*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 $S_{20,w} = 6{\cdot}5$–$8{\cdot}1$ 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

elutions 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, $\beta$ conglycinin and $\gamma$ 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 ($\times$ 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, $\beta$-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 AGl-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] álso 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($S^0_{20}$) | 396000 | 0·72 | 3·2 | 1·38 | [115] |
| | 1·25 | 340000 | | | | [116] |
| | 13·45 | | | | | [117] |
| | 13·45 | | | | | [50] |
| | 13·35 | | | | | [85] |
| | 14·6($S_{20}$) | 330000 | 0·72 | 3·86($D^0_{20}$) | 1·216 | [118] |
| | 12·0($S_{20,w}$) | | | | | [86] |
| | 14·7($S^0_{20,w}$) | 350000 | 0·721 | 3·67($D^0_{20,w}$) | 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($S^0_{20,w}$) | 309000–322000 | 0·715 | 3·48($D_{20,w}$) | 1·40 | [69] |
| | 12·3 | | | | | [121] |
| | 12·8 | 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($S^0_{20}$) | 393000 | | 2·99($D^0_{20}$) | 1·46 | [124] |
| *L. angustifolius* | 13·05($S_{20}$) | | | | | [16] |
| | 11·6($S^0_{20}$) | 336000 | | 3·16($D^0_{20}$) | 1·46 | [35] |
| *L. luteus* | 11·34 | | | | | [16] |
| | 11·6 | | | | | [125] |
| | 11·48 | | | | | [126] |
| *L. polyphyllus* | 13·24 | | | | | [16] |
| *Medicago sativa* | 11·41($S_{20}$) | | | | | [16] |
| *Phaseolus aureus* | 11·3 | | | | | [80] |
| *P. coccineus* | 12·1 | | | | | [16] |
| *P. nanus* | 11·1 | | | | | [16] |
| *P. vulgaris* | 11·3 | | | | | [16] |
| | 11·8 | 340000 | | | | Derbyshire, E. (unpubl.) |
| *Pisum sativum* | 12·6 | | | | | [79] |
| | 12·4 | 330000 | 0·735 | 3·49($D_{20}$) | | [16] |
| | 13·1 | 398000 | 0·733 | 2·99($D^0_{20}$) | 1·43 | [127] |
| | 12·1($S^0_{20}$) | 388000 | | | | [116] |
| | 12·3($S_{20}$) | | | | | [128] |
| | 13·7($S^0_{20,w}$) | 410000 | | 3·02($D^0_{20}$) | 1·43 | [129] |
| *Trifolium hybridum* | 12·90($S_{20}$) | | | | | [16] |
| *T. pratense* | 11·22($S_{20}$) | | | | | [16] |
| *Vicia faba* | 11·4 | 328000 | | | | [64] |
| | 11·8 | | | | | [130] |
| *V. sativa* | 12·8 | 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($S^0_{20}$) | 377000 | | 3·15($D^0_{20}$) | 1·41 | [112] |
| *Aesculus hyppocastanum* | 12·9($S_{20}$) | 430000 | 0·748 | 2·9 | | [113] |
| *Anacardium occidentale* | 12·85($S^0_{20,w}$) | 238000–260000 | 0·737 | 4·75–5·0 ($D_{20,w}$) | 1·07 | [114] |
| *Balanites aegyptica* | 12·3($S_{obs}$) | | 0·685 | | | [133] |
| *Bertholletia excelsa* | | 303000 | | | | [134] |
| | 13·3($S^0_{20}$) | 295000 | | 4·26 | | [135] |
| | 11·78($S^0_{20}$) | 212000 | 0·743 | | 1·02 | [136] |
| *Beta vulgaris* | 13·9($S_{20}$) | ~250000 | | 5·3 | | [137] |
| *Brassica napus* | 12($S^0_{20}$) | | | | | [109] |
| *B. juncea* | 12($S^0_{20,w}$) | | | | | [138] |
| *B. nigra* | 11·8 | | | | | [138] |
| *B. hirta* | 11·5 | | | | | [138] |
| *Cannabis sativa* | 13·2 | 334000 | | | | [116] |
| | 12·8 | 212000 | | 5·6($D^0_{20,w}$) | | [139] |
| | 12·8 | 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 ($cm^2sec^{-1} \times 10^{-7}$) | Frictional ratio | Reference |
|---|---|---|---|---|---|---|
| *Cucurbita maxima* | 12·1 | | | | | [143] |
| | | 340000 | | | | [144] |
| *Fagopyrum esculentum* | 13·0($S^0_{20,w}$) | 270000 | | | | [145] |
| *Gossypium barbadense* | 13·0($S_{20,w}$) | | | | | [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($S^{\circ}_{20,w}$) | 24300 | 6 | Gly | Basic | [106, 130] |
| | *B* | 2·28($S^{\circ}_{20,w}$) | 37600 | 4 | Leu | Acidic | |
| | *C* | 2·25($S^{\circ}_{20,w}$) | 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. $\alpha$ and $\beta$, with MW's of 35000, and $\gamma$ and $\delta$ 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 ($cm^2sec^{-1} \times 10^{-7}$) | 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($S^{\circ}_{20}$) | 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($S^{\circ}_{20,w}$) | 330000 | 0·729 | | 2·17 | [76] |
| | 7·9($S^{\circ}_{20,w}$) | 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($S^{\circ}_{20}$) | 204000 | | 3·80 | 1·43 | [124] |
| *L. angustifolius* | 8·2($S_{20}$) | | | | | [16] |
| | 7·8($S^{\circ}_{20}$) | 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($S^{\circ}_{20}$) | | | | | [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($S^{\circ}_{20,w}$) | 140000 | | 5·5 | | [87] |
| | 6·5($S^{\circ}_{20}$) | | | | | [130] |
| *Pisum sativum* | 8·1($S_{20}$) | | | | | [16] |
| | 7·1($S^{\circ}_{20}$) | | | | | [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($S^{\circ}_{20}$) | | | | | [130] |
| | 7·1($S_{20,w}$) | 150000 | | | | [178] |
| *V. sativa* | 7·1($S_{20}$) | | | | | [16] |
| | 7·5($S^{\circ}_{20,w}$) | 193000 | | | | [88] |
| *Vigna unguiculata* | 7·3($S^{\circ}_{20}$) | | | | | [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, Podalyrieae, 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 phaseolin [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] |
| | *Trifollum 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 ½-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 $\beta$ conglycinin, one of four globulins from *Glycine max* separated by immunoelectrophoresis [208], and Koshiyama (personal communication) has identified his own protein with $\beta$ conglycinin. Catsimpoolas and Ekenstam [208] reported identity between Koshiyama's protein and $\gamma$ conglycinin, an immunoelectrophoretic component different from $\beta$ conglycinin; however, Koshiyama has isolated $\gamma$ 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 $\gamma$ 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 $\gamma$ 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 conphaseolin, 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 12S 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, $\gamma$ 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 $\gamma$ 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 $\gamma$ 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*<br>*Glycine max*<br>*Lupinus angustifolius*<br>*Phaseolus aureus*<br>*Phaseolus vulgaris*<br>*Pisum sativum* | *Canavalia ensiformis*<br>*Phaseolus vulgaris*<br>*Pisum sativum*<br>*Vicia faba* | *Phaseolus aureus*?<br>*Phaseolus vulgaris* | *Canavalia ensiformis*?<br>*Lupinus angustifolius*<br>*Phaseolus vulgaris*?<br>*Pisum sativum*?<br>*Vicia faba* | *Canavalia ensiformis*?<br>*Phaseolus vulgaris*?<br>*Phaseolus aureus*?<br>*Pisum sativum*?<br>*Vigna unguiculata* |

