Volume 222, number 1, 175–180 FEB 05154 September 1987

Construction of a hybrid cDNA encoding a major legumin precursor polypeptide and its expression and localization in *Saccharomyces cerevisiae*

Jennifer N. Yarwood, Nick Harris, Ashton Delauney, Ron R.D. Croy, John A. Gatehouse, Martin D. Watson and Donald Boulter

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

Received 10 August 1987

A hybrid cDNA encoding a major legumin precursor polypeptide was constructed. Using a phosphoglycerate kinase expression vector levels of expression of 1–2% total protein were obtained in *Saccharomyces cerevisiae*. Immunocytochemical localization demonstrated that the legumin was accumulated within the Golgi vesicles.

cDNA; Legumin; Expression vector; Golgi vesicle

1. INTRODUCTION

Legumin, a seed storage protein in pea (*Pisum sativum* L.), is synthesised as precursor polypeptides which undergo co- and post-translational proteolysis [1–3]. In the case of 'major' legumin species, derived from the *Leg*A gene subfamily, the precursor polypeptide is of 517 amino acids (M_r 58995, previously estimated as approx. 60000); co-translational proteolysis removes a signal peptide of 21 amino acids, and post-translational proteolysis ('processing') occurs on the C-terminal side of Asn 332 to generate the mature, disulphide-linked a- and b-polypeptides (M_r approx. 38000 and 21000). Processing occurs in the protein bodies in the seed, within 2 h of synthesis [2,4], the amount of precursor present in vivo being thus small.

Study of these proteolytic events, and other aspects of legumin biosynthesis and structure, such as molecular assembly, would be facilitated by the availability of precursor polypeptides. Here, the construction of a hybrid cDNA sequence encoding a *Leg*A-derived precursor polypeptide, and its expression in yeast (*Saccharomyces cerevisiae*), are described.

Correspondence address: J.N. Yarwood, Botany Department, Science Laboratories, University of Durham, South Road, Durham DH1 3LE, England

2. MATERIALS AND METHODS

2.1. *Materials*

Tricine, phenylmethylsulphonyl fluoride (PMSF) and leupeptin were from Sigma. Peroxidase-linked goat anti-rabbit (GAR) antibody was from BioRad. Affinity-purified anti-pea legumin rabbit antibodies were purified as described [5]. Oligonucleotides were synthesised using a DNA synthesiser (model 318A, Applied Biosystems).

2.2. *Bacterial and yeast strains and culture conditions*

JM83 *ara*, (*lac-pro* A,B), *rps*L (= *str*A) 080, *lac*Z M15 (Bethesda Research Labs) was used as the bacterial host in all experiments and *S. cerevisiae* MC 16 *ade* 2-1, *leu* 2-3, *lys* 2-1, *his* 4-712^{FS}, SUF2 [6] was the yeast host. Yeast

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

transformants were grown under selection conditions for leucine independence in YNB (Difco) minimal media with appropriate supplements. All bacterial growth conditions, media, and transformation procedures were as detailed by Maniatis et al. [7]. Yeast cells were transformed using lithium acetate [8].

2.3. *Plasmids*

pBR322, pUC8 and pUC18 were used for cDNA construction; the yeast expression plasmid pMA91 was a gift from Drs A.J. and S.M. Kingsman [9]. Recombinants were selected by phenotypic change on X gal plates or by colony hybridisation using appropriately labelled fragments. Fragment orientation was determined by restriction analysis. cDNA inserts were sequenced by the dideoxy 'forward and backwards' technique [10,11].

2.4. *Yeast extracts and protein gels*

Yeast cells (20 ml) were grown for 48 h at 30°C in minimal media and appropriate supplements to stationary phase. Cell pellets were washed with extractant and suspended in 200 μl extractant [50 mM Tricine, 200 mM NaCl, pH 8, 3 μg/ml leupeptin, 10 μl/ml PMSF (36 mg/ml in dry ethanol)]. Soluble proteins were extracted by vortex-mixing with an equal volume of acid-washed 450–500 μm glass beads for 3 × 1 min at 0°C. Insoluble material was removed by centrifugation for 10 min in an Eppendorf microfuge. To the supernatant was added 1/4 vol. of 5 × SDS buffer and samples boiled at 100°C for 2 min. Proteins associated with the remaining pellet were extracted by boiling for 20 min with 100 μl of 1 × SDS buffer. Material remaining was removed by centrifugation for 10 min in an Eppendorf microfuge. Samples were applied to a 17% SDS-polyacrylamide gel electrophoresis (PAGE), overlaid with 1 μl mercaptoethanol and subjected to PAGE according to Laemmli [12].

Proteins were electroblotted onto a nitrocellulose filter as described in the Sartorius manual. Nitrocellulose filters were blocked using 5% non-fat dry milk in 20 mM Tris, 0.9% NaCl, pH 7.2. Blots were then incubated with affinity-purified anti-legumin diluted 1:250 in blocking buffer, extensively washed in the same buffer, and incubated with peroxidase-conjugated GAR IgG diluted in blocking buffer and finally extensively washed. The immune complex was visualised using diaminobenzidine as described in the Amersham manual or by detection with ^{125}I-labelled *S. aureus* protein A followed by autoradiography [13]. Quantitative Western blots were performed using a pure legumin standard of 1 mg/ml. Protein was estimated by method of Warburg and Christian [14] using purified bovine serum albumin (BSA) and legumin standards. Appropriate levels of extractant were added to all blanks.

2.5. *Construction of cDNA library*

cDNA library construction was as described previously [15].

2.6. *Immunogold labelling*

Yeast colonies were fixed for 1 h at room temperature on agar plates in 1.25% glutaraldehyde, 3% paraformaldehyde in 0.1 M phosphate buffer, 0.1 M sucrose and incubated overnight in the previous buffer with 1 M sucrose. Individual colonies were frozen in freon slush on metal stubs and cut on a Sorvall HT2-B ultramicrotome and transferred using 'transfer drops' (4% paraformaldehyde in phosphate-buffered saline (PBS) containing 2.4 M sucrose) to carbon-coated formavar grids (200-mesh copper). The grids were processed through 5 changes of PBS.

Grids were washed with water and placed on drops of 1% aqueous BSA for 10 min, again washed and incubated on drops of anti-legumin (1:10 in PBST) for 30 min at room temperature. Control sections were incubated in PBST. Sections were washed 10 × 1 min on drops of PBST before being incubated with a 1:20 dilution of GAR IgG conjugated to 20 nM gold particles in PBST (Bio Clin) for 30 min at room temperature. Sections were washed 10 × 1 min in drops of PBST, being finally washed in distilled water before staining for 10 min with uranyl acetate. The washed sections were examined using a Philips 400 TEM.

3. RESULTS

A hybrid coding and 3′-untranslated sequence equivalent to a full-length *Leg*A-derived cDNA was assembled from several shorter cDNA species and the 5′-end of the coding sequence from the *Leg*A gene. The construction of the hybrid se-

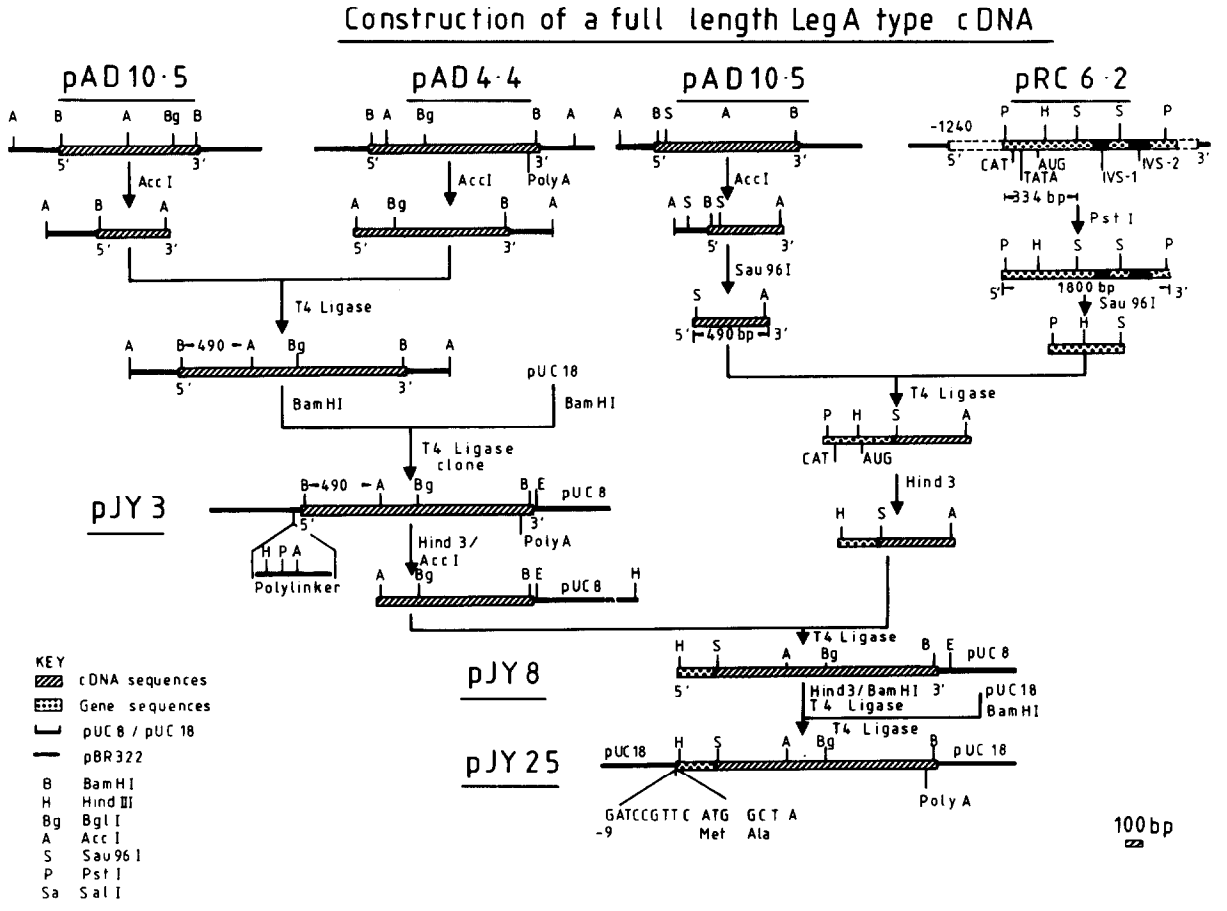

Fig.1. Construction of hybrid legumin cDNA.

quence, designed pJY25 is shown in fig.1. The region from amino acid 3 to 67 was derived by excision of an appropriate fragment from a genomic subclone of the *Leg*A gene pRC6.2 [16]; the region from amino acid 68 to 231 was derived from part of the insert of a cDNA, pAD10.5 [17], and the extreme 3′-region of coding sequence, plus 3′-untranslated sequence including the polyadenylation signal, was derived by excision of part of the insert from a further cDNA species pAD4.4 [17]. These fragments were assembled in pairs prior to a final ligation to produce the construct pJY8, a pUC8 subclone (see fig.1). This construct lacked the first 7 bases of the *Leg*A coding sequence, which was supplied by the addition of a suitable pair of oligonucleotide sequences, containing the missing sequence plus a *Bam*HI linker in the appropriate reading frame. This final construct, pJY25, was obtained as a clone in pUC18. The integrity of the coding sequence across the ligation points was checked by DNA sequencing.

The insert from pJY25 was excised with *Bam*HI, and ligated into the *Bgl*II expression site of the yeast expression vector pMA91 [9] to produce pJY28, which contains the yeast phosphoglycerate kinase (PGK) promoter and 5′- and 3′-untranslated regions and 94 base pairs of 3′-coding sequence. pJY28 was grown in *E. coli*, checked by restriction analysis and then transformed into yeast. Transformants were identified by selection on yeast minimal media in the absence of additional leucine, tested for correct auxotrophy and single colonies grown and tested for legumin expression.

The expression of legumin in yeast containing the pJY28 construct was assayed by Western blot

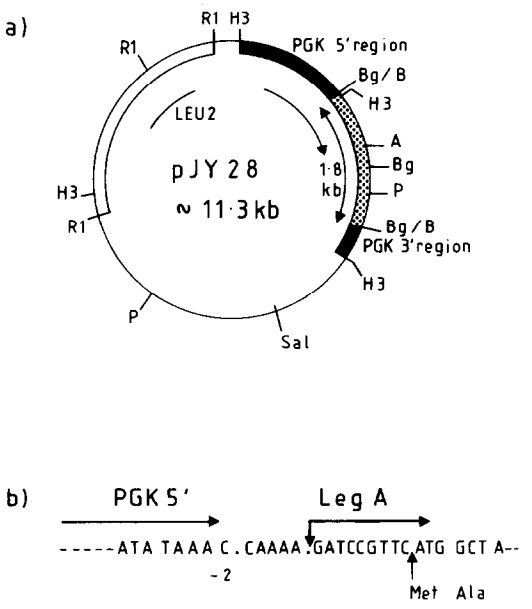

Fig.2. (a) Structure of pJY28. (Black box) PGK sequence; (hatched box) *Leg*A cDNA insert; (open box) yeast DNA sequence containing the *LEU2* gene and the 2μ origin of replication; (thin line) pBR322 sequence containing the *Amp*^{R} gene. The 1.8 kb *Bam*HI insert from pJY25 is inserted at the *Bgl*II site of pMA91 [9] oriented 5′-3′ with respect to the 5′-untranslated PGK sequence. The arrow indicates the direction of transcription. (b) Nucleotide sequence around the *Bgl*II expression site. The sequence to the left of the dot is the 5′-untranslated region of PGK extending to include the nucleotide −2 upstream from the PGK initiator codon. The sequence between the dots is the remainder of the *Bgl*II linker and the sequence to the right of the dot is the synthetic oligonucleotide on pJY25; the arrow represents the initiator codon of the *Leg*A cDNA. A, *Acc*I; B, *Bam*HI; Bg, *Bgl*II; H3, *Hin*dIII; P, *Pst*I; Sal, *Sal*I; R1, *Eco*RI.

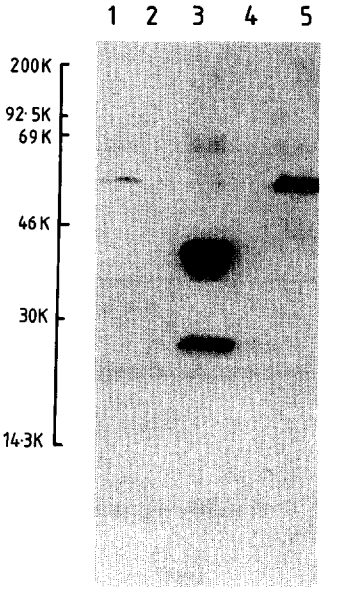

Fig.3. Western blot analysis of legumin synthesised in *S. cerevisiae*. Lanes: (1) supernatant protein from *S. cerevisiae* MC 16 transformed with pJY28; (2) supernatant protein from untransformed *S. cerevisiae* MC 16; (3) 20 μl legumin in 1 × SDS sample buffer; (4) pellet proteins from untransformed *S. cerevisiae*; (5) pellet proteins from *S. cerevisiae* transformed with pJY38. All samples were run in the presence of β-mercaptoethanol, and probed as described in section 2.4. The substrate was 3,3′-diaminobenzidine tetrahydrochloride.

analysis of proteins extracted under non-denaturing and denaturing conditions. The results of such an analysis are shown in fig.3. Polypeptides reacting with anti-legumin IgG are found only in transformed yeast containing the pJY28 construct. The indicated M_r of the legumin polypeptide is approx. 60000, and the presence of the reducing agents does not generate polypeptides of M_r ~38000 and 21000, as in legumin extracted from pea. Protein extraction was routinely carried out in the presence of the protease inhibitors, leupeptin and PMSF, to prevent non-specific proteolysis by endogenous yeast proteases. However even when limited proteolysis was allowed to occur the legumin synthesised in yeast was not cleaved into the polypeptides of M_r 21000 and 38000 but instead generated a random spectrum of fragments in the range M_r 60000–17000 (not shown).

The legumin was found predominantly in the proteins extracted under denaturing conditions from the 'pellet' fraction, with only small or almost undetectable amounts being present in the 'supernatant' protein fraction isolated under non-denaturing conditions. By calibrating the Western blots with known amounts of pea legumin, the level of expression of legumin in yeast could be estimated as 1–2% of the total yeast protein.

The distribution of legumin within the transformed yeast was shown by immunocytochemical localisation in sections cut by cryo-

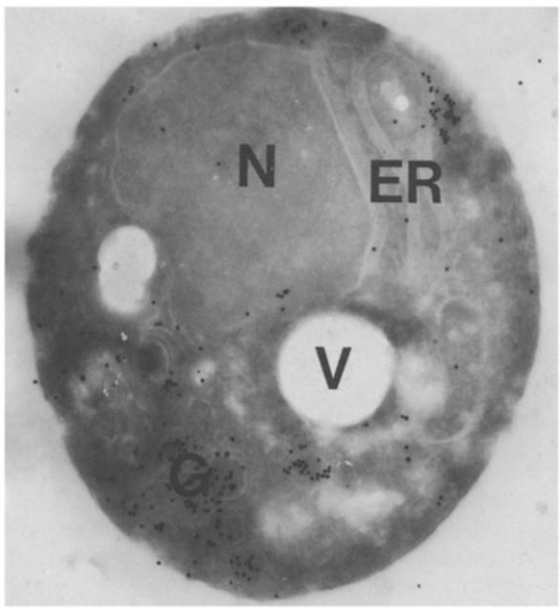

Fig.4. Immunogold localization of legumin synthesis in transformed *S. cerevisiae*. Labelling was carried out as detailed in section 2.6. Note the labelling associated with the Golgi vesicles. N, nucleus; V, vacuole; G, Golgi apparatus. Magnification × 50000.

ultramicrotomy. After incubation with an affinity-purified primary rabbit anti-legumin IgG, the antigen distribution was visualised using GAR-gold (20 nM). Most label was associated with the clustered, densely staining residues of the Golgi apparatus, with some additional staining associated with the endoplasmic reticulum (fig.4). The main vacuoles were unstained, as were the nuclear and cytoplasmic compartments, where staining was no greater than the background level.

4. DISCUSSION

Previous studies [18,19] have failed to produce full-length *Leg*A-derived cDNA species, despite the fact that this was possible for vicilin cDNA, the other major seed storage protein in pea. The reason for this failure is unclear but we have consistently noted an apparent instability of almost complete *Leg*A cDNA species. As a consequence, the construction of a yeast expression vector containing the complete legumin coding sequence was not a straightforward task, and the strategy described above had to be adopted.

The level of expression of legumin in yeast transformed with the pJY28 construct is comparable with those of other heterologous proteins that have been expressed using this and related vectors, e.g. chymosin [9], interferon-α-2 [20]. In comparison, expression levels of up to 50% of total protein have been obtained for yeast PGK in a comparable vector [21]. Conditions for optimising legumin expression have not been explored in the present study but it is apparent that legumin expression does not significantly interfere with the normal growth of yeast.

The failure of yeast to carry out the specific proteolytic cleavage of legumin may be due to a variety of causes. First, yeast may lack a protease comparable to that which carries out the legumin processing in plant seeds. The protease responsible for processing in plants is highly specific, in that cleavage only occurs at one asparagine residue in the entire legumin precursor polypeptide, the same enzyme is thought to be involved in specific cleavages C-terminal to asparagine residues in some vicilin polypeptides [22], and in separating the a- and b- polypeptides of the seed lectin [23], again by cleaving C-terminal to asparagine. The commonly described yeast proteases do not have comparable specificities [24]. Secondly, the newly synthesised legumin may be sequestered into an intracellular compartment in yeast where it is inaccessible to the yeast proteases. The legumin construct pJY28 directs the synthesis of a legumin precursor polypeptide complete with signal peptide, and this would be expected to direct the nascent legumin polypeptide into the lumen of the endoplasmic reticulum, where it would be protected from cytoplasmic and vacuolar proteases. In agreement with this hypothesis the electron-microscopic localisation studies suggest that legumin is synthesised at the endoplasmic reticulum and is sequestered within the Golgi apparatus, in a manner similar to that occurring within pea cotyledonary tissue. Thirdly, it is likely that the legumin polypeptide synthesised in yeast associates within the endoplasmic reticulum into an insoluble oligomeric form [2], in which susceptible bonds are protected from attack. Indeed, using vicilin cDNA expressed in the same yeast host we have demonstrated the presence of a spectrum of partially assembled vicilin molecules [25]. Transport of the newly synthesised legumin into

the lumen of the endoplasmic reticulum would be expected to involve removal of the signal peptide; experiments to determine whether this occurs are currently in progress.

In pea cotyledon the legumin is transferred to, and accumulated within, the vacuolar protein bodies; this step of the post-translational processing and packaging does not appear to have occurred within the transformed yeast.

ACKNOWLEDGEMENTS

We should like to thank N. Appelby and H. Grindley for technical assistance and Drs A.J. and S.M. Kingsman for the plasmid, pMA91. The oligonucleotides were supplied by the University of Durham and the work was supported by an AFRC LINK grant to the Botany Department, Durham University and the IFR Norwich.

REFERENCES

[1] Evans, I.M., Croy, R.R.D., Hutchinson, P., Boulter, D., Payne, P.I. and Gordon, M.E. (1979) Planta 144, 455–463.
[2] Chrispeels, M.J., Higgins, T.J.V. and Spenser, D. (1982) J. Cell Biol. 93, 306–313.
[3] Chrispeels, M.J. (1984) Phil. Trans. R. Soc. B. 309–322.
[4] Chrispeels, M.J., Higgins, T.J.V., Craig, S. and Spenser, D. (1982) J. Cell Biol. 93, 5–14.
[5] Croy, R.R.D., Gatehouse, J.A., Evans, I.M. and Boulter, D. (1980) Planta 148, 57–63.
[6] Beggs, J. (1978) Nature 275, 104–109.
[7] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY.
[8] Ito, H., Fukuda, Y., Murata, K. and Kimura, A. (1983) J. Bacteriol. 153, 163–168.
[9] Mellor, J., Dobson, M.J., Roberts, N.A., Tuite, M.F., Emtage, J.S., White, S., Lowe, P.A., Patel, T., Kingsman, A.J. and Kingsman, S.M. (1983) Gene 24, 1–14.
[10] Maat, J. and Smith, A.J.H. (1978) Nucleic Acids Res. 5, 4537–4546.
[11] Seif, I., Khoury, G. and Dhar, R. (1980) Nucleic Acids Res. 8, 2225–2239.
[12] Laemmli, U. (1970) Nature 227, 680–685.
[13] Burnette, N.V. (1981) Anal. Biochem. 112, 195–203.
[14] Layne, E. (1957) Methods Enzymol. 3, 447–454.
[15] Lycett, G.W., Delauney, A.J., Gatehouse, J.A., Gilroy, J., Croy, R.R.D. and Boulter, D. (1983) Nucleic Acids Res. 11, 2367–2380.
[16] Lycett, G.W., Croy, R.R.D., Shirsat, A.H. and Boulter, D. (1984) Nucleic Acids Res. 12, 4493–4506.
[17] Lycett, G.W., Delauney, A.J., Zhao, W., Gatehouse, J.A., Croy, R.R.D. and Boulter, D. (1984) Plant Mol. Biol. 3, 91–96.
[18] Delauney, A.J. (1984) PhD Thesis, University of Durham.
[19] Gatehouse, L.N. (1985) MSc Thesis, University of Durham.
[20] Tuite, M.F., Dobson, M.J., Roberts, N.A., King, R.M., Burke, D.C., Kingsman, S.M. and Kingsman, A.J. (1982) EMBO J. 1, 603–608.
[21] Mellor, J., Dobson, M.J., Roberts, N.A., Kingsman, A.J. and Kingsman, S.M. (1985) Gene 33, 215–226.
[22] Gatehouse, J., Lycett, G.W., Delauney, A.J., Croy, R.R.D. and Boulter, D. (1983) Biochem. J. 212, 427–432.
[23] Higgins, T.J.V., Chandler, P.M., Spenser, D., Chrispeels, M.J. and Zurawski, G. (1983) NATO Adv. Sci. Inst. Ser. A.63, 93–96.
[24] Pringle, J.R. (1975) Methods Cell Biol. 12, 149–184.
[25] Watson, M.D., Lambert, N., Delauney, A., Yarwood, J.N., Croy, R.R.D., Gatehouse, J.A., Wright, D.J. and Boulter, D. (1987) in press.