

Isolation and characterization of cDNAs encoding oat 12S globulin mRNAs

Gregory Walburg¹ & Brian A. Larkins²

Department of Botany and Plant Pathology, Purdue University, West Lafayette, IN 47907, U.S.A.

¹*Current address: Dept. of Biology, Washington University, St. Louis, MO 63130, U.S.A.*

Summary

A cDNA library was made from poly(A⁺) RNA isolated from developing oat seeds, and oat globulin cDNA clones were identified by hybridization with synthetic oligonucleotides. Globulin clones were characterized by restriction enzyme mapping and cross-hybridization analysis. Based on these comparisons, four classes of globulin clones were distinguished. These clones hybridized to multiple DNA fragments in restriction enzyme digests of oat genomic DNA, indicating that the genes exist in a multigene family. The nucleotide sequence of one of the globulin cDNA clones was determined. The amino acid sequence derived from the DNA sequence verified its identity as an oat globulin and confirmed that the protein is synthesized as a precursor similar to legume 11S storage globulins. The basic polypeptide encoded at the 3' end of the mRNA was found to be homologous to the basic polypeptides of other 11S seed globulins.

Introduction

Oat (*Avena sativa* L.) is unusual among cereal species in that its major seed storage protein is a globulin rather than a prolamin (21). The oat globulin, which comprises 50–80% of the seed protein, shares many structural characteristics with the 11–12S storage globulins of dicot seeds (9). These proteins appear to be hexamers of Mr 60 000–65 000 subunit polypeptides, with each subunit in turn being composed of a Mr 36 000–40 000 polypeptide with an acidic pI and a Mr 20 000–23 000 polypeptide with a basic pI (4, 27). The acidic and basic polypeptides are cross-linked by disulfide bonds. The oat 12S globulin differs from the 11–12S globulins of legumes by requiring much higher concentrations of salt for maximum solubility (4).

As is true of the legume storage proteins, oat

globulin is synthesized by membrane-bound polyribosomes (1) and is deposited within vacuole membranes (22). Based on *in vitro* and *in vivo* labeling experiments, the Mr 60 000 subunits are synthesized as precursors that consist of one acidic and one basic polypeptide (5, 18, 27). The subunits are synthesized as precursors which are post-translationally cleaved to give acidic and basic polypeptides joined by a disulfide linkage (4, 27).

In order to investigate the molecular basis for the high level of globulin synthesis in oat seed, we synthesized and identified cDNA clones corresponding to oat 12S globulin proteins. These clones were used to determine the extent of homology among these sequences and the number of genes encoding 12S globulins in the oat genome. The nucleotide sequence of one clone was determined, and from this we were able to deduce the amino acid sequence of the corresponding protein.

²Author to whom correspondence should be addressed. Journal paper number 10460 of the Purdue Agricultural Experiment Station.

Abbreviations: ds, double stranded; kb, kilobase.

Materials and methods

DNA restriction endonucleases were purchased from Bethesda Research Laboratories (Gaithersburg, MD) or from New England Biolabs (Beverly, MA), Calf intestinal alkaline phosphatase was from Boehringer Mannheim (Indianapolis, IN) and terminal transferase and polynucleotide kinase were from P-L Biochemicals (Milwaukee, WI). [γ - ^{32} P]ATP and [α - ^{32} P]dCTP were obtained from New England nuclear (Boston, MA) or from Amersham (Chicago, IL). Nitrocellulose was purchased from Schleicher and Schuell (Keene, NH). The apparatus used for dot hybridizations was purchased from Bethesda Research Laboratories (Gaithersburg, MD).

DNA isolation and analysis

Preparation of plasmid DNA, restriction enzyme digestion, nick translation of DNA, DNA-end labeling, and bacterial transformation were performed by standard methods as described by Maniatis *et al.* (16). DNA sequencing was by the method of Maxam and Gilbert (20) with the modification of Krayev *et al.* (13).

Synthesis of oligonucleotides

Deoxyoligonucleotides were synthesized manually by a solid phase method with dimethoxytrityl nucleoside phosphoramidites (19) with materials purchased from Applied Biosystems Inc. (Foster City, CA). Polymers of 17 nucleotides were separated from incomplete products in a 20% polyacrylamide gel that contained 7 M urea in TBE (0.01 M Tris, pH 8.3, 0.01 M boric acid, and 0.01 mM EDTA). The oligonucleotides were eluted in 0.5 M NH_4 acetate, 1 mM EDTA (20).

Construction of an oat cDNA library

Poly (A^+) was prepared from developing oat seeds as previously described (27). The RNA was size-fractionated on 5–20% sucrose density gradients. Fractions determined to be most active in globulin synthesis by *in vitro* translation sedimented at approximately 18S and these were used for cDNA synthesis. First strand and second strand cDNA synthesis was as described by Buell *et al.* (6)

and Wickens *et al.* (28). After S1 nuclease treatment, the double-stranded cDNAs were separated by electrophoresis in 5% polyacrylamide gels and molecules longer than 800 bp were excised and eluted. Following homopolymer tailing with dTTP, the cDNAs were inserted into the Pst I site of pUC8 that has been homopolymer tailed with dATP. The recombinant plasmids were used to transform JM83 cells. The resulting white colonies were selected for further analyses.