Table 13. Amino acid composition of the subunit with MW 81000 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 81000 was reported (Table 13). Comparison between this and the composition of the parent protein suggests that the amino acid composition of the 51000 subunit is different from that of the 81000 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.

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 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($S^{\circ}_{20}$) |
| *Cytisus supinus* | 1·8($S_{20}$) |
| *Lupinus albus* | 2·7($S^{\circ}_{20}$) |
| *L. luteus* | 2·0($S^{\circ}_{20}$) |
| *Glycine max* | 2·8($S^{\circ}_{20,w}$) |
| | 2·3($S^{\circ}_{20,w}$) |
| *Phaseolus coccineus* | 4·3($S^{\circ}_{20}$) |
| *P. vulgaris* | 4·9($S^{\circ}_{20}$) |
| Non-Legumes | |
| *Beta vulgaris* | 1–2S |
| *Brassica hirta* | 1·8($S^{\circ}_{20,w}$) |
| *B. juncea* | 1·8($S^{\circ}_{20,w}$) |
| *B. nigra* | 1·8($S^{\circ}_{20,w}$) |
| *Helianthus annuus* | 1·7($S^{\circ}_{20}$) |
| *Ricinus communis* | 1·6($S^{\circ}_{20}$) |
| *Telfairea pedata* | 1·6($S^{\circ}_{20}$) |
| *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($S_{20}$) |

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 | 32600 | 18200 | 27600 | 25400 |
| 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).

11*S* and 7*S* 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.