to disrupt disulphide bridges and the other treatments result in cleavage of hydrogen bonds. Disulphide bonds are not always located near the surface of the molecule and prior disruption of hydrogen bonds is usually necessary to ensure that all disulphide bonds are exposed to the thiol reagent. The total cleavage of hydrogen bonds in large proteins also offers difficulties and high concentrations of dissociating agent are necessary to complete and maintain separation of subunits. When guanidine hydrochloride is employed, a final reagent concentration of 4 M is adequate but a minimum concentration of 6 M urea is often essential. Sodium dodecylsulphate (SDS) can be bound by protein to form 2 types of complex which differ in their stability; the more stable complex requires a binding ratio of 1·4 g/g protein [103]. It is necessary, therefore, to employ SDS under conditions which ensure this binding ratio is achieved, for example by introduction of SDS at a concentration (w/w) some ten times that of the protein to be dissociated or by dialysis of the protein against a much larger volume of a medium of lower SDS concentration [104]. Even at the high binding ratio, reassociation of dissociated subunits has been reported, although these complexes usually involve only a small proportion of the protein. Separation of the dissociated subunits can be accomplished by utilising differences in either their charge or molecular weight. In general, our experience has been that better resolution and higher recoveries are obtained when any disulphide–sulphydryl groups of the subunits are permanently blocked, for example, by carboxymethylation. Modifications of this kind however, may preclude their subsequent use for some physical and biological studies.

Ion-exchange chromatography conducted in dissociating medium has been used successfully for the preparation of the component subunits of several 11S proteins. Dlouha *et al.* [105] fractionated *S*-sulphoedestin prepared from *Cannabis sativa* into 2 components on DEAE-cellulose using phosphate buffers of varying molarities in 6 M urea. The same ion-exchange resin was used in 4 M urea with a salt gradient to separate the subunits of legumin of *Vicia sativa* [106] and glycinin [107]. Separation of the subunits of CM-legumin from *Vicia faba* was achieved by chromatography in 6 M urea on the resin AGI-X2 [64]. Masaki and Soejima [108] separated three fractions from urea-dissociated 7S globulin of *Glycine max* on DEAE-cellulose. Ghetie and Buzila [70] separated two components from cryoprecipitated 7S globulin of *Pisum sativum* by using Sephadex G-200 in urea, and Grant and Lawrence [79] also using *P. sativum* 7S protein (vicilin) isolated a series of urea-dissociated subunits from polyacrylamide gels and determined their amino acid compositions and *N*-termini.

Wright and Boulter [64] successfully applied the technique of preparative SDS polyacrylamide gel electrophoresis to the separation of the acidic and basic subunits of *Vicia faba*, but even though this work demonstrates the feasibility of subunit separation on the basis of molecular weight, little has been attempted in this area. Goding *et al.* [109] have prepared a glycoprotein subunit from the 12S globulin of *Brassica napus* using gel filtration on Sephadex G-100 in 2 M urea at pH 2·8. This component represented the largest subunit, on the basis of its elution volume, and had a $S^0_{20,w}$ value of 2·7 S and *N*-terminal glycine; three other components were also separated by this method. Lastly, Catsimpoolas [93] isolated 6 subunits, 3 acidic and 3 basic, from glycinin by isoelectric focusing in urea-dithiothreitol medium.

Characteristics and distribution

Since the biological role of seeds is much the same throughout the flowering plants, it would appear likely that specific storage proteins have evolved and that equivalent homologously related proteins might exist in a wide range of plants. These proteins would be distinguishable from other proteins by their size, subunit composition and dissociation behaviour, which is related to the need to have a structure adapted to the drying out and wetting up of the seed, and by their amino acid composition, which is related to their function as storage compounds. Furthermore, their existence in different plants should be revealed by serological cross-reactivity. Thus, serological studies have shown that proteins immunologically related to vicilin and legumin of *Pisum sativum* and *Vicia faba* occur in other members of the Fabeae and Trifolieae. On the other hand, lack of cross-reactivity between *Vicia faba* and members of the Phaseoleae does not imply that homologous proteins do not occur there also, since a single change in the amino acid sequence of a protein can have a drastic effect serologically [110, 111]. This section, therefore, presents a representative selection of the data on the structure and composition of legume storage globulins, in order to itemise their general basic characteristics and to assess the extent of their distribution in higher plants.

LEGUMIN

Many dicotyledonous seeds have been shown to contain proteins which occur in large amounts and which have sedimentation coefficients of approximately 11S and/or MW's in the region of 300000–400000. A comprehensive list is given in Table 1, together with their known physical characteristics. Sedimentation coefficients of supposedly identical globulins sometimes differ considerably, e.g. 10·8–14·6 for the peanut globulin, arachin; however, equivalent globulins have not been examined always at the same protein concentration(s) or in the

Table 1. The distribution of legumin-like seed proteins

Species	Sedimentation coefficient	MW (daltons)	Partial specific volume	Diffusion constant ($cm^2 sec^{-1} \times 10^{-7}$)	Frictional ratio	Reference
Legumes						
Acacia alata	11·63(S_{20})					[16]
A. decipiens	12·70(S_{20})					[16]
A. saligna	13·67(S_{20})					[16]
Arachis hypogaea	13·05(S_{20})					[16]
	14·6(S_{20})	396000	0·72	3·2	1·38	[115]
	13·2(S_{20})	340000				[116]
	13·3($S_{20,w}$)					[117]
	15·5(S_{obs})*					[50]
	13·36($S^{\circ}_{20,w}$)					[85]
	14·6(S°_{20})	330000	0·72	3·86(D_{20})	1·216	[118]
	12·0($S_{20,w}$)					[86]
	14·7($S^{\circ}_{20,w}$)	350000	0·721	3·67($D_{20,w}$)	1·252	[119]
Astragalus galegiformis	13·17(S_{20})					[16]
Cytisus laburnum	14·02(S_{20})					[16]
C. supinus	13·38(S_{20})					[16]
Dolichos lablab	11·66(S_{20})					[16]
	12·67(S_{20})					[120]
Ervum lens	13·18(S_{20})					[16]
Genista tinctoria	13·34(S_{20})					[16]
Glycine max	13·1(S_{20})					[16]
	12·2($S_{20,w}$)	345000–363000	0·719	2·91($D_{20,w}$)	1·55	[55]
	12·2($S_{20,w}$)	309000–322000	0·715	3·48($D_{20,w}$)	1·40	[69]
	12·39($S_{20,w}$)					[121]
	14·0($S_{20,w}$)	380000		3·57		[122, 123]
	11·80($S_{20,w}$)					[73]
Lathyrus clymenum	13·00(S_{20})					[16]
L. odoratus	12·00(S_{20})					[16]
L. sativus	13·04(S_{20})					[16]
L. silvestris	12·97(S_{20})					[16]
Lotus tetragonolobus	13·07(S_{20})					[16]
Lupinus albus	12·29(S_{20})					[16]
	12·6(S°_{20})	393000		2·99(D_{20})	1·46	[124]
L. angustifolius	13·05(S_{20})					[16]
	11·6(S°_{20})	336000		3·16(D_{20})	1·46	[35]
L. luteus	11·53(S_{20})					[16]
	11·6(S°_{20})					[125]
	11·4($S_{20,w}$)					[126]
[illegible]	12·20(S_{-})					[16]
Medicago sativa	11·41(S_{20})					[16]
Phaseolus aureus	11·3*					[80]
P. coccineus	12·16(S_{20})					[16]
P. nanus	10·10(S_{20})					[16]
P. vulgaris	11·02(S_{20})					[16]
	11·6($S_{20,w}$)	340000				Derbyshire, E. (unpubl.)
Pisum sativum	12·6($S_{20,w}$)					[79]
	12·64(S°_{20})	330000	0·735	3·49(D_{20})		[16]
	13·1(S_{20})	398000	0·733	2·99(D_{20})	1·43	[127]
	12·1(S_{20})	388000				[116]
	12·3(S_{20})					[128]
	13·7($S_{20,w}$)	410000		3·02(D°_{20})	1·43	[129]
Trifolium hybridum	12·90(S_{20})					[16]
T. pratense	11·22(S_{20})					[16]
Vicia faba	11·4($S^{\circ}_{20,w}$)	328000				[64]
	11·5(S_{20})					[130]
V. sativa	12·9($S^{\circ}_{20,w}$)	360000	0·722	3·20		[88, 131]
	11·48(S_{20})	208000				[132]
Vigna unguiculata	11·2($S_{20,w}$)	320000				Derbyshire, E. (unpubl.)
Non-Legumes						
Acanthosicyos horrida	12·6(S_{20})	377000		3·15(D_{20})	1·41	[112]
Aesculus hyppocastanum	12·9(S_{20})	430000	0·748	2·9		[113]
Anacardium occidentale	12·85($S_{20,w}$)	238000	0·737	4·75–5·0 ($D_{20,w}$)	1·07	[114]
		260000				
Balanites aegyptica	12·3(S_{obs})		0·685			[133]
Bertholletia excelsa		303000				[134]
	13·3(S_{20})	295000		4·26		[135]
	11·78(S_{20})	212000	0·743		1·02	[136]
Beta vulgaris	13·9(S_{20})	~250000		5·3		[137]
Brassica napus	12(S°_{20})					[109]
B. juncea	12($S^{\circ}_{20,w}$)					[138]
B. nigra	11·8($S^{\circ}_{20,w}$)					[138]
B. hirta	11·7($S^{\circ}_{20,w}$)					[138]
Cannabis sativa	13·2(S°_{20})	334000				[116]
	12·8($S_{20,w}$)	212000		5·6($D^{\circ}_{20,w}$)		[139]
	12·8(S_{20})	309000	0·744	3·93(D_{20})		[135]
		360000				[134]
		335000				[140]
		300000				[105]
Citrullus vulgaris		343000				[141]
Citrus aurantiaca	11·38(S_{20})	210000				[142]

Continued—

Table 1—*cont.*

Species	Sedimentation coefficient	MW (daltons)	Partial specific volume	Diffusion constant ($cm^2 sec^{-1} \times 10^{-7}$)	Frictional ratio	Reference
Cucurbita maxima	12·1					[143]
		340000				[144]
Fagopyrum esculentum	13·0($S^0_{20,w}$)	270000				[145]
Gossypium barbadense	13·0($S^0_{20,w}$)					[146]
Helianthus annuus	11·9(S^0_{20})	343000		3·24(D^0_{20})	1·42	[147]
	11·8($S^0_{20,w}$)					[68]
Nicotiana sp.		325000				[148]
		350000				[134]
Prunus avium, *P. cerasus*, *P. domestica*		~300000				[149]
P. amygdalus	12·3($S_{20,w}$)		0·733			[84]
	12·5(S_{20})	329000	0·746	3·62(D_{20})		[135]
		330000				[150]
		~300000				[151]
		206000				[152]
	11·41(S^0_{20})	208000	0·746		1·03	[136]
Ricinus communis	12·9(S^0_{20})	332000		3·62(D^0_{20})	1·28	[153]
Sesamum indicum	13·39($S^0_{20,w}$)					[154]
	12·7($S^0_{20,w}$)	450000	0·735	2·6($D^0_{20,w}$)	1·5	[155]
Sinapis alba	12·7(S_{20})					[25]
Telfairea pedata	13·3(S^0_{20})	311000		3·99	1·19	[156]

* Not corrected to water of 20°. † Determined by density gradient centrifugation.

same media, and the discrepancy is due, in part, to the inconsistent reporting of the data, i.e. as S_{20}, $S_{20,w}$, or $S^0_{20,w}$ values.

A globulin component, δ-globulin, of similar size to legumin, is also present in some monocotyledons. Both Quensel [157] and Danielsson [16] found a component with sedimentation constant of 12·0*S* and MW of 300000, in the embryos of *Hordeum vulgare*. Of the monocotyledons he investigated, Danielsson found this δ-globulin only in *H. vulgare* and then in small and variable amounts. Since then, Pence and Elder [158] have observed a δ-globulin in *Triticum aestivum*, which has an $S_{20,w}$ value of 10·64*S*, and Morita and Yoshida [159] have reported the presence of a δ-globulin in the embryos of *Oryza sativa*. *Cocos nucifera*, another monocotyledon, has a seed protein with S^0_{20} value of 11·37 [160].

The amino acid composition of a representative sample of these 11S proteins (Table 2), suggests they have a storage role and may be equivalent proteins, since they all have a high content of amides (glutamic acid–glutamine, aspartic acid–asparagine and arginine). A wide variation is seen in the reported values for their cysteine and methionine contents, but the most likely causes of this variation are the low level of these amino acids and the difficulties associated with their accurate determination. This may account for the fact that certain globulins have been reported as being completely devoid of cysteine [2, 50]. The glycoprotein nature of the 11S globulin of *Glycine max* [162–164], *Arachis hypogaea* [86, 165], *Phaseolus aureus* [80], *Vicia faba* [57], *Phaseolus vulgaris* (Derbyshire, E., unpublished results) and *Brassica* spp. [109], has been investigated. Generally, the carbohydrate content is low (<1%) and is mainly in the form of neutral sugars, although glucosamine [80] and galactosamine [109] have geen identified in two cases.

