Phytochemistry, Vol. 26, No. 1, pp. 47–50, 1987.
Printed in Great Britain.

0031-9422/87 $3.00 + 0.00
Pergamon Journals Ltd.

CONSTANCY OF COMPOSITION OF STORAGE PROTEINS DEPOSITED IN *PISUM SATIVUM* SEEDS

JEAN-CLAUDE HUET, JACQUES BAUDET and JACQUES MOSSÉ*

Laboratoire d'Etude des Protéines, Département de Physiologie et Biochimie Végétales, Centre INRA, 78000, Versailles, France

(*Revised received* 17 *April* 1986)

Key Word Index—*Pisum sativum*; Leguminosae; seeds; storage proteins; accumulation; amino acid composition.

Abstract—Proteins from *Pisum sativum* seeds with a wide range of protein contents were compared with regard to amino acid composition. The amino acids were determined with a high degree of accuracy in six different hydrolysates for each of the 33 samples studied. The results showed that the composition of the extra proteins deposited in the medium to high protein seeds remains perfectly constant regardless of the protein level, growth conditions or genotype of the peas.

INTRODUCTION

The rates of proteosynthesis and accumulation of seed storage protein can change widely in *Pisum* species depending on growth conditions [1, 2], seed location in the plant [3] and possibly genotype [2, 4]. These variations result in seeds with very different protein contents. It has been shown [1–4] that even for pea seeds at the same stage of development, protein content (6.25 N) can vary from 15 to 40% of the seed dry weight. Despite several detailed studies of the major storage fractions accumulated in pea seeds (see refs. [5, 6] for reviews), the composition of the extra proteins deposited in high protein seeds, when compared to that of low protein seeds, has not, until now, been investigated for any species. In order to characterize the supplement of proteins deposited in high protein seeds, the amino acid compositions of 33 seed samples of round garden pea selected for their regular distribution over a wide range of protein content were determined with the maximum accuracy now available. This was done for 14 samples of the same genotype which differed in protein content, as well as for one to five samples from each of ten other varieties. The variation in total seed protein composition allowed us to determine the composition of the extra proteins deposited in high protein seeds.

RESULTS AND DISCUSSION

The 11 different varieties investigated for their seed proteins (ten from *Pisum sativum* L. and one from *P. transcaucasicum* L., the number of samples analysed and the total nitrogen content (N) of the samples are set out in Table 1. This table shows that the N range of the first variety (Amino) overlapped the ranges of the six following varieties, each represented by one to five different samples. The last four samples (and varieties) gave the highest N values.

For each protein amino acid (i) determined, its level (D_i) in the seed (in g of amino acid nitrogen per 100 g of seed dry wt) was plotted as a function of N. The results for lysine are shown in Fig. 1. It can be seen that D_{Lys} is a linear function of N. The same result (not shown) was found for each of the other protein amino acids of pea seeds, meaning that equations such as $D_i = \alpha_i N + \beta_i$ account for variations of D_i with N, α_i being the slope of the regression line of the i-th amino acid and β_i its intercept with the ordinate. Rather than giving an extensive table of the amino acid compositions of the 33 seed samples analysed, the values found for these coefficients are reported in Table 2, with the corresponding coefficients of correlation r_i. It should be noted that the latter have values of *ca* 0.97 or greater with the exceptions of tyrosine, tryptophan, methionine and cysteine, which are typically difficult to analyse. In other words, the values of r_i are an index of the accuracy of the amino acid determination rather than an index of the linearity between D_i and N.

As far as free amino acids and non-protein compounds are concerned, according to the literature data [7], these correspond to a low percentage of total nitrogen in pea

Table 1. Varieties investigated, number of seed samples analysed and range of nitrogen percentage (on a seed dry matter basis)

Variety	No. of samples	Nitrogen content from	to
Amino	14	2.84	4.36
Vendevil	3	2.99	3.88
Finale	5	3.29	4.26
Frimas	3	3.49	4.13
Miranda	2	3.58	3.67
570	1	4.36	
Midiver	1	4.40	
H4	1	4.40	
Transcaucasicum	1	4.72	
Gullivert	1	4.86	
Fin des gourmets	1	5.15	

*To whom correspondence should be addressed.

