Plant Physiol. (1988) 87, 698–704
0032-0889/88/87/0698/07/$01.00/0

# Molecular Characterization of Oat Seed Globulins[1]

Received for publication January 11, 1988 and in revised form March 10, 1988

MARK A. SHOTWELL[2], CLAUDIO AFONSO, ERIC DAVIES, RUTH S. CHESNUT, AND BRIAN A. LARKINS*
*Department of Botany and Plant Pathology, Purdue University, West Lafayette, Indiana 47907* (M.A.S., R.S.C., B.A.L.); *and School of Biological Sciences, University of Nebraska, Lincoln, Nebraska 68508* (C.A., E.D.)

## ABSTRACT

**We have isolated full-length cDNA clones that encode oat (*Avena sativa*) seed storage globulin mRNAs from a cDNA library in the expression vector lambda gtll. The longest of these clones, pOG2, has an 1840-base pair insert that encodes a complete precursor subunit with a signal peptide of 24 amino acids followed by an acidic polypeptide of 293 amino acids and a basic polypeptide of 201 amino acids. Near the C terminus of the acidic polypeptide are four repeats of a highly conserved, glutamine-rich octapeptide. Other oat globulin cDNA clones contain five of these repeats. Nucleotide sequence comparisons between these clones indicate that the genes encoding these proteins are highly conserved. We estimate there to be 7 to 10 genes for the oat globulin per haploid genome. Comparisons of amino acid sequences show that the oat globulin is 30 to 40% homologous with storage globulins of legumes and about 70% homologous with the rice seed storage globulin (glutelin).**

During their development, plant seeds accumulate large amounts of storage proteins that serve as sources of nitrogen, sulfur, and carbon compounds during seed germination (25). Two major classes of storage proteins can be distinguished: globulins, which are found principally in the cotyledons and axis of the embryo, and prolamines, which are found in the endosperm of cereal seeds (13). Two major types of storage globulins have been described that have sedimentation coefficients of about 7S and 11S. The proportion of the two types of globulins is variable among dicots; monocot embryos contain only the 7S globulin, which is present in the scutellum.

Storage globulins account for most of the protein in dicot seeds, but they generally make up only a small fraction of the protein found in cereal seeds. Instead, most cereals contain predominantly prolamine-type storage proteins. Oats and rice are exceptions. These two contain only small amounts of prolamine (5–10%), and most of their storage protein is an 11 to 12S globulin that is synthesized in the endosperm. Both of these proteins are structurally related to the 11S globulins found in dicots (25), but both are much less soluble than the dicot 11S globulins. The oat globulin requires 0.8 to 1.0 M NaCl for solubility, whereas the characteristics of the rice globulin (glutelin) are such that denaturing solvents are required for solvation.

Previous studies in our laboratory (29) and elsewhere (5, 7) have partially characterized the structure of the oat 12S globulin. This protein is a hexamer of apparent $M_r$ 320,000 (21). The subunits are $M_r$ 54,000, and each consists of an $M_r$ 32,000 acidic polypeptide that is disulfide bonded to an $M_r$ 22,000 basic polypeptide. We also reported the isolation of a cDNA clone corresponding to one of these subunits that contains the coding sequence for all of the basic polypeptide but only a small portion of that for the acidic polypeptide (30). We now report the isolation of cDNA clones containing the entire coding sequence of the oat 12S globulin. DNA sequence analysis of the longest of these cDNA clones shows that it codes for a polypeptide of the size and amino acid composition of purified oat globulin. Structural comparison of this protein with those from legumes suggests that the reduced solubility of the oat globulin is due to the presence of repeated peptides of neutral isoelectric point that are found near the C terminus of the acidic polypeptide. We also show that the amino acid sequence predicted by the oat globulin cDNA is more similar to that of the globulin (glutelin) of rice seeds than to the 11S globulins of legumes.

## MATERIALS AND METHODS

**Materials.** Restriction endonucleases, *Eco*RI linkers, and DNA ligase were purchased from Bethesda Research Laboratories (Gaithersburg, MD). [$\alpha$-$^{32}$P]dCTP, [$\alpha$-$^{32}$P]dATP, and [$\alpha$-$^{35}$S]dATP were purchased from New England Nuclear (Boston, MA) or from Amersham (Arlington Heights, IL). Nick translation kits were purchased from Amersham. Nitrocellulose was from Schleicher and Schuell (Keene, NH). Lambda gt11 arms and *in vitro* packaging extract were obtained from Promega (Madison, WI).

**Construction of an Oat cDNA Library in λ-gt11.** Total RNA was isolated from frozen immature oat (*Avena sativa*) grains (10–15 DAF) of cultivar Gary by the procedure of Hall *et al.* (10). The RNA was suspended in 10 mM Tris-HCl (pH 7.5) and 500 mM KCl, and poly(A)-containing RNA was obtained by one cycle of oligo(dT)-cellulose chromatography (3). After precipitation in ethanol, the poly(A$^+$) RNA was dissolved in sterile water.

cDNA was synthesized, treated with *Eco*RI methylase, and tailed with synthetic *Eco*RI linkers as described by Huynh *et al.* (12). The cDNA was size-fractionated by chromatography on a Bio-Gel A-50 m column (12), and fractions containing the largest cDNAs were collected and used in the construction of the library. These fractions contained about 0.5 μg of cDNA of average size 2.2 to 2.5 kb[3]. One hundred ng of the *Eco*RI-linked cDNA was ligated with 2 μg of λ-gt11 arms and assembled into phage particles *in vitro*. A total of $2.9 \times 10^5$ recombinant clones were obtained, and the library was stored at 4°C.