It is apparent from the data assembled in Tables 1 and 2 that there are many similarities between the 11S protein isolated from these different sources. However, this information is insufficient to decide if most legume seeds have an homologous counterpart to the legumin

Table 2. The amino acid composition of some 11S seed globulins, recalculated as mol % from the original data where necessary

	Species														
	Arachis hypogaea	*Cicer arietinum*	*Glycine max*		*Phaseolus vulgaris* Derbyshire, E. (unpubl)	*Pisum sativum*			*Vicia faba*			*Brassica napus*	*Citrullus vulgaris*	*Cucurbita maxima*	*Helianthus annuus*
Reference / Amino acid	[85]	[161]	[167]	[162]		[161]	[79]	[64]	[2]	[161]	[57]	[109]	[141]	[144]	[68]
Asp	13·3	9·84	11·71	12·01	9·5	12·70	11·9	12·86	12·3	10·6	11·27	10·04	9·9	10·1	10·80
Thr	2·8	4·33	3·82	5·04	4·9	3·23	3·5	3·25	3·7	4·28	4·16	4·91	4·0	2·9	3·59
Ser	6·0	6·79	5·97	8·50	7·3	5·90	6·8	6·60	7·4	6·50	6·00	5·54	6·6	7·1	4·96
Glu	19·4	13·87	21·43	19·17	13·1	16·50	20·1	18·03	19·9	16·40	14·94	18·00	15·6	16·1	20·07
Pro	5·5	nd	6·46	5·71	5·1	nd	5·4	6·16	5·4	nd	8·08	6·21	3·6	4·0	5·12
Gly	7·4	7·52	7·48	7·24	8·0	6·20	7·5	6·91	7·7	7·40	7·35	9·64	8·0	7·5	7·51
Ala	6·2	6·36	6·20	5·37	6·9	5·78	6·0	5·91	6·3	6·10	5·51	6·80	7·2	6·9	6·45
Val	5·1	5·52	5·22	4·93	7·0	4·77	5·1	5·21	5·1	4·91	5·39	6·21	6·1	6·4	6·22
½ Cys	0·7	0·82	0·63	1·50	0·6	0·72	0·6	nd	0·0	0·80	1·52	0·57	0·8	0·7	2·22
Met	tr	1·03	1·30	1·09	1·5	0·53	nd	0·76	0·3	0·59	0·36	1·54	2·5	2·1	1·97
Ile	3·9	4·41	4·14	5·06	4·9	4·06	4·0	4·13	4·3	3·98	4·65	4·74	4·7	5·1	4·91
Leu	7·1	7·99	7·05	6·30	8·7	7·98	7·6	8·08	8·5	7·84	7·84	8·36	7·1	8·2	6·86
Tyr	3·3	2·54	2·66	2·69	2·9	2·56	1·7	2·67	2·1	2·61	3·67	2·23	2·5	2·8	2·13
Phe	4·9	4·90	4·63	3·80	3·6	4·22	3·6	3·40	3·2	3·56	3·80	3·89	5·1	4·7	4·99
Lys	2·1	5·19	3·93	3·76	7·8	4·40	5·2	5·03	4·2	4·57	4·53	3·80	2·3	2·9	1·91
His	2·0	2·43	1·73	1·88	3·0	2·45	2·6	1·96	2·4	2·44	2·57	1·77	2·1	1·9	2·33
Arg	10·2	7·55	5·64	5·12	4·8	7·50	8·6	9·05	8·0	7·95	7·59	5·03	11·9	10·7	6·94
Trp	nd	nd	nd	0·83	0·7	nd	nd	nd	nd	nd	0·75	0·71	nd	nd	1·04

nd = not determined. tr = trace amounts.

Table 3. Characteristics of subunits of 11S globulins

Species	Subunit designation	Sedimentation coefficient	MW	No. of subunits in 11S protein (Calculated)	N-terminal amino acids	pI	Reference
Cannabis sativa	*A*	2·45($S_{20,w}$)	27000	6	—	—	[105]
	B	1·85($S_{20,w}$)	23000	6	Gly	—	
Glycine max	$A_{1,2,3}$	—	37200	6	{Gly, Leu/Ile,}	4·75, 5·15, 5·40	[82, 93, 166]
	$B_{1,2,3}$	—	22300	6	Phe	8·0, 8·25, 8·50	
	Alkaline	1·21($S_{20,w}$)	{30000–*}	—	Gly	8–9	[121, 167]
	Acidic$_{\alpha,\beta}$	—	{35000}	—	Leu, Phe	4·8–5·2	
Vicia faba	$\alpha_{1,2}$	—	36200	6	Leu, Thr	Acidic	[64, 65]
	$\beta_{1,2,3}$	—	{20100; 20900–23800}	6	Gly	Basic	
Vicia sativa	*A*	1·4($S^0_{20,w}$)	24300	6	Gly	Basic	[106, 130]
	B	2·28($S^0_{20,w}$)	37600	4	Leu	Acidic	
	C	2·25($S^0_{20,w}$)	32600	2	Thr	Acidic	

* Average MW.

of *Pisum sativum* and it is, therefore, necessary to examine the evidence of their subunit structures.

The molecular size of the globulins has always been indicative of the existence of a subunit structure and this has been substantiated by data from dissociation studies and *N*-terminal amino acid analysis. However, there are only a few instances in which subunit structure has been examined in detail, or where subunits have been isolated for further characterisation; the relevant data have been collected together in Table 3. This is not an exhaustive list but gives those subunit structures which have been studied in most detail; the data demonstrate the striking similarities in the properties of the subunits of the 11S globulin from different sources. Each globulin has both acidic and basic subunits with MW's between 27000–37000 and 20000–24000 respectively. The subunit MW's presented in Table 3 may not be directly comparable however, as they were obtained by two different techniques, namely SDS gel electrophoresis and ultracentrifugation, and the range of MW's may prove to be narrower than that reported in the table. The variation in the number and characteristics of both acidic and basic subunits reported by different workers (see Table 3), is probably attributable to the different resolving powers of the analytical techniques they employed. Glycine is the *N*-terminal amino acid of the basic subunits of all the globulins and leucine is one of the *N*-termini of the acidic subunits of three of them (Table 3). These two amino acids also figure predominantly in the *N*-terminal amino acids of other 11S globulins [79, 109, 119, 161], and consequently this may represent a general property of legumin-like storage globulins.

An indication of how closely the subunits of these globulins resemble one another can be gauged from amino acid composition data. This is presented in Tables 4a and b for the acidic and basic subunits respectively, of legumin of *Vicia faba* and *Vicia sativa*, edestin of *Cannabis sativa* and glycinin of *Glycine max*. There is a remarkable similarity in the amino acid composition of subunits of the same type and significant differences between the compositions of the acidic and basic subunits. Thus, the acidic subunits are all richer in glutamic acid–glutamine than the basic subunits, while the latter contain more alanine, valine and leucine. The variable amino acid composition of both the acidic and basic subunits from different sources may indicate that only parts of the molecule have been conserved, presumably those parts most intimately involved in maintaining the tertiary and quaternary structure of the protein. Confirmation of this must await primary sequence investigations.

Table 4a. The amino acid composition (mol %) of acidic subunits of some 11S seed globulins

Amino acid	Species: *Glycine max** [166]	*Vicia faba* [64]	*Vicia sativa*† [131]	*Cannabis sativa* [168]
Asp	12·66	13·00	12·15	11·85
Thr	3·36	3·07	2·79	3·36
Ser	5·92	6·47	6·52	6·71
Glu	24·53	22·11	20·92	19·10
Pro	7·08	5·46	5·36	3·48
Gly	7·74	7·75	7·63	8·37
Ala	3·77	3·78	4·90	5·51
½ Cys	nd	nd	1·05	1·28
Val	3·91	3·41	3·55	5·93
Met	0·49	0·74	0·68	1·33
Ile	4·05	4·49	5·01	5·06
Leu	5·71	6·52	6·34	6·22
Tyr	2·15	2·39	2·63	2·82
Phe	3·42	2·93	3·55	3·48
His	2·57	2·62	2·71	2·11
Lys	6·27	4·90	3·90	1·66
Arg	6·37	10·33	9·21	11·19
Trp	nd	nd	nd	0·54

nd = not determined. * Average values, assuming A_1, A_2 and A_3 are present in equimolar amounts. † Calculated from values for *B* and *C* subunits using the *B*–*C* ratio of 2:1.

Table 4b. The amino acid composition (mol %) of basic subunits of some 11S seed globulins

Amino acid	Species: *Glycine max* [166]*	[167]	*Vicia faba* [64]	*Vicia sativa* [131]	*Cannabis sativa* [168]
Asp	13·77	12·96	11·60	12·57	12·13
Thr	4·27	4·68	4·25	4·39	4·30
Ser	7·05	6·94	6·86	6·45	5·87
Glu	15·41	14·40	10·21	9·51	11·53
Pro	5·45	5·85	4·66	5·12	3·81
Gly	6·64	6·86	6·03	6·17	6·14
Ala	7·35	6·34	9·78	9·92	8·86
½ Cys	nd	1·62	nd	0·91	0·98
Val	6·15	6·81	9·57	8·73	8·10
Met	0·87	1·10	0·46	0·46	2·56
Ile	4·44	4·80	3·54	3·98	4·51
Leu	8·99	9·24	10·94	10·15	7·99
Tyr	2·77	2·79	3·14	3·06	3·10
Phe	4·63	5·35	3·29	3·52	4·73
His	1·73	1·51	1·36	1·60	1·63
Lys	4·84	3·14	5·95	4·16	3·10
Arg	5·64	5·43	8·35	8·28	9·79
Trp	nd	nd	nd	1·01	0·87

nd = not determined. * Average values, assuming B_1, B_2 and B_3 are present in equimolar amounts.

The only information so far is that of Dlouha *et al.* [169], who examined the structure of peptides isolated from a tryptic hydrolysate of the A-chain of edestin. Their results indicate localisation of groups of acidic, neutral and basic amino acid residues.

A universal legumin-like storage globulin could be envisaged as having a structure compatible with the above characteristics, i.e. it would consist of twelve subunits with equimolar amounts of two types, acidic and basic, which differ in their MW's. these being 27000–37000 and 20000–24000 respectively. However, there is some evidence of greater complexity. Thus, Catsimpoolas and Wang [170] observed numerous bands when the acidic or basic subunits of glycinin were subjected to analytical scanning isoelectric focusing in urea–dithiothreitol media. They attributed the multiplicity of components mainly to differences in the primary structure of the subunits, although no conclusive evidence was presented to substantiate this statement. Similarly, isoelectric focusing of either acidic or basic subunits of CM-legumin of *Vicia faba* in 6 M urea revealed approximately 8 components [65] (Wright, D. J., unpublished data), although in this case variation in charge modification as a result of carboxymethylation may have been the cause.

Recently, Yotsuhashi and Shibasaki [119] published details of an arachin structure in which the latter was comprised of 6 kinds of subunits having the average MW of 29000 ($S^0_{20,w} = 1{\cdot}8$) and which possessed 3 glycine, 2 leucine (isoleucine) and 1 valine residue as the *N*-terminal amino acids. Whilst this model obviously has many similarities with those subunit structures listed in Table 3, other studies on arachin have revealed a more complex subunit structure. Thus, whilst Singh and Dieckert [85] reported substantial amounts of 6 subunits with MW's of approximately 15000, 24000, 29000, 37000, 41000 and 43000, and Shetty and Rao [86] found the predominant subunits had MW's of 18000, 25000, 33000 and 38000, both groups of workers observed additional subunits with MW's between 15000 and 70000. Although components with MW's approximating to those of the acidic and basic subunits listed in Table 3 were found, further investigations are required to explain the apparent multiplicity of subunit species, and the differences in the proposed structures of the arachin molecule.

The heterogeneity observed in the acidic and basic subunits poses an interesting question, namely, is there a unique legumin molecule composed of all these subunits or does legumin exhibit polymorphism, i.e. are there several types of legumin comprised of different subunits? It has already been established that the peanut protein, arachin, exists in nature in three variant forms which differ in their subunit compositions and also in their relative proportions in different seeds [165, 171]. Four distinct subunits were identified, viz. α and β, with MW's of 35000, and γ and δ with MW's of 10000; the most probable structures for the 3 variant forms were given as $\alpha_4\beta_4\gamma_2\delta_2$, and $\beta_8\gamma_2\delta_2$ and $\alpha_8\gamma_2\delta_2$.

Although detailed structural studies on the remaining seed globulins given in Table 1 do not exist, the fact that many of them are known to undergo dissociation in a comparable manner to legumin, edestin and glycinin, indicates that their subunit structures may be similar. Dissociation proceeds in a stepwise fashion as follows

$$\sim 11S \rightarrow \sim 7S \rightarrow \sim 3S \rightarrow \sim 2S \text{ (acidic and basic subunits)}$$

The conditions and the extent of dissociation vary from one globulin to another. Thus, arachin forms a reversible dissociating system between the dimeric (14·6S) and monomeric (9S) species [118], the latter being favoured at alkaline pH and low ionic strength. This has also been demonstrated for the 11S globulins of 3 species of lupin seed (*Lupinus angustifolius*, *L. albus* and *L. luteus* [37, 124] and, more recently, the 11S globulin of seeds of *Helianthus annuus* was reported to dissociate into a 7S component at ionic strengths below 0·3 [68]. Wolf and Briggs [172] observed that glycinin underwent reversible dissociation to 7S and 3S species when the ionic strength of its solutions was lowered. On the other hand, some 11S globulins have been reported not to undergo association–dissociation reactions with changes in ionic strength in the pH range which is generally regarded as being non-denaturing for seed globulins, i.e. pH values between ~4·5 and 9·5 [173]. These include the globulins from *Vicia faba* [124], *Vicia sativa* [88] and *Pisum sativum* [128, 129]. Dissociation of these globulins and also more extensive dissociation of the aforementioned globulins can be accomplished by employing acidic or alkaline pH's outside this 'stability range', or, alternatively, by the use of detergents, guanidine hydrochloride or urea [79, 125, 172, 174–177]. Thus, treatment of legumin of *Vicia faba* with ascorbic acid at pH 2·2 caused irreversible dissociation [178], at pH 2·8 legumin of *Vicia sativa* dissociated completely to a component with an $S_{20,w}$ value of 2·4 [88] and Johnson and Richards [127] obtained 3S species from the legumin of *Pisum sativum* at pH 2·05. Similarly, the 11S globulin of *Brassica napus* dissociated into components with sedimentation coefficients of 3S and 2·3S at pH 3·6 and in 6 M urea respectively [83]. On the other hand, legumin of *Vicia sativa* was reported as undergoing dissociation to components of roughly the same sedimentation coefficient (2·48S and 2·49S), in 4 M guanidine hydrochloride [106] and at pH below the isoelectric point [88] respectively, although the MW's of these two components calculated by the method of Trautman [179], were 30000 and 58000 respectively. Ionic strength has been shown to have a marked effect on acid denaturation. For example, at pH 2·2 the slowest sedimenting fraction of glycinin was observed to change from 5S at 0·2 ionic strength to 4S at 0·1 ionic strength, and finally with time to 2S [180]. Although the dissociation of the 11S molecule at acidic pH was counteracted by increasing ionic strength, it precipitated at 1·0 ionic strength on standing, indicating that some conformational change had occurred. Koshiyama [69] observed that addition of 0·1 M NaCl to a solution of glycinin in 0·1 N HCl, pH 2, altered the sedimentation coefficient from 2·48S to 3·75S. The latter author suggested that acid denaturation not only caused dissociation of the protein into subunits by electrostatic repulsion of charged groups, but also resulted in the unfolding of the polypeptide chains. The question of whether the 3S dissociation product is an artefact or whether it exists as such in the native molecule or indeed whether its structure involves disulphide cross-links or some other type of bonding, e.g. non-polar, hydrophobic interactions, remains to be answered. Nevertheless, the establishment of 3S and 2S subunits as separate entities now seems to be beyond doubt and this removes some of the confusion engendered by earlier work. Thus, the 3S subunits produced by acid treatment of the globulins of *Cucurbita* spp. [181] and the 13S globulin from *Fagopyrum esculen-*

tum [145] probably do not represent the monomeric subunits as suggested by the corresponding authors, but they may undergo further dissociation by selection of the appropriate conditions.

