Plant Physiol. (1989) 91, 636–643
0032-0889/89/91/0636/08/$01.00/0

Received for publication February 27, 1989
and in revised form May 19, 1989

Molecular Characterization of the Major Maize Embryo Globulin Encoded by the *Glb1* Gene[1]

Faith C. Belanger and Alan L. Kriz*

Department of Agronomy, University of Illinois, Urbana, Illinois, 61801

ABSTRACT

One of the most abundant proteins in maize (*Zea mays* L.) embryos is the molecular weight 63,000 globulin encoded by the *Glb1* gene. To obtain DNA clones corresponding to *Glb1*, a cDNA library corresponding to RNA from developing maize embryos was constructed in a lambda expression vector and screened with antibodies specific for *Glb1*-encoded proteins. Here we report the complete nucleotide sequence, as determined from two overlapping clones, of pcGlb1S, a 2009 base pair clone containing the entire translated region of *Glb1*. The deduced amino acid sequence of pcGlb1S shows similarities to 7S-type seed storage proteins of wheat and legumes. Southern blot analysis of maize DNA confirms previous genetic studies which had indicated the presence of a single copy of *Glb1* per haploid genome. Northern blot analysis indicates that *Glb1* transcripts are present throughout most of embryo development and that expression of this gene is limited to seed tissues. Embryos homozygous for a *Glb1* null allele, in which *Glb1*-encoded proteins are not detectable, contain low levels of *Glb1* transcripts which are a different size from those encoded by functional alleles. This suggests that the defect in the null allele is at the level of gene transcription or RNA processing.

Maize embryos contain large amounts of saline-soluble, water-insoluble proteins called globulins. The major globulin component, a mol wt 63,000 protein designated GLB1 (formerly PROT) (10), is one of the most abundant proteins in mature embryos. GLB1 has no known enzymatic function and is currently believed to serve as a storage protein (8). Genetic analysis of GLB1 variants indicated that this protein is encoded by a single gene, *Glb1* (for *Globulin-1*), on the long arm of chromosome 1 (22). Several alleles of this gene, including a CRM^- null,[2] have been described (17, 22). Allelic polymorphism of *Glb1*-encoded proteins is observed as a function of mobility in SDS-PAGE. The three most commonly occuring *Glb1* alleles have the designations *L*, *I*, and *S*, for *Large, Intermediate,* and *Small* proteins, respectively. By convention, the protein product of each allele is indicated in the upper case (*e.g. Glb1-S* encodes GLB1-S). Pulse-chase radiolabeling and *in vitro* translation experiments have indicated that at least three protein-processing steps occur in the formation of the mature protein from the primary translation product (10). The final processing step is controlled by the unlinked gene *Mep* (chromosome 5, near *Pr*), and embryos homozygous for the recessive *mep* allele accumulate the processing intermediate GLB1′ (10, 22). We have recently reported on additional characterization of maize embryo globulins (8).

Here we report the isolation and characterization of a full-length[3] cDNA clone, pcGlb1S,[4] corresponding to the S allele of the *Glb1* gene. Using this clone as a probe, we have determined that expression of the gene appears to be limited to seed tissues and that embryos homozygous for the null allele contain small amounts of transcript corresponding to *Glb1*.

MATERIALS AND METHODS

Materials

Embryos homozygous for the *Glb1-L* and *Glb1-S* alleles were obtained from field-grown plants of the maize (*Zea mays* L.) inbred lines W64A and Va26, respectively, as previously described (8). The *Glb1-O* null allele was originally identified in a Black Beauty popcorn line (22). Nitrocellulose and Magnagraph nylon membranes were obtained from Micron Separations, Inc. (Westboro, MA). Random priming kits, a cDNA synthesis kit, *Eco*RI, T4 DNA ligase, S1 nuclease, and protein and RNA mol wt standards were purchased from Bethesda Research Laboratories (Gaithersburg, MD). LambdaZAP vector arms, Gigapack packaging extracts, and exonuclease III/mung bean nuclease deletion kits were from Stratagene (La Jolla, CA). [α-^{32}P]-labeled dATP (3000 Ci/mmol) was obtained from New England Nuclear (Boston, MA). Oligo(dT)-cellulose (type 7) was from Pharmacia, Inc. (Piscataway, NJ).