**Screening the cDNA Library with Oat Globulin Antiserum.** A total of $1.2 \times 10^4$ phage clones were grown on *Eschericia coli*. Antiserum against oat globulin (29) was used for serological

[1] Supported by a grant from the United States Department of Agriculture Competitive Research Grants Office (87-CRCR-1-2356) to B. A. L. This is Journal Paper No. 11,423 of the Purdue Agricultural Research Station.

[2] Recipient of a postdoctoral fellowship from the National Science Foundation.

[3] Abbreviations: kb, kilobase pairs; DPA, days postanthesis.

screening as described by Huynh *et al.* (12). Immunoperoxidase staining of the filters was as described by Barbara *et al.* (4). This involved incubating the filters for 16 h at room temperature with rabbit anti-oat globulin serum diluted 1:500 in HST buffer (10 mM Tris [pH 7.5], 1 M NaCl, 0.5% Tween-20). The primary antiserum was then removed, and the filters were washed extensively in TBS buffer (20 mM Tris [pH 7.5], 0.5 M NaCl) containing 0.15% Tween-20. The filters were then incubated with goat anti-rabbit IgG-horseradish peroxidase conjugate (GAR-HRP, Bio-Rad) diluted 1:2000 in HST buffer. The unbound GAR-HRP was removed by extensive washing with TBS buffer containing 0.15% Tween-20, and after a final wash with TBS without Tween, the HRP color development reaction was carried out according to the manufacturer's instructions.

**Screening the cDNA Library by Nucleic Acid Hybridization.** A total of $7.2 \times 10^4$ clones were transferred to nitrocellulose filters as described by Maniatis *et al.* (17). Those clones containing oat globulin cDNA sequences were identified by hybridization with a 301-bp *Eco*RI-*Hin*cII fragment from cDNA clone pOG1, which was previously isolated from the cDNA library by antibody screening. The DNA fragment was radioactively labeled by nick translation and hybridized in 6× SSC (1× SSC = 0.15 M NaCl, 0.015 M sodium citrate [pH 7.5]), 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 0.1% bovine serum albumin, 0.01 M EDTA, 0.5% SDS, and 0.1 mg/ml sheared calf thymus DNA for 18 h at 65°C. The filters were washed twice for 15 min with 2× SSC, 0.5% SDS at room temperature and then twice for 15 min with 1× SSC, 0.1% SDS at 65°C.

**DNA Sequence Analysis.** Oat globulin cDNA fragments were isolated from λ-gt11 clones and subcloned as *Eco*RI fragments into either pUC18 or pGem3 (Promega, Madison, WI). After further subcloning of smaller restriction fragments, nucleotide sequences were determined by the dideoxy chain termination method (24) according to the instructions accompanying sequencing kits from Promega and United States Biochemicals (Cleveland, OH).

**Estimation of Globulin Sequences in the Oat Genome.** Oat genomic DNA from cultivar Gary was provided by Dr. Mike Murray, Agrigenetics, Madison, WI. This DNA was digested to completion with *Eco*RI, fractionated by electrophoresis in a 0.8% agarose gel, and transferred to nitrocellulose according to Meinkoth and Wahl (19). Conditions for hybridization were similar to those described for screening the cDNA library, except that the 1840-bp insert from cDNA clone pOG2 was labeled by nick translation and used as the probe.

## RESULTS

**Isolation of Oat Globulin cDNA Clones.** We previously constructed a cDNA library of oat seed RNA from which we recovered clones containing oat 12S globulin sequences (30). The clone pOG77 contained sequence coding for the basic polypeptide at the C-terminus of the oat globulin subunit, but lacked most of the coding sequence for the acidic polypeptide. The amino acid sequence of the acidic polypeptide is unknown because its N terminus is blocked (29). Therefore, in order to derive the sequence of the acidic polypeptide, it was necessary to isolate full-length cDNA clones corresponding to oat globulin mRNAs.

In attempts to isolate longer oat globulin cDNA clones, we rescreened our cDNA library with pOG77, but this failed to yield full-length clones. As an alternative approach, we constructed a cDNA library in the expression vector λ-gt11 using size-selected poly($A^+$) RNA isolated from developing oat seeds. This library was initially screened with polyclonal antibodies against purified oat 12S globulin (29). From $1.2 \times 10^4$ clones, we identified seven that gave a positive reaction with the anti-oat globulin serum. The cDNA inserts in these clones ranged from about 350 to 1650 bp. After subcloning into pUC18, portions of the cDNA sequences were determined. We found that the longest of the seven cDNAs, plasmid pOG1, was incomplete at the 5′ end; it coded for the entire acidic and basic polypeptides, but it contained sequence for only the last 10 amino acids of the signal peptide and totally lacked 5′ noncoding sequence. A partial restriction map of the 1650-bp pOG1 cDNA insert is presented in Figure 1, along with that of pOG77, the previously isolated clone (30).

In order to obtain full-length cDNA clones, we isolated a 301-bp *Eco*RI-*Hin*cII restriction fragment from the 5′ end of the pOG1 insert (Fig. 1) and used it as a hybridization probe to rescreen the λ-gt11 cDNA library. In this way, we isolated a number of clones with inserts larger than that in pOG1 (Fig. 1). Of these, clone pOG2 contained the largest insert and was selected for further analysis.

