Genes affecting starch biosynthesis exert pleiotropic effects on the protein content and composition of pea seeds

Richard K Hughes,[1]* Neil Desforges,[2] Christopher Selwood,[2] Roger Smith,[2] Charles I Speirs,[2] Georges Sinnaeve,[3] Peter G Gorton,[4] Julian Wiseman,[4] Kornelia Jumel,[5] Stephen E Harding,[5] Sandra E Hill,[6] Val Street,[6] Trevor L Wang[1] and Cliff L Hedley[1]

[1]*John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK*

[2]*Central Nutritional Laboratory, Pedigree Masterfoods, Melton Mowbray LE13 1BB, UK*

[3]*Ministère des Classes Moyennes et de l'Agriculture, Département Qualité des Productions Agricoles, Chaussée de Namur 24, B-5030 Gembloux, Belgium*

[4]*Division of Agriculture and Horticulture, University of Nottingham, Sutton Bonington, Loughborough LE12 5RD, UK*

[5]*National Centre for Macromolecular Hydrodynamics, University of Nottingham, Sutton Bonington, Loughborough LE12 5RD, UK*

[6]*Division of Food Sciences, School of Biological Sciences, University of Nottingham, Sutton Bonington, Loughborough LE12 5RD, UK*

Abstract: We have used a series of near–isogenic mutant pea lines defective in starch biosynthesis to identify genes that exert pleiotropic effects on the protein content and composition of the pea seed. The total protein contents of the flours obtained from the dry seed of these lines, determined using a new method based on amino acid analysis and validated by near-infrared reflectance spectroscopy, ranged from 19 to 28% of dry matter. Modifications in the legumin (11S)/vicilin (7S) ratios were confirmed by analytical ultracentrifugation and accompanied by compensatory increases in the albumin (2S–4S) content. Improving the albumin fraction of plant protein can be used to improve its nutritional value. There is additional variability in the solubility profiles for the protein in these lines. This variability has implications for the functional properties, processing behaviour, potential industrial applications, digestibility and nutritional value of the different pea meal and the protein fractions derived therefrom. Digestibility of the protein in meal and in air-classified and extruded products from selected lines will be compared in nutritional evaluation trials with poultry to assess the suitability of these products for incorporation into animal diets.
© **2001 Society of Chemical Industry**

Keywords: pea; starch; protein; near-isogenic; mutants; extrusion processing

INTRODUCTION

Pea seeds, like those of other legumes, contain a relatively high proportion of total protein, but there are numerous methods to determine total protein. Traditionally, this has been determined by applying a conversion factor to the total nitrogen level of the seed or meal, and a factor of 6.25 has been applied. This is inappropriate for many vegetable proteins and does not account for the proportion of non-protein nitrogen.[1]

Quantitation of total protein requires complete extraction of all protein. For this reason, determination of protein in hydrolysates is preferable to extraction with a particular buffer, which will selectively extract proteins on the basis of their solubility. Since all protein is composed of amino acids, determining amino acid composition to calculate total protein content is a logical corollary.[2] Quantitation of individual protein amino acids should provide a direct estimate of total protein content. Moreover, since each amino acid has known nitrogen content, this should also enable calculation of the nitrogen-to-protein conversion factor. However, unless an amino acid analysis is carried out rigorously, the estimates of protein content are invalid and tend to underestimate the true protein content.[3] Some amino acids are very unstable (eg tryptophan), are partially destroyed (eg serine and threonine) or are incompletely released (eg isoleucine and valine) following acid hydrolysis of the protein. Ultrafiltration can be used to determine the proportions of low-molecular-weight non-protein nitrogen compounds and free amino acids. For industrial purposes there is also a need to validate a more rapid, physical technique such as near-infrared reflectance spectroscopy (NIRS) with chemical data for estimating total protein contents.