Preparation of GLB1-Specific Antiserum

GLB1-S was fractionated from whole globulin by cryoprecipitation (8) and stored at 4°C as a suspension in distilled water at a concentration of 2 mg/mL. A volume of the suspension equivalent to 600 μg protein was emulsified with an equal volume of Freund's complete adjuvant and injected

[1] This contribution from the University of Illinois Agricultural Experiment Station was supported by grants from The Standard Oil Company of Ohio, a wholly owned subsidiary of BP America, Inc., and the U.S. Department of Agriculture Competitive Research Grants Office (No. 88-37262-3427) to A. L. K.

[2] Abbreviations: CRM, cross-reacting material; DAP, days after pollination; ds, double-stranded.

[3] The term full-length as used here indicates that the cDNA clone contains the entire translated region of the gene.

[4] The GenBank accession number for pcG1b1S is M24845.

subcutaneously into multiple sites of a 10-week old New Zealand White rabbit. A second inoculation of 300 μg GLB1-S in Freund's incomplete adjuvant was administered 2 weeks later. Blood was collected from the marginal ear vein 3 weeks after the secondary inoculation and serum was prepared from the blood by standard methods.

Protein Extraction and Immunoblot Analysis

For analysis of total proteins from imbibed mature embryos (8), kernels were soaked overnight in water and embryos excised from the softened kernels with the aid of a sharpened spatula. For developmental analyses, field-grown ears were transported to the laboratory and the crowns of the kernels were sliced off with a sharp knife or scalpel blade. Embryos were removed from the kernels with a small spatula, immediately frozen in liquid nitrogen, and stored at −70°C until needed. Individual mature embryos were macerated in 1 mL SDS sample buffer (2% [w/v] SDS, 50 mM Tris [pH 6.8] 5% [v/v] 2-mercaptoethanol, 10% [v/v] glycerol) at a ratio of 50 mg/mL (fresh weight) directly in 1.5 mL microcentrifuge tubes with the aid of a fitted pellet pestle (Kontes Scientific Glassware, Morton Grove, IL). Extracts were heated to 100°C for 2 min, centrifuged at 13,500*g* for 5 min, and immediately subjected to SDS-PAGE as previously described (8). Five μL of each sample was applied to a 12.5% polyacrylamide gel and subsequently electroblotted onto nitrocellulose by using a semidry blotting apparatus (PolyBlot; American Bionetics, Hayward, CA) according to the manufacturer's specifications. Immunoblot analysis, in which GLB1-specific antiserum was used at 1:2000 dilution, was performed as previously described for analysis of maize zeins (13).

Nucleic Acid Isolation and Gel Blot Analysis

Total RNA was isolated from frozen tissue by using the guanidine-HCl method described by Cox (3). For Northern blot analysis, 10 μg of total RNA was subjected to electrophoresis in formaldehyde agarose gels and transferred to either nitrocellulose or nylon (Magnagraph) membranes as described by Selden (23). Hybridization conditions were as previously described (9).

Isolation of maize nuclear DNA from unfertilized ears and Southern blot analysis were performed as previously described (9) except that Magnagraph membranes were used as the transfer medium. Samples for gene copy-number reconstructions contained *Eco*RI-cleaved DNA from a genomic clone corresponding to *Glb1-L* (our unpublished data). For both RNA and DNA blots, the transferred nucleic acids were UV-linked to the membrane by using a Stratalinker 1800 apparatus (Stratagene).

For use in hybridizations, the cDNA insert was isolated from the plasmid clone by *Eco*RI digestion, separation on a 1% agarose gel, and binding of the fragment to NA-45 paper (Schleicher and Schuell, Keene, NH) as recommended by the manufacturer. The isolated insert was labeled with [α-^{32}P] dATP by using a commercial random priming kit (BRL).

Construction and Screening of an Embryo-Specific cDNA Library

To obtain cDNA clones corresponding to *Glb1*, a library of cDNA sequences was prepared in a bacteriophage λ-expression vector. Total RNA was prepared from 27 DAP embryos of the maize inbred line Va26 *(Glb1-S/S)*. Polyadenylated RNA was fractionated from total RNA by oligo(dT)-cellulose chromatography (1). First and second strand cDNA synthesis was performed by using a commercial cDNA synthesis kit based on the RNase H procedure of Gubler and Hoffman (4). The ds cDNA was protected from *Eco*RI cleavage by treatment with *Eco*RI methylase, and *Eco*RI oligonucleotide linkers (both from New England Biolabs, Beverly, MA) were added to the ends of the ds cDNA with T4 DNA ligase. Subsequent to linker cleavage with *Eco*RI, excess linkers were removed by passage through a NACS Prepac column (BRL) as described by the manufacturer. The ds cDNA was ligated to the arms of the expression vector LambdaZAP, packaged by using the Gigapack system, and the resultant recombinant phage was used to infect the *Escherichia coli* host BB4. The primary library consisted of approximately 500,000 clones which were obtained from an estimated 0.1 μg of ds cDNA.

The initial plating of the library was screened by preparing nitrocellulose replicas of the plates by the plaque-lift technique as described by Huynh *et al.* (7). Filters were agitated in a solution of TTBS/1% (w/v) nonfat dry milk (TTBS: 50 mM Tris [pH 7.4] 0.15% [v/v] Tween-20, 140 mM NaCl) for 1 h, then transferred to a solution of GLB1-specific antibody (1:500 dilution in TTBS) and agitated overnight. Immunoreactive clones were detected as for the protein immunoblots. Plaques were picked, plated, and subjected to three additional rounds of screening. Of 10 initial plaque picks, all continued to yield positive reactions after the final screen. The cDNA inserts were excised from LambdaZAP as recombinant pBluescript SK(-) plasmids according to the manufacturer's protocols. Rescreening of the library with a radiolabeled probe was essentially as described by Huynh *et al.* (7).

DNA Sequencing

For nucleotide sequence analysis, the cDNA inserts in pBluescript were subcloned into M13 mp18 and mp19 (30) to obtain inserts in opposite orientations. Overlapping unidirectional deletions corresponding to either strand were prepared from the appropriate M13 clone RF by using a commercial exonuclease III/mung bean nuclease deletion kit (Stratagene). The deletions were sized by using a rapid S1 nuclease procedure (16). Dideoxynucleotide sequencing (21) of single-stranded templates with a modified T7 DNA polymerase (Sequenase) was performed by using a commercial sequencing kit (United States Biochemical Corp., Cleveland, OH). The deoxyguanine triphosphate (dGTP) analog 7-deaza dGTP was used to resolve GC compressions. Analysis of DNA sequences was performed on an IBM PC AT with either IBI/Pustell Sequence Analysis software (International Biotechnologies Inc., New Haven, CT) or DNAStar programs (DNAStar, Inc., Madison, WI).

RESULTS

Characterization of GLB1-Specific Antibodies

The specificity of GLB1-specific polyclonal antibodies was determined in an immunoblot of proteins from individual embryos of varying allelic composition (Fig. 1). Lanes 1 and 2 of Figure 1 are from embryos of the inbred lines W64A and Va26 which are homozygous for the *Glb1-L* and *Glb1-S* alleles, respectively. The major proteins recognized by the antiserum are GLB1-L and GLB1-S and their respective immediate precursors GLB1′-L and GLB1′-S (10). There were, however, several other polypeptide bands of lower mol wt which also reacted with the antiserum. Since these are absent in embryos homozygous for the *Glb1-O* null allele (Fig. 1, lane 3), we inferred that these low mol wt polypeptides were derived from Glb1-encoded proteins. Such polypeptides were present in *Glb1-L/O* and *Glb1-S/O* embryos (Fig. 1, lanes 4 and 5, respectively), but at levels lower than those observed for *Glb1-L/L* or *S/S* embryos. This is apparently due to the presence of only a single functional *Glb1* allele in the heterozygotes. We therefore concluded that the antiserum was specific for *Glb1*-encoded proteins and was appropriate for use in screening a cDNA expression library.

Isolation and Characterization of pcGlb1S

In preliminary experiments we observed that embryos of the maize inbred line Va26 contained large amounts of GLB1-S protein. We therefore chose this material as a source of RNA for production of a cDNA library as described above. A screen of 250,000 clones yielded 10 immunoreactive clones with inserts ranging in size from 700 to 1800 bp. Southern blot analysis revealed that all of these clones hybridized with a radiolabeled probe prepared from the longest insert.

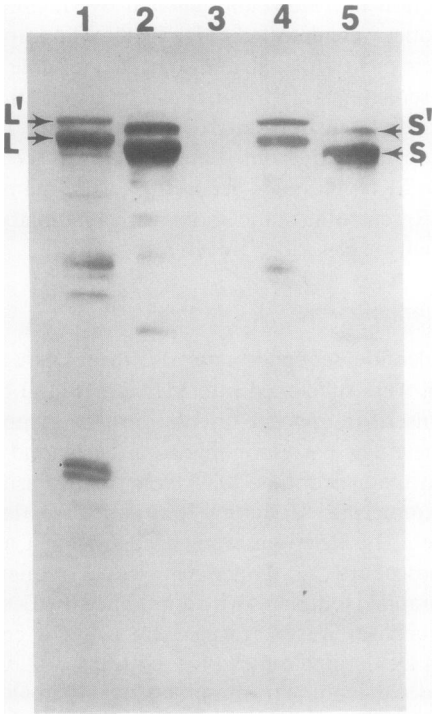

Figure 1. Specificty of GLB1 antiserum. Total protein extracts from embryos of varying *Glb1* genotypes were subjected to immunoblot analysis as described in "Materials and Methods." Lane 1, *Glb1-L/L* (W64A); lane 2, *Glb1-S/S* (Va26); lane 3, *Glb1-O/O;* lane 4, *Glb1-L/O* (W64A x *Glb1-O/O);* lane 5, *Glb1-S/O* (Va 26 x *Glb1-O/O).*

The 1800 bp clone, designated pcA1, was chosen for further characterization and subjected to nucleotide sequence analysis. A 300 bp *Eco*RI/*Bst*EII restriction fragment from the 5′ region of this clone was used as a radiolabeled probe to rescreen the cDNA library for a full length cDNA clone, which we have designated pcGlb1S. The 5′ region of the full length clone was sequenced to provide 55 nucleotides of overlap with the original 1800 bp clone pcAl. The nucleotide sequences of the two clones were identical in this region of overlap. Because of the high G + C content (66%) of these clones, it was necessary to use the dGTP analog deaza dGTP to minimize GC compressions and obtain unambiguous sequence data. The nucleotide sequence was determined from both strands for the appropriate regions of each clone.

The combined sequence of the two cDNA clones and the deduced amino acid sequence are shown in Figure 2. The sequence surrounding the presumed initiator methionine corresponds to a relaxed consensus (27) of Kozak's rule, ANNATGG (11, 14). An AT-rich region between nucleotide positions +1937 and +1949 is considered to be a potential polyadenylation signal and contains the consensus sequence AATAAA (18). The nucleotide sequence predicts a polypeptide of 573 amino acids with a calculated mol wt of 65,025 D. From SDS-PAGE analysis of immunoprecipitated *in vitro* translation products, the mol wt for the primary translation product of *Glb1-S* has been estimated as 67,600 D (10). The mol wt calculated from the deduced amino acid sequence is 2,575 D lower than the experimentally determined estimate. We do not consider this discrepancy to be significant since it is not uncommon for proteins to exhibit anomalous migration in SDS-PAGE (2). The presence of a region of deduced amino acid sequence *(underlined* in Fig. 2) identical to the N-terminal region of experimentally determined GLB1-S protein sequence (8) confirms that these cDNA clones correspond to *Glb1*. Previous studies demonstrated that GLB1′ and GLB1 differ at only one terminus (10). The present sequence analysis indicates that the final processing of GLB1′ to GLB1, controlled by the *Mep* gene (22), is localized to the N-terminal portion of GLB1′. In GLB1′-S, this cleavage probably occurs between the aspartate and glutamate residues at positions 86 and 87 (Fig 2, arrowhead). The single internal methionine at amino acid position 355 (Fig. 2, boxed) is consistent with previous cyanogen bromide cleavage analysis (10) which yielded two cleavage fragments corresponding in size to those predicted from this sequence analysis.

The amino acid composition deduced from the pcGlb1S sequence in the region corresponding to GLB1-S (Table I) is in close agreement with that obtained from direct analysis of the GLB1-S protein (8). The protein contains high amounts of glutamate, arginine, serine, and glycine, and low amounts of cysteine, methionine, and tryptophan.

A hydropathy plot (12) of the deduced protein sequence is depicted in Figure 3. The sequence is extremely hydrophilic with the exception of a strongly hydrophobic region corre-

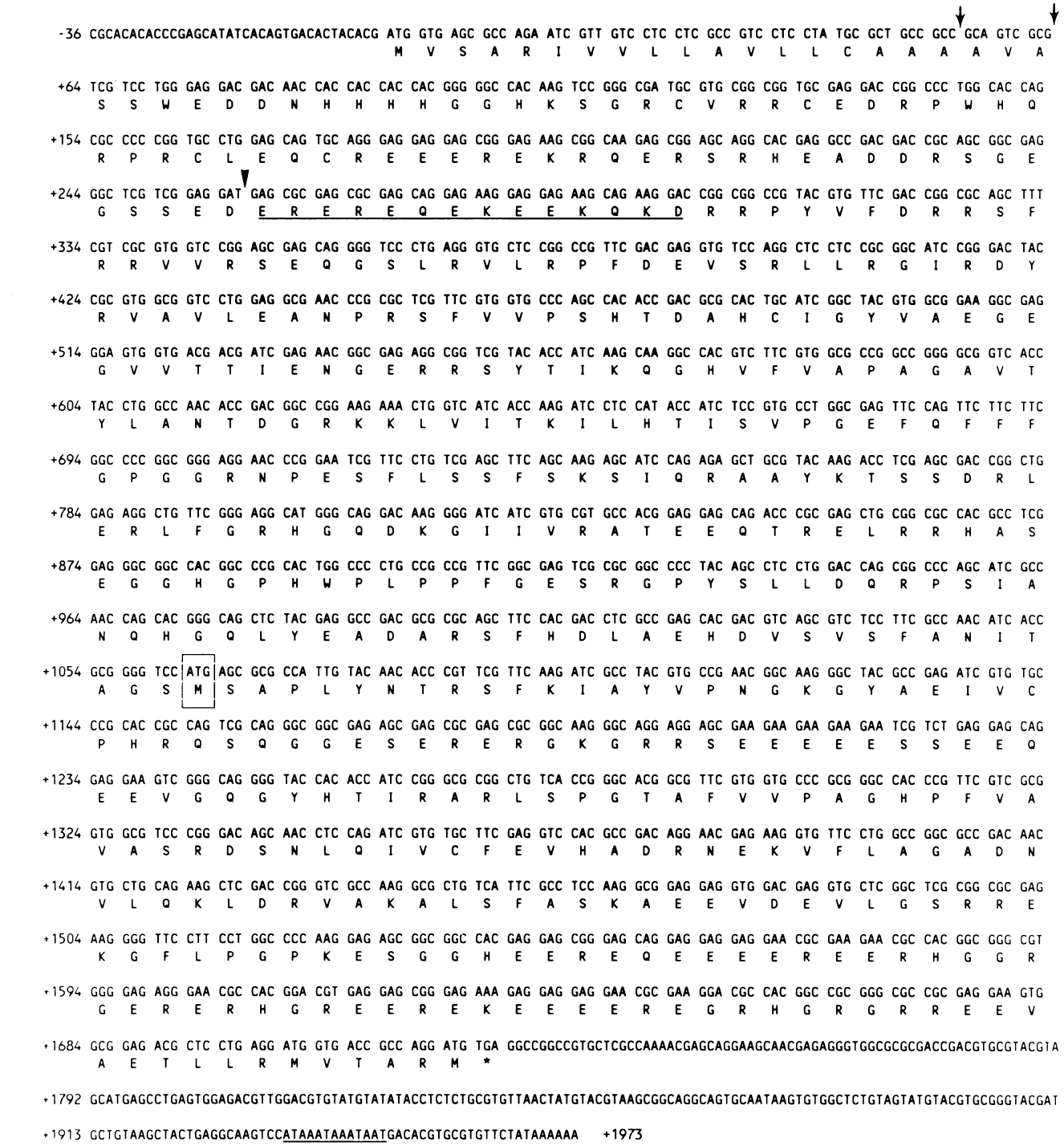

Figure 2. Complete nucleotide and deduced amino acid sequence of pcGlb1S. The sequence from nucleotide positions −36 to +189 was determined directly from pcGlb1S, and the sequence from positions +190 to +1973 was determined from pcA1. The region of overlap described in the text extends from position +190 to +245. The open reading frame extends from nucleotide positions +1 to +1719. The TGA termination codon is indicated by an asterisk. The deduced amino acid sequence from positions +259 to +300 (amino acids 87–100, *underlined)* corresponds to the N-terminal amino acid sequence determined directly from GLB1-S (8). An AT-rich region at positions +1937 to +1949 *(underlined)* is considered to be a potential polyadenylation site. Proposed sites for potential signal peptide cleavage are indicated by the two arrows, and the proposed site for *Mep*-catalyzed cleavage is indicated by an arrowhead. The internal methionine residue detected by CNBr cleavage (10) is boxed.

Table I. *Deduced Amino Acid Compositions of Glb1-S Gene Products*

Values were determined from the pcGlb1S nucleotide sequence in the indicated regions

	Amino Acid Position			
	87-573[a]		1-573[b]	
	mol%	*No. aa*	*mol%*	*No. aa*
Glu	13.35	65	13.26	76
Arg	12.11	59	12.57	72
Gly	9.45	46	8.90	51
Ser	7.60	37	7.85	45
Val	7.39	36	7.53	42
Ala	7.19	35	7.50	43
Leu	5.75	28	5.76	33
Pro	4.52	22	4.19	24
Phe	4.52	22	3.84	22
Lys	4.11	20	3.84	22
His	3.90	19	4.54	26
Gln	3.70	18	3.66	21
Asp	3.49	17	4.01	23
Thr	3.49	17	2.97	17
Ile	3.29	16	2.97	17
Tyr	2.46	12	2.09	12
Asn	2.26	11	2.09	12
Trp	0.21	1	0.52	3
Cys	0.62	3	1.40	8
Met	0.62	3	0.70	4
		487		573

[a] Corresponds to GLB1-S. [b] Corresponds to primary translation product.

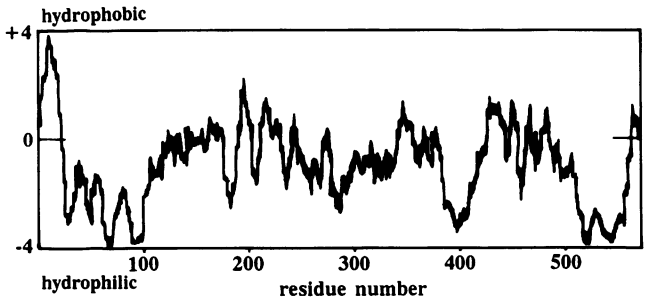

Figure 3. Hydropathy plot of deduced protein sequence of GLB1-S. This plot was generated by using the algorithm of Kyte and Doolittle (12), as supplied with DNAStar sequence analysis programs, using a window of seven amino acids.

sponding to the 21 N-terminal amino acids. This hydrophobic region is similar to the general characteristics of transmembrane peptide signal sequences (28). In particular, this region has a high degree of homology with the presumed signal sequence of aleuraine, a secreted thiol protease from barley aleurone (20). The two most likely sites for cleavage of the presumed signal sequence (Fig. 2, arrows) were predicted by the weighted matrix method of von Heijne (26). To conform with standard designations for processing intermediates which contain signal sequences, the primary translation product will subsequently be referred to as preproGLB1′ and the short-lived *in vivo* precursor, previously referred to as preGLB1′ (10), will be designated proGLB1′.

Genomic Representation of *Glb1*

Genetic analysis demonstrated that GLB1 is encoded by a single gene (22). In Southern blot analysis of maize DNA from the inbred line W64A *(Glb1-L/L)*, pcGlb1S hybridized to a single *Eco*RI fragment of 3.4 kb (Fig. 4, lane 1) as would be expected for a single gene since there are no *Eco*RI sites within the clone. Gene copy-number reconstruction analysis confirms the presence of a single copy of *Glb1* per haploid genome (Fig. 4). An identically sized *Glb1*-specific *Eco*RI fragment is present in DNA from plants homozygous for the

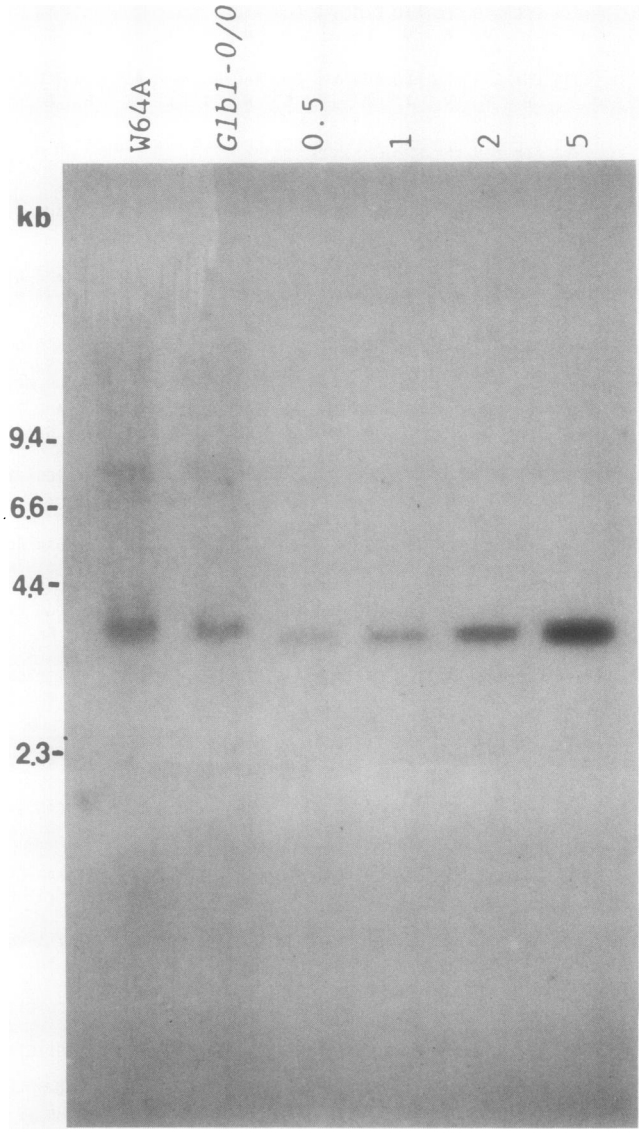

Figure 4. Southern blot analysis of maize DNA from W64A *(Glb1-L/L)* and *Glb1-O/O* plants. The first two lanes each contained 5.3 μg *Eco*RI-cleaved DNA from either of the two genotypes, as indicated. This amount of DNA corresponds to 1 x 10^6 haploid genomes. Gene copy-number reconstruction lanes contained *Eco*RI-cleaved DNA from a *Glb1-L* genomic clone (our unpublished data) corresponding to 0.5 x 10^6, 1 x 10^6, 2 x 10^6, or 5 x 10^6 copies of the cloned fragment, as indicated. The blot was probed with the radiolabeled cDNA insert from pcGlb1S.

null allele (Fig. 4, lane 2). This indicates that the lack of GLB1 protein in the null is not due to simple deletion of the gene. In addition, no differences in the hybridization pattern in *Bam*HI and *Hind*III digests were observed for the two genotypes (data not shown).

Northern Blot Analysis of Normal and *Glb1-O/O* Embryos

Embryos homozygous for the *Glb1* null allele lack detectable GLB1 protein (22) and translatable *Glb1* mRNA as determined by immunoprecipitation of *in vitro* translation products (10). To determine if *Glb1* transcripts are present in *Glb1* null embryos, total RNA from *Glb1-L/L* and *Glb1-O/O* embryos at 24 and 27 DAP was subjected to Northern blot analