The purification and partial amino acid sequence of a polypeptide from the glutelin fraction of rice grains; homology to pea legumin

Wen-Ming Zhao*, John A. Gatehouse+ and Donald Boulter

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

Received 4 August 1983

The glutelin fraction was extracted from grain meals of rice (*Oryzea sativa*) with 50 mM Tris–HCl buffer (pH 8.8) containing 6 M urea and 10 mM 2-mercaptoethanol. Polypeptides of glutelin were separated and purified by ion-exchange chromatography under denaturing conditions. Analysis by two-dimensional gel electrophoresis showed that 2 major polypeptides of the rice glutelin fraction, M_r 36000 and 22000, were linked in disulphide bonded pairs containing one M_r 36000 and one M_r 22000 subunit. A partial amino acid sequence of the purified M_r 22000 glutelin subunit showed it to be homologous to the β-subunit of pea legumin, a storage protein which also contains disulphide-linked subunit pairs (M_r 38000 and M_r 22000). It is therefore proposed that the major component of rice glutelin is a legumin-like protein.

*Rice (*Oryzea sativa*)* *Storage protein* *Amino acid sequence* *Legumin* *Glutelin*

1. INTRODUCTION

The major storage proteins of most cereal grains are of the glutelin or prolamin types [1]. In rice, unlike most other cereals, glutelin is the only major fraction in the endosperm, comprising $\geqslant 80\%$ of the total protein with prolamin accounting for $<5\%$. In [2] a rice glutelin preparation was reduced and alkylated and shown by sodium dodecylsulphate–polyacrylamide gel electrophoresis to consist of 3 major subunits (M_r 38000, 25000 and 16000) present in the ratio of 16:3:1. However, since the rice glutelin fraction conforms to the general properties of cereal glutelin, in being insoluble in non-denaturing media it is difficult to characterize [1]; that it is not known how many different proteins make up this fraction, since it is only possible to purify polypeptides from it rather than protein molecules. In order to try to elucidate the nature of the rice glutelin fraction, we have purified several polypeptides from it and present partial amino acid sequence data for one of them.

2. MATERIALS AND METHODS

Rice seeds from the 1981 harvest of round-shaped rice (Guang-ji 9) were obtained from the Crop Breeding Research and Teaching Group of the Agricultural Department (Northwestern College of Agriculture, People's Republic of China). Standard proteins and cyanogen bromide were obtained from Sigma Chemical Co. (Poole, Dorset). Other chemicals were from British Drug House Ltd. (Poole, Dorset) and were of analytical grade wherever possible.

2.1. *Extraction of the glutelin fraction*

Air-dried mature rice seeds were used and grains were dehulled in a mortor with a pestle and milled

* Permanent address: Department of Basic Course, Northwestern College of Agriculture, Yangling Town, Shaanxi Province, People's Republic of China

+ To whom correspondence should be addressed

Published by Elsevier Science Publishers B.V.

00145793/83/$ 3.00 © 1983 Federation of European Biochemical Societies

for 30 s in a ball mill. The flour was seived through a 365 μm mesh and material not passing through was discarded. The meal was extracted twice with hexane at 4°C for 30 min (10 ml hexane/g meal) and air-dried. A portion (25 g) of defatted rice meal was extracted with 250 ml 10 mM Tris–HCl buffer (pH 7.8), containing 0.5 M NaCl and 1 mM EDTA at 4°C for 20 h [2,10,11] centrifuged at 23000 × g for 20 min and the supernatant (containing albumin and globulin proteins) discarded. The residues were extracted again with 300 ml 50 mM Tris–HCl buffer (pH 8.8) containing 6 M urea and 10 mM β-mercaptoethanol at 4°C for 27 h, centrifuged and the precipitate discarded.

2.2. *Purification of glutelin fraction polypeptides*

If required the supernatant was dialysed against distilled water and lyophilised to give glutelin fraction polypeptides; however, this material was difficult to re-solubilise. The supernatant from glutelin extraction was loaded and washed on to a column (1.5 cm diam., 100 ml vol.) of DEAE-cellulose, equilibrated with 50 mM Tris–HCl buffer (pH 8.8) containing 6 M urea and 10 mM β-mercaptoethanol, at a flow rate of 50 ml/h; 10 ml fractions were collected. After all non-bound material had washed off, the column was eluted with a gradient. Fractions containing protein of NaCl (0–0.5 M) in equilibration buffer were pooled by peaks, dialysed against distilled water, and then freeze-dried. They were assayed for polypeptide content by SDS–polyacrylamide gel electrophoresis.

Materials unretarded by the DEAE-cellulose column (700 mg) was dissolved in 50 ml 8 M urea/50 mM β-mercaptoethanol/10 mM sodium orthophosphate buffer (pH 6.5). The solution was flushed with N_2, incubated at 20°C for 3 h, and then dialysed 3-times against 250 ml of the same buffer, containing 25 mM β-mercaptoethanol. The reduced, denatured glutelin was loaded onto a column (30 × 1.8 cm) of CM-cellulose equilibrated with the same buffer as used for dialysis. The column was eluted with the starting buffer until all non-bound material had washed off, and was subsequently eluted with a gradient of NaCl (0–0.6 M) at 30 ml/h. Fractions (5 ml) were collected; fractions containing protein peaks pooled, dialysed against water, lyophilised and analysed by SDS–PAGE.

2.3. *Gel electrophoresis*

SDS–polyacrylamide gel electrophoresis, and two-dimensional SDS–polyacrylamide gel electrophoresis were carried out in 12% and 17% acrylamide gel slabs using the buffer system in [12] and the methods in [3]. Molecular mass calibration of gels was performed using standard proteins (M_r): cytochrome c (12500); β-lactoglobulin (17500); chymotrypsinogen (25000); ovalbumin (43000); bovine serum albumin (67000); and phosphorylase (97000).

2.4. *Cyanogen bromide cleavage and tryptic digestion*

CNBr cleavage was carried out on protein dissolved in 70% formic acid using a 100 × molar excess of CNBr for 24 h at room temperature [13]. Tryptic digestion of reduced and carboxymethylated protein [14] was carried out at an enzyme: protein ratio of 1:50 (w/w) for 18 h at 20°C in 0.1 M HN_4HCO_3 (pH 8.0).

2.5. *Sequencing*

Peptides were separated by gel filtration on a column of Bio-Gel P-4; peaks from this column were pooled, lyophilised and subjected to reverse-phase HPLC as in [17]. Peptides were sequenced by the manual DABITC method [15]; sequences were checked by semi-quantitative amino acid composition analysis of hydrolysed peptides by dansylation and separation by thin-layer chromatography [16]. N-terminal sequences of intact polypeptides were obtained by using the manual DABITC method.

3. RESULTS

3.1. *Extraction of glutelin fraction*

The glutelin fraction was extracted from rice meal according to the scheme shown in the flow diagram (fig.1). After extraction of albumin and globulin fractions in aqueous salt solution, the glutelin fraction was extracted with pH 8.8 buffer (50 mM Tris–HCl) containing 6 M urea and 10 mM mercaptoethanol. When analysed by SDS–polyacrylamide gel electrophoresis, under reducing conditions the glutelin fraction consisted of 3 major bands at estimated M_r of 36000, 22000 and 14000, respectively (fig.2). These polypeptides are equivalent to those identified as major com-

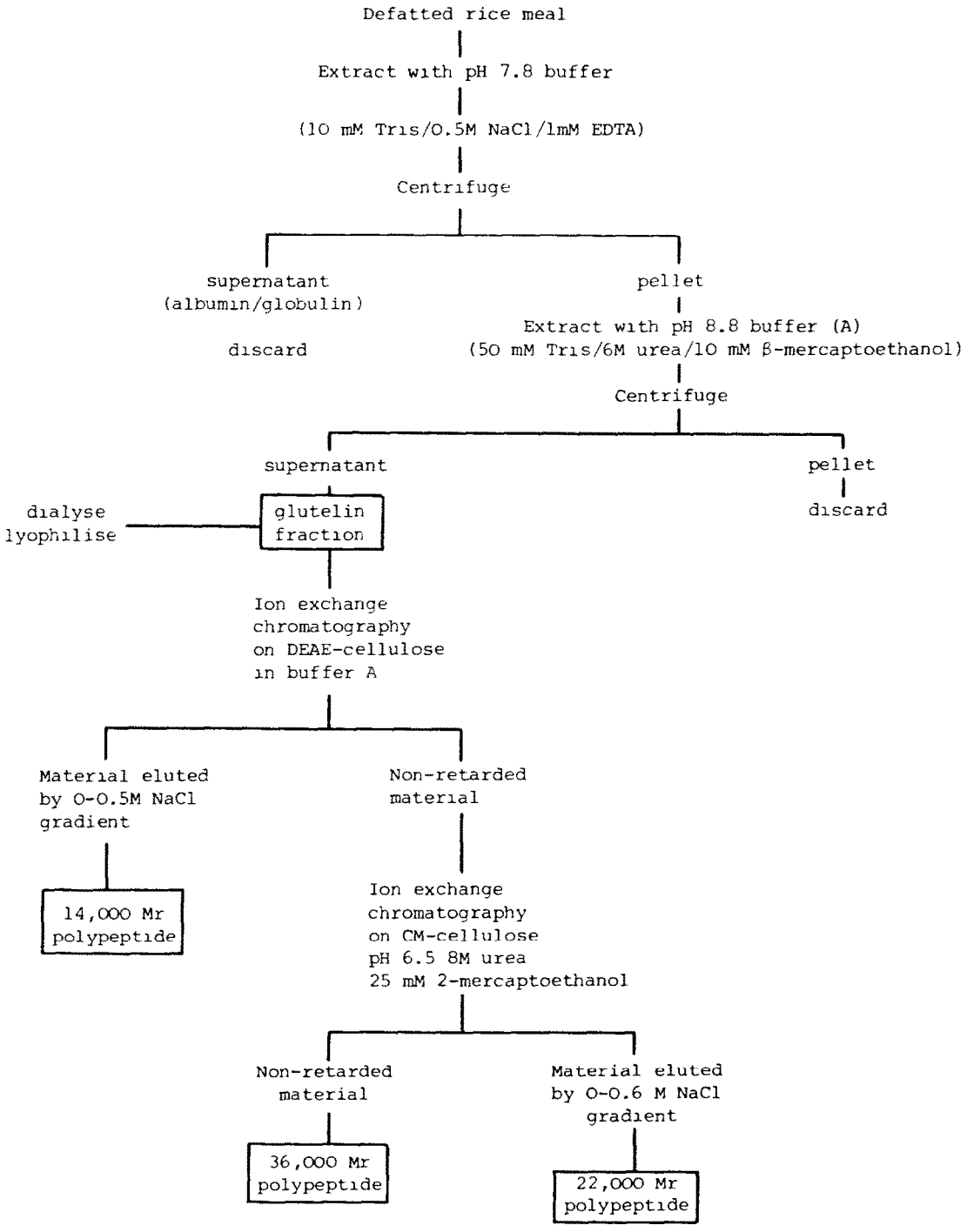

Fig.1. Flow diagram for purification of rice glutelin polypeptides.

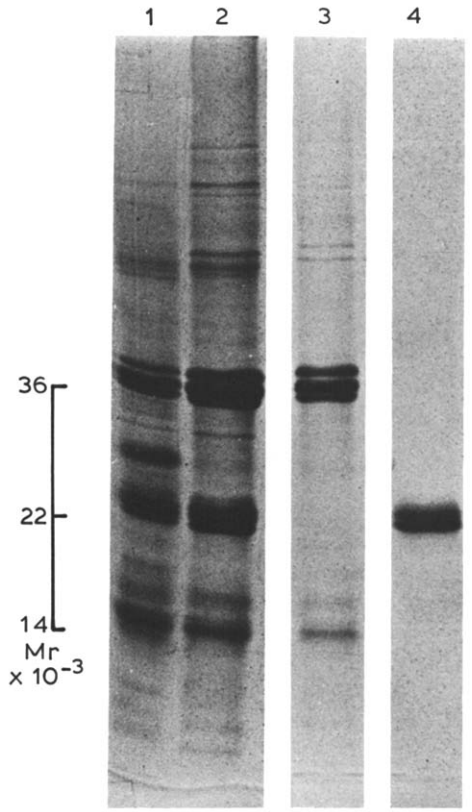

Fig.2. SDS–polyacrylamide gel electrophoresis of rice proteins and purified polypeptide (17% acrylamide gel slab, reducing conditions): (1) total protein extracted from rice grains by SDS-sample buffer; (2) glutelin fraction; (3) M_r 36000 polypeptide from glutelin fraction; (4) M_r 22000 polypeptide from glutelin fraction.

ponents of rice glutelin in [2]. The glutelin fraction extracted by 5% NaOH and subsequently analysed on gels gave similar polypeptide pattern except that in this case the M_r 14000 polypeptide formed a higher proportion of the protein in the glutelin fraction as assessed by Coomassie blue staining (not shown).

3.2. *Two-dimensional gel electrophoresis of glutelin fraction*

Rice glutelin fraction and rice total protein extracts were analysed by two-dimensional SDS–polyacrylamide gel electrophoresis employing non-reducing conditions in the first dimension and reducing conditions in the second dimension (fig.3). The results clearly show that the major polypeptides of M_r 36000 and 22000 in the glutelin fraction are associated in disulphide-bonded subunit pairs, each subunit pair (M_r ~52000) containing one larger and one smaller subunit). Other protein subunits do not cleave into two components on reduction, although one subunit present in the albumin/globulin fraction increases its app. M_r on reduction (M_r ~26000 when reduced) indicating a high level of intrachain disulphide bonding [3].

3.3. *Purification of glutelin fraction polypeptides*

The glutelin fraction was initially chromatographed on an ion-exchange column of DEAE-cellulose eluted with a linear concentration gradient of NaCl in glutelin extraction buffer. The non-bound material contained the M_r 36000 and M_r 22000 polypeptides, whereas the M_r 14000 polypeptides bound to the column and were eluted by the salt gradient (not shown).

The unbound fraction was further purified by ion-exchange chromatography on a column of CM-cellulose, equilibrated with 10 mM sodium phosphate buffer (pH 6.5) containing 8 M urea and 25 mM β-mercaptoethanol. Bound polypeptides were eluted by a linear concentration gradient of NaCl in the same buffer. The elution profile is shown in fig.4. The final peak (b4) of protein eluted at 0.22–0.28 M NaCl gradient phosphate buffer was shown to be pure M_r 22000 polypeptide by SDS–polyacrylamide gel electrophoresis (fig.2). The unbound material was predominantly M_r 36000 polypeptide.

3.4. *Properties of and partial amino acid sequence data for* M_r *22000 glutelin fraction polypeptide*

The M_r 22000 polypeptide had glycine as N-terminal amino acid. The subunit was cleaved by treatment with trypsin and the peptides produced were separated by gel filtration and reverse-phase HPLC, and sequenced. The N-terminal region of the intact polypeptide was also determined directly. The combined sequence data are presented in fig.5, matched to the known sequence of the pea legumin β-subunit [4,5] (W.-M.Z., unpublished). When peptides from the rice glutelin M_r 22000 polypeptides are fitted optimally to the legumin β-subunit sequence (involving one break of one

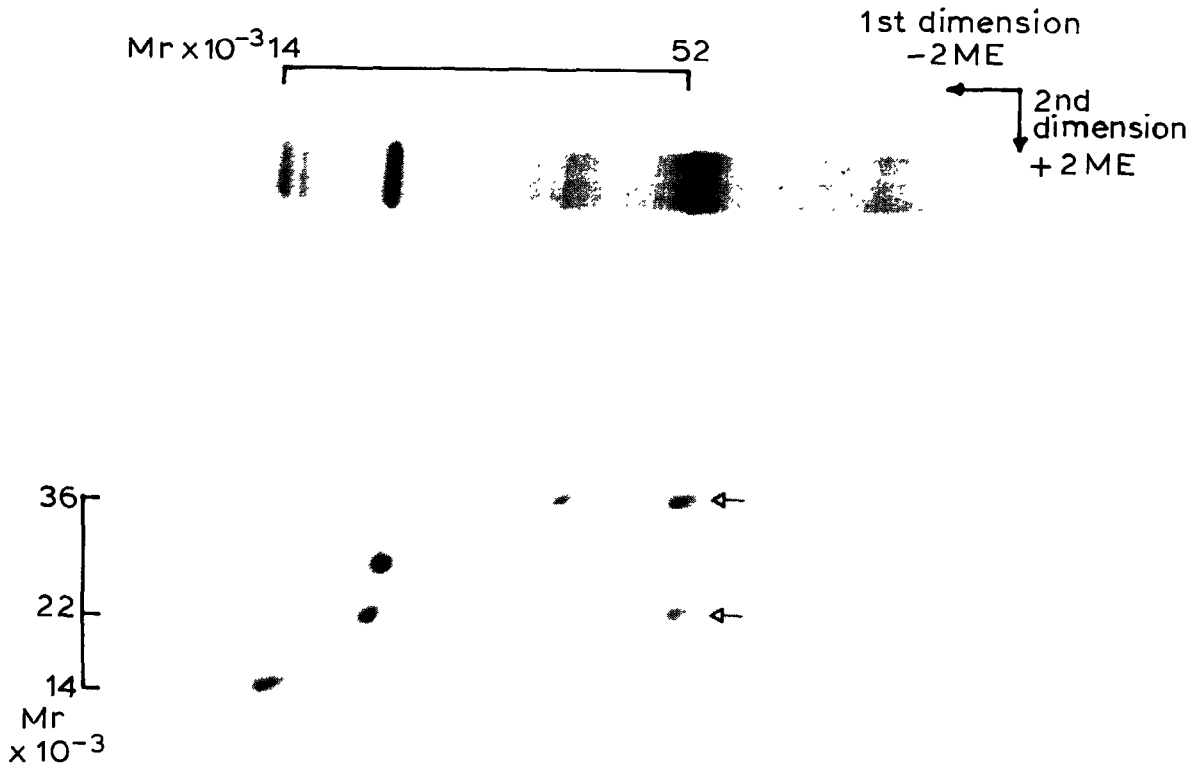

Fig.3. Two-dimensional SDS–polyacrylamide gel electrophoresis of total protein extracted from rice grains by SDS-sample buffer. First dimension (horizontal), non-reducing conditions, second dimension (vertical), reducing conditions. Arrows indicate glutelin subunits joined in disulphide-bonded pairs (dimers of the subunit pair separated in the first dimension are also shown in the second dimension separation).

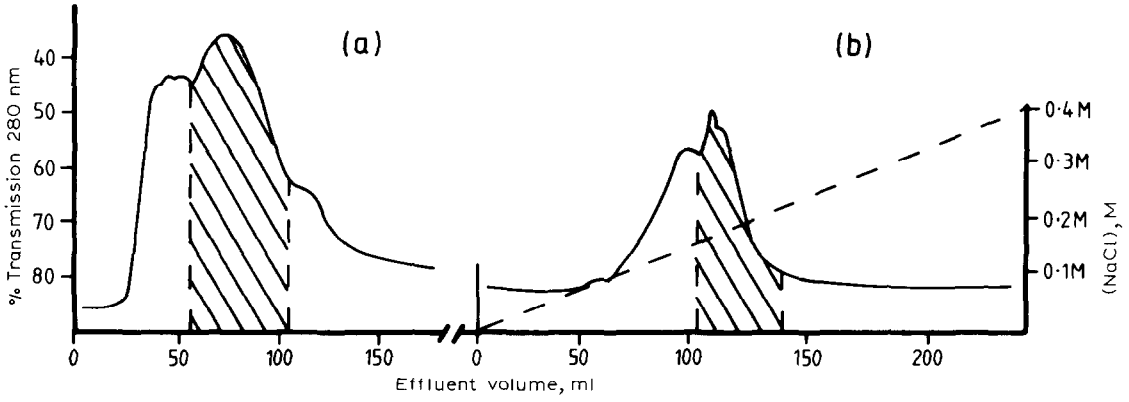

Fig.4. Purification of glutelin polypeptide by ion-exchange chromatography on CM-cellulose under denaturing conditions (10 mM sodium phosphate buffer (pH 6.5) containing 8 M urea, 25 mM 2-mercaptoethanol). Hatched areas show fractions pooled to give M_r 36000 polypeptide (a) and M_r 22000 polypeptide (b) as shown in fig.2.

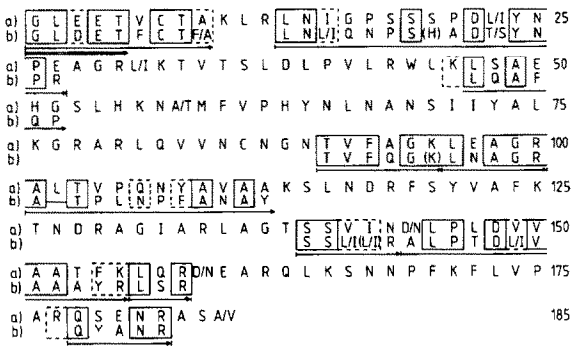

Fig.5. Sequence homology between pea legumin β-subunit (a) and rice glutelin fraction M_r 22000 polypeptide (b). Pea legumin β-subunit sequence from [4] with extra residue heterogeneity from [5] and unpublished data (W.-M.Z.). All the rice glutelin peptides that could be fitted to the pea legumin β-subunit sequence are shown; some sequence heterogeneity, and L/I uncertainty, is shown for rice glutelin peptides. Boxes indicate residue identity, dotted lines partial identity or residues of similar functional properties (inferred from start positions of glutelin tryptic peptides in some cases): horizontal bars, gaps introduced to avoid overlapping peptides as to improve homology; arrows, tryptic peptides ⟶ or N-terminal sequence ⟹; vertical bars, termini of peptides.

amino acid only in the rice glutelin sequences to improve homology) 41 residues match exactly and a further 12 residues match closely; i.e., they are amino caids of similar functional properties. A further 25 residues overlap but do not match. Other tryptic peptides of the rice glutelin M_r 22000 polypeptide could not be fitted to the legumin β-subunit sequence. The overall homology between the two sequences is therefore about 25% based on matching of glutelin peptides to the legumin β-subunit sequence. The homology is significant and is particularly marked at the N-terminal regions of the 2 sequences; 6 out of the first 9 amino acids are identical and a further 2 are similar. Evidence of sequence heterogeneity in the rice glutelin M_r 22000 polypeptide was shown by the isolation of peptides varying in sequence at single positions; this resembles the sequence heterogeneity found in the legumin β-subunit (not shown).

Although the rice glutelin M_r 22000 polypeptide partial sequence does not include any methionine residues, treatment of this polypeptide with cyanogen bromide gave multiple bands at M_r ~15000, indicating the presence of at least one methionine residue in the sequence.

4. DISCUSSION

The glutelin fraction from rice seeds has been shown to contain a polypeptide (M_r 20000) which has sequence homology to pea legumin β-subunits, and which is linked to another glutelin fraction polypeptide (M_r 36000) by one or more disulphide bonds. It is thus reasonable to suppose that the M_r 36000 glutelin fraction polypeptide is homologous to the legumin α-subunit (which is of similar M_r and exhibits a similar band pattern of several closely spaced bands on SDS–PAGE) and that these 2 polypeptides form subunit pairs in legumin-like molecules. In [5], a polypeptide of M_r 57000 was observed as a prominent component both in the polypeptides synthesised in developing rice seeds and in the in vitro translation products of mRNA from developing rice seeds. They have also observed post-translational cleavage of this precursor to glutelin fraction polypeptides of M_r 37000–39000 and 22000–23000 (corresponding to the M_r 33000 and 20000 glutelin polypeptides here); these data support the above suggestion, since pea legumin is also synthesised as an M_r 60000 precursor molecule containing both α- and β-subunits in a single polypeptide which is subsequently cleaved to produce the observed M_r 38000 α- and M_r 22000 β-subunits [7]. This is not the first cereal storage protein to be identified as homologous to legumin, since a globulin storage protein in oat has similar disulphide bonded subunit pairs and has been shown to be synthesised in a similar manner; in this case partial immunological identity with legumin was also demonstrated [8,9]. Although there was no reaction by 'Western' blotting, of the rice glutelin fraction M_r 36000 or M_r 22000 polypeptides with antibodies raised against pea legumin, the antibodies used were raised against non-denatured protein and reacted only weakly with pea legumin β-subunits under the same conditions (not shown).

The finding that the glutelin fraction of rice contains a legumin-like protein (referred to subsequently as 'rice legumin') is novel in that the difference in sequence between pea legumin and rice legumin affects the solubility properties of the protein, so that it is no longer soluble in aqueous non-

denaturing conditions. This is possibly due to the rice legumin aggregating in higher multimers of subunit pairs than the soluble hexameric pea legumin molecule, although this cannot at present be determined. The identification of rice legumin widens the distribution of this storage protein type to the glutelin fraction of proteins, and suggests that it may be present but unrecognised in other cereals also. However, a legumin-type storage protein could be a minor component in other cereals, since although rice legumin is the major protein in rice glutelin, other proteins, such as that represented by the rice M_r 14000 polypeptide in this fraction, may prove to be more related to the 'classical' glutelin fraction proteins of those other cereals.

ACKNOWLEDGEMENTS

We thank Mr John Gilroy for invaluable technical advice, assistance and for confirming much of the amino acid sequence data presented. We also thank Drs M. Richardson and M. Evans for advice and Mr D. Bown and Mr P. Preston for their technical assistance.

REFERENCES

[1] Shewry, P.R., Miflin, B.J., Forde, B.G. and Bright, S.W.J. (1981) Sci. Prog. Oxf. 67, 575–600.
[2] Juliano, B.D. and Boulter, D. (1976) Phytochemistry 15, 1601–1606.
[3] Matta, N.K., Gatehouse, J.A. and Boulter, D. (1981) J. Exp. Bot. 32, 182–197.
[4] Croy, R.R.D., Lycett, G.W., Gatehouse, J.A., Yarwood, J.N. and Boulter, D. (1982) Nature 295, 76–79.
[5] Casey, R., March, J.F. and Sanger, E. (1981) Phytochemistry 20, 161–163.
[6] Yamagata, H., Sugimoto, T., Tanaka, K. and Kasal, Z. (1982) Plant Physiol. 70, 1094–1100.
[7] Croy, R.R.D., Gatehouse, J.A., Evans, I.M. and Boulter, D. (1980) Planta 148, 49–56.
[8] Matlashewski, G.J., Adeli, K., Altosaar, I., Shewry, P.R. and Miflin, B.J. (1982) FEBS Lett. 145, 208–211.
[9] Brinegar, A.C. and Peterson, D.M. (1982) Plant Physiol. 69 suppl., 156.
[10] Pascual, C.G., Juliano, B.O. and Tanaka, Y. (1981) Phytochemistry 20, 2471–2475.
[11] Villareal, R.M. and Juliano, B.O. (1981) Phytochemistry 20, 1785–1789.
[12] Laemmli, U.K. (1970) Nature 237, 680–685.
[13] Kasper, C.B. (1970) in: Protein Sequence Determination (Needleham, S.B. ed) pp.137–140, Chapman and Hall, London.
[14] Glazer, A.N., Delange, R.J. and Sigman, D.S. eds (1975) Chemical Modification of Proteins; Selected Methods and Analytical Procedures, Elsevier, Amsterdam, New York.
[15] Chang, J.Y., Brauer, D. and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[16] Woods, K.R. and Wang, K.-T. (1967) Biochim. Biophys. Acta 133, 369–370.
[17] Hirano, H., Gatehouse, J.A. and Boulter, D. (1982) FEBS Lett. 145, 99–102.