Identifying the genetic and physiological factors responsible for modifying the protein content and

composition (including amino acid balance) and hence nutritional value of crops is of great interest. In economic terms the potential of peas has not been fully exploited or realised; nevertheless, as a model system they are an invaluable source, as their genetics is well characterised. Most of the carbon entering the pea seed during development is partitioned into either protein or starch. There are therefore two methods for modifying the protein content of the dry seed: firstly, to manipulate those genes which directly affect the synthesis of the proteins themselves; secondly, to shift partitioning of carbon towards protein by reducing the synthesis of starch.

A number of near-isogenic mutant pea lines that are affected in starch biosynthesis have become available.[4,5] They have mutations at single genetic loci and include *r, rb, rrb* (double mutant), *rug3* (sim1, low starch and sim32, starchless), *rug4, rug5* and *lam*. It has become important to know the true protein content of wild-type peas in order to establish if the mutant lines have modified protein contents. Preliminary data have indicated that the protein content for some of these lines was different from wild-type, but this was determined as total nitrogen multiplied by 6.25.[6] Since their protein and nitrogen compositions might be variable, their total protein contents and nitrogen-to-protein conversion factors are best calculated individually.

Mutations at these genetic loci have already been shown to affect protein composition by modifying the amount of legumin.[5] For mutation at the *r* locus, this has been shown to possibly result through a reduction in the stability of legumin mRNA following an increase in the soluble carbohydrate content and, hence, osmotic pressure of the cellular environment.[7] Reductions in legumin have been associated with increases in the total albumin fraction;[6,8] however, the methods used to determine the relative proportion of this fraction are questionable.[9] Improvements in the albumin fraction are an important determinant of nutritional value.[10] They are richer in sulphur amino acids and contain protease inhibitors.

Peas are a quality raw material for animal feeding (pigs, poultry and ruminants), increasingly so as reliance on animal protein feedstuffs is being increasingly questioned, and the protein digestibility of 'feed' peas (round, tannin-less peas) has been shown to be comparable to that of wheat and soybean. A complete comparison of the protein digestibility and nutritional value of the meal from the mutant lines described above, and the processed products derived therefrom has not been determined; the expected variability in their protein content and composition suggested that they should be investigated as a potential source of improved raw materials.

The present work will serve four purposes: (1) validate a chemical and physical method to determine the true protein content of pea seeds; (2) identify the nature of the modifications, if any, in protein content and composition and the mutations responsible/key enzymes (and their structural genes and/or promoters) involved in the regulation of carbon partitioning from starch to protein synthesis; (3) determine the effects of the mutations on seed nitrogen distribution; (4) determine the effects of the mutations and processing (air classification and extrusion) on the digestibility of the protein from selected mutants *in vivo* and their suitability for use in animal diets.

EXPERIMENTAL
Seed material
The pea seed material used within this study is a selection from a wider group of mutants developed at the John Innes Centre, described earlier[4]. The correct allelic description of the lines is as follows (in parenthesis) –*RRRbRb* (wild type line); *rrRbRb* (*r* allele at the *r* locus); *RRrbrb* (*rb* allele at the *rb* locus); *rrrbrb* (*rrb*, double mutant combining both the *r* and *rb* alleles); *rug3* sim1 (*rug3-a* allele at the *rug3* locus); *rug3* sim32 (*rug3-b* allele at the *rug3* locus); *rug4* (*rug4-b* allele at the *rug 4* locus); *rug5* (*rug5-a* allele at the *rug 5* locus) and *lam* (*lam-c* allele at the *lam* locus).

Production of pea flour and meal
Pea flour or meal was obtained by grinding dry seeds (10–50 g fresh weight) for 3 min in a Shatterbox (Spex, USA) or a coffee grinder (Moulinex, UK) respectively. Flour was used for NIRS and for analysis of protein amino acids to calculate total protein contents. Meal was used for the production of protein concentrates, the filtrates for free amino acid analysis and for the determination of protein solubility profiles (Fig 1).