Identification of 12S globulin clones

To identify globulin sequences among the population of cDNA clones, two pools of synthetic oligonucleotides were synthesized corresponding to the previously determined NH_2 -terminal amino acid sequence of the globulin basic polypeptide (27). The mixtures of synthetic oligonucleotides were end-labeled with ^{32}P -ATP and hybridized to bacterial colonies that were lysed onto nitrocellulose as previously described (11). Plasmids were purified from colonies showing positive hybridization signals and analyzed by restriction enzyme digestion. A clone containing the largest insert (pOG77) was used for nucleotide sequence analysis.

Northern analysis of endosperm RNA

Messenger RNA or total RNA was separated by electrophoresis in denaturing gels containing 5 mM methyl mercury hydroxide and 1.4% agarose (3). After the gels were stained with ethidium bromide, they were blotted onto nitrocellulose and hybridized in a solution containing $5 \times \text{SSC}$, 20 mM NaPO_4 , pH 6.8, 1% SDS, 0.02% Ficoll, polyvinylpyrrolidone, and bovine serum albumin, and 5% dextran sulfate in 50% formamide (16). The hybridization probe was prepared by isolating the cDNA insert from pOG77 and labeling the DNA by nick translation with ^{32}P -dCTP (18). After hybridization the filters were washed with $1 \times \text{SSC}$ containing 0.1% SDS at room temperature, followed by a wash at 68°C with the same solution.

Southern hybridization analysis of oat DNA

DNA was isolated from 7-day oat shoots by a rapid preparation procedure (8). The DNA was digested with Bam HI, Eco RI, and Hind III, sepa-

rated by electrophoresis in a 0.8% agarose gel, and transferred to nitrocellulose. Hybridization with the nick-translated cDNA insert from pOG77 was in a solution containing 40% formamide, $5 \times$ SSC, $5 \times$ Denhart's, 1% SDS, 250 mg/ml denatured calf thymus DNA, and 20 mg/ml poly(A) at 62°C for 18 h. The filters were washed in $0.06 \times$ SSC, 0.1% SDS at 62°C for 2 h and autoradiographed.

Dot blot hybridizations

Filters for dot hybridization were prepared as described by Kafatos *et al.* (12). Plasmids were digested with Bam HI, denatured in 0.4 N NaOH for 10 min at temperature, and neutralized with NH_4 acetate. Nitrocellulose filters were wetted with H_2O and rinsed with 1 M NH_4 acetate before applying 40 μl samples containing 0.5 μg of plasmid DNA in 0.1 M NaOH and 1 M NH_4 acetate. The filter was then rinsed in $3 \times$ SSC and dried.

The probes for hybridization were made by nick-translating inserts isolated from the cDNA clones. The hybridization and wash solutions were the same as described for Southern analysis.

Results

Construction and identification of oat 12S globulin clones

The mRNAs encoding oat 12S globulin precursors sediment at approximately 18S (22). To increase the proportion of globulin sequences among cDNA clones, we isolated mRNAs of the appropriate size from total poly(A) RNA of developing oat endosperm. The 18S mRNAs directed the synthesis of oat globulins, as determined by immunoprecipitation with a polyclonal antibody generated against purified 12S globulin (27). These mRNAs were used as templates for synthesis of double stranded (ds) cDNAs as previously described (6, 28). To insure the isolation of large cDNA inserts, the ds cDNA was separated by electrophoresis in 5% polyacrylamide gels. Molecules of 800 bp or larger were excised from the gel and recovered by electroelution. The ds cDNAs were homopolymer tailed and annealed with the plasmid pUC8 that had been homopolymer tailed with the complementary nucleotide. After transformation of competent

bacterial cells, we selected 560 recombinant plasmids for subsequent analysis. Restriction enzyme analysis of these plasmids indicated that 95% of the recombinants contained cDNA inserts of 800 or more nucleotides.

To identify clones containing oat globulin sequences, we synthesized oligonucleotides for use as hybridization probes. The sequences of these probes were derived from the NH_2 -terminal sequence of the basic polypeptides (27). It was not possible to identify a peptide that is encoded by a unique nucleotide sequence, so we selected two regions that had minimal codon degeneracy. These two sequences corresponded to amino acids 3–8 and 13–18 of the basic polypeptide. The mixture of oligonucleotides corresponding to each of the regions consisted of 32 sequences (oligo I) and 48 sequences (oligo II), respectively (Fig. 1).

Hybridization of ^{32}P -labeled oligonucleotide II to the cDNA library led to the identification of 20 putative globulin clones. Plasmids from colonies showing positive hybridization were prepared and hybridized separately with labeled oligo I and oligo II to confirm their identification. Eleven of the 20 clones hybridized with both oligonucleotides and these were subsequently purified for detailed analysis.

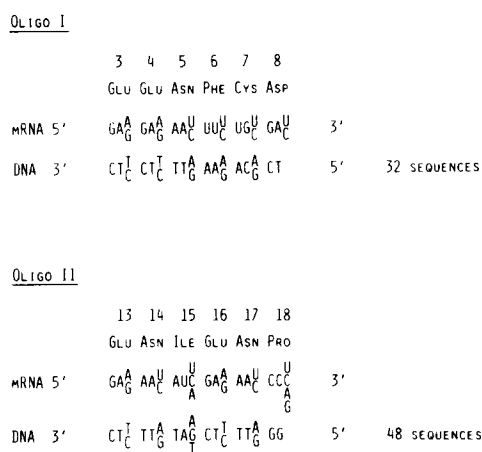


Fig. 1. Synthetic deoxyoligonucleotides used as hybridization probes for the detection of oat globulin cDNA clones. Two 17-mers were synthesized that corresponded to the predicted mRNA sequences encoding residues 3–8 (oligo I) and 13–18 (oligo II) of the oat globulin basic polypeptide. Based on codon degeneracy pools of 32 and 48 sequences were synthesized for oligo I and II, respectively.

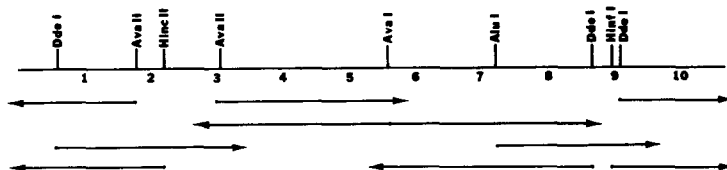


Fig. 2. Restriction map and sequencing strategy of pOG77. Only restriction enzyme sites used for end labeling are shown. The scale of the map is in hundreds of base pairs. The horizontal arrows indicate the direction and extent of sequence analysis. All regions of the cDNA insert were sequenced at least twice; both strands were sequenced for the majority of the clone.

Characterization of globulin cDNA clones

As an initial step to confirm the identity of the cloned cDNAs as 12S globulin sequences, the recombinant plasmids were hybridized to total RNA or poly(A⁺) RNAs from developing oat endosperm. Following separation in agarose gels containing methylmercury hydroxide (3), the RNA was transferred to nitrocellulose and hybridized to cloned cDNAs that were labeled by nick translation. Each of the recombinant plasmids was found to hybridize to a mRNA of approximately 1800 nucleotides based on its migration relation to the marker 18S rRNA (data not shown).

Analysis of the size of the cDNA inserts of these 11 clones indicated that the largest one, designated pOG77, was approximately 1200 nucleotides in length. To verify that this clone corresponded to a 12S globulin protein, we prepared a detailed restriction enzyme map (Fig. 2) and determined its nucleotide sequence. The DNA sequence of pOG77 showed that the cDNA insert extends to the 3' end of the mRNA and includes a portion of the poly(A) tail (Fig. 3). There is a single open reading frame of 313 amino acids that precedes a TAA termination codon. Two potential polyadenylation sequences (AATAAA) are found in the non-coding region following the stop codon. Amino acids 112 through 129 of the predicted sequence were found

to correspond to the NH₂-terminus of the 12S globulin basic polypeptide (27), thus confirming the identity of the clone as a storage globulin se-

```

... GGT AAC AAC AAG AGA GAG CAA CAG TTT GGA CAA AAC ATA TTC AGT GGA TTC AGT      54
   Gly Asn Asn Lys Arg Glu Gln Gln Phe Gly Gln Asn Ile Phe Ser Gly Phe Ser

GTC CAA CTT CTT AGT GAG GGC CTT GGT ATA AGT CAG CAA GCA CAA AAG ATT CAG      111
   Val Gln Leu Leu Ser Glu Ala Leu Gly Ile Ser Gln Gln Ala Ala Gln Lys Ile Gln

AGT CAA AAT GAC CAA AGA GGT GAG ATA ATT CGT GTG AGT CAA GGC CTT CAA TTC TTG      168
   Ser Gln Asn Asp Gln Arg Gly Glu Ile Ile Arg Val Ser Gln Gly Leu Gln Phe Leu

AAG CCT TTT GTT TCC CAA CAA GGA CCA GTA GAG CAT CAA GGC TAC CAA CCA ATT CAA      225
   Lys Pro Phe Val Ser Gln Gln Gly Pro Val Glu His Gln Ala Tyr Gln Pro Ile Gln

AGT CAA CAA GAA CAA TCA ACC CAA TAC CAG GTA GGG CAA TCA CCA CAA TAT CAA CAA      282
   Ser Gln Gln Glu Gln Ser Thr Gln Tyr Gln Val Gly Gln Ser Pro Gln Tyr Gln Glu

GGA CAA TCA ACT CAA TAC CAG TCA GGA CAG TCA TGG GAC CAA AGT TTC AAT GGT TTG      339
   Gly Gln Ser Thr Gln Tyr Gln Ser Gly Gln Ser Trp Asp Gln Ser Phe Asn Gly Leu
   (NH2)G L

GAG GAG AAT TTC TGT TCA TTG GAG GCA AGG CAA AAC ATC GAA AAC CCG AAA CGT GGC      396
   Glu Glu Asn Phe Cys Ser Leu Glu Ala Arg Gln Asn Ile Glu Asn Pro Lys Arg Ala
   E E N F C D L E A R E N I E N P

GAC ACG TAC AAC CCA CGT GGT GGC AGG ATA ACA CAT CTC AAT AGC AAG AAT TTT CCC      453
   Asp Thr Tyr Asn Pro Arg Ala Gly Arg Ile Thr His Leu Asn Ser Lys Asn Phe Pro

ACC CTT AAC TTG GTG CAA ATG AGT CCT ACA AGA GTA AAT TTA TAC CAG AAT GCT ATT      510
   Thr Leu Asn Leu Val Gln Met Ser Ala Thr Arg Val Asn Leu Tyr Gln Asn Ala Ile

CTT TCA CCA TAC TGG AAC ATT AAT GCT CAC AGT GTC ATG CAC ATG ATC CAA GGA CGT      567
   Leu Ser Pro Tyr Trp Asn Ile Asn Ala His Ser Val Met His Met Ile Gln Gly Arg

GCT CGA GTT CAA GTT GTC AAT AAC CAT GGT CAG ACC GTA TTC AAT GAC ATT CTT CGT      624
   Ala Arg Val Gln Val Val Asn Asn His Gly Gln Thr Val Phe Asn Asp Ile Leu Arg

CGC GGA CAA CTA ATA ATC CCA CAA CAC TAT GTT GTT CTC AAG AAG GCA GAG CGT      681
   Arg Gly Gln Leu Leu Ile Ile Pro Gln His Tyr Val Val Leu Lys Lys Ala Glu Arg

GAA GGA TGC CAA TAT ATT TCA TTC AAG ACC ACC CCC AAC TCT ATG GTT AGC TAC ATC      738
   Glu Gly Cys Gln Tyr Ile Ser Phe Lys Thr Thr Pro Asn Ser Met Val Ser Tyr Ile

GCA GGA AAG ACC TCC ATC CTA CGT GCA TTG CCC GTT GAT GTC CTC GGC AAT GCA TAC      795
   Ala Gly Lys Thr Ser Ile Leu Arg Ala Leu Pro Val Asp Val Leu Ala Asn Ala Tyr

CGC ATT TCT AGG CAG GAA TCC CAA AAC CTC AAA AAT AAT CGT GGA GAA GAG TTT GGT      852
   Arg Ile Ser Arg Gln Glu Ser Gln Asn Leu Lys Asn Asn Arg Gly Glu Glu Phe Gly

GCA TTC ACC CCT AAG TTT GCA CAA ACG GGC TCC CAG AGT TAC AGG ACG ACG GAG AGT      909
   Ala Phe Thr Pro Lys Phe Ala Gln Thr Gly Ser Gln Ser Tyr Arg Thr Arg Glu Ser

CAT CTT CGA CTG AGA AGG CAT CCG AGT GAA TAA GTGAG TGTAAATGAA ACTAGATATG      967
   His Leu Arg Leu Arg Arg His Pro Ser Glu End

TGAAATAAAG GCATAGCATG TTTCAGCCT AGTGGTATAT AACCGCTTAC TCAATAAATA      1027
   1065
   AGTTTCCTGGT GTTA24

```

Fig. 3. Nucleotide and predicted amino acid sequence of pOG77. The NH₂-terminal sequence of purified basic polypeptides is shown in single letter designation below the predicted amino acid sequence; this also indicates the junction of the acidic and basic polypeptides. The amino acid sequence repeats in the acidic polypeptide between nucleotides 238 and 315 are underlined. The polyadenylation signals in the 3' untranslated regions are overlined.

quence. The length of the basic polypeptide is 202 amino acids, and corresponds to a protein of Mr 26744.

We were unable to determine the sequence of the acidic polypeptide of the 12S globulin because the NH₂-terminus is blocked (27). Although the 111 amino acids preceding the basic polypeptide would only correspond to a portion of the acidic polypeptide, the amino acid composition of this sequence is similar to that previously determined (21). Furthermore, the sequence of this portion of the clone is homologous to amino acid sequences of acidic polypeptides from other 11S seed storage globulins (14, 15, 17, 24). Within this region of the protein there is a short repeated peptide of 10 amino acids that has the sequence QSTQYQVGQS. This peptide begins at amino acid 79 and extends through amino acid 89. An imperfect repeat of the sequence, QSTQYQSGQS, lies between amino acids 96 and 105, and between these 10 amino acid repeats, at positions 91–95, is a five amino acid variant of this repeat QYQEG.

Comparison of sequence homology among globulin cDNA clones

Restriction enzyme mapping of the 11 globulin cDNA clones revealed considerable sequence heterogeneity. To further examine the relationships among these clones, we analyzed the extent of their sequence homology by cross-hybridization. Samples of each plasmid were immobilized on nitrocellulose filters and hybridized at several criteria to the ³²P-labeled insert from pOG77 (Fig. 4). At 45°C in

5 × SSC the pOG77 sequence hybridized to all of the other cDNA clones. However, at 62°C in 5 × SSC pOG77 hybridized to only four of the 10 clones.

To examine the relationships between the clones that did not hybridize with pOG77 at the stringent criterion, a similar experiment was performed. Each of the 11 clones was immobilized on nitrocellulose, and replicate samples were hybridized with ³²P-labeled inserts from those clones that did not cross-hybridize with pOG77. At the stringent criterion these clones can be divided into four distinct groups (Fig. 5). Group 1 contains those which hybridize with pOG77 and includes clones pOG107, pOG234, pOG258, and pOG302 (Fig. 4). Clones pOG250 and pOG270 did not hybridize with any other clones, and thus form two distinct classes. The fourth group contains plasmids pOG252, pOG277, and pOG438 (Fig. 5). Thus, while there is a significant degree of homology among the globulin sequences, by using a more stringent hybridization criterion it is possible to distinguish among them.

Southern hybridization analysis of genomic DNA

To investigate the number of genes encoding globulin subunits in the oat genome, we hybridized pOG77 to oat genomic DNA at a criterion that should allow cross-hybridization to all the characterized globulin sequences. DNA was isolated from seven-day-old germinating oat shoots and digested with the restriction enzymes Bam HI, Eco RI, and Hind III, which do not cut in the cDNA clones.

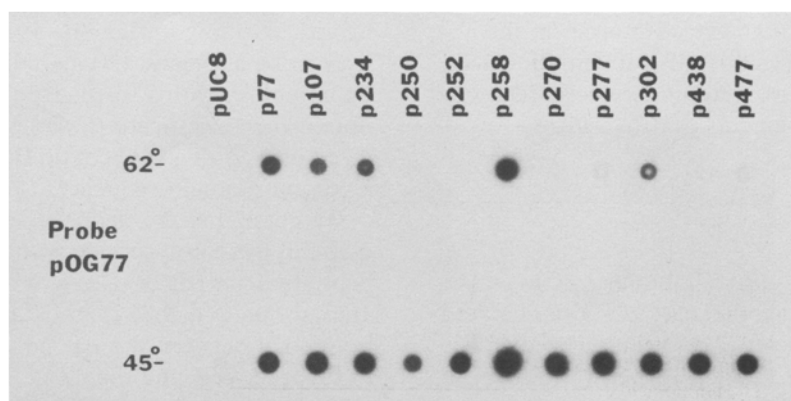


Fig. 4. Cross-hybridization analysis of pOG77 with oat 12S globulin cDNA clones. Recombinant plasmids were fixed on nitrocellulose and hybridized with the ³²P labeled insert from pOG77 in 5 × SSC, 40% formamide at 42°C and 62°C.

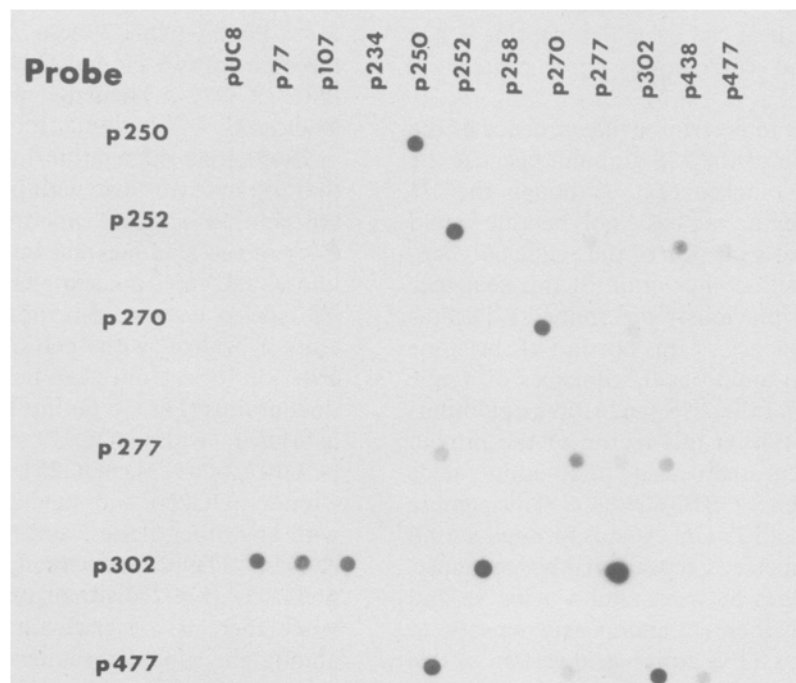


Fig. 5. Cross-hybridization analysis of homology among oat 12S globulin cDNA clones. Recombinant plasmids corresponding to various globulin clones were fixed on nitrocellulose and hybridized at 62 °C in $5 \times$ SSC, 40% formamide with ^{32} P-labeled inserts from each of the clones that did not cross-hybridize with pOG77.

The DNA was separated in an 0.8% agarose gel and transferred to nitrocellulose by the procedure of Southern (26). At a criterion of $T_m - 27^\circ\text{C}$, the cDNA insert hybridized to multiple DNA fragments in each of the digests (Fig. 6). The hybridization intensity of some of these fragments was more intense than others, suggesting that some of the genes may be clustered or that they are present on similar sized DNA fragments. However, in the absence of restriction maps of full-length cDNA clones, it is difficult to determine precise relationships among the fragments that hybridize.

Discussion

Because of the sequence homology between the basic polypeptides of the oat 12S globulin and those of legume species (27), we attempted to use cDNA clones corresponding to pea (15), soybean (24), and rape seed 11S globulins (25) as heterologous probes to identify oat 12S globulin sequences. However, none of these were found to hybridize to

oat endosperm poly(A) RNA at even a low hybridization stringency. We therefore used synthetic oligonucleotides to identify oat globulin sequences.

Although the oligonucleotide probes we synthesized were complex mixtures, i.e. 32 and 48 sequences (Fig. 1), they were nevertheless effective hybridization probes. Eleven of the 20 clones that hybridized to oligonucleotide II were found to also hybridize to oligonucleotide I. All 11 of these appear to be authentic 12S globulin sequences based on homology with a verified globulin clone. We did not further investigate the nine clones that hybridized only to oligonucleotide II, but they may also represent globulin sequences.

We compared the deduced sequence of the oat globulin basic polypeptide with that of basic polypeptides from dicot 11S globulins (Fig. 7). This analysis showed that there is approximately 40% homology between the oat polypeptide and those in soybean, pea, and rape seed. The basic polypeptide of the oat globulin was more homologous to that of the soybean Gy₄ subunit than it was to the other sequences that we compared.

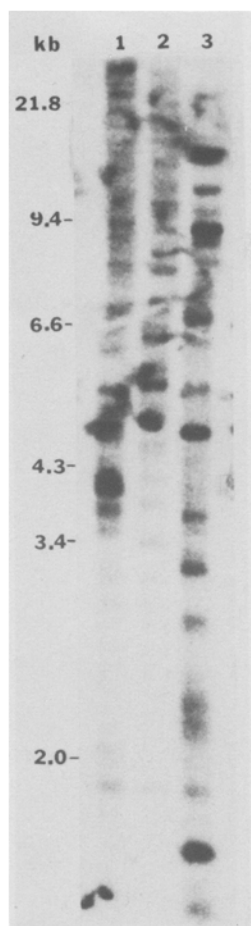


Fig. 6. Analysis of genes encoding oat globulin by Southern hybridization. Oat DNA was digested with the restriction enzymes Bam HI (lane 1), Eco RI (lane 2), and Hind III (lane 3), and separated by electrophoresis in a 0.8% agarose gel. The DNA was transferred to nitrocellulose and hybridized with nick translated insert from pOG77. The position of the mol wt markers is shown at the left of the figure.

Although we have no sequence information for the acidic polypeptide, the protein deduced from this region of the clone is similar to the acidic polypeptides of dicot 11S seed globulins (14, 15, 24). In addition to having an amino acid composition consistent with the acidic polypeptide of the oat globulin, this region of the protein has several structural features of other globulin precursors. In legume 11S globulins (14, 15, 24) repeated peptides of 10–30 amino acids were found to precede the NH_2 -terminus of the basic polypeptide. We identified similar repeats between amino acids 79–105 in

pOG77. The region where these repeats occur appears to be a highly variable portion of the acidic polypeptide (2). The sequence of the first 79 amino acids in pOG77 is similar to the comparable region in legume 11S globulins.

Processing of the soybean 11S globulin precursor includes the removal of a peptide of four amino acids between the acidic and basic polypeptides (17); however, in pea the COOH -terminal amino acid of the acidic polypeptide immediately precedes the NH_2 -terminus of the basic polypeptide (14). Since the identity of the COOH -terminal amino acid of the oat acidic polypeptide is unknown, we cannot determine the presence or absence of a linker peptide in this region of the molecule. However, we found an asparagine residue precedes the NH_2 -terminal glycine of the basic polypeptide in the oat globulin precursor. This Asn-Gly sequence is conserved at the site of proteolytic cleavage for all globulin sequences that have been examined (7, 14, 15, 17, 24, 25).

Two-dimensional gel analysis of oat 12S globulins revealed a large number of acidic and basic polypeptides (4, 27), suggesting that these proteins are encoded by a complex mixture of mRNAs. Comparison of restriction enzyme maps of globulin cDNA clones showed considerable heterogeneity among these sequences, and this was supported by results of experiments comparing sequence homology (Figs. 4, 5). At a criterion calculated to be $T_m - 27^\circ\text{C}$ there was sufficient homology among the clones that they all cross-hybridized. It was possible to distinguish four groups of clones when hybridization was at a more stringent criterion ($T_m - 8^\circ\text{C}$). Although we do not know the amount by which the coding sequences of these clones overlap, it is likely that they have a considerable region of common coding sequence. This conclusion is based upon the fact that each clone hybridized to oligonucleotides corresponding to the NH_2 -terminus of the basic polypeptide, which is in the central region of the mRNA. Since the size of the cDNA inserts in these clones ranged from 850–1200 bp, they should overlap a significant proportion of the sequence at the juncture of the acidic and basic polypeptides.

Southern hybridization analysis of oat genomic DNA demonstrated that the globulin genes are present in the genome in multiple copies. The clone pOG77 hybridized to a number of DNA fragments,

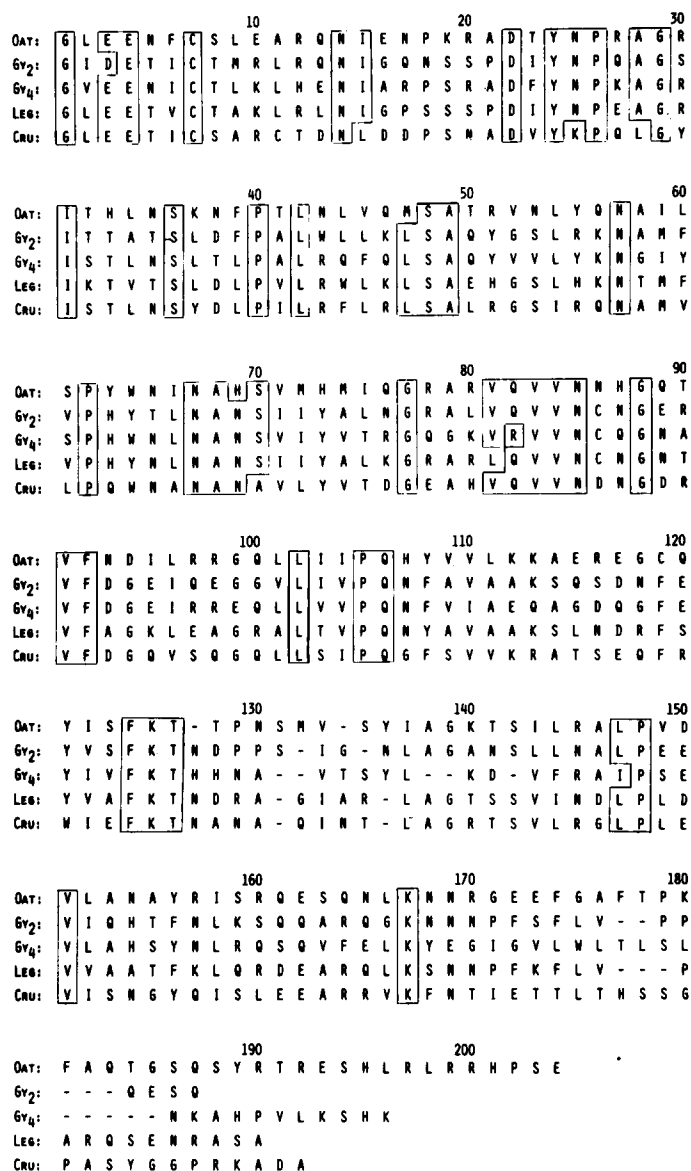


Fig. 7. Comparison of the amino acid sequences of basic polypeptides from five different 12S globulins. The sequences that are compared are from oat, soybean (17, 24), pea (7), and rape seed (25). Gaps were introduced in the sequence to maximize homology; identical amino acids are enclosed to highlight them.

some of which showed hybridization intensity consistent with the presence of multiple gene copies. At the criterion used in this analysis we would expect to detect all sequences similar to the group of cDNA clones we have described. However, it is possible that there are other globulin genes in the genome that are less homologous to pOG77. There

are at least two sub-families of genes encoding the soybean 11S globulin (24), and each contains two or three genes. Likewise, in pea the 11S globulins are encoded by a relatively small number of genes (7, 10). Our results indicate that oat 12S globulins are probably encoded by a larger multigene family than the 11S globulins of legumes.

Acknowledgements

We wish to thank Dr Robert Hanau for help with oligonucleotide synthesis, and Dr Karl Pedersen and Dr David Marks for assistance with DNA sequencing methods. Technical assistance from Pamela Gutay and Judith Lindell was very much appreciated.

References

- Adeli K, Altosaar I: Role of endoplasmic reticulum in biosynthesis of oat globulin precursors. *Plant Physiol* 73:949–955, 1983.
- Argos P, Narayana SVL, Nielsen NC: Structural similarity between legumin and vicilin storage proteins from legumes, *EMBO J* (in press).
- Bailey JM, Davidson N: Methyl mercury as a reversible denaturing agent for agarose gel electrophoresis *Anal Biochem* 70:75–85, 1976.
- Brinegar AC, Peterson DM: Separation and characterization of oat globulin polypeptides *Arch Biochem Biophys* 219:71–79, 1982.
- Brinegar AC, Peterson DM: Synthesis of oat globulin precursors *Plant Physiol* 70:1767–1769, 1982.
- Buell GN, Wickens MP, Payvar F, Schimke RT: Synthesis of full length cDNAs from four partially purified oviduct mRNAs *J Biol Chem* 253:2471–2482, 1978.
- Croy RRD, Lycett GW, Gatehouse JA, Yarwood JN, Boulter D: Cloning and analysis of cDNAs encoding plant storage protein precursors *Nature* 295:76–79, 1982.
- Dellaporta SL, Wood J, Hicks JB: A plant DNA miniprep: Version II. *Plant Mol Biol Reporter* 1:19–21, 1983.
- Derbyshire E, Wright DJ, Boulter D: Legumin and vicilin, storage proteins of legume seeds. *Phytochemistry* 15:3–24, 1976.
- Domoney C, Casey R: Measurement of gene number for seed storage proteins in *Pisum*. *Nucleic Acids Res* 13:687–699, 1985.
- Gough NM, Gough J, Metcalf D, Kelso A, Nicola NA, Burgess AW, Dunn AR: Molecular cloning of cDNA encoding a murine haematopoietic growth regulator, granulocyte-macrophage colony stimulating factor. *Nature* 309:763–767, 1984.
- Kafatos FC, Jones CW, Efstratiadis A: Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure, *Nucleic Acids Res* 7:1541–1552, 1979.
- Krayev AS, Kramerow DA, Skryabin KG, Ryslov AP, Bayer AA, Georgien CP: The nucleotide sequence of the ubiquitous repetitive DNA sequence B1 complementary to the most abundant class of mouse fold-back RNA, *Nucleic Acids Res* 8:1201–1215, 1980.
- Lycett GW, Croy RRD, Shirgat AH, Boulter D: The complete nucleotide sequence of a legumin gene from pea (*Pisum sativum* L). *Nucleic Acids Res* 12:4493–4506, 1984.
- Lycett GW, Delauney AJ, Zhao W, Gatehouse JA, Croy RRD, Boulter D: Two cDNA clones coding for the legumin protein of *Pisum sativum* L. contain sequence repeats. *Plant Mol Biol* 3:91–96, 1984.
- Maniatis T, Fritsch EF, Sambrook J: *Molecular Cloning: A laboratory Manual*. Cold Spring Harbor Laboratory, New York, 1982.
- Marco YA, Thanh VH, Tumer NE, Scallion BJ, Nielsen NC: Cloning and structural analysis of DNA complementary to A₂B_{1a} subunit of glycinin. *J Biol Chem* 259:13436–13441, 1984.
- Matlashewski GJ, Adeli K, Altosaar I, Shewry PR, Mifflin BJ: *In vitro* synthesis of oat globulin. *FEBS Lett* 145:2085–2092, 1982.
- Matteucci MD, Caruthers MH: The synthesis of oligodeoxypyrimidines on a polystyrene support. *Tetrahedron Lett* 21:719–722, 1980.
- Maxam A, Gilbert W: Sequencing DNA with base-specific chemical cleavages. *Methods Enzymol.* 65:499–559, 1980.
- Peterson DM: Subunit structure and composition of oat seed globulin. *Plant Physiol* 62:506–509, 1978.
- Rossi HA, Luthe DS: Isolation and characterization of oat globulin messenger RNA. *Plant Physiol* 72:578–582, 1983.
- Saigo RH, Peterson DM, Holy J: Development of protein bodies in oat starch endosperm, *Can J Bot* 61:1206–1215, 1983.
- Scallion BJ, Thanh VH, Floener LA, Nielsen NC: Identification and characterization of DNA clones encoding Group-II glycinin subunits. *Theor Appl Genet* (submitted) 1985.
- Simon AE, Tenbarger KM, Scotfield SR, Finkelstein RR, Crouch ML: Nucleotide sequence of a cDNA clone of *Brassica napus* 12S storage protein shows homology with legumin from *Pisum sativum*. *Plant Mol Biol* (submitted) 1985.
- Southern EM: Detection of specific DNA sequences among DNA fragments separated by electrophoresis, *J Mol Biol* 98:503–517, 1975.
- Walburg G, Larkins BA: Oat Seed Globulin. Subunit characterization and demonstration of its synthesis as a precursor, *Plant Physiol* 72:161–165, 1983.
- Wickens MP, Buell GN, Schimke RT: Synthesis of double-stranded DNA complementary to lysozyme, ovomucoid and ovalbumin mRNAs, *J Biol Chem* 253:2483–2495, 1978.

Received 27 August 1985; in revised form 13 November 1985; accepted 19 November 1985.