**Sequence Analysis of pOG2.** The insert from pOG2 was restriction-mapped and then subcloned into pGem3 for nucleotide sequence analysis. The pOG2 insert is 1840 bp long, including 29 adenosine residues at the 3′ end, and contains one open reading frame beginning with the ATG codon 146 bp from the 5′ end (Fig. 2). The sequence contains 518 codons with a TGA translation stop. Where they overlap, the nucleotide sequence of the pOG2 insert is identical to that of the pOG77 clone (30), indicating that these two cDNA clones are copies of the same mRNA sequence.

Earlier studies demonstrated that oat globulin subunits are synthesized as precursors with N-terminal signal peptides (1, 6, 29). In the pOG2 sequence, there are grouped hydrophobic amino acid residues following the initiation codon, as would be expected for a signal peptide (Fig. 2). Because the amino acid sequence at the N-terminus of the acidic polypeptide is blocked (29), we cannot identify with certainty the site of cleavage of the signal peptide. Nevertheless, we believe that the cleavage occurs between the alanine at position 24 and the glutamine at position 25 on the basis of signal peptide cleavage specificities described by Von Heijne (28) (Fig. 2). A cyclized glutamine residue at the N terminus could then account for its resistance to Edman degradation.

The amino acid sequence at the N terminus of the basic polypeptide of two oat globulins has been determined (29). This

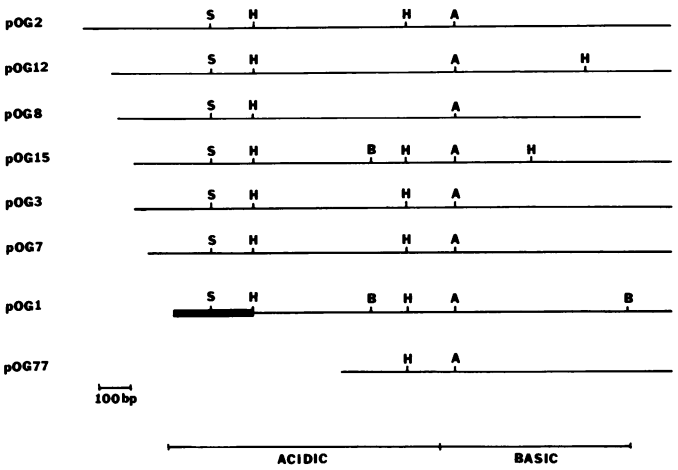


FIG. 1. Restriction enzyme maps of oat globulin cDNA clones. Shown are the restriction maps of the inserts of pOG1, the longest clone from the antibody screening, and six clones from the subsequent screening by nucleic acid hybridization (pOG2, pOG12, pOG8, pOG15, pOG3, and pOG7). The 5′ terminal restriction fragment from pOG1 used as probe in this hybridization is denoted by a heavier line. Included for comparison is the restriction enzyme map of pOG77, the partial oat globulin cDNA clone isolated previously (30). Restriction enzymes: A, *Ava*II; B, *Bam*HI; H, *Hin*cII; S, *Sca*I.

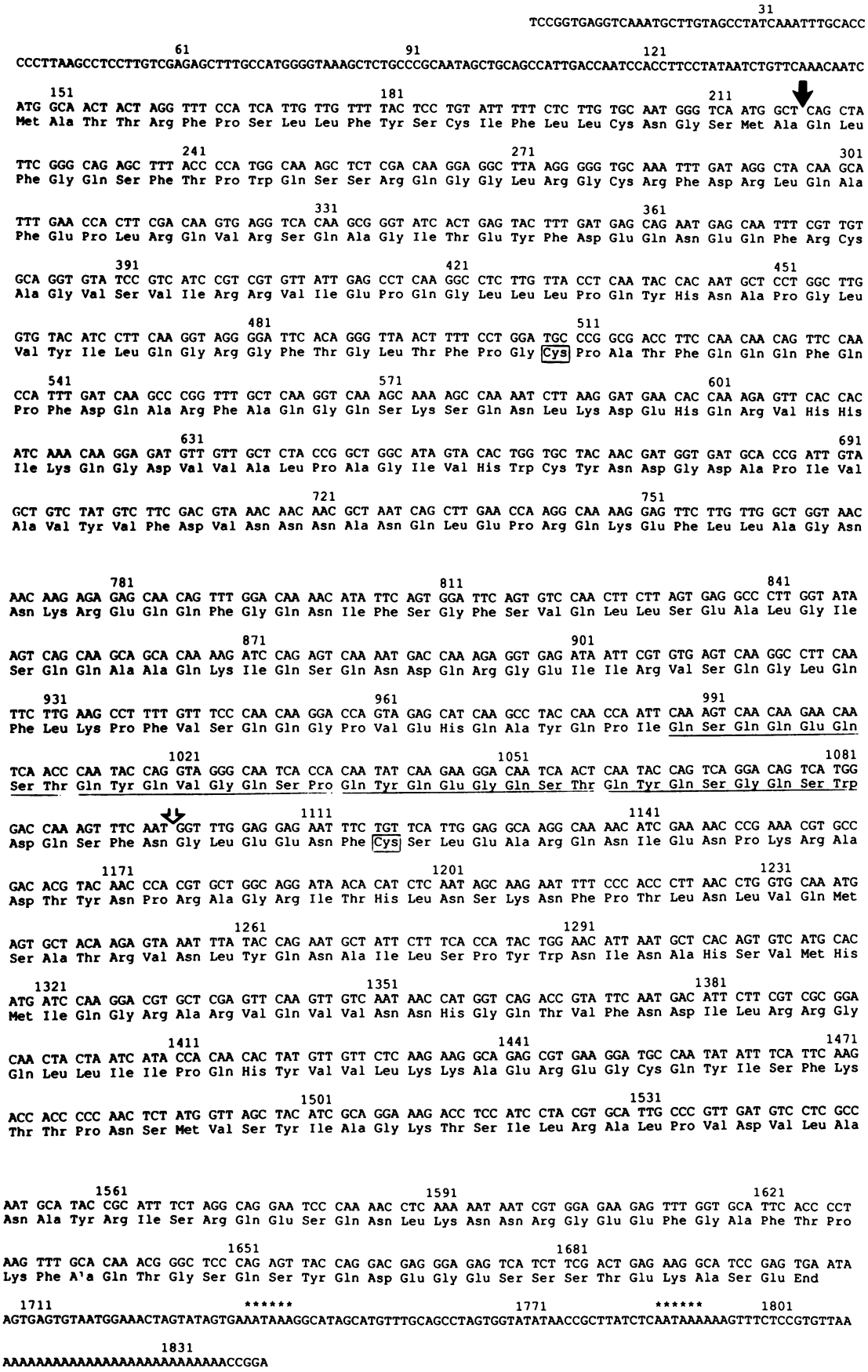


FIG. 2. Nucleotide and deduced amino acid sequence of oat globulin cDNA clone pOG2. The solid arrow indicates the presumed site of cleavage of the signal peptide. The open arrow indicates the site of proteolytic processing to yield the acidic and basic polypeptides. The boxed cysteine residues are those believed to be involved in interchain disulfide bonding. The four octapeptide repeats near the C terminus of the acidic polypeptide are underlined. The asterisks indicate the AATAAA poly(A) addition sequences.

allowed us to identify the site of proteolytic cleavage into acidic and basic polypeptides between the asparagine at position 317 and the glycine at position 318 (assuming a single cleavage site) (Fig. 2). Thus, the pOG2 cDNA predicts a signal peptide of 24 amino acids followed by an acidic polypeptide of 293 amino acids with an $M_r$ of 33,100 and a basic polypeptide of 201 amino acids with an $M_r$ of 22,700. These mol wt predictions are in close agreement with estimates made earlier on the basis of the mobility of isolated acidic and basic polypeptides in polyacrylamide gels (29).

The pOG2 insert sequence predicts that the amidated amino acids, asparagine and glutamine, make up 20.4% of the total amino acid content, consistent with the role of seed storage proteins as nitrogen stores. The oat globulin is characteristically deficient in the sulfur-containing amino acids, and the pOG2 sequence predicts a content of cysteine plus methionine of only 1.8%. The pOG2 insert sequence contains no instances of the tripeptide Asn-×-Ser/Thr, which serves as a signal for N-linked glycosylation. The absence of such sequences is consistent with earlier conclusions that the oat globulin is not glycosylated (1, 21).

Downstream of the stop codon, the pOG2 insert contains two AATAAA poly(A)-addition sequences, a common feature of plant mRNAs. Figure 2 also shows that the pOG2 cDNA encodes four imperfect repeats of an eight amino acid sequence near the C terminus of the acidic polypeptide. Sequence analysis of four other oat globulin cDNA clones isolated by the two screening procedures revealed that they also contain repeats of this eight amino acid sequence (Fig. 3). Interestingly, whereas pOG1, pOG2, and pOG21 contain four repeats of this sequence, clones pOG4 and pOG5 have five repeats. The consensus of this eight amino acid sequence from these five cDNAs is Gln-Tyr-Gln-(Val/Glu)-Gly-Gln-Ser-Thr.

**Homologies between Oat Globulin cDNA Clones.** The oat globulin cDNA clones we have isolated have similar but not always identical maps for four selected restriction endonucleases (Fig. 1). More extensive mapping showed that no two clones have the same restriction map, except for pOG2 and pOG77 in their region of overlap (data not shown).

We immobilized insert DNA from eight cDNA clones onto nitrocellulose filters and probed these filters with four contiguous restriction fragments that span the pOG2 sequence. This 'dot blot' hybridization analysis showed that each of these cDNA clones is highly homologous to pOG2 along its entire length (data not shown). Cross-hybridizations were at a criterion of $T_m$-10. Limited nucleic acid sequence comparisons showed that these cDNA clones are between 89 and 94% identical.

**Estimation of Oat Globulin Gene Number.** As described in "Materials and Methods," we digested oat genomic DNA with *Eco*RI, separated the products by electrophoresis, and transferred the DNA to nitrocellulose. Oat globulin gene sequences were then detected by hybridization with isolated pOG2 insert that was labeled with $^{32}P$ by nick translation (Fig. 4). By comparison

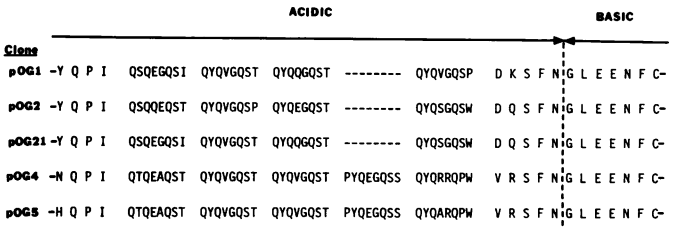


FIG. 3. Sequences of octapeptide repeats deduced from five oat globulin cDNA clones. Shown are the residues at the C terminus of the acidic polypeptide and the first seven residues at the N terminus of the basic polypeptide. The dashed line indicates the site of cleavage between acidic and basic polypeptides.

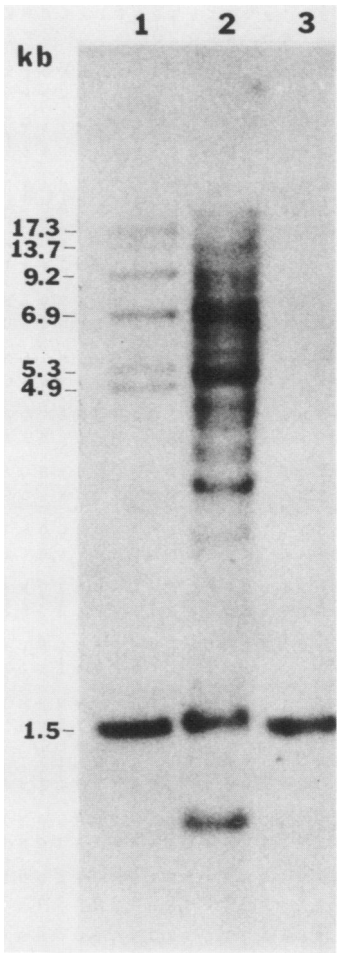


FIG. 4. Oat genomic Southern blot. Oat genomic DNA, 8.8 μg, equivalent to $4 \times 10^6$ haploid genomes, was digested with *Eco*RI, fractionated by agarose gel electrophoresis, and transferred to nitrocellulose as described in "Materials and Methods." Hybridization was with the 1840-bp cDNA insert from pOG2 labeled by nick translation with [$^{32}P$]dATP. Lane 1, end-labeled λ-DNA size markers; lane 2, *Eco*RI-cut oat genomic DNA; lane 3, 0.5 gene copy reconstruction (5.5 pg of *Eco*RI-cut pOG2 DNA).

with a reconstruction lane, we estimate there to be 6 to 8 copies of the oat globulin gene per haploid genome; since cultivated oats is hexaploid, there may be 35 to 50 globulin genes per nucleus.

**Comparison of the Oat Globulin with Other 11S Globulins.** Previous work has shown that the overall structure of the oat 12S globulin is quite similar to that of the 11S storage globulins of legumes and other dicots (5, 7, 29). In comparing the amino acid sequence deduced from oat globulin cDNA clone pOG2 with some of these other storage globulins, we find 31% sequence identity with soybean glycinin (Fig. 5, left) and 38% with pea legumin, but 70% sequence identity with the rice globulin (glutelin) (Fig. 5, right) (Table I). Comparisons of the predicted hydropathy of these amino acid sequences confirm the closer relationship between the proteins of oats and rice than between the proteins of oats and soybean. The hydropathy profiles of soybean glycinin (Fig. 6A) and the oat globulin (Fig. 6B) are quite similar, but the profiles of the oat globulin (Fig. 6B) and the rice globulin (glutelin) (Fig. 6C) are nearly identical. The most striking difference between the soybean profile and those of oats and rice is the extremely hydrophilic character of the C-terminal residues of the acidic polypeptide of soybean glycinin,

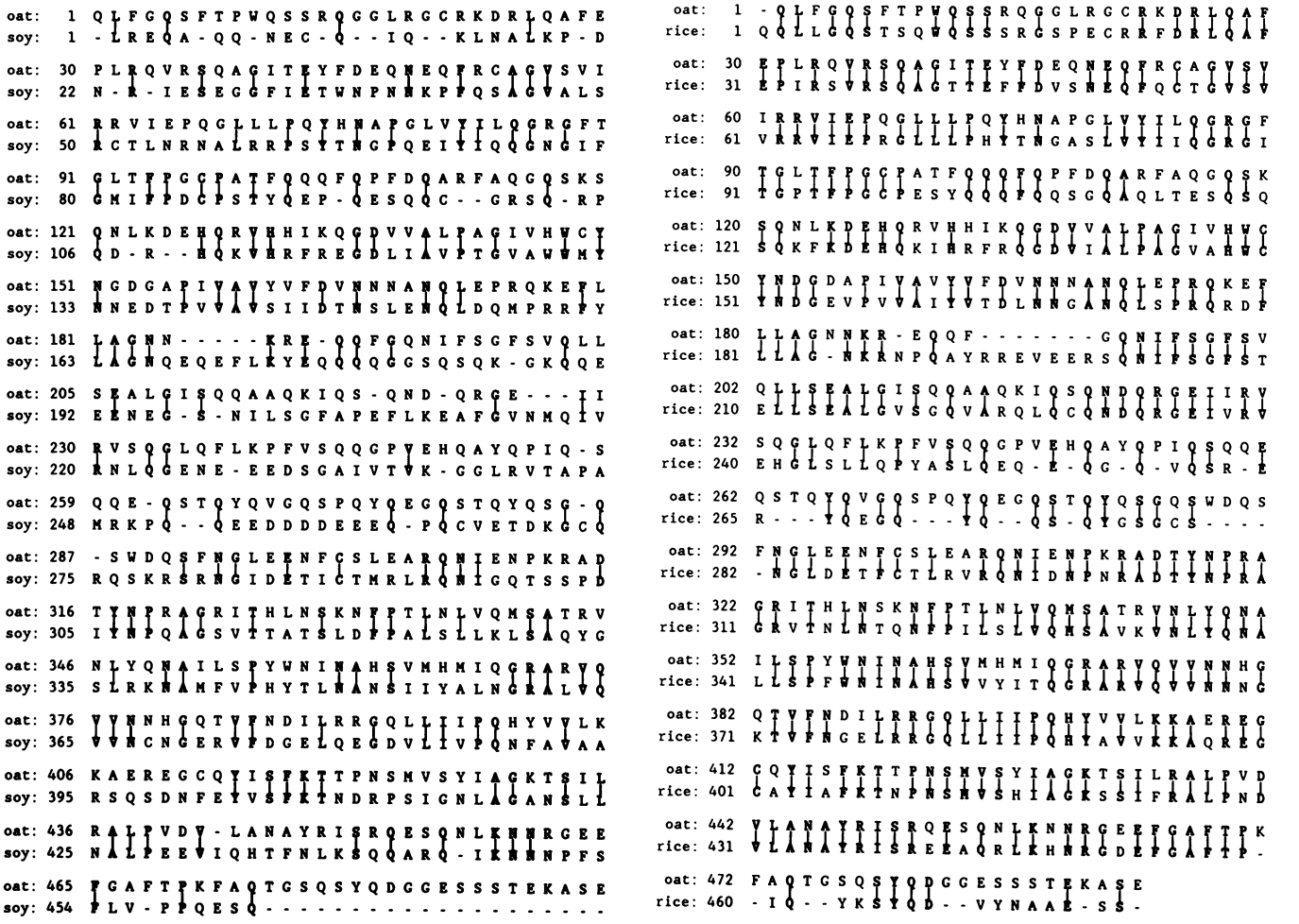

FIG. 5. Amino acid sequence comparison between oat globulin and soybean glycinin (left) and between oat globulin and rice globulin (glutelin) (right). The oat globulin sequence is from cDNA clone pOG2, the soybean glycinin sequence is from the $Gy_2$ gene (18), and the rice globulin (glutelin) sequence is from cDNA clone pREE61 (27). The sequences were aligned using the Microgenie computer program which inserts gaps to maximize homology.

Table I. *Percent Amino Acid Homology between 11-12S Seed Storage Globulins of Different Plants*

Deduced amino acid sequences were compared using the Microgenie sequence analysis program. Amino acid sequences were derived as follows: oat from cDNA clone pOG2; rice from cDNA clone pREE61 (27); pea from gene *1egA* (15); soy from gene $Gy_2$ (18).

| | Oat | Rice | Pea |
|---|---|---|---|
| Rice | 70.4 | | |
| Pea | 38.1 | 39.9 | |
| Soy | 31.2 | 42.8 | 67.8 |

the hypervariable region, which consists of stretches of acidic residues (Fig. 6A). The hypervariable regions of the oat and rice globulins are much less highly charged, consisting largely of neutral amino acids.

## DISCUSSION

Oat globulin is a collective term referring to a group of closely related saline-soluble proteins that are deposited in large amounts in the endosperm of the developing oat seed and serve as stores of nitrogen, sulfur, and carbon. Biochemical studies established that the oat globulin has many structural features in common with the well-characterized storage globulins from the cotyledons of legumes and other dicots (5, 7, 29). Like those proteins, the oat globulin is synthesized as a precursor polypeptide with an N-terminal signal peptide. After removal of the signal peptide, the protein is proteolytically processed into a larger polypeptide with an acidic isoelectric point and a smaller polypeptide with a basic isoelectric point; the two chains remain linked by a disulfide bond.

A comparison between the amino acid sequence predicted by the cDNA clone pOG2 and that specified by the soybean glycinin gene $Gy_2$ reveals 31% sequence identity (Fig. 5, left). The homology is 28% between acidic polypeptides but 37% between basic polypeptides. This difference may reflect a somewhat more highly conserved secondary structure of the basic polypeptides, which are believed to be tightly folded in the interior of the molecule and surrounded by the more hydrophilic acidic polypeptides, shielding them from the solvent (22). Furthermore, the regions of homology in the less highly conserved acidic polypeptides are centered on proline residues whose positions are strictly maintained in the oat and soybean sequences (Fig. 5, left). This arrangement suggests that regions containing these conserved proline residues are important in the folding of the acidic polypeptide around the basic polypeptide and that considerable sequence divergence is tolerated in other regions of the acidic chain.

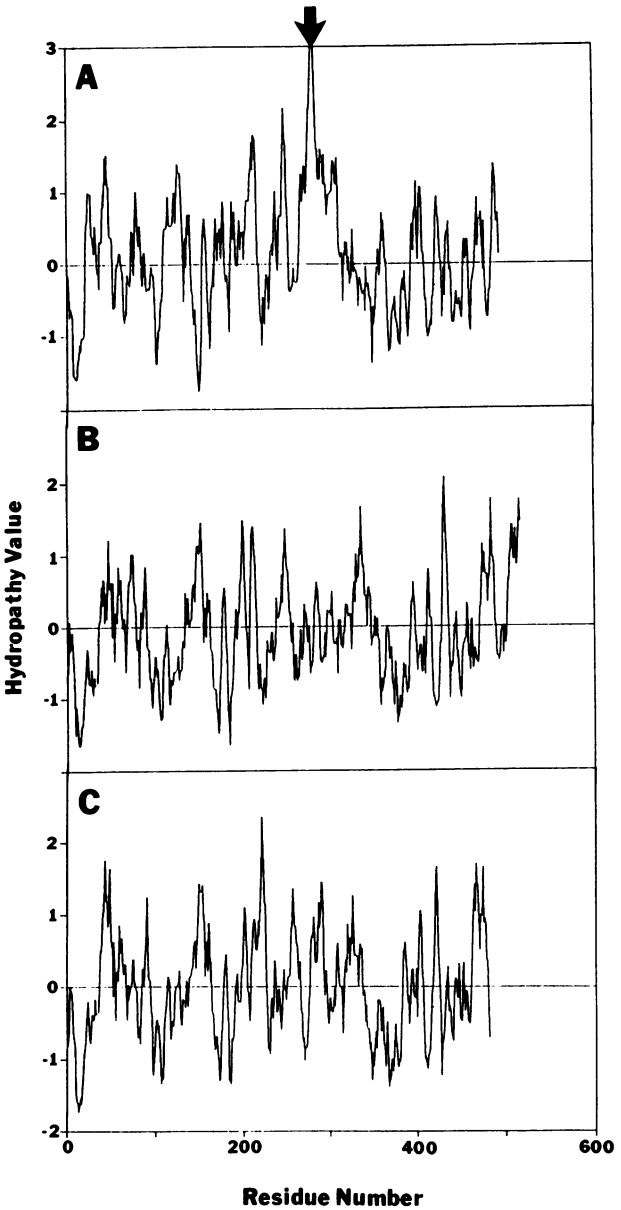


FIG. 6. Comparison of hydropathy of oat globulin with soybean glycinin and rice globulin (glutelin). Hydropathy was determined by the algorithm of Hopp and Woods (11) using sequence analysis programs on the VAX computer at the University of Wisconsin, Madison (8). The soybean sequence (A) is from gene $Gy_2$ (18), the oat sequence (B) is from cDNA clone pOG2, and the rice sequence (C) is from cDNA clone pREE61 (27). The arrow indicates the position of the hypervariable region, which falls roughly between residues 270 and 310.

Interestingly, of the six cysteine residues in the pOG2 sequence and seven in the $Gy_2$ sequence, only two lie in identical positions, one in the acidic polypeptide and one in the basic polypeptide. These two cysteine residues have been identified as those involved in the formation of the single interchain disulfide bond in the glycinins (26). It seems likely that the disulfide linkage in the oat globulin also occurs between these two cysteine residues.

The partial cDNA clone pOG77, which has the same sequence as the full-length clone pOG2, has previously been used in RNA blot hybridizations to quantitate oat globulin message during seed development (14). In this analysis, globulin mRNA was detectable by 3 DPA, increased to a maximum at 15 DPA, and then decreased in amount as the seeds reached maturity. The peak of globulin mRNA abundance at 15 DPA coincided with the peak of globulin protein synthesis as assayed by incorporation of [$^{35}$S]sulfate into immunoprecipitable polypeptides.

Our analysis showed that the amino acid sequence specified by pOG2 is 70% identical to that of the rice globulin (glutelin) reported by Takaiwa *et al.* (27) (Fig. 5, right). This high sequence similarity is perhaps not surprising in light of the cross-reactivity of oat globulin antibodies with the rice protein (23) and is consistent with the closer evolutionary relationship between oats and rice than between oats and the legumes.

The oat globulin and the rice globulin (glutelin) are highly homologous but with three differences. First, the rice sequence contains a seven-amino acid insertion (Arg-Arg-Glu-Val-Glu-Glu-Arg) near the middle of the acidic polypeptide (Fig. 5, right). Second, the oat protein contains several additional residues at the C terminus of the basic polypeptide. Third, the oat globulin amino acid sequence differs from the rice protein, and all other storage globulins reported to date, in having glutamine-rich repeats of eight amino acids near the C terminus of the acidic polypeptide. Some oat globulin polypeptides have four repeats of this highly conserved octapeptide and others five (Fig. 3). The part of the molecule in which these repeats occur has been termed the hypervariable region (2). In soybean glycinins (20) and pea legumins (16), the hypervariable region consists of stretches of acidic residues of different lengths. In the rice protein (27), this region is rich in glutamine (10 out of 30 residues), but these amino acids are not organized into repeats as in the oat globulin.

Plietz and Damashun (22) proposed on the basis of extensive physical measurements that the hypervariable region resides at the surface of the globulin subunit molecule in contact with the solvent. If this is so, the hydropathy of the amino acids in the hypervariable region would disproportionately influence the solubility properties of the protein. Thus, the less hydrophilic hypervariable region of the oat globulin compared with soybean glycinin (*cf.* Fig. 6, A and B) may largely explain the higher salt concentration required for solubility of the oat protein than for glycinin and other legume proteins (1.0 M NaCl *versus* 0.4 M NaCl).

Aside from their probable effect on solubility, we cannot suggest any possible functional significance for the unique, highly conserved octapeptide repeats in the oat globulin. Our observation that some cDNA clones contain four repeats and others five, however, is consistent with the size heterogeneity of purified oat globulin acidic polypeptides in polyacrylamide gels (29). This variability in the number of repeats may also indicate that unequal crossing over events have occurred at the hypervariable region of the oat globulin genes during the course of evolution.

Storage globulins are encoded by small multigene families in the legumes. There are, for example, six genes for soybean glycinin (RL Fischer, TL Sims, GN Drews, RB Goldberg, personal communication) and eight genes for pea legumin (9) in the cultivars chosen for study. We estimate from the genomic blot shown in Figure 4 that the gene family for the oat 12S globulin is about the same size, consisting of 6 to 8 genes. In contrast, the gene families coding for prolamine storage proteins in monocots may be much larger (25). It is interesting to speculate whether the prolamines of oats, the avenins, are likewise encoded by large numbers of genes that are not expressed at high levels. Perhaps the regulatory sequences that cause high levels of prolamine gene transcription in the endosperm of the developing cereal seed have been altered in oats leading to the predominant expression of the globulin genes. Alternatively, amplification of genes encoding prolamine storage proteins may not have occurred during the evolution of wild oats and the derivation of cultivated varieties of oats, as has been speculated for most other cereals, leaving oats with the more 'primitive' globulin genes to encode the majority of its storage protein. Analysis of gene sequences coding

for both globulins and prolamines will be the first step in answering these questions.

*Acknowledgments*—We thank J. S. Lindell and C. Logan for their help in preparing figures for this paper. We also wish to express our gratitude to Dr. N. C. Nielsen for his help in preparing the hydropathy plots.

## LITERATURE CITED

1. ADELI K, I ALTOSAAR 1983 Role of endoplasmic reticulum in biosynthesis of oat globulin precursors. Plant Physiol 73: 949–955
2. ARGOS P, SVL NARAYANA, NC NIELSEN 1985 Structural similarity between legumin and vicilin storage proteins from legumes. EMBO J 4: 1111–1117
3. AVIV H, P LEDER 1972 Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc Natl Acad Sci USA 69: 1408–1412
4. BARBARA DJ, EE KAWATA, PP UENG, RM LISTER, BA LARKINS 1987 Production of cDNA clones from the MAV isolate of barley yellow dwarf virus. J Gen Virol 68: 2419–2427
5. BRINEGAR AC, DM PETERSON 1982 Separation and characterization of oat globulin polypeptides. Arch Biochem Biophys 219: 71–79
6. BRINEGAR AC, DM PETERSON 1982 Synthesis of oat globulin precursors: analogy to legume 11S storage protein synthesis. Plant Physiol 70: 1767–1769
7. BURGESS SR, PR SHEWRY, GJ MATLASHEWSKI, I ALTOSAAR, BJ MIFLIN 1983 Characteristics of oat (*Avena sativa* L.) seed globulins. J Exp Bot 34: 1320–1332
8. DEVEREUX J, P HAEBERLI, O SMITHIES 1984 A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12: 387–395
9. DOMONEY C, R CASEY 1985 Measurement of gene numbers for seed storage proteins in *Pisum*. Nucleic Acids Res 13: 687–699
10. HALL TC, Y MA, BU BUCHBINDER, JW PYNE, SM SUN, FA BLISS 1978 Messenger RNA for G1 protein of French bean seeds: cell-free translation and product characterization. Proc Natl Acad Sci USA 75: 3196–3200
11. HOPP TP, KR WOODS 1981 Prediction of protein antigenic determinants from amino acid sequences. Proc Natl Acad Sci USA 78: 3824–3828
12. HUYNH TV, RA YOUNG, RW DAVIS 1985 Construction and screening of cDNA libraries in λgt10 and λgt11. *In* DM Glover, ed, DNA Cloning: A Practical Approach, Vol 1. IRL Press, Washington, DC, pp 49–78
13. LARKINS BA 1987 Modification of proteins encoded by seed storage protein genes. *In* G Bruening, J Harada, T Kosuge, A Hollaender, eds, Tailoring Genes for Crop Improvement: An Agricultural Perspective, Plenum Press, New York, pp 163–167
14. LUTHE DS 1987 Storage protein synthesis during oat (*Avena sativa* L.) seed development. Plant Physiol 84: 337–340
15. LYCETT GW, RRD CROY, AH SHIRSAT, D BOULTER 1984 The complete nucleotide sequence of a legumin gene from pea (*Pisum sativum* L.). Nucleic Acids Res 12: 4493–4505
16. LYCETT GW, AJ DELAUNEY, W ZHAO, JA GATEHOUSE, RRD CROY, D BOULTER 1984 Two cDNA clones coding for the legumin protein of *Pisum sativum* L. contain sequence repeats. Plant Mol Biol 3: 91–96
17. MANIATIS T, EF FRITSCH, J SAMBROOK 1982 Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
18. MARCO YA, VH THANH, NE TUMER, BJ SCALLON, NC NIELSEN 1984 Cloning and structural analysis of DNA encoding an $A_2B_{1a}$ subunit of glycinin. J Biol Chem 259: 13436–13441
19. MEINKOTH J, G WAHL 1984 Hybridization of nucleic acids immobilized on solid supports. Anal Biochem 139: 267–284
20. NIELSEN NC 1985 The structure and complexity of the 11S polypeptides in soybeans. J Am Oil Chem Soc 62: 1680–1686
21. PETERSON DM 1978 Subunit structure and composition of oat seed globulin. Plant Physiol 62: 506–509
22. PLIETZ P, G DAMASCHUN 1986 The structure of the 11S seed globulins from various plant species: comparative investigations by physical methods. Studia Biophys 116: 153–173
23. ROBERT LS, C NOZZOLILLO, I ALTOSAAR 1985 Homology between rice glutelin and oat 12S globulin. Biochim Biophys Acta 829: 19–26
24. SANGER F, S NICKLEN, AR COULSON 1977 DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467
25. SHOTWELL MA, BA LARKINS 1988 The biochemistry and molecular biology of seed storage proteins. *In* A Marcus, ed, The Biochemistry of Plants, A Comprehensive Treatise, Vol 15. Academic Press, New York. In press
26. STASWICK PE, MA HERMODSON, NC NIELSEN 1984 Identification of the cystines which link the acidic and basic components of the glycinin subunits. J Biol Chem 259: 13431–13435
27. TAKAIWA F, S KIKUCHI, K OONO 1986 The structure of the rice storage protein glutelin deduced from cDNA. FEBS Lett 206: 33–35
28. VON HEIJNE G 1983 Patterns of amino acids near signal-sequence cleavage sites. Eur J Biochem 133: 17–21
29. WALBURG G, BA LARKINS 1983 Oat seed globulin: subunit characterization and demonstration of its synthesis as a precursor. Plant Physiol 72: 161–165
30. WALBURG G, BA LARKINS 1986 Isolation and characterization of cDNAs encoding oat 12S globulin mRNAs. Plant Mol Biol 6: 161–169