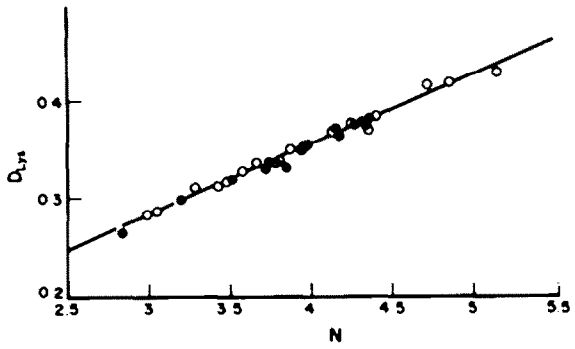

Fig. 1. Variation of amino acid content D_{Lys} (g of lysine nitrogen/100 g seed dry matter) as a function of total seed nitrogen percentage N (on a dry matter basis). ● —● variety Amino; O O, other varieties.

seed and can be neglected as a first approximation. Thus the existence of relationships between D_i and N becomes apparent for the protein amino acids in pea seed: as protein accumulates, the level of each protein amino acid can only increase. Theoretically, the level could remain constant, provided that all the polypeptide chains deposited were devoid of one or more amino acid(s). But, realistically, such a deficiency cannot occur in more than a very small number of polypeptides. Hence, as protein accumulates, seed nitrogen levels increase and are inevitably bound to protein levels. However, the nature of the relationships between D_i and N is not obvious. In a study dealing with a single pea variety, Eppendorfer and Bill [1] could not decide between a first or second degree equation for these relationships, whereas others [4] were favourably inclined towards a second degree equation and possible variations in coefficients depending on variety. Relationships with protein content have already been published for the sulphur amino acids [8, 9]. The present results show that the relationships between D_i and N are perfectly linear for all the protein amino acids. Among the set of seeds analysed, calculations were made in order to determine whether significant deviations could be found between regression lines for the subset of the 14 samples of variety Amino (Fig. 1) and regression lines for the subset of the 19 other samples corresponding to ten other varieties (Fig. 1). These calculations showed that deviations between the two subsets were not significant for any of the protein amino acids. Therefore, contrary to what has so far been assumed or suggested, any condition able to modify N, whether or not genetic in origin, corresponds to the same changes in amino acids. Furthermore, since the relationships are linear, the same is true for any mixture of grains of either the same variety with different N or even of different varieties: in any case, amino acid composition can be predicted from N.

Table 2. Slope (α_i), intercept (β_i) (± s.d.) and correlation coefficient (r_i) of regression lines representing amino acid nitrogen content (D_i) against nitrogen content (N), both expressed in g/100 g of seed dry matter

Amino acid	α_i ± s.d. × 10^4	β_i ± s.d. × 10^4	r_i* × 10^3
Gly	404 ± 10	407 ± 40	990
Ala	335 ± 19	349 ± 37	988
Val	316 ± 12	217 ± 49	977
Leu	448 ± 12	115 ± 46	990
Ile	250 ± 12	189 ± 48	966
Ser	357 ± 11	197 ± 44	985
Thr	194 ± 7	350 ± 29	978
Tyr	109 ± 7	224 ± 28	939
Phe	196 ± 8	242 ± 32	975
Trp	66 ± 5	13 ± 21	913
Pro	290 ± 12	150 ± 49	973
Met	44 ± 3	61 ± 12	935
Cys	43 ± 7	256 ± 27	753
Lys	716 ± 17	685 ± 66	992
His	413 ± 12	1 ± 49	986
Arg	3126 ± 135	− 4813 ± 537	972
Asx†	815 ± 19	− 272 ± 74	991
Glx†	1068 ± 19	− 272 ± 74	995
NH_3‡	839 ± 32	− 250 ± 126	979

*Significance level of r_i: 0.55 ($P = 0.001$).
†Calculated as Asp and Glu, respectively.
‡Amide nitrogen calculated as NH_3.

This linearity allowed us to determine the amino acid composition of the extra proteins deposited when passing from low to high protein seeds. Let us consider two seed samples differing in their nitrogen content, N_1 and N_2 ($> N_1$). The corresponding amounts of the i-th amino acid in the seeds are $D_i(N_1)$ and $D_i(N_2)$, respectively Hence the amino acid content of extra proteins accumulated when passing from the first to the second sample and expressed in g of amino acid nitrogen per g of extra protein nitrogen is equal to the ratio $[D_i(N_2) - D_i(N_1)]/(N_2 - N_1)$. This ratio is itself equal to α_i, which means that it depends neither on N_1 nor on N_2. This shows that the amino acid composition of extra proteins accumulated in pea seeds higher in protein than the lowest protein seeds remains constant regardless of the seed protein content. As far as we know, this is the first time that such a constancy of composition has been demonstrated.

It can be assumed that these extra proteins consist of a mixture of the three major protein fractions of pea seed (albumins, legumins and vicilins). On the one hand, the ratios of these fractions in round pea seeds can be calculated from literature data with N values ranging from 3 to 5%: *ca* 40% legumins, 35% vicilins and 25% albumins [10]. On the other hand, the amino acid compositions of these fractions have recently been determined for albumins [11]. For the two globulin fractions the compositions can be deduced from the nucleotide sequences of legumin [12] and vicilin [13] cDNAs. Although the deduced compositions concern single polypeptide chains, they are in excellent agreement with those directly determined from total legumins and total vicilins [14]. They have thus been used for the present discussion and reported with albumin in Table 3 together with the calculated composition of a mixture reconstituted (CM) from the three major storage protein fractions, alongside extra protein (EP) composition as determined in the present study. The latter two compositions are similar with only small deviations for a few amino acids. Only arginine shows a significant difference which can be explained by its very high level in the free amino acid pool, corresponding to *ca* 7% of total seed nitrogen (unpublished results). This confirms that extra proteins

Table 3. Compositions of the three major fractions of storage proteins from the pea seed: legumins (Leg), vicilins (Vic) and albumins (Alb) calculated from refs. [12], [13] and [11], respectively. Comparison of a calculated mixture (CM)* of these three fractions with extra proteins (EP). All data are given in residues per 1000 residues

Amino acid	Major fractions Leg	Vic	Alb	CM	EP
Gly	75	47	101	72	66
Ala	71	47	93	68	54
Val	50	59	56	55	51
Leu	73	99	49	76	73
Ile	42	54	39	45	40
Ser	59	88	66	71	58
Thr	30	26	62	37	32
Tyr	24	24	31	26	18
Phe	42	52	38	45	32
Trp	6	0	11	5	5
Pro	48	43	53	48	47
Met	8	0	11	6	7
Cys	10	0	n.d.	4†	7
Lys	44	66	82	61	58
His	20	12	20	17	22
Arg	97	62	37	70	126
Asx	123	141	122	129	132
Glx	177	180	129	166	173

*Calculated mixture = 0.4 × Leg + 0.35 × Vic + 0.25 × Alb.
†Minimum value.
n.d., not determined.

consist chiefly of a mixture of the major storage protein fractions of pea seed, with ratios practically remaining constant, contrary to total seed proteins.

In spite of the limited number of varieties involved in the present study, we suggest that the values of the coefficients α_i and β_i are quite general and can be extended to all round seeded varieties of *P. sativum*. This is all the more plausible because homologous results have recently been found with maize seeds [15]. There are, of course, some rare exceptions corresponding to two different kinds of events. Sulphur deficiency has been shown to result in significant modification of seed storage protein composition [16–18]. It is then difficult to predict such a modification. In addition to this phenotypic occurrence, biochemical mutants of *P. sativum* will probably fail to conform to the relationships discovered in the present work just as do opaque-2 or floury-2 mutants of maize when compared to normal maizes [19].

Two conclusions can be drawn on the deposition of storage proteins. First, as protein accumulation in legume seeds displays sigmoidal kinetics, and the time courses for each of the storage proteins/g dry wt of seed [20] exhibit maxima which differ according to the storage protein fraction as well as to the final amount of protein accumulated, the present results show that the ratios between the maxima of each storage protein fraction remain constant regardless of the growth conditions and genotype of *P. sativum*. This means that a very tight developmental regulation of storage proteins occurs, by mechanisms already discussed by Boulter and co-workers [5, 21]. Second, it can be noted that extra proteins accumulated in pea seed are particularly rich in high nitrogen amino acids. It is known that most proteins have a nitrogen content of *ca* 16% (g of nitrogen per 100 g of protein, i.e. 100 g of amino acid residues). In fact, this percentage varies widely depending on the amino acid in question. It is equal to 16% for serine (e.g. seryl residues) but is lower for 12 other amino acids, dropping to 8.6% for tyrosine, while it is higher than 19.5% for seven amino acids: arginine (35.9%), histidine (30.6%), asparagine and glycine (24.6%), glutamine and lysine (21.9%) and alanine (19.7%). Significantly, six out of these seven (histidine excepted) are those that exhibit the highest molar ratios in extra proteins (with the exception of leucine). In other words, everything occurs as if the pea seed was storing the highest possible amount of nitrogen in the minimum number of moles of amino acids. This trend is curiously convergent with the suggestion that seed storage proteins are packed together in the smallest possible volume, as a direct consequence of their tertiary and quaternary structures [6, 22].

EXPERIMENTAL

Material. The seed samples investigated were taken from field-grown garden peas cultivated in diverse locations and conditions under the control of the French Technical Institute of Cereals and Forage (I.T.C.F.). They were selected for their distribution over a wide range of protein content, for a given variety (such as Amino) as well as for different varieties (see Table 1). The conditions of seed sampling and milling for analysis have been previously detailed [23].

Analytical methods. Dry matter content and nitrogen content (microkjeldahl) were determined in triplicate. Amide nitrogen corresponding to half the amount of asparagine plus glutamine nitrogen was determined as free NH_3 obtained from a short hydrolysis (3 hr in 2 M HCl at 115°). Amino acids were determined from five separate analyses made by single CC of five different hydrolysates: four of these were acid (6 M HCl) in order to account for losses due either to degradation or to incomplete release (15, 24 and 48 hr plus an 18 hr hydrolysis of a sample previously oxidized by performic acid); an extra alkaline hydrolysis in $Ba(OH)_2$ enabled the tryptophan to be determined in triplicate. All the details of the methods used have been described in a previous paper [23].

REFERENCES

1. Eppendorfer, W. H. and Bille, S. W. (1974) *Plant Soil* **41**, 33.
2. Matthews, P. and Arthur, E. (1985) *The Pea Crop* (Hebblethwaite, P. D., Heath, M. C. and Dawkins, T. C. K., eds) p. 369. Butterworth, London.
3. Baudet, J., Cousin, R. and Mossé, J. (1977) *Protein Quality from Leguminous Crops*, p. 316. Commission of the European Communities, EUR. 5686 EN.
4. Holt, N. W. and Sosulski, F. W. (1979) *Can. J. Plant Sci.* **59**, 653.
5. Boulter, D. (1981) *Adv. Bot. Res.* **9**, 1.
6. Mossé, J. and Pernollet, J. C. (1982) *Chemistry and Biochemistry of Legume Seeds* (Arora, S. K., ed.) p. 111. Edward Arnold, London.
7. Holt, N. W. and Sosulski, F. W. (1981) *Can. J. Plant Sci.* **61**, 515.
8. Jermyn, W. A., MacKenzie, S. L. and Slinkard, A. E. (1979) *Can. J. Plant Sci.* **59**, 231.
9. Evans, I. M. and Boulter, D. (1980) *J. Sci. Food Agric.* **31**, 238.

10. Casey, R., Sharman, J. E., Wright, D. J., Bacon, J. R., Guldager, P. (1982) *Qual. Plant. Plant Foods Hum. Nutr.* **31**, 333.
11. Bhatthy, R. S. (1982) *J. Agric. Food Chem.* **30**, 620.
12. Lycett, G. W., Croy, R. R. D., Shirsat, A. H. and Boulter, D. (1984) *Nucleic Acids Res.* **12**, 4493.
13. Lycett, G. W., Delauney, A. J., Gatehouse, J. A., Gilroy, J., Croy, R. R. D. and Boulter, D. (1983) *Nucleic Acids Res.* **11**, 2367.
14. Gueguen, J., Vu, A. T. and Schaeffer, F. (1984) *J. Sci. Food Agric.* **35**, 1024.
15. Mossé, J., Huet, J. C. and Baudet, J. (1986) *J. Plant Physiol.* **125**, 199.
16. Randall, P. J., Thomson, J. A. and Schroeder, H. E. (1979) *Aust. J. Plant Physiol.* **6**, 11.
17. Chandler, P. M., Spencer, D., Randall, P. J. and Higgins, T. J. V. (1984) *Plant Physiol.* **75**, 651.
18. Evans, I. M., Gatehouse, J. A. and Boulter, D. (1985) *Biochem. J.* **232**, 261.
19. Rao, R. and Pernollet, J. C. (1981) *Agronomie* **1**, 909.
20. Dure, L. S., III, (1975) *Annu. Rev. Plant Physiol.* **26**, 259.
21. Evans, I. M., Gatehouse, J. A., Croy, R. R. D. and Boulter, D. (1984) *Planta* **160**, 559.
22. Pernollet, J. C. and Mossé, J. (1983) *Seed Proteins* (Daussant, J., Mossé, J. and Vaughan, J., eds) p. 155. Academic Press, London.
23. Mossé, J., Huet, J. C. and Baudet, J. (1985) *J. Cereal Sci.* **3**, 115.