Amino acid profiling
Amino acid profiles were determined after acid hydrolysis by HPLC with pre-column derivatisation and fluorescence detection. Sample containing approximately 100 mg protein was acidified with 50 ± 2 ml 6 M HCl, the internal standard nor-leucine (2.5 ml, 10 mM) was added, refluxed for 23 h (just boiling) and adjusted to pH 3.0 ± 0.1. After filtration (Whatman No 4), phenol (5 ml, 5% w/v) was added to a final volume of 250 ml. An aliquot (10 μl) was derivatised with 6-aminoquinolyl-*N*-hydroxysuccinimidyl carbamate (ACCQ, Fluor). Tryptophan was determined by alkaline hydrolysis in the presence of barium hydroxide. For determination of cysteine (as cysteic acid) and methionine, samples were oxidised with performic acid at 2–8 °C for 16 h prior to acid hydrolysis. Amino acid recovery in the hydrolysates after HPLC analysis was determined from the internal standard; the proportion of amino acids recovered from the protein during digestion was assumed to be 95%. Free amino acids were determined in the filtrate obtained at the ultrafiltration step used for producing protein concentrates. Total protein contents were calculated from the sum of the amounts of each amino acid (expressed as a percentage of dry matter), less free amino acids.
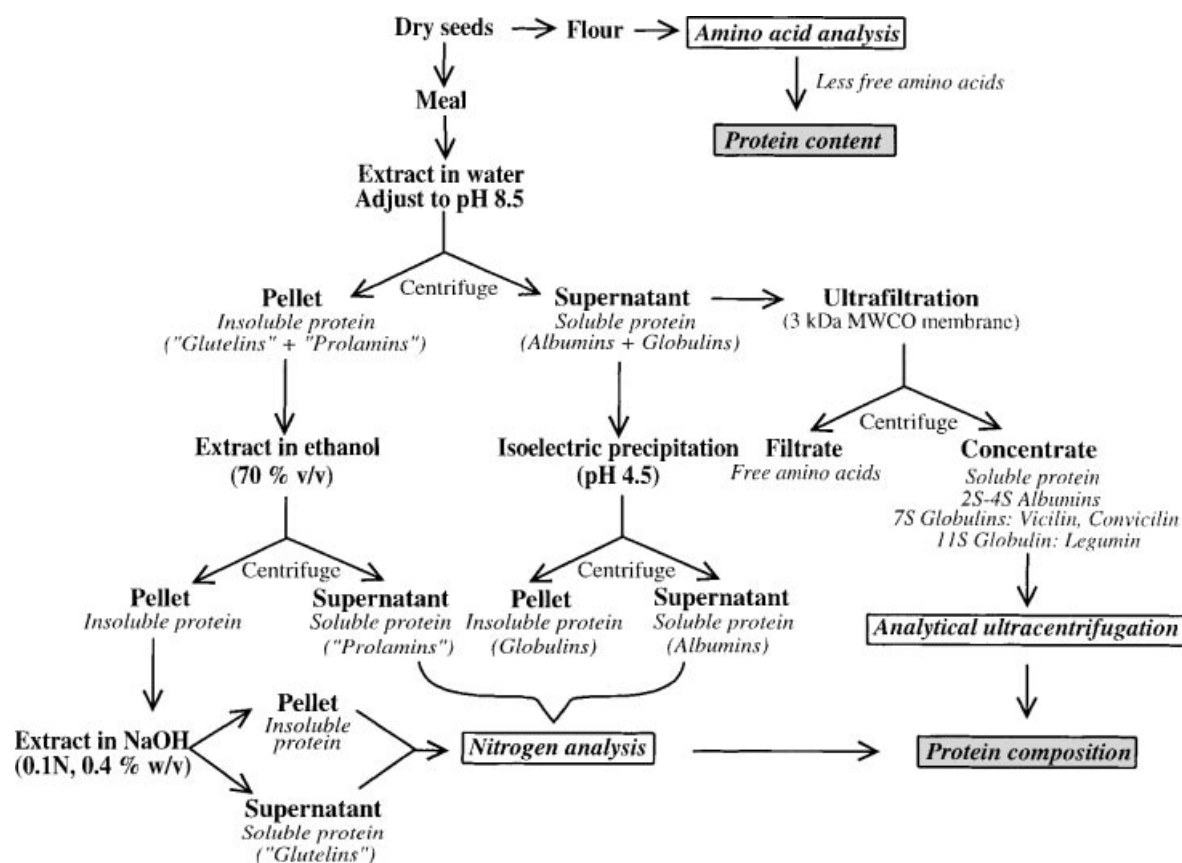
Figure 1. Extraction procedure for analysis of protein content and composition of dry wild-type and mutant pea seeds.

Production of protein concentrates for analytical ultracentrifugation

Pea meal (10 g) was dispersed by slow addition to water (100 ml), while stirring to avoid clumping, for 30 min at room temperature; pH 8.5 was maintained by the addition of 50 mM NaOH. The mixture (total volume 150 ml) was then centrifuged at $40\,000 \times g$ for 15 min at 20 °C. An aliquot (15 ml) of the supernatant was then centrifuged at $5000 \times g$ in a Centriplus-3 (Amicon, Millipore, UK) and washed on the filter with 32.5 mM K_2HPO_4/2.6 mM KH_2PO_4 buffer, pH 7.6, $\mu = 0.1$ (buffer B, total volume 30 ml). The volume was maintained between 10 and 15 ml. Concentrates were frozen in liquid nitrogen, lyophilised and stored at −20 °C. The products were redissolved in buffer B containing 0.4 M NaCl, $\mu = 0.5$ (buffer C) at a concentration of 16 mg ml^{-1}, centrifuged at $12\,000 \times g$ for 10 min at room temperature. Supernatants were used for analytical ultracentrifugation measurements.

Analytical ultracentrifugation

Sedimentation velocity experiments were performed on a Beckman Model E analytical ultracentrifuge equipped with a Schlieren optical system and coupled on-line to a CCD camera. 700 μl of each supernatant were filled into 20 mm pathlength ultracentrifuge cells prior to loading into a four-hole rotor. All measurements were made at 25 °C. Sedimentation was measured at a rotor speed of 48 000 rpm to trace the movement of the sedimenting boundary.

Determination of protein solubility profiles

Protein fractionation of pea meal was carried out according to a modification of the method of Saharan and Khetarpaul.[11] Pea meal was dispersed in water at pH 8.5 as for the production of protein concentrates, and the mixture was centrifuged at $40\,000 \times g$ for 15 min at 20 °C. The supernatant was adjusted slowly to pH 4.5 by the addition of 0.1 M HCl, stirred slowly for 15 min and centrifuged at $5000 \times g$ at 20 °C; the supernatant and pellet were described as the albumin and globulin fractions respectively. The pellet obtained after water extraction at pH 8.5 was extracted in ethanol (70% v/v) and centrifuged at $40\,000 \times g$ for 15 min at 20 °C. The supernatant was described as the 'prolamin' fraction. The pellet was extracted with 0.1 M NaOH and centrifuged at $40\,000 \times g$ for 15 min at 20 °C; the supernatant was described as the 'glutelin' fraction and the pellet as the insoluble fraction. The purified protein fractions were subjected to total nitrogen analysis; the nitrogen content of the fractions was assumed to be directly proportional to their protein content.

Near-infrared reflectance spectroscopy (NIRS)

NIRS measurements were carried out on a Foss-NIRSystems 5000 scanning monochromater (1100–

2500 nm). The spectra acquisition and data treatment were achieved using the Foss-Infrasoft International package (ISIV 4.0). The calibrations were obtained by partial least squares fittings. The calibrations were based on the Kjeldahl method for total nitrogen analysis using 6.25 as a conversion factor. Best results were achieved with the first derivative of the spectra using the segment and gap method.

RESULTS AND DISCUSSION

The total protein contents of the meal from all mutant lines were highly different from each other and from wild-type (Fig 2). Using the new method based on amino acid analysis, *rug4* and *lam* had lower protein contents than the other mutant lines and wild-type. The starchless mutant *rug3* sim32 had a higher protein content than the rest of the group. The general distribution of protein contents among the lines was similar when protein contents were determined using NIRS, a conversion factor of 6.25, and using the new method based on amino acid analysis. Using 6.25 as a conversion factor and NIRS tended to overestimate the proportions of total protein, but, since the NIRS calibrations were based on nitrogen analysis, this was not surprising. NIRS tended to underestimate the protein contents of the mutants *r* and *rrb*; for mutant *r*, this was also true when a conversion factor of 6.25 was used.

Analysis of the composition of the protein concentrates from wild-type and mutant lines during analytical ultracentrifugation generally confirmed previous observations using crossed immunoelectrophoresis of large differences in the $7S/11S$ ratios (Fig 3);[5] differences between the two studies in the $7S/11S$ ratio for the *lam* mutant were probably due to differences in the number of backcrosses for this material. Previously, $7S/11S$ ratios were determined

using meal extracted with 0.5 M NaCl;[5] for ultracentrifugation analysis, meal was extracted without salt, which suggested that there were no differences in the selective extraction of $7S$ or $11S$ between the two methods.

There was generally a negative correlation between the relative amount of $11S$ (legumin) and the amount of total protein (compare Figs 2 and 3). Mutants with legumin contents that appeared to be lower than wild-type (*r, rb, rrb, rug5, rug3* sim1 and *rug3* sim32) tended to have total protein contents higher than wild-type. Those mutants with similar profiles to wild-type (*rug4* and *lam*), and possibly higher legumin content in the case of *lam*, had lower total protein contents.

The relative proportions of the $2S$–$4S$ (albumin) fraction in wild-type and mutant lines were also highly different. Those mutants with reduced levels of legumin appeared to be associated with compensatory increases in the albumin fraction. The starchless mutant *rug3* sim32, for example, appeared to have virtually no legumin but the highest proportion of albumin. It was not judged suitable to estimate the proportions of the different fractions from the area under the curves by peak fitting, since this would have tended to overestimate the proportion of total protein due to albumin.

The solubility profiles for the total protein from wild-type and mutant lines were also highly different (Fig 4). Interestingly, all mutants had a higher albumin content than wild-type. However, since some mutants had a lower total protein content than wild-type, the relationship between albumin content and total protein (or legumin) content was less clear and this prompted an investigation of the effects of the mutations on other protein fractions.

The proportion of total globulin, as well as total 'insoluble' protein (ie that which could not be described as albumin or globulin), in wild-type and mutant lines was also different. With the exception of mutant *r*, there was generally a correlation between total protein (or legumin) and total 'insoluble' protein. Thus mutants *rug4* and *lam* which had lower total protein also had a lower proportion of total 'insoluble' protein. Conversely, mutants *rb, rrb, rug3* sim1 and *rug3* sim32 which had higher total protein (lower legumin) had a higher proportion of 'insoluble' protein. Mutant *r* (and to a smaller extent *rug5*) was unusual in that it had high total protein and low legumin contents, but a low level of 'insoluble' protein and a higher total globulin content than predicted; the globulin content of the other mutants was generally as predicted from the results obtained using the analytical ultracentrifuge.

No major differences in the composition of the albumin fractions from wild-type and mutant lines have been detected (Welham T *et al*, unpublished). This suggested that mutations at single genetic loci which directly affect starch biosynthesis would appear to exert pleiotropic, indirect, effects on the synthesis and/or accumulation of a broad range of albumin and
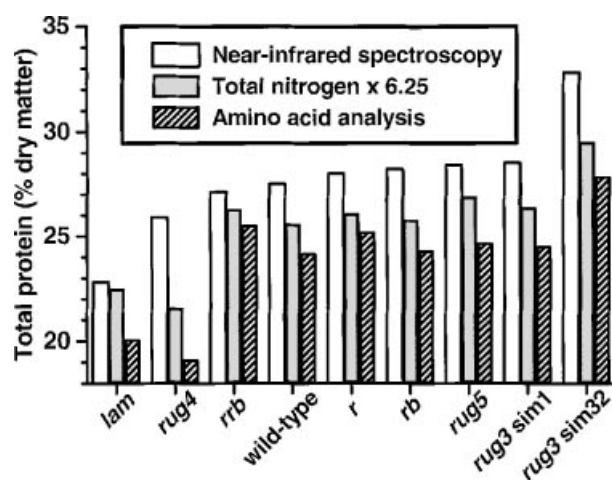

Figure 2. Comparison of methods for estimating total protein content of dry wild-type and mutant pea seeds.. The total protein contents of pea flours were determined by NIRS, total nitrogen scaled by the traditionally used conversion factor of 6.25, and using the new method based on amino acid analysis. The pea lines are ranked for protein content using the NIRS data.
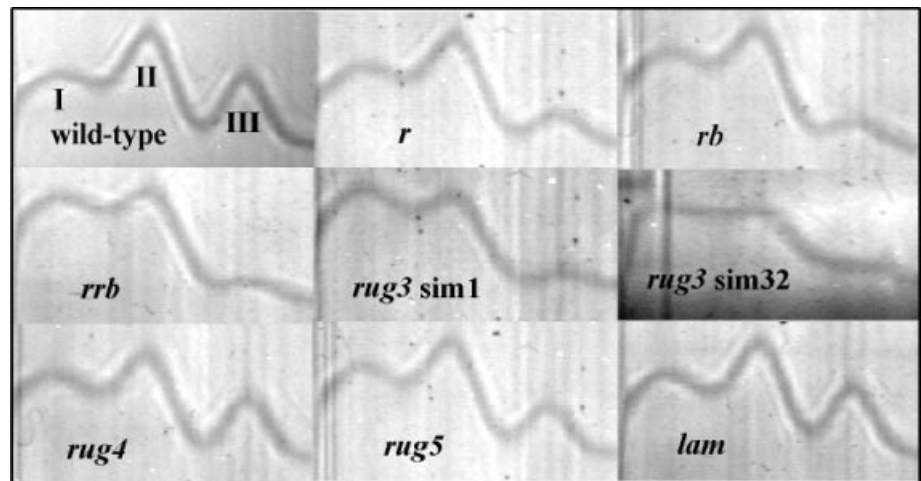
Figure 3. Protein composition of dry wild-type and mutant pea seeds determined by analytical ultracentrifugation. Schlieren patterns of protein concentrates are shown to compare the relative proportions of albumin (2*S*–4*S*, peak I), 7*S* globulin (vicilin + convicilin, peak II) and 11*S* globulin (legumin, peak III). Peak positions are shown for wild-type only.

other 'insoluble' proteins, possibly in response to, or accompanied by, changes in the synthesis and/or accumulation of legumin (11*S*) globulin.

Extrusion processing of meal from mutants *r* and *rug3* sim32 resulted in products that were visibly different from each other and from wild-type (Fig 5). The product obtained using the *rug3* sim32 mutant resembled more closely that which was typically obtained using soya bean meal. The origins of the differences in the properties of the extruded materials are being investigated by differential scanning calorimetry (DSC), but are most likely to have arisen from an interaction between the starch and lipid or protein,

or the lack of such an interaction in the case of the *rug3* sim32 mutant which has no starch. Poultry feeding trials with extruded and air-classified products and non-processed meal from wild-type and mutants *r* and *rug3* sim32 will allow us to determine the effects of processing and mutation at the *rug3* or *r* genetic locus on the digestibility of pea protein.

We are currently using the pea mutants to explore the relationships between protein and lipid, starch or soluble carbohydrate content, which are also affected by the mutations, in order to understand better the basis for the effect of starch synthesis on protein content and composition. Nevertheless, it is clear that the genetic manipulation of starch synthesis by conventional techniques can be used to modify the amount and quality of the protein in the pea seed.
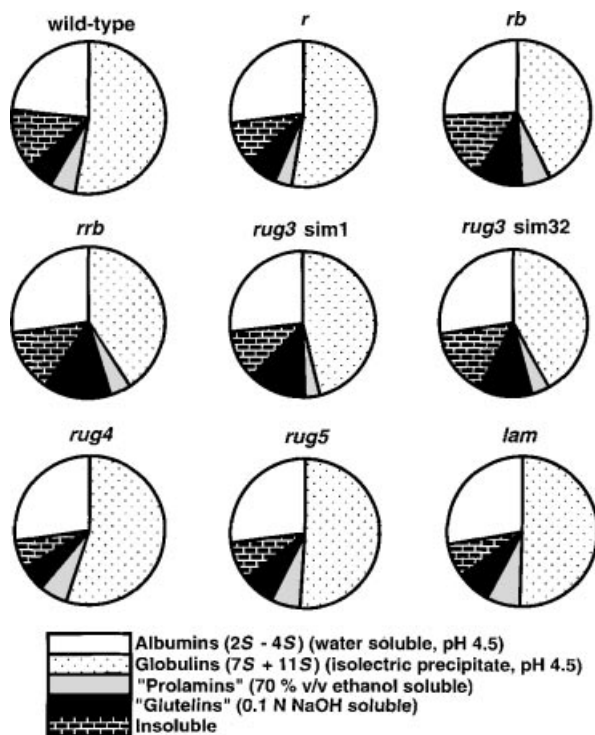

Figure 4. Protein composition of dry wild-type and mutant pea seeds determined by nitrogen analysis of purified protein fractions. The relative proportions of the different protein fractions isolated from wild-type and mutant pea meal are shown.

Figure 5. Extrusion products from wild-type and mutant pea flours. Pea flours of (clockwise from top) wild-type and mutants *r* and *rug3* sim32 were extruded using a Clextral BC21 twin-screw extruder (screw speed 200 rpm). Moisture contents in the extruder (assuming 10% moisture in the original flours) were 18.8, 19.6 and 24.1% for wild-type, *r* and *rug3* sim32 respectively. The maximum temperatures recorded at the die (3 mm) were 86, 87 and 86 °C respectively. A cutter (4500 rpm) was fitted to generate small, intact pieces suitable for feeding to poultry.

REFERENCES
1 Boisen S, Bech-Andersen S and Eggum BO, A critical view on the conversion factor 6.25 from total nitrogen to protein. *Acta Agric Scand* **37**:299–304 (1987).
2 Mossé J, Nitrogen to protein conversion factor for ten cereals and six legumes or oilseeds. A reappraisal of its definition and determination. Variation according to species and to seed protein content. *J Agric Food Chem* **38**:18–24 (1990).
3 Mossé J and Baudet J, Crude protein content and amino acid composition of seeds: variability and correlations. *Qual Plant Plant Foods Human Nutr* **32**:225–245 (1983).
4 Wang TL, Bogracheva TY and Hedley CL, Starch: as simple as A, B, C? *J Exp Bot* **49**:481–502 (1998).
5 Casey R, Domoney C, Forster C, Hedley C, Hitchin E and Wang T, The effect of modifying carbohydrate metabolism on seed protein gene expression in peas. *J Plant Physiol* **152**:636–640 (1998).
6 Perez MD, Chambers SJ, Bacon JR, Lambert N, Hedley CL and Wang TL, Seed protein content and composition of near-isogenic and induced mutant pea lines. *Seed Sci Res* **3**:187–194 (1993).
7 Turner SR, Barratt DHP and Casey R, The effect of different alleles at the *r* locus on the synthesis of seed storage proteins in *Pisum sativum*. *Plant Mol Biol* **14**:793–803 (1990).
8 Schroeder HE, Quantitative studies on the cotyledonary proteins of the genus *Pisum*. *J Sci Food Agric* **33**:623–633 (1982).
9 Chambers SJ, Bacon JR and Lambert N, The quantitative analysis of seed proteins from peas using high performance liquid chromatography. *Phytochem Anal* **3**:49–54 (1992).
10 Tabe L and Higgins TJV, Engineering plant protein composition for improved nutrition. *Trends Plant Sci* **3**:282–286 (1998).
11 Saharan K and Khetarpaul N, Protein quality traits of vegetable and field peas: varietal differences. *Plant Foods Human Nutr* **45**:11–22 (1994).