The relative ease of disruption of the molecular structure of these globulins indicates that most, if not all, of the interchain bonding is non-covalent in nature, e.g. hydrogen bonding. There are, however, a number of instances where covalent links, i.e. disulphide bridges, have been implicated in the bonding between subunits. Dlouha *et al.* [105] reported that the A and B chains of edestin were joined by disulphide bonds, and more recently, Wright and Boulter [64] published a subunit structure for legumin of *Vicia faba*, in which the acidic and basic subunits formed 'intermediary' subunits via disulphide bridges. These 'intermediary' subunits could correspond to the 3S component depicted in the dissociation scheme above. Disulphide-bonded subunits have also been shown to exist in solutions of both dissociated arachin [85, 171, 182] and dissociated glycinin [166]. In all these cases disulphide-bridged subunits were shown to be present by conducting dissociation studies first in the absence and then in the presence of a reducing agent. However, as Tombs and Lowe [16] and Wright [65] point out, it is possible that these subunits represent artefacts produced by disulphide bond formation between sulphydryl groups, exposed as a result of disruption of secondary and tertiary structure by the dissociating agent; this phenomenon has been observed in studies of other protein systems, e.g. myosin [183]. Clarification of the position with regard to disulphide bridges is complicated by the fact that most isolation and purification procedures include a reducing agent at some stage to minimize the possibility of intermolecular disulphide bond formation leading to protein aggregation [184], and it has been reported that even in a native molecule, disulphide groups are apparently accessible to reducing agents [185]. Because of the non-covalent forces operating between subunits, any reduction of internal disulphide bonds would only become apparent after dissociation of the molecule. It may thus have to be recognised that the true state of disulphide–sulphydryl groups in the native molecule can only be ascertained from a preparation free from reducing agent.

The size of the seed globulins would appear to make the elucidation of their 3-dimensional structure a formidable problem. Nevertheless, some progress has been made with the aid of electron microscopy. Thus, edestin was observed to be a spherical particle with a diameter of ~80–85 Å [186–188], and more recently, two 3-dimensional models, based on electron microscopy experiments were proposed for glycinin [93, 189]. In one model [93], glycinin consists of two annular–hexagonal structures packed one on top of another and each composed of 6 subunits. It is suggested that alternation of acidic and basic subunits within the structure contributes by ionic interactions to the stability of the molecule. Such a structure in which all the binding forces are equivalent and non-covalent in character does not explain, however, the apparent stability of the 3S moiety produced by acid denaturation of glycinin [123, 172, 190] and of other 11S globulins [83, 88, 127, 145, 191]. If one assumes, on the other hand, that these 3S components contain disulphide crosslinks, whether inherent or as artefacts, it then becomes difficult to explain the production of smaller subunits (2S) by the action of urea or guanidine hydrochloride, neither of which is capable of reducing disulphide bonds. Thus, Vaintraub and Shutov [190] observed that at pH 2·6 glycinin dissociated to a 3·5S component, whereas treatment with 4 M urea, even in the absence of 2-mercaptoethanol, resulted in dissociation to a species with an $S_{20,w}$ value of 2·04S.

The wide distribution of 11S seed protein (see Table 1) which, when investigated in more detail from different sources, shows considerable chemical and structural similarities (see Tables 2–4), is insufficient evidence to equate this protein with the legumin of *Pisum*. The information to date gives some idea of the properties we may expect of a protein which is equivalent to legumin but the data are too incomplete to make a decision as to the extent of its distribution. Serological studies also have been used to determine the extent to which legumin occurs in different legumes and the results of these studies are gathered together in Table 5.

Dudman and Millerd [31] found that all representatives of the tribes Fabeae (Vicieae), Trifolieae and Ononideae examined, contain proteins which gave identical cross-reactions with legumin of *Vicia faba*, except for *Cicer* and two species of Ononideae, where there was only partial identity. Legumin of partial immunological identity was also established to be present in *Daviesia mimosoides* and *Swainsonia stipularis* and some evidence for a legumin-like protein was obtained in the Loteae and Coronilleae, but this will need confirmation. Legumin was not detected immunologically in the tribes Sophoreae, Dalbergieae, Genisteae and Phaseoleae. Whereas Kloz and Turkova [95] found essentially the same results with the Vicieae, Genisteae and Phaseoleae, contrary to the finding of Dudman and Millerd [31] they did not detect legumin in the 3 members of the Trifolieae which they examined. Kloz and Turkova [95] used antibodies prepared against the protein of *Pisum sativum*, whereas Dudman and Millerd [31] used those of *Vicia faba*. The latter workers however, showed immunological identity between the legumin of *Pisum sativum* and *Vicia faba*.

Table 5. Distribution of proteins immunologically related to legumin

Tribe	Species	Reference
Astragaleae	*Swainsonia stipularis*	[31]
Fabeae	*Cicer arietinum*	[31, 95]
	Lathyrus clymenum	[95]
	L. odoratus	[95]
	L. sativus	[31, 95]
	L. sylvestris	[95]
	Lens culinaris	[31, 95]
	Pisum sativum	[31, 95]
	Vicia faba	[31, 95]
	V. sativa	[95]
Ononideae	*Ononis pubescens*	[31]
	O. serrata	[31]
Podalyrieae	*Daviesia mimosoides*	[31]
Trifolieae	*Medicago sativa*	[31]
	M. scutellata	[31]
	M. truncatula	[31]
	Melilotus alba	[31]
	Trifolium fragiferum	[31]
	T. hirtum	[31]
	T. incarnatum	[31]
	T. subterraneum spp.	
	brachycalycinum	[31]
	subterraneum	[31]
	yanninicum	[31]
	Trigonella foenumgraecum	[31]

Table 6. The distribution of vicilin-like seed globulins

Species	Sedimentation coefficient	MW (daltons)	Partial specific volume	Diffusion constant ($cm^2 sec^{-1} \times 10^{-7}$)	Frictional ratio	Reference
Acacia alata	7·9(S_{20})					[16]
A. decipiens	8·0(S_{20})					[16]
A. farnesiana	8·0(S_{20})					[16]
A. longifolia	7·6(S_{20})					[16]
A. penninervis	7·4(S_{20})					[16]
A. saligna	7·8(S_{20})					[16]
A. verticillata	7·8(S_{20})					[16]
Arachis hypogaea	8·4(S_{20})					[16]
	8·7(S^0_{20})	190000	0·72			[192]
	7·8	142000				[193]
Astragalus galegiformis	8·3(S_{20})					[16]
Canavalia ensiformis	6·4(S_{20})					[194]
Cytisus laburnum	8·1(S_{20})					[16]
C. supinus	8·0(S_{20})					[16]
Dolichos lablab	7·3(S_{20})					[16]
	7·2($S_{20,w}$)					[120]
	7·5($S_{20,w}$)					[120]
	7·8($S_{20,w}$)					[120]
Ervum lens	7·3(S_{20})					[120]
Genista tinctoria	8·5(S_{20})					[16]
Glycine max	8·0(S_{20})					[16]
	8·0($S^0_{20,w}$)	330000	0·729		2·17	[76]
	7·9($S^0_{20,w}$)	193000	0·725			[66]
	6·7($S_{20,w}$)	105000				[195]
	8·0($S_{20,w}$)					[121]
	7·2($S_{20,w}$)					[67]
	7·5($S_{20,w}$)					[67]
	7·8($S_{20,w}$)					[67]
Lathyrus clymenum	7·6(S_{20})					[16]
L. odoratus	7·6(S_{20})					[16]
L. sativus	7·5(S_{20})					[16]
L. silvestris	7·5(S_{20})					[16]
Lotus tetragonolobus	8·3(S_{20})					[16]
Lupinus albus	8·2(S_{20})					[16]
	8·3(S^0_{20})	204000		3·80	1·43	[124]
L. angustifolius	8·2(S_{20})					[16]
	7·8(S^0_{20})	181000		4·20	1·34	[35]
L. luteus	8·3(S_{20})					[16]
	7·4(S_{20})					[36]
L. polyphyllus	8·7(S_{20})					[16]
Medicago sativa	6·8(S_{20})					[16]
Phaseolus aureus	8·0*					[80]
P. coccineus	7·4(S_{20})					[16]
P. lunatus	6·3(S^0_{20})					[130]
P. nanus	6·6(S_{20})					[16]
P. vulgaris	7·3(S_{20})					[16]
	6·8($S_{20,w}$)	151000				Derbyshire, E (unpubl.)
	7·6($S^0_{20,w}$)	140000		5·5		[87]
	6·5(S^0_{20})					[130]
Pisum sativum	8·1(S_{20})					[16]
	7·1(S^0_{20})					[128]
Trifolium hybridum	7·7(S_{20})					[16]
T. pratense	7·7(S_{20})					[16]
T. repens	7·3(S_{20})					[16]
Vicia faba	7·1(S_{20})					[16]
	6·8(S^0_{20})					[130]
	7·1($S_{20,w}$)	150000				[178]
V. sativa	7·1(S_{20})					[16]
	7·5($S^0_{20,w}$)	193000				[88]
Vigna unguiculata	7·3(S^0_{20})					[130]

* Determined by density gradient centrifugation.

VICILIN

Legumes in which seed globulins with sedimentation coefficients of ~7S have been identified, are listed in Table 6. The chemical compositions of 7S globulin fractions are very similar to one another and this is true whether we are considering pure 7S proteins or vicilin preparations which are known to contain more than one vicilin type protein (see later). Representative data are given in Table 7, which shows that all fractions contain substantial quantities of the dicarboxylic acids and/or their amides and small amounts of methionine; cysteine, except in *Lupinus* and *Arachis*, and tryptophan, are either absent or present in very low amounts.

In those preparations where amide nigrogen has been determined, the amide content is very high (Table 7), reflecting their role as storage proteins. The level of arginine is not as high as in legumin, except in *Lupinus* spp. and *Arachis hypogaea*, but is higher generally than in the standard protein of Smith [196].

Carbohydrate has been reported in preparations from *Vicia faba* [200], *Phaseolus aureus* [80], *Phaseolus vulgaris* [87] and *Glycine max* [66]. The latter two preparations are very similar in their content of neutral sugars (4·5 and 4·8%) and hexosamine (1·1 and 1·2%); the preparations from *Vicia faba* and *Phaseolus aureus* contained less than 2% neutral sugars and 0·2% hexosamine.

Table 7. Amino acid composition of 7S globulins

Amino acid	Species								
	*Arachis hypogaea** [198]	*Cicer arietinum** [161]	*Glycine max* [66]	*Lupinus angustifolius*† [199]	*Lupinus luteus*† [199]	*Phaseolus vulgaris*† [87]	*Pisum sativum** [196]	*Victa faba** [161]	*Vicia sativa*† [197]
Asp	11·6	12·3	14·1	13·3	12·2	12·4	12·0	11·9	11·2
Thr	2·4	2·9	2·8	1·7	2·2	3·4	3·4	2·9	2·7
Ser	4·8	6·2	6·8	3·4	3·6	6 7	5·8	5·1	7·1
Glu	19·9	15·2	20·5	24·2	21·2	15·1	19·3	17·6	18·0
Pro	4·1	nr	4·3	2·7	3·6	2·9	3·5	nr	3·9
Gly	5·5	4·0	2 9	1·4	1·3	2·7	3·1	2·5	2·9
Ala	3·6	3·0	3·7	1·3	1·5	3·0	3·0	3·1	3·0
½ Cys	2·1	1·2	0·3	1·5	1·4	0·3	0·4	0 3	0
Val	4·5	4·6	5·1	2·5	2·9	5·2	4 6	4·3	3·7
Met	1·4	1·0	0·3	0	0	0·7	0·2	0·4	0·6
Ile	3·3	4 4	6·4	4 9	4·7	5·6	5·1	5·2	5·7
Leu	6·3	8·9	10 3	6 3	7·6	9·1	9·2	9·3	9·3
Tyr	3·6	2·8	3·6	5·7	6·2	3·5	3·0	3·8	4·0
Phe	4·6	7·2	7·4	5·2	5·6	6·6	6·2	6·8	5·8
His	2·4	2·8	1·7	2·3	1·7	2·6	2·1	2·4	2·7
Lys	3·7	7 1	7·0	4·0	3·5	5·6	7·9	8·1	8·1
Arg	11·6	9·5	8·8	15·2	13·5	5·0	7·3	7·8	10 7
Trp	nr	nr	0·3	0	0	0·8	0·1	nr	0
NH_3	2·3	nr	1·7	2·3	2·5	1·8	nr	nr	nr

* g/16 g N. † g/100 g protein. nr = not reported.

N-terminal analysis of 7S globulin fractions results in the labelling of several (up to 9) amino acids in the same preparation (Table 8) and often only the more heavily labelled amino acids are reported. Serine is common to all preparations and aspartate and/or glutamate is usually labelled. The relatively large number of *N*-termini may be due to the presence of an equal number of different polypeptide chains in the globulin, or, alternatively, it could reflect the difficulties which are inherent in the successful application of the technique. Serine, for example, is often present as a contaminant and a control reaction should be carried out using dansylation without subsequent hydrolysis to detect such contaminants.

Vicilin-type proteins have been identified immunologically in the Fabeae (Vicieae) and Trifolieae (Table 9), but only with *Vicia* and *Pisum* was identical cross-reactivity obtained [31]. They also obtained some evidence for vicilin-type proteins in the Ononideae, Podalyricae, Loteae, but not in the Sophoreae, Dalbergieae, Genisteae, Astragaleae, Coronilleae and Phaseoleae. Kloz and Turkova [95] using antisera to *Pisum* vicilin also identified vicilin-type proteins in the Vicieae and Trifolieae, and failed to detect them in the Phaseoleae and Genisteae; the vicilin-like proteins of *Vicia*, *Lens* and *Lathyrus* gave cross-reactivity of identity. Klozova and Kloz [203] prepared anti-bodies to phascolin [204] (7S globulin preparation of *Phaseolus*) and tested these against various other legume extracts. Whilst extracts of some members of the genus *Phaseolus* contained a protein which was identical immunologically with phaseolin, others did not cross-react; other genera of the Phaseoleae and those of the tribes Coronilleae and Vicieae which were examined, also gave negative results. Bourdillon's phaseolin has properties similar in many respects to those of glycoprotein II, the major 7S globulin of *Phaseolus vulgaris* [87]. However, since these proteins were purified by different procedures, identity between them cannot be assumed.

The "classical" legumin preparations of Osborne and Danielsson are not pure, but contain some vicilin as well as other contaminating proteins. More sophisticated fractionation procedures have led to the preparation of pure legumin, but modern methods of characterisation have also shown that this protein accounted for most of the protein in the legumin fractions of earlier investigators. The "classical" vicilin fraction, on the other hand, when subjected to modern separatory techniques, has been shown to contain more than one major protein. The vicilins prepared from *Pisum sativum* and *Vicia sati-*

Table 8. *N*-Terminal amino acids of 7S preparations

Species	Method	*N*-terminal amino acids	Reference
Cicer arietinum	DNP	serine, aspartate, threonine	[161]
Glycine max	DNP	serine, aspartate, glutamate, glycine, alanine, valine, leucine, tyrosine	[66]
	DNP	serine, aspartate, glutamate, glycine, alanine, valine, leucine, tyrosine	[76]
	DNP	serine, aspartate, glutamate, glycine, alanine, valine	[108]
	nr*	serine, aspartate, glutamate, glycine alanine, valine, leucine, phenylalanine, tryptophan	[121]
Phaseolus vulgaris	DNS	serine, leucine, threonine	Barker, R. D. J. (unpubl.)
Pisum sativum	DNP	serine, aspartate, threonine	[161]
	PTH	serine, aspartate 7 others	[79]
	DNP	serine, aspartate, glutamate	[201]
Vicia faba	DNP	serine, aspartate, glutamate	[161]
	DNS	serine, aspartate, glutamate, leucine, threonine, lysine	[200]
Vicia ervilia	DNP	serine, glutamate, lysine, valine	[202]
Vicia sativa	DNP	serine, glutamate, lysine	[202]

The most prominently labelled amino acids are underlined. * nr = not reported.

Table 9. Distribution of proteins immunologically related to vicilin

Tribe	Species	References
Fabeae	*Cicer arietinum*	[31, 95]
	Lathyrus clymenum	[95]
	L. odoratus	[95]
	L. sativus	[31, 95]
	L. silvestris	[95]
	Lens culinaris	[31, 95]
	Pisum sativum	[31, 95]
	Vicia faba	[31, 95]
	V. sativa	[95]
Trifolieae	*Medicago sativa*	[31, 95]
	M. scutellata	[31]
	M. truncatula	[31]
	Melilotus alba	[31]
	Trifolium fragiferum	[31]
	T. hirtum	[31]
	T. incarnatum	[31, 95]
	T. pratense	[95]
	T. subterraneum spp.	
	brachycalycinum	[31]
	subterraneum	[31]
	yanninicum	[31]
	Trigonella foenumgraecum	[31]

va by Danielsson's procedures have been separated into 2 and 3 fractions respectively, by DEAE cellulose chromatography [78, 79]; that of *Dolichos lablab* has been shown to contain three proteins by chromatography on hydroxylapatite [120], and those of *Vicia faba* [5, 65] *Arachis hypogaea* [205], *Glycine max* [206] and *Phaseolus aureus* [207], contain at least two components as shown by a variety of methods.

The vicilin fraction of *Pisum sativum* separates, at low ionic strength in the ultracentrifuge, into 2 molecular species, 7S and 11S [128]. The first of these proteins probably corresponds to the component which eluted from a cellulose column with the starting buffer, i.e. was not adsorbed, when Grant and Lawrence fractionated *Pisum* vicilin by ion-exchange chromatography on DEAE cellulose [79]. The second protein, which at low ionic strength sedimented as an 11S species, probably corresponds to the adsorbed component of Grant and Lawrence since this component, at 0·3 ionic strength existed in both 7S and 11S forms. Ghetie and Buzila [70, 71] used the same technique to obtain corresponding fractions which were then shown to be immunologically identical. These latter authors showed that the unadsorbed protein from the DEAE cellulose column is smaller, with a MW of 150000, and is less soluble at low temperature than the adsorbed protein, which has a MW of 190000. Different proportions of the two proteins were obtained from different batches of seed. When the pH is lowered from 7·0 to 6·2 at low ionic strengths, the two proteins do not associate further (Derbyshire, E., unpublished) and in this respect they differ from the major protein of *Phaseolus vulgaris*, Glycoprotein II (see later). We conclude that the vicilin fraction of *Pisum sativum* contains 2 major proteins. One of these is cold soluble, has a MW of approximately 190000, and associates to an 11S form at low ionic strength at pH values of 7·0 and 6·2 and the other is a cryoprotein of MW 150000, which does not associate at low ionic strength at pH values of 7·0 and 6·2.

Thus, although the 7S fraction from different legumes shows substantial chemical similarities (Tables 7 and 8), the heterogeneity of this fraction from *Pisum sativum* raises the question, is the major 7S protein of different legumes always the same equivalent protein? This question will only be answered when the proteins have been fully characterised. So far only three 7S proteins, two from *Glycine max* [66, 74, 99] (Koshiyama, I., personal communication) and one from *Phaseolus vulgaris* [87] have been purified and characterised. The major 7S protein of *Glycine max* corresponds in at least three properties to the larger of the two vicilins of *Pisum sativum*, but the minor 7S protein of *Glycine max* and the 7S protein of *Phaseolus vulgaris* have different properties to both of the vicilins of *Pisum sativum*.

The major 7S protein from *Glycine max* isolated by Koshiyama [74, 99] is cold soluble, has a MW of 180000–210000 at high ionic strength ($S^0_{20,w} = 7{\cdot}92$, $I = 0{\cdot}5$), and associates to a larger molecular species, the size of which (9S–12S) is dependent on protein concentration, at low ionic strength ($I = 0{\cdot}1$) (Tables 6 and 10); its isoelectric point is pH 4·9 and it contains 15·9% Kjeldahl nitrogen and 5% carbohydrate. Organic phosphorus is absent. It has an amino acid composition which is similar to that of other seed 7S proteins, i.e. a high content of aspartate and glutamate residues, and a low content of $\frac{1}{2}$-cystine and methionine. Amide ammonia accounts for 1·7% of the protein and eight different *N*-terminal amino acids have been reported. It dissociates to 5S and 2S forms in 0·01 M HCl, to 3S in detergent and to 1S–2S in urea. The 7S globulin of *Glycine max* has been isolated and characterised also by Roberts and Briggs [76], who employed ammonium sulphate precipitation. Only 4% of the 7S globulin present initially was recovered as the pure protein, but it was regarded as typical of the initial material as judged by its association to a 9S–12S form at low ionic strength, its dissociation to a 3S form in the presence of detergent and its behaviour during chromatography on calcium phosphate. Koshiyama regards his own preparation as a more highly purified form of the Roberts and Briggs protein and discounts the remarkably high MW (300000) determined for the latter. The major protein of the Roberts and Briggs preparation is identical with β conglycinin, one of four globulins from *Glycine max* separated by immunoelectrophoresis [208], and Koshiyama (personal communication) has identified his own protein with β conglycinin. Catsimpoolas and Ekenstam [208] reported identity between Koshiyama's protein and γ conglycinin, an immunoelectrophoretic component different from β conglycinin; however, Koshiyama has isolated γ conglycinin and has shown that it differs from his major 7S globulin. Thus, it is probable that Catsimpoolas and Ekenstam had not reproduced the preparative conditions employed by Koshiyama and had, in fact, isolated γ conglycinin and not the major 7S globulin. Gamma conglycinin has a MW of 104000 ($S^0_{20,w} = 6{\cdot}6$), an isoelectric point at pH 5·8 and a carbohydrate content of 5% (w/w). The protein does not associate at low ionic strength, at pH 7·6, and it is not an agglutinin. The concentration of γ conglycinin in the crude 7S globulin fraction is approximately 10%.

The major 7S protein (Glycoprotein II) from *Phaseolus vulgaris* has a MW of 140000 ($S^0_{20,w} = 7{\cdot}6$) (Table 6) and associates to a 19S form in the pH range 3·4–6·4. It does not associate to an 11S form at low ionic strength in the pH range, pH 2·2–8·0. Glycoprotein II has an isoelectric point of pH 5·4 and contains 15·5% Kjeldahl nitrogen, 5·5% carbohydrate and only trace amounts of phosphorus. It has a lower content of glutamate and a higher content of methionine and tryptophan than the

Table 10. Sedimentation coefficients of 7S globulins in media of high ($I = \geqslant 0.3$) and low ($I = 0.1$) ionic strength

Species	$I = \geqslant 0.3$	$I = 0.1$	Reference
Arachis hypogaea	8·7	*12·6 20*	[192]
Canavalia ensiformis	7	*11* (part)	Derbyshire, E. (unpubl.)
Glycine max	7	*10·5*	[76]
	7·5	*11·5*	[66]
	6·7	*10·4*	[195]
	7·9	9–12	[99]
Lupinus albus	8·3	*12·2*	[124]
L. angustifolius	7·8	7·8	[35]
L. luteus	7·3	>7·9	[125]
Phaseolus lunatus	6·3	6·5	[130]
P. vulgaris	6·9	6·9	[87]
	6·5	6·8	[130]
Pisum sativum	7·1	*10·8* (part)	[130]
Vicia faba	6·8	7·6	[130]
V. sativa	8·0	8·0	[88]
Vigna unguiculata	7·3	*11·1*	[130]

major 7S protein of *Glycine max*: amide ammonia accounts for 1·8% of the protein. It dissociates to a 2S form in guanidine hydrochloride. Pusztai and Watt [87] recovered only 9% of the total protein as pure glycoprotein II; however, a protein fraction which accounts for 35% of the total protein is almost identical in properties to glycoprotein II [209]. The 7S globulin (Gl) isolated from *Phaseolus vulgaris* by acid extraction [100] associates to 18S in the pH range 3·8–5·4 and dissociates to 3S at pH 12 and is equivalent to glycoprotein II.

A disconcerting discrepancy in the literature, is the assignment by Danielsson [25] of sedimentation coefficients of 11·0S to phaseolin and 7·3S to conphascolin, the major and minor globulins respectively of *Phaseolus vulgaris* [210]. However, reference to the text and Fig. 11B of Danielsson's original report [16] shows clearly that the major globulin sediments as a 7S species and is equated with vicilin. McLeester *et al.* [211] also report a high (12S) sedimentation coefficient for the major protein of *Phaseolus vulgaris*, but in a subsequent publication [100] a 12S form is not recorded and the same globulin is shown to sediment as a 7S globulin and to associate to an 18S species at acid pH, i.e. it is equivalent to glycoprotein II.

The data available for 7S proteins of other species are very sparse and this limitation precludes their positive identification with any of the 7S globulins already described. The fact that there are usually several components in the 7S fraction, none of which has been properly separated and characterised, is a further complication.

7S proteins which associate at low ionic strength ($I = 0.1$) (Table 10) and thus resemble the major 7S globulin of *Glycine max*, have been identified as major globulins in *Arachis hypogaea* [192], *Vigna unguiculata* and *Lupinus albus* [124, 130]. Major 7S proteins which do not associate at low ionic strength have been identified in preparations from *Vicia faba* [130], *Vicia sativa* [88], *Lupinus angustifolius*, *Lupinus luteus* and *Phaseolus lunatus* [35, 125, 130] (see Table 6). However, the latter proteins were not examined at pH 6 and, therefore, it is not possible to discriminate between those which may be equivalent to Glycoprotein II and those which may be equivalent to the smaller of the 7S vicilins of *Pisum*. The major protein fraction, canavalin, of *Canavalia ensiformis* [212–214] resembles the 7S globulin fraction of *Pisum*, since approximately 50% of the protein associates to 11S at low ionic strength at pH 7·0 and an equal quantity remains as a 7S protein (Derbyshire, E., unpublished); at pH 6·4 and high ionic strength both proteins exist as 7S species, whereas at low ionic strength both 7S and 11S species occur.

On the basis of the ultracentrifuge data 3 different types of 7S globulin can be recognised. One of these does not associate at low ionic strength, the second associates to an 11S species at pH 7 at low ionic strength and the third associates, to an 18S species, only at pH values below pH 7.

Globulins with sedimentation coefficients of approximately 7S have been identified in a few other dicotyledons, including *Helianthus annuus* [147], *Beta vulgaris* [137], *Telfairea pedata* [156], *Cucurbita maxima* [215] and *Gossypium barbadense* [216]. Except in *Gossypium*, they represent only minor components of the globulin fractions which consist mainly of 12S protein. Since, for example, the 11S protein of soyabean is often accompanied by a small quantity of its monomer, the 7S form of glycinin [208, 217], it is possible that the 7S globulin of the non-legume dicotyledons may be a form of legumin; however, in the absence of other data the possibility that they are vicilins cannot be overlooked.

An 8S globulin, γ globulin, is the major globulin of embryos of *Hordeum vulgare* [16] and *Oryza sativa* [159], and a similar protein is also found as the major seed globulin of 7 other genera of the Gramineae. The γ globulin of rice has been resolved into 3 proteins and at least one of these is a glycoprotein [218, 219]. A 12S globulin was not detectable in these monocotyledons, except in barley, rice and wheat [158], in which it occurred as a minor component. However, it would be premature to equate the γ globulins with vicilin. In *Cocos nucifera* an 8S globulin is accompanied by a much greater concentration of 12S globulin [160].

Table 11. The subunit composition of 7S globulin fractions of legumes

Species	MW ($\times 10^{-3}$)	Reference
Canavalia ensiformis	56	Derbyshire, E.,
	43	(unpubl.)
Glycine max	23	[108]
	81	
	51	
	35	[121]
	24	[220]
	22·5	[220]
Lupinus angustifolius	56	[221]
	32	
	20	
Phaseolus aureus	63·5	[80]
	50	
	29·5	
	24	
Phaseolus vulgaris	56	Derbyshire, E., (unpubl.)
	50	
	47	
	43	
	23	
	50	Barker, R. D. J., (to be publ.)
	47	
	35	[87]
	43	
Pisum sativum	56	Derbyshire, E., (unpubl.)
	43	
	23	
Vicia faba	55·5	[65]
	46	
	33·3	
	31·5	
	43·4	[65]
Vigna unguiculata	56	Culkeen, J. and Carasco, J., (unpubl.)
	53	

A greater uncertainty surrounds the subunit structures of the 7S storage globulins than those of the 11S globulins, and this is a reflection of the existence of several vicilin-like proteins and also the heterogeneity of most 7S preparations. The data available is limited almost completely to the apparent MW's of the subunits (Table 11) and the different methods used make it difficult to compare the values, especially those of *Glycine max*. However, it is apparent in most cases that the subunits are not linked by disulphide bridges since the inclusion of 2-mercaptoethanol in the dissociating media does not alter the size of the dissociation products.

Different MW's have been assigned to the subunits of the 7S globulin of soyabean isolated in three laboratories, possibly due to lack of identity between the preparations used. The MW of the subunits in 8 M urea was determined in the ultracentrifuge as 23000 by Koshiyama [222] and as 35000 by Okubo *et al.* [121]; both workers used an assumed frictional ratio value. Higher values for the MW's of the subunits were obtained by Masaki and Soejima [108] who isolated urea dissociated globulin components with MW's of 81000 and 51000 as determined on SDS gels. Koshiyama [220] estimated, on the basis of *N*-terminal amino acids, that there were 9 subunits in his major 7S protein fraction; however, the MW's of the 2 components isolated by Masaki and Soejima suggest that there are 3 subunits and the *N*-terminal amino acid analysis of the 2 components gave aspartate (and serine) and serine respectively, which is consistent with this number.

The estimated MW's of subunits of the *Phaseolus vulgaris* 7S globulin are higher than those determined for the soyabean subunits by Koshiyama [220] and Okubo *et al.* [121]. Thus, Pusztai and Watt [87] obtained a range of values between 35000 and 43000 for Glycoprotein II by a variety of methods and they suggest that there are three or four subunits in this protein. Barker, R. D. J. (unpublished) prepared the major 7S globulin of *Phaseolus vulgaris* by cryoprecipitation and resolved the subunits into two components in the molar ratio 2:1 with MW's 50000 and 47000. Although these values are different from those of Pusztai and Watt [87], the methods of preparation of the 7S globulin differed and, in any case, absolute reliance cannot be placed on values of MW's, particularly of glycoproteins, determined on SDS gels. Subunits with MW's 50000 and 47000 are the major subunits also of the cold soluble 7S globulin and of the 7S globulin prepared by ammonium sulphate precipitation but in these preparations other subunits were identified as minor components which suggests that other globulins were present with Glycoprotein II (Derbyshire, E., unpublished). Similar subunits, including the minor ones, are found in other cultivars and species of *Phaseolus* including *P. formosus* and *P. acutifolius*, although the proportions of the subunits relative one to another are different in different cultivars (Derbyshire, E., unpublished). One of the minor subunits of *Phaseolus vulgaris* referred to above, has a MW of approximately 30000 and has been subsequently identified as a 6S agglutinin (Derbyshire, E., unpublished). Recently, Pusztai and Watt [223] have also isolated a globulin agglutinin with subunit MW 30000 from *Phaseolus vulgaris*, although the MW of the native protein was not reported. Subunits similar in size to those of the 6S agglutinin occur in 7S globulin fractions of other species and it is unfortunate that these fractions have not been monitored for agglutinating activity.

Subunits with MW's similar to those of the other minor subunits of the 7S globulin fraction of *Phaseolus vulgaris*, i.e. 56000, 43000 and 23000, are major components of the 7S globulin fractions of *Pisum sativum* and *Canavalia ensiformis*. In *Vicia faba* where 2 vicilins have been separated, one dissociates with major subunits with MW's 55000 and 46000 and the other dissociates with subunits with MW 43400 [65]. The 7S globulin isolated from *Vigna unguiculata* represents another form of globulin, different from that of the vicilins of *Vicia faba*, since it dissociates to major subunits with MW 56000 and 53000 (Carasco, J. and Culkeen, J., unpublished); also, 7S globulin of *Vigna unguiculata* associates to an 11S form at low ionic strength in contrast to the vicilins of *Vicia faba* which remain as 7S protein under these conditions [130]. In this laboratory we assign a MW of 53000 to the major subunit of the 7S globulin of *Phaseolus aureus*. The subunit of MW 23000 of *Phaseolus vulgaris* is found in several other species (see Table 11). In *Lupinus angustifolius* it is greatly enriched in one of the 7S globulin fractions [221], suggesting it is a subunit of a specific globulin possibly equivalent to the major 7S globulin of *Glycine max*, since Koshiyama [220] estimates the MW of the latter's subunits as 23000. This subunit is absent in the 2 vicilins of *Vicia faba* and it is not a subunit of glycoprotein II.

Whilst there is sufficient similarity in the subunit patterns of different taxa to suggest that equivalent vicilin proteins are involved, it is not possible to be sure of how many, or whether the subunit MW's as determined on gels are the minimum subunit MW's of the native proteins. It must be emphasized that neither the action of dissociating agents nor the bonding between seed globulin polypeptides is fully understood, and it is possible that complete separation of the individual polypeptide chains is not achieved by the techniques of dissociation usually employed. Taken at face value the subunit patterns suggest that there may be as many as 5 different forms of 7S globulin, vicilins, each of which may be recognised by 1 or 2 characteristic subunits (Table 12). The ratio of these forms may vary from species to species

Table 12. The distribution of the characteristic subunits of 7S globulins of legumes

	Subunit (MW)				
	I (23000)	II (43000)	III (50000, 47000)	IV (56000, 46000)	V (56000, 53000)
Species	*Canavalia ensiformis* *Glycine max* *Lupinus angustifolius* *Phaseolus aureus* *Phaseolus vulgaris* *Pisum sativum*	*Canavalia ensiformis* *Phaseolus vulgaris* *Pisum sativum* *Vicia faba*	*Phaseolus aureus?* *Phaseolus vulgaris*	*Canavalia ensiformis?* *Lupinus angustifolius* *Phaseolus vulgaris?* *Pisum sativum?* *Vicia faba*	*Canavalia ensiformis?* *Phaseolus vulgaris?* *Phaseolus aureus?* *Pisum sativum?* *Vigna unguiculata*

Table 13. Amino acid composition of the subunit with MW 81 000 of soyabean 7S globulin compared to that of the parent protein

Amino acid	g amino acid residue 100 g protein: Subunit	7S globulin
Asp	10·18	11·18
Thr	1·80	3·14
Ser	4·24	4·79
Glu	20·10	17·54
Pro	1·44	5·21
Gly	2·96	3·37
Ala	4·39	3·66
½ Cys	0	1·52
Val	3·25	4·68
Met	0·39	0·43
Ile	4·07	4·99
Leu	6·76	8·15
Tyr	2·68	3·51
Phe	2·68	5·55
His	2·05	2·32
Lys	3·32	6·26
Arg	8·25	7·37
NH_3	2·05	2·32

and even in extracts prepared in different ways from the same species. Normally, it would appear that all the forms do not occur in a single species except possibly in *Phaseolus vulgaris*. The determination of the *N*-terminal regions of the amino acid sequences is an important next step in order to clarify the situation.

Only in the case of the soyabean proteins have subunits actually been isolated [108] and only the amino acid composition of the component with MW 81 000 was reported (Table 13). Comparison between this and the composition of the parent protein suggests that the amino acid composition of the 51 000 subunit is different from that of the 81 000 subunit; the 2 components also differ in their *N*-terminal amino acids, serine and aspartate (and serine) respectively.

OTHER GLOBULIN PROTEINS

Although 11S and 7S globulins are the major storage proteins of legume seeds, other globulins occur in extracts. A 15S protein often accompanies purified legumin as a minor component in the ultracentrifuge and its presence correlates with a minor component of low electrophoretic mobility. Association of legumin via disulphide bridges during isolation possibly explains the occurrence of the 15S form. Similarly, the observation of 18S globulins in several legumes [16] could be due to the association of a 7S globulin, since a 7S globulin is usually the major globulin of those species with an 18S component.

Smaller globulins, 2S 4S, are also found in several legumes [16, 36, 55, 75, 92, 153, 192, 208], other dicotyledons [37, 83, 137, 138, 147, 156], and monocotyledons [16, 157, 224–226] (Table 14), and they account for more than 5% of the protein of one or more species from each of the 3 taxa. Two small globulins with sedimentation values of 2·8S and 2·3S have been isolated from soyabean meal and characterised [227], and a 2S globulin has been identified in protein body preparations of soyabean [69, 208]. 2S proteins, which are water soluble during isolation [35, 228, 229] and a 2S protein present in undialysed preparations [130] have been reported but their relationships, if any, to the 2S globulins have not been investigated. The amino acid compositions of the 2S globulins from 5 species are given in Table 15, and generally they show the major characteristics of those of storage protein, i.e. high concentrations of amides and concentrations of arginine greater than that in a standard protein [196]. There are substantial differences in amino acid composition between the 2S globulins from different genera and even between those from species of the same genus, e.g. *Brassica* spp. The isolation of two 2S globulins from soyabean suggests that additional 2S globulins may eventually be shown to be present in other preparations.

Table 14. Sedimentation coefficients of 2S globulins

Species	Sedimentation coefficient
Dicotyledons	
Legumes	
Acacia alta	1·3(S_{20})
A. longifolia	2·7(S_{20})
A. penninervis	1·5(S_{20})
Arachis hypogaea	2·0(S_{20})
	2·0(S_{20})
Cytisus supinus	1·8(S_{20})
Lupinus albus	2·7(S_{20})
L. luteus	2·0(S_{20})
Glycine max	2·8($S_{20,w}$)
	2·3($S_{20,w}$)
Phaseolus coccineus	4·3(S_{20})
P. vulgaris	4·9(S_{20})
Non-Legumes	
Beta vulgaris	1 2S
Brassica hirta	1·8($S_{20,w}$)
B. juncea	1·8($S_{20,w}$)
B. nigra	1·8($S_{20,w}$)
Helianthus annuus	1·7(S_{20})
Ricinus communis	1·6(S_{20})
Telfairea pedata	1·6(S_{20})
Monocotyledons	
Avena sativa	2·6(S_{20})
Festuca rubra	2·4(S_{20})
Hordeum vulgare	2·5(S_{20})
Oryza sativa	1·6(S)
Secale cereale	2·6(S_{20})
Triticum aestivum	2·5(S_{20})
Zea mays	2·6(S_{20})

Recently, a heterogeneous globulin fraction, γ conglutin, which includes a 10S molecular species, has been isolated from *Lupinus angustifolius* [221, 230]. The γ conglutin dissociates to subunits with MW 40000 in the absence of reducing agents and 30000 and 17000 when reduced. It represents 10–25% of the total globulin in different cultivars and species of *Lupinus* and it was not

Table 15. Amino acid compositions (mol %) of 2S globulins (recalculated where appropriate)

Species	*Brassica hirta*	*Brassica nigra*	*Glycine max*		*Lupinus luteus*	*Oryza sativa*
Reference	[138]	[138]	[227]		[199]	[226]
Sedimentation coefficient	1·8S	1·8S	2·8S	2·3S	2·0S	1·6S
MW	nr	nr	32 600	18 200	27 600	25 400
Amino acid						
Asp	10·9	2·1	15·4	13·3	10·0	3·2
Thr	4·2	3·1	4·4	5·0	1·4	1·9
Ser	5·5	4·6	5·9	5·5	6·2	10·8
Glu	19·9	17·1	9·7	11·4	35·5	22·6
Pro	10·0	19·5	6·2	6·1	2·9	4·9
Gly	13·0	6·6	9·1	9·1	3·8	8·5
Ala	7·9	6·2	4·7	5·6	1·9	5·6
Val	6·3	6·5	7·1	7·3	1·9	3·8
½ Cys	Trace	Trace	1·0	1·5	8·1	4·5
Met	Trace	Trace	0·6	1·9	0·5	4·6
Ile	6·1	4·5	9·4	6·4	3·8	1·3
Leu	10·2	9·4	7·1	9·1	11·5	6·0
Tyr	2·3	0·7	1·8	2·9	0·5	5·4
Phe	4·2	2·7	5·9	5·1	3·4	2·4
Lys	3·8	5·8	5·7	5·1	0·9	0·1
His	1·7	5·6	0·6	0·7	0·5	Trace
Arg	5·7	4·8	5·5	4·2	7·2	13·7
Trp	nr	nr	1·7	nr	nr	0·6
NH_3	13·5	16·8	13·2	nr	31·2	12·8

nr = not reported.

observed previously because it is not resolved from the 11S and 7S globulins in the ultracentrifuge. The relatively high concentration of this fraction in the seed extracts, its globulin nature and its occurrence in the protein body fraction, strongly suggest that it is storage protein. It is of especial importance nutritionally, since it has a high content of sulpho-amino acids (4·2 g/16 g N) and its concentration, relative to that of other globulins, in the seed ranges from 10 to 25% in different cultivars [92, 230]. The importance of these characteristics in relation to a grain legume improvement programme has been emphasised and discussed elsewhere [91, 92, 231].

Biosynthesis and degradation

It is not our intention to discuss the biosynthesis and degradation of legume storage protein in any detail, since this has been done elsewhere [232, 234]. However, perspective requires a brief mention of some aspects and these are given below.

Typical protein bodies from seeds of legumes do not contain ribosomes (e.g. see [1]) and do not contribute significantly to protein synthesis [235]. The storage protein which accumulates in protein bodies is synthesised elsewhere, on membrane-bound or free cytoplasmic polysomes [236], and is then channelled through the endoplasmic reticulum prior to packaging into the protein bodies [237].

Harris and Boulter [238] suggest that some of the cytoplasmic vesicles into which the storage protein of *Vigna unguiculata* is transferred may originate from the Golgi apparatus, although some may be formed by subdivision of the main vacuole, which is the mechanism implicated in *Phaseolus vulgaris* [1]. The involvement of the Golgi apparatus in the deposition and transport of protein, including plant cell wall glycoprotein and secretory proteins in animals has been demonstrated in a variety of tissues (see [239] for references).

11S and *7S* globulins are found together in protein body fractions isolated from soyabean [208, 240, 241], peanut [205, 242] and broadbean [243], and legumin and vicilin are both laid down in the same protein bodies [9], although the protein bodies of a seed may not all have identical protein compositions.

The composition of the storage protein fraction of legumes changes during the course of accumulation of protein during seed development [5, 25, 28, 244–249], and these changes suggest that the rates of synthesis of individual proteins differ and may even vary during ripening. Several investigations have employed electrophoresis of total seed protein or globulin to follow the changes which occur in individual proteins [4, 244, 247] and these results demonstrate that changes do occur and they give some indication of which electrophoretic components are involved. However, different globulins may have similar electrophoretic mobilites and this, together with the possible presence of both 'monomeric' and 'dimeric' forms of a protein [208, 248, 250], prevents more detailed interpretation of the data. Similarly, results obtained by use of the ultracentrifuge alone [25, 28] must also be interpreted with caution. In order to follow the changing protein pattern during seed development, proteins must be monitored by one or more of their unique characteristics, for example, immunological determinants and subunit composition. The limited critical data which are available confirm that in *Vicia faba* [5, 9] and in *Glycine max* [249] the ratio of the major globulins (11S and 7S) change during ripening, and that at least one of the 7S proteins of these species is synthesised, like that of *Pisum sativum* [25, 28], earlier than the 11S protein. The 7S protein of mature *Vicia faba* seeds has been separated into 2 vicilins, and it is probable that the different vicilins are synthesised at different rates, since the subunit composition of the vicilin fraction of *V. faba* changes throughout the development of the seed [5]. The ratio of three individual electrophoretic components of the soyabean 7S protein also varies during ripening [249].

With the onset of germination the protein bodies increase in volume and late in germination have increased fifty-fold [12, 13]. The expansion is accompanied by proteolysis and it has been suggested [13, 243, 251] that latent enzymes, including protease and acid phosphatase, laid down in the protein bodies during ripening, are activated as part of the germination process.

Vicilin of *Pisum sativum*, phaseolin, arachin and the 11S and 7S globulins of soyabean are degraded during germination to proteins of greater electrophoretic mobility as demonstrated by use of immunoelectrophoresis [248, 250, 252–254]. Legumin and vicilin are degraded at the same rate in *Pisum sativum* [255] but in other species, for example *Lupinus luteus* [256] and *Glycine max* [250, 257], different relative rates of breakdown have been reported. Daussant [252, 258] suggests that a progressive deamidation of storage protein takes place as a first step in degradation, and that this is followed by cleavage of disulphide bonds when these are present. Further decrease in the size of the proteins is brought about by protease systems, the activity of which may be regulated by pH changes in the seed as a result of an accumulation of amides, and Ghetie [259] has proposed that separation of protein and non-protein moieties of conjugated storage proteins may be necessary before proteolysis can proceed. The polypeptides released by these successive steps are then available for progressive hydrolysis by the endo- and exo-peptidases of the seed [260] to their constituent amino acid [14].

CONCLUSIONS

The storage globulins are complex proteins and the technology of their isolation, purification and characterisation is only just becoming well enough understood for different laboratories to obtain comparable results.

For this reason we have excluded from this review much of the work carried out in the first half of the century and for references to those investigations the reader is referred to the earlier reviews [10, 15, 54, 173, 193, 261–263].

Usually globulins are extracted within the pH range 7–9; however, Hall and colleagues [4] have proposed a new method for the isolation of storage proteins using acid extraction to produce a G1 and G2 globulin fraction. The G1 globulin fraction from *Phaseolus vulgaris* which McLeester *et al.* [211] called legumin, was subsequently equated with the major storage protein, Glycoprotein II [100, 264]. In our view, G1 should not be referred to as legumin but as an impure preparation of Glycoprotein II. The G2 fraction which sediments in the ultracentrifuge with a sedimentation coefficient of 6·6 and dissociates into subunits of MW 32000, they equated

with vicilin. However, Pusztai and Watt [223] have extracted 6S agglutinins from seeds of *Phaseolus vulgaris* with subunit MW's between 30000 and 34000, and one of these is a globulin. The agglutinating potential of the G2 fraction was not determined by Hall *et al.* [4] but comparable preparations in this laboratory examined by us, agglutinate red blood cells (Derbyshire, E., unpublished); this evidence suggests that the G2 fraction which is relatively pure, does not represent a typical vicilin.

The use of the acid extraction method with *Vicia faba* led to the production of much lower MW polypeptides than those obtained with *Phaseolus* [178, 211], and in view of the general tendency of storage globulins to dissociate at low pH values, it cannot be recommended for general use: an additional disadvantage of the method is the possible acceleration of deamidation in the presence of ascorbate [265].

On the basis of biological, structural and chemical composition data, it is concluded that proteins with at least some of the properties of legumin have been found in dicotyledons, other than legumes, and even in monocotyledonous plants. The similarity of this protein in different plants could be due either to convergent evolution in response to a common functional need, or to common ancestry. Only an investigation into amino acid sequence differences, similar to that conducted by Boulter [266] for cytochromes *c*, could provide the evidence to decide which of these two possibilities occurred.

Conclusions as to extent of the distribution of legumin in plants depends on the criteria used to identify the protein. Dudman and Millerd [31] investigated eleven tribes of the Leguminosae serologically, and found that a legumin-like protein was restricted to members of the Fabeae, Trifolieae, Podalyrieae, Astralageae and Ononideae. However, it is clear from other homologous protein data sets such as those of plastocyanin, that lack of cross-reactivity between taxa cannot be taken as evidence of the absence of an homologous protein [267]. Jackson *et al.* [268] using maps of tryptic peptides of globulin preparations, found little resemblance between those of *Vicia faba* (Vicieae) and *Phaseolus vulgaris* (Phaseoleae), which again suggests the absence of homologous storage proteins in these 2 tribes. However, there is a considerable amount of structural data [16] (Derbyshire, E., unpublished), indicating that an 11S protein equivalent to legumin does occur in low concentration in *Phaseolus vulgaris*.

We conclude that it would be premature to adopt the suggestion of Hall *et al.* [4] and McLeester *et al.* [211] that the term 'legumin' should be dropped, or that of Millerd [232] that it should only be applied to this protein in the Vicieae and Trifolieae, even though in the past legumin preparations have not always been strictly equivalent due to the presence of different impurities. The retention of trivial names for legumin, such as edestin, should also be retained for the present since they are widely used in the literature.

The position with regard to vicilin is less clear-cut and much more work will be needed before a full understanding of this fraction from different plants is available. When sufficiently investigated, the vicilin fraction has been shown to contain more than one protein and as many as 5 forms of vicilin may exist. A similar situation to that already described for legumin exists with regard to the distribution of vicilin-like proteins. Danielsson [16] has shown the presence of a 7S peak in the ultracentrifuge with salt extracts of seeds from various tribes of the Leguminosae, including the Acacieae in the subfamily Mimosoideae, and on this evidence alone has equated the protein responsible with his vicilin preparation from *Pisum*. However, other proteins with sedimentation coefficients close to those of the vicilins exist and hence ultracentrifugal evidence of the presence of vicilin must be treated with caution. In contrast, Dudman and Millerd [31] using a vicilin fraction from *Vicia faba*, found immunological evidence of a vicilin-like protein in members of the Fabeae and Trifolieae, but not in those of the other nine legume tribes investigated. Even so, the physical and chemical data which are available, suggest that 7S globulins are of widespread occurrence.

Dudman and Millerd [31], on the basis of serological data, have claimed, in contrast to Kloz and co-workers who also used serology, that legumin is more "primitive" than vicilin. Blagoveschenskii [269] came to a similar conclusion on the basis of studies on legume seed weight, percentage nitrogen and type of seed protein, although the evidence for the statement is not fully substantiated as there are inconsistencies in the correlations suggested. In our view, the question as to whether or not one of these proteins is phylogenetically older than the other, cannot be decided on the present evidence since the criteria used are insufficient.

Although vicilin and legumin have been shown to be distinct serologically [9, 31], it is possible that they may have subunit(s) in common since the separated subunits of vicilin and legumin have not been tested serologically. Analysis of SDS gel patterns to determine whether or not vicilin and legumin have common subunits, is made difficult by the presence on gels of weak bands; thus, although the bands due to the major subunits are not common, some of the minor component bands may be. An alternative approach to this problem is that of Jackson *et al.* [161] who used tryptic peptide maps. These indicated that there was considerable overlap in the peptide patterns of legumin and vicilin, ranging from about 80% with *Pisum sativum* to 40% with *Cicer arietinum*, and it is probable that the substantial similarity between fingerprint patterns can safely be interpreted as indicative of some degree of common structure. However, the preparations of vicilin and legumin used by Jackson *et al.* [161] were cross-contaminated, although the legumin preparations contained very little vicilin and it is unlikely that it contributed significantly to the tryptic maps of legumin.

Storage proteins are synthesised by the usual ribosomal template mechanism [232, 233]. Since the products of this process are clearly defined we can expect that the storage proteins are also carefully prescribed and this is borne out by the fingerprint data of Jackson *et al.*, who showed that different species of the same genus gave virtually identical globulin fingerprints. However, since storage proteins each consist of several subunits, it is possible that several forms of the same protein may occur within an individual or in a population of individuals, and this has been shown to be the case with the storage protein, arachin, of peanuts [165].

In *Pisum* the 2 immunologically identical vicilins (see earlier) occur in different ratios in different populations [70]. Also, the subunit composition of the vicilin fraction from different varieties of pea varied in such a way as to suggest that *Pisum* vicilin is polymorphic [270]. In

breeding tests carried out by Hynes vicilin was not maternally inherited [270] and Davis [271] found no evidence of maternal inheritance when he analysed the electrophoretic patterns of the total globulins of some hybrids of *Pisum* cultivars, but did in the case of others.

Polymorphism of storage proteins may be of frequent occurrence and the possibility exists for a complex interplay between the different gene loci responsible for the subunits, not only of a specific vicilin but perhaps of different proteins. It is important, therefore, that genetic studies be carried out to ascertain the way in which different subunits may combine to give polymorphic proteins.

The fact that the relative proportions of the different storage proteins can differ considerably in different legume species [91], and also in different varieties of the same species [229, 230, 269], suggests that there are possibilities for considerable change in the storage protein without loss of seed viability. However, constraints will operate on changes in the composition of storage protein, since it must satisfy the following requirements: (a) act as a suitable nitrogen store, i.e. have an amino acid composition which on hydrolysis will supply all the nitrogen compounds of the seedling using the enzymic machinery of the latter; (b) be synthesised on and pass into the rough endoplasmic reticulum; and (c) contain the subunit interaction groups to allow correct packaging during seed maturation and subsequently unpackaging during germination (the latter requirements relate to the specificity of the proteolytic system of the germinating seed).

Acknowledgements—We wish to thank Miss M. Berryman for help in preparing the manuscript, and the Agricultural Research Council for financial support.

REFERENCES

1. Opik, H. (1968) *J. Exp. Botany* **19**, 64.
2. Millerd, A., Simon, M. and Stern, H. (1971) *Plant Physiol.* **48**, 419.
3. Beevers, L. and Poulson, R. (1972) *Plant Physiol.* **49**, 476.
4. Hall, T. C., McLeester, R. C. and Bliss, F. A. (1972) *Phytochemistry* **11**, 647.
5. Wright, D. J. and Boulter, D. (1972) *Planta (Berl.)* **105**, 60.
6. Smith, D. L. (1973) *Ann. Botany* **37**, 795.
7. Hill, J. E. and Breidenbach, R. W. (1974) *Plant Physiol.* **53**, 747.
8. Briarty, L. G., Coult, D. A. and Boulter, D. (1969) *J. Exp. Botany* **20**, 358.
9. Graham, T. A. and Gunning, B. E. S. (1970) *Nature, London* **228**, 81.
10. Chibnall, A. C. (1939) *Protein Metabolism in the Plant*, Yale University Press.
11. Bagley, B. W., Cherry, T. H., Rollins, M. L. and Altschul, A. M. (1963) *Am. J. Botany* **50**, 523.
12. Opik, H. (1966) *J. Exp. Botany* **17**, 427.
13. Briarty, L. G., Coult, D. A. and Boulter, D. (1970) *J. Exp. Botany* **21**, 513.
14. Boulter, D. and Barber, J. T. (1963) *New Phytol.* **62**, 301.
15. Osborne, T. B. (1924) *The Vegetable Proteins*, 2nd edn., Longmans-Green, London.
16. Danielsson, C. E. (1949) *Biochem. J.* **44**, 387.
17. Osborne, T. B. and Campbell, G. F. (1898) *J. Am. Chem. Soc.* **20**, 348–362, 362–375, 393–405.
18. Osborne, T. B. (1894) *J. Am. Chem. Soc.* **16**, 633.
19. Osborne, T. B. (1894) *J. Am. Chem. Soc.* **16**, 703.
20. Osborne, T. B. (1894) *J. Am. Chem. Soc.* **16**, 757.
21. Osborne, T. B. and Campbell, G. F. (1898) *J. Am. Chem. Soc.* **20**, 419.
22. Svedberg, T. and Lysholm, A. (1927) *Nova Acta Regiae Soc. Sci. Upsaliensis. Volumen Extra Ordinem Editum*, p. 1, Ser. IV, Vol. 7(4), 3.
23. Tiselius, A. (1930) *Nova Acta Regiae Soc. Sci. Upsaliensis.*
24. Svedberg, T. (1937) *Nature, London* **139**, 1051.
25. Danielsson, C. E. (1952) *Svensk. Kem. Tid.* **64**, 43.
26. Danielsson, C. E. (1949) *Acta Chem. Scand.* **3**, 41.
27. Danielsson, C. E. (1950) *Acta Chem. Scand.* **4**, 762.
28. Danielsson, C. E. (1952) *Acta Chem. Scand.* **6**, 149.
29. Heywood, V. (1971) *Chemotaxonomy of the Leguminosae* (Harborne, J. B., D. Boulter and B. L. Turner, eds.), pp. 1–29. Academic Press, London.
30. Kloz, J. (1971) *Chemotaxonomy of the Leguminosae* (Harborne, J. B., D. Boulter and B. L. Turner, eds.), pp. 309–365. Academic Press, London.
31. Dudman, W. F. and Millerd, A. (1975) *Biochem. Systematics Ecol.* **3**, 25.
32. Schulze-Menz, G. K. (1964) *A. Engler's Syllabus der Pflanzenfamilien*, II. Band (Herausgegeben von Prof. Dr. Hans Melchior), 12th edn., pp. 230–242. Gebruder Borntraeger, Berlin.
33. Nash, A. M. and Wolf, W. J. (1967) *Cereal Chem.* **44**, 183.
34. Nash, A. M., Kwolek, W. F. and Wolf, W. J. (1971) *Cereal Chem.* **48**, 360.
35. Joubert, F. J. (1955) *Biochim. Biophys. Acta* **16**, 370.
36. Joubert, F. J. (1955) *Biochim. Biophys. Acta* **17**, 444.
37. Joubert, F. J. (1956) *Biochim. Biophys. Acta* **19**, 172.
38. Smith, A. K. and Rackis, J. J. (1957) *J. Am. Chem. Soc.* **79**, 633.
39. McKerrow, J. H. and Robinson, A. B. (1971) *Analyt. Biochem.* **42**, 565.
40. Nash, A. M., Kwolek, W. F. and Wolf, W. J. (1974) *Cereal Chem.* **51**, 220.
41. Hulme, A. C., Jones, J. D. and Wooltorton, L. S. (1964) *Phytochemistry* **3**, 173.
42. Loomis, W. D. and Battaille, J. (1966) *Phytochemistry* **5**, 423.
43. Anderson, J. W. (1968) *Phytochemistry* **7**, 1973.
44. Lazarus, N. R., Ramel, A. H., Rustum, Y. M. and Barnard, E. A. (1966) *Biochemistry, N.Y.* **5**, 4003.
45. Clark, J. F. and Jakoby, W. B. (1970) *J. Biol. Chem.* **245**, 6065.
46. Lederer, F. and Jacq, C. (1971) *European J. Biochem.* **20**, 475.
47. Diezel, W., Bohme, H.-J., Nissler, K., Freyer, R., Heilmann, W., Kopperschlager, G. and Hofmann, E. (1973) *European J. Biochem.* **38**, 479.
48. Gray, J. C. and Kekwick, R. G. O. (1973) *Biochem. Soc. Trans.* **1**, 455.
49. Wolf, W. J. and Sly, D. A. (1967) *Cereal Chem.* **44**, 653.
50. Neucere, N. J. (1969) *Analyt. Biochem.* **27**, 15.
51. Pusztai, A. (1966) *Biochem. J.* **99**, 93.
52. Kjeldahl, J. (1896) *Biederm. Zbl.* **25**, 197.
53. Jaffé, W. G., Brücher, O. and Palozzo, A. (1972) *Z. Immun.-Forsch.* **142**, 439.
54. Danielsson, C. E. (1956) *Ann. Rev. Plant Physiol.* **7**, 215.
55. Wolf, W. J. and Briggs, D. R. (1959) *Arch. Biochem. Biophys.* **85**, 186.
56. Hartman, R. J. and Cheng, L. T. (1936) *J. Chin. Chem. Soc., Peiping* **4**, 149.
57. Bailey, C. J. and Boulter, D. (1970) *European J. Biochem.* **17**, 460.
58. Circle, S. J. (1950) *Soybeans and Soybean Products* (K. S. Markley, ed.), pp. 275–370. Vol. I, Interscience, New York.
59. Goring, D. A. I. and Johnson, P. (1955) *Arch. Biochem. Biophys.* **56**, 448.
60. Kelley, J. J. and Pressey, R. (1966) *Cereal Chem.* **43**, 195.
61. Shutov, A. D. and Vaintraub, I. A. (1965) *Ukr. Biokhim. Zh.* **37**, 177.

62. Muntz, K., Horstmann, C. and Scholz, G. (1972) *Kulturpflanze* **20**, 277.
63. Scholz, G., Richter, J. and Manteuffel, R. (1974) *Biochem. Physiol. Pflanzen* (*BPP*) **166**, 163.
64. Wright, D. J. and Boulter, D. (1974) *Biochem. J.* **141**, 413.
65. Wright, D. J. (1973) Ph.D. Thesis, University of Durham.
66. Koshiyama, I. (1968) *Cereal Chem.* **45**, 394.
67. Hasegawa, K., Kusano, T. and Mitsuda, H. (1963) *Agr. Biol. Chem.* **27**, 878.
68. Schwenke, K. D., Schultz, M., Linow, H.-J., Uhlig, J. and Franzke, C. (1974) *Nährung* **18**, 709.
69. Koshiyama, I. (1972) *Intern. J. Peptide Protein Res.* **4**, 167.
Koshiyama, I. (1972) *J. Sci. Food Agr.* **23**, 853.
70. Ghetie, V. and Buzila, L. (1968) *Rev. Roum. Biochim.* **5**, 271.
71. Buzila, L. (1967) *Rev. Roum. Biochim.* **4**, 103.
72. Agrawal, B. B. L. and Goldstein, I. J. (1967) *Biochim. Biophys. Acta* **147**, 262.
73. Mitsuda, H., Kusano, T. and Hasegawa, K. (1965) *Agr. Biol. Chem.* **29**, 7.
74. Koshiyama, I. (1968) *Cereal Chem.* **45**, 405.
75. Vaintraub, I. A. (1965) *Biokhimiya* **30**, 628.
76. Roberts, R. C. and Briggs, D. R. (1965) *Cereal Chem.* **42**, 71.
77. Kitamura, K., Okubo, K. and Shibasaki, K. (1974) *Agr. Biol. Chem.* **38**, 1083.
78. Vaintraub, I. A. and Shutov, A. D. (1964) *Biokhimiya* **29**, 863.
79. Grant, D. R. and Lawrence, J. M. (1964) *Arch. Biochem. Biophys.* **108**, 552.
80. Ericson, M. C. and Chrispeels, M. J. (1973) *Plant Physiol.* **52**, 98.
81. Okubo, K. and Shibasaki, K. (1966) *Agr. Biol. Chem.* **30**, 939.
82. Catsimpoolas, N., Rogers, D. A., Circle, S. J. and Meyer, E. W. (1967) *Cereal Chem.* **44**, 631.
83. Bhatty, R. S., Mackenzie, S. L. and Finlayson, A. J. (1968) *Can. J. Biochem.* **46**, 1191.
84. Steenkamp, D. J. and Joubert, F. J. (1967) *J. S. African Chem. Inst.* **20**, 81.
85. Singh, J. and Dieckert, J. W. (1973) *Prep. Biochem.* **3**, 73.
86. Shetty, K. J. and Rao, M. S. N. (1974) *Analyt. Biochem.* **62**, 108.
87. Pusztai, A. and Watt, W. A. (1970) *Biochim. Biophys. Acta* **207**, 413.
88. Shutov, A. D. and Vaintraub, I. A. (1966) *Biokhimiya* **31**, 726.
89. Johnson, P., Shooter, E. M. and Rideal, E. K. (1950) *Biochim. Biophys. Acta* **5**, 376.
90. Jaffe, W. G. and Hannig, K. (1965) *Arch. Biochem. Biophys.* **109**, 80.
91. Boulter, D., Evans, I. M. and Derbyshire, E. (1973) *Qualitas Plantarum* **23**, 239.
92. Gillespie, J. M. and Blagrove, R. J. (1975) *Aust. J. Plant Physiol.* **2**, 29.
93. Catsimpoolas, N. (1969) *FEBS Letters* **4**, 259.
94. Catsimpoolas, N. and Leuthner, E. (1969) *Biochim. Biophys. Acta* **181**, 404.
95. Kloz, J. and Turkova, V. (1963) *Biol. Plant.* (*Praha*) **5**, 29.
96. Ghetie, V. and Buzila, L. (1964) *Biol. Plant.* (*Praha*) **6**, 202.
97. Osserman, E. F. (1960) *J. Immunol.* **84**, 93.
98. *Methods in Enzymology* (1953), Vol. 1. In progress (S. P. Colowick and N. O. Kaplan, eds.), Academic Press, New York.
99. Koshiyama, I. (1968) *Agr. Biol. Chem.* **32**, 879.
100. Sun, S. M., McLeester, R. C., Bliss, F. B. and Hall, T. C. (1974) *J. Biol. Chem.* **249**, 2118.
101. Reithel, F. J. (1963) *Adv. Protein Chem.* **18**, 123.
102. Cleland, W. W. (1964) *Biochemistry* **3**, 480.
103. Reynolds, J. A. and Tanford, C. (1970) *Proc. Natl. Acad. Sci. U.S.A.* **66**, 1002.
104. Weber, K., Pringle, J. R. and Osborn, M. (1972) *Methods in Enzymology* 26 (S. P. Colowick and N. O. Kaplan, eds.), Academic Press, New York, pp. 3–27.
105. Dlouha, V., Keil, B. and Sorm, F. (1963) *Coll. Czech. Chem. Commun.* **28**, 2969.
106. Vaintraub, I. A. and Nguyen Thanh Thien (1968) *Dokl. Akad. Nauk SSSR* **180**, 1239.
107. Vaintraub, I. A. (1967) *Mol. Biol.* **1**, 807.
108. Masaki, T. and Soejima, M. (1972) *Science Rep. Fac. Agr. Ibaraki Univ.* **20**, 35.
109. Goding, L. A., Bhatty, R. S. and Finlayson, A. J. (1970) *Can. J. Biochem.* **48**, 1096.
110. Margoliash, E., Nisonoff, A. and Reichlin, M. (1970) *J. Biol. Chem.* **245**, 931.
111. Reichlin, M., Nisonoff, A. and Margoliash, E. (1970) *J. Biol. Chem.* **245**, 947.
112. Joubert, F. J. and Cooper, D. C. (1954) *J. S. African Chem. Inst.* **7**, 99.
113. Saverborn, S. and Danielsson, C. E. (1943) *Svensk Kem. Tid.* **55**, 155.
114. Ventura, M. M. and Filho, J. X. (1964) *Anais Acad. Brasil Cienc.* **36**, 21.
115. Johnson, P. (1946) *Trans. Faraday Soc.* **42**, 28.
116. Brand, B. P., Goring, D. A. I. and Johnson, P. (1955) *Trans. Faraday Soc.* **51**, 872.
117. Naismith, W. E. F. and McDavid, H. M. R. (1958) *J. Appl. Chem. London* **8**, 605.
118. Johnson, P. and Shooter, E. M. (1950) *Biochim. Biophys. Acta* **5**, 361.
119. Yotsuhashi, K. and Shibasaki, K. (1973) *Nihon Shokuhin Kogyo Gakkai-Shi* **20**, 519.
120. Vysokos, T. Y., Klivanskaya, V. V. and Sayanova, V. V. (1972) *Rast. Belki* **10**, 42.
121. Okubo, K., Sagara, G. and Shibasaki, K. (1969) *Tohoku J. Agr. Res.* **20**, 222.
122. Kretovich, V. L., Smirnova, T. I. and Frenkel, S. I. (1956) *Biokhimiya* **21**, 842.
123. Kretovich, V. L., Smirnova, T. I. and Frenkel, S. I. (1958) *Biokhimiya* **23**, 128.
124. Joubert, F. J. (1957) *J. S. African Chem. Inst.* **10**, 21.
125. Joubert, F. J. (1955) *J. S. African Chem. Inst.* **8**, 68.
126. Naismith, W. E. F. (1955) *J. Sci. Food Agr.* **6**, 721.
127. Johnson, P. and Richards, E. G. (1962) *Arch. Biochem. Biophys.* **97**, 260.
128. Joubert, F. J. (1955) *J. S. African Chem. Inst.* **8**, 75.
129. Brand, B. P. and Johnson, P. (1958) *Trans. Faraday Soc.* **54**, 1911.
130. Joubert, F. J. (1957) *J. S. African Chem. Inst.* **10**, 16.
131. Vaintraub, I. A. and Nguyen Thanh Thien (1971) *Mol. Biol.* **5**, 59.
132. Sjogren, B. and Svedberg, T. (1930) *J. Am. Chem. Soc.* **52**, 3279.
133. Naismith, W. E. F. (1955) *Nature, London* **175**, 692.
134. Drenth, J. and Wiebenga, E. H. (1955) *Rec. Trav. Chim.* **74**, 813.
135. Polson, A. G. (1939) *Kolloid-Z.* **87**, 149.
136. Svedberg, T. and Sjogren, B. (1930) *J. Am. Chem. Soc.* **52**, 279.
137. Danielsson, C. E. and Ingelman, B. (1947) *Svensk Kem. Tid.* **59**, 162.
138. Mackenzie, S. L. and Blakely, J. A. (1972) *Can. J. Botany* **50**, 1825.
139. Svedberg, T. and Stamm, A. J. (1929) *J. Am. Chem. Soc.* **51**, 2170.
140. Beeckmans, M. L. and Lontie, R. (1946) *Bull. Soc. Chim. Biol.* **28**, 509.
141. Mourgue, M., Barbe, J., Campenio, S., Lanet, J., Savary, J. and Vinet, L. (1971) *Ann. Pharm. Fr.* **29**, 583.
142. Krejci, L. and Svedberg, T. (1934) *J. Am. Chem. Soc.* **56**, 1706.

143. Fuerst, C. R., McCalla, A. G. and Colvin, J. R. (1954) *Arch. Biochem. Biophys.* **49**, 207.
144. Campenio, S., Barbe, J., Lanet, J., Lasry, S. and Savary, J. (1970) *Ann. Pharm. Fr.* **28**, 165.
145. Belozerskii, M. A., Viktorova, L. N. and Shpikiter, V. O. (1968) *Biokhimiya* **33**, 97.
146. Naismith, W. E. F. (1956) *J. Appl. Chem.* **6**, 283.
147. Joubert, F. J. (1955) *Biochim. Biophys. Acta* **16**, 520.
148. Crowfoot, D. and Fankuchen, I. (1938) *Nature, London* **141**, 522.
149. Putzeys, P. and Beeckmans, M. L. (1946) *Bull. Soc. Chim. Biol.* **28**, 503.
150. Putzeys, P., Rondelet-Minet, J. and Rutgeerts, M. J. (1963) *Bull. Soc. Chim. Biol.* **45**, 1145.
151. Barre, R. (1953) *Bull. Soc. Chim. Biol.* **35**, 899.
152. Norval, F. B. (1937) *J. Biol. Chem.* **120**, 63.
153. Joubert, F. J. (1956) *J. S. African Chem. Inst.* **9**, 86.
154. Sinha, N. K. and Sen, A. (1962) *Trans. Bose Res. Inst.* **25**, 37.
155. Ventura, M. M. and Lima, I. H. (1963) *Anais Acad. Brasil Cienc.* **35**, 55.
156. Joubert, F. J. (1955) *J. S. African Chem. Inst.* **8**, 43.
157. Quensel, O. (1942) Dissertation, University of Uppsala, Sweden.
158. Pence, J. W. and Elder, A. H. (1953) *Cereal Chem.* **30**, 275.
159. Morita, Y. and Yoshida, C. (1968) *Agr. Biol. Chem.* **32**, 664.
160. Sjogren, B. and Spychalski, R. (1930) *J. Am. Chem. Soc.* **52**, 4400.
161. Jackson, P., Boulter, D. and Thurman, D. A. (1969) *New Phytol.* **68**, 25.
162. Shvarts, V. S. and Vaintraub, I. A. (1967) *Biokhimiya* **32**, 162.
163. Wolf, W. J., Sly, D. A. and Kwolek, W. F. (1966) *Cereal Chem.* **43**, 80.
164. Fukushima, D. (1968) *Cereal Chem.* **45**, 203.
165. Tombs, M. P. (1965) *Biochem. J.* **96**, 119.
166. Catsimpoolas, N., Kenney, J. A., Meyer, E. W. and Szuhaj, B. F. (1971) *J. Sci. Food Agr.* **22**, 448.
167. Okubo, K., Asano, M., Kimura, Y. and Shibasaki, K. (1969) *Agr. Biol. Chem.* **33**, 463.
168. Zmrhal, Z. (1967) *Coll. Czech. Chem. Commun.* **32**, 2337.
169. Dlouha, V., Keil, B. and Sorm, F. (1964) *Coll. Czech. Chem. Commun.* **29**, 1835.
170. Catsimpoolas, N. and Wang, J. (1971) *Analyt. Biochem.* **44**, 436.
171. Tombs, M. P. and Lowe, M. (1967) *Biochem. J.* **105**, 181.
172. Wolf, W. J. and Briggs, D. R. (1958) *Arch. Biochem. Biophys.* **76**, 377.
173. Brohult, S. and Sandegren, E. (1954) *The Proteins* (H. Neurath and K. Bailey, eds.), Vol. IIA, pp. 487–512. Academic Press, New York.
174. Brand, B. P. and Johnson, P. (1956) *Trans. Faraday Soc.* **52**, 438.
175. Burk, N. F. and Greenberg, D. M. (1930) *J. Biol. Chem.* **87**, 197.
176. Cater, C. W. and Naismith, W. E. F. (1958) *Arch. Biochem. Biophys.* **77**, 98.
177. Evans, W. J. (1957) *Arch. Biochem. Biophys.* **72**, 226.
178. Wright, D. J. and Boulter, D. (1973) *Phytochemistry* **12**, 79.
179. Trautman, R. (1956) *J. Physiol. Chem.* **60**, 1211.
180. Wolf, W. J., Rackis, J. J., Smith, A. K., Sasame, H. A. and Babcock, G. E. (1958) *J. Am. Chem. Soc.* **80**, 5730.
181. Anderson, D. G. and McCalla, A. G. (1960) *Can. J. Biochem. Physiol.* **38**, 275.
182. Shetty, K. J. and Rao, M. S. N. (1973) *Indian J. Biochem. Biophys.* **10**, 149.
183. Paterson, B. and Strohman, R. C. (1970) *Biochemistry* **2**, 4094.
184. Briggs, D. R. and Wolf, W. J. (1957) *Arch. Biochem. Biophys.* **72**, 127.
185. Kolthoff, I. M., Anastasi, A. and Tan, B. H. (1958) *J. Am. Chem. Soc.* **80**, 3235.
186. Hall, C. E. (1950) *J. Biol. Chem.* **185**, 45.
187. De Robertis, E., Franchi, C. M. and Podolsky, N. (1953) *Biochim. Biophys. Acta* **11**, 507.
188. St. Angelo, A. J., Yatsu, L. Y. and Altschul, A. M. (1968) *Arch. Biochem. Biophys.* **124**, 199.
189. Saio, K., Matsuo, T. and Watanabe, T. (1970) *Agric. Biol. Chem.* **34**, 1851.
190. Vaintraub, I. A. and Shutov, A. D. (1971) *Biokhimiya* **36**, 1086.
191. Bailey, K. (1942) *Biochem. J.* **36**, 140.
192. Johnson, P. and Naismith, W. E. F. (1953) *Disc. Faraday Soc.* **13**, 98.
193. Altschul, A. M., Yatsu, L. Y., Ory, R. L. and Engelman, E. M. (1966) *Ann. Rev. Plant Physiol.* **17**, 113.
194. Sumner, J. B., Graten, N. and Eriksson-Quensel, I. (1938) *J. Biol. Chem.* **125**, 45.
195. Naismith, W. E. F. (1955) *Biochim. Biophys. Acta* **16**, 203.
196. Boulter, D. and Derbyshire, E. (1971) *Chemotaxonomy of the Leguminosae* (Harborne, J. B., D. Boulter and B. L. Turner, eds.), pp. 285–308. Academic Press, London.
197. Shvarts, V. S. (1968) *Trudy po Khimii Prirodnykh Soedinenii* **7**, 134.
198. Dawson, R. (1971) *Analyt. Biochem.* **41**, 305.
199. Gerritsen, T. (1956) *Biochim. Biophys. Acta* **22**, 269.
200. Bailey, C. J. and Boulter, D. (1972) *Phytochemistry* **11**, 59.
201. Vaintraub, I. A. and Gofman, Y. Y. (1961) *Biokhimiya* **26**, 13.
202. Vaintraub, I. A., Shutov, A. D. and Klimenko, V. G. (1962) *Biokhimiya* **27**, 349.
203. Klozova, E. and Kloz, J. (1972) *Biol. Plant. (Praha)* **14**, 379.
204. Bourdillon, J. (1949) *J. Biol. Chem.* **180**, 553.
205. Daussant, J., Neucere, N. J. and Yatsu, L. Y. (1969) *Plant Physiol.* **44**, 471.
206. Catsimpoolas, N. (1970) *Cereal Chem.* **47**, 70.
207. Klivanskaya, V. and Sayanova, V. (1973) *Belki Semyan Kul't Rast*, pp. 48–61.
208. Catsimpoolas, N. and Ekenstam, C. (1969) *Arch. Biochem. Biophys.* **129**, 490.
209. Racusen, D. and Foote, M. (1971) *Can. J. Botany* **49**, 2107.
210. Waterman, H. C., Johns, C. O. and Jones, D. B. (1923) *J. Biol. Chem.* **55**, 93.
211. McLeester, R. C., Hall, T. C., Sun, S. M. and Bliss, F. A. (1973) *Phytochemistry* **12**, 85.
212. Jones, D. B. and Johns, C. O. (1916) *J. Biol. Chem.* **28**, 67.
213. Sumner, J. B. (1919) *J. Biol. Chem.* **37**, 137.
214. Sumner, J. B. and Somers, G. F. (1944) *Laboratory Experiments*. Academic Press.
215. Mourgue, M., Baret, R., Renai, J. and Savary, J. (1968) *C. R. Soc. Biol.* **162**, 1128.
216. Rossi-Fanelli, A., Antonini, E., Brunori, M., Bruzzesi, M., Caputo, A. and Satriani, F. (1964) *Biochem. Biophys. Res. Commun.* **15**, 110.
217. Wolf, W. J. (1972) *Fd Technol.* **26**, 44.
218. Sawai, H. and Morita, Y. (1970) *Agr. Biol. Chem. (Tokyo)* **34**, 53.
219. Sawai, H. and Morita, Y. (1970) *Agr. Biol. Chem. (Tokyo)* **34**, 61.
220. Koshiyama, I. (1971) *Agr. Biol. Chem.* **35**, 385.
221. Blagrove, R. J. and Gillespie, J. M. (1975) *Aust. J. Plant Physiol.* **2**, 13.
222. Koshiyama, I. (1970) *Agr. Biol. Chem.* **34**, 1815.
223. Pusztai, A. and Watt, W. B. (1974) *Biochim. Biophys. Acta* **365**, 57.
224. Saverborn, S., Danielsson, C. E. and Svedberg, T. (1944) *Svensk Kem. Tidskr.* **56**, 75.
225. Wall, J. S. (1964) *Proteins and Their Reactions* (H. W. Schultz and A. F. Anglemier, eds.). Symposium on Foods pp. 315–341. Avi Westport, Connecticut.

226. Houston, D. F. and Mohammad, A. (1970) *Cereal Chem.* **47**, 5.
227. Vaintraub, I. A. and Shutov, A. D. (1969) *Biokhimiya* **34**, 984.
228. Finlayson, A. J., Bhatty, R. S. and Christ, C. M. (1969) *Can. J. Botany* **47**, 679.
229. Catsimpoolas, N., Ekenstam, C. and Meyer, E. W. (1969) *Cereal Chem.* **46**, 357.
230. Gillespie, J. M. and Blagrove, R. J. (1974) *Search* **5**, 600.
231. Boulter, D., Evans, I. M., Thompson, A. and Yarwood, A. (1973) *Nutritional Improvement of Food Legumes by Breeding* (M. Milner, ed.), pp. 205–215. Protein Advisory Group of the United Nations System, United Nations, New York 10017.
232. Millerd, A. (1975) *Ann. Rev. Plant Physiol.* **26**, (in press).
233. Dure, L. (1975) *Ann. Rev. Plant Physiol.* **26**, (in press).
234. Zalik, S. and Jones, B. L. (1973) *Ann. Rev. Plant Physiol.* **24**, 47.
235. Wheeler, C. T. and Boulter, D. (1966) *Biochem. J.* **100**, 53P.
236. Payne, P. I. and Boulter, D. (1969) *Planta (Berlin)* **84**, 263.
237. Bailey, C. J., Cobb, A. and Boulter, D. (1970) *Planta (Berlin)* **95**, 103.
238. Harris, N. and Boulter, D., to be published.
239. Northcote, D. H. (1971) *Endeavour* **31**, 26.
240. Wolf, W. J. (1970) *J. Am. Oil Chem. Soc.* **47**, 107.
241. Koshiyama, I. (1972) *Agric. Biol. Chem.* **36**, 62.
242. Neucere, N. J. and Ory, R. L. (1970) *Plant Physiol.* **45**, 616.
243. Morris, G. F. I., Thurman, D. A. and Boulter, D. (1970) *Phytochemistry* **9**, 1707.
244. Kondo, K., Mori, S. and Kojima, M. (1954) *Bull. Res. Inst. Food Sci. Kyoto Univ.* **15**, 37.
245. Ghetie, V. (1959) *Studii Cerc. Biochim.* **2**, 293.
246. Klimenko, V. G. and Berezovikov, A. D. (1963) *Biokhimiya* **28**, 238.
247. Klimenko, V. G. and Pinegina, R. I. (1964) *Biokhimiya* **29**, 377.
248. Kloz, J., Turkova, V. and Klozova, E. (1966) *Biol. Plant. (Praha)* **8**, 164.
249. Hill, J. E. and Breidenbach, R. W. (1974) *Plant Physiol.* **53**, 747.
250. Catsimpoolas, N., Campbell, T. G. and Meyer, E. W. (1968) *Plant Physiol.* **43**, 799.
251. Morris, G. F. I. (1968) Ph.D. Thesis, University of Liverpool.
252. Daussant, J. (1968) *Serol. Mus. Bull.* **40**, 6.
253. Buzila, L. (1969) *Studii Cerc. Biochim.* **12**, 229.
254. Buzila, L. (1969) *Studii Cerc. Biochim.* **12**, 327.
255. Danielsson, C. E. (1951) *Acta Chem. Scand.* **5**, 541.
256. Morawiecka, B. (1961) *Acta Biochim. Polon.* **8**, 313.
257. Catsimpoolas, N., Ekenstam, C., Rogers, D. A. and Meyer, E. W. (1968) *Biochim. Biophys. Acta* **168**, 122.
258. Daussant, J. (1975) *The Chemistry and Biochemistry of Plant Proteins* (J. B. Harborne and C. F. Van Sumere, eds.), pp. 31–69. Academic Press, London.
259. Ghetie, V. (1966) *Rev. Roum. Biochim.* **3**, 353.
260. Ryan, C. A. (1973) *Ann. Rev. Plant Physiol.* **24**, 173.
261. Vickery, H. B. (1945) *Physiol. Rev.* **25**, 347.
262. Lugg, J. W. H. (1949) *Adv. Protein Chem.* **5**, 229.
263. Stahman, M. A. (1963) *Ann. Rev. Plant Physiol.* **14**, 137.
264. Sun, S. M. and Hall, T. C. (1975) *J. Agr. Food Chem.* **23**, 184.
265. Robinson, A. B., Irving, K. and McCrea, M. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 2122.
266. Boulter, D. (1974) *Nobel 25. Chemistry in Botanical Classification* (G. Bendz and J. Santesson, eds.), pp. 211–216. Almqvist & Wiksell, Uppsala.
267. Wallace, D. G. and Boulter, D., in press.
268. Jackson, P., Milton, J. M. and Boulter, D. (1967) *New Phytol.* **66**, 47.
269. Blagoveshchenskii, A. V. (1967) *Biochemistry of Leguminous Plants*, pp. 3–88. Biokhimiya Bobovykh Rastonii. Akad. Nauk. SSSR.
270. Hynes, M. J. (1968) *Aust. J. Biol. Sci.* **21**, 827.
271. Davis, D. R. (1973) *Nature, New Biology* **245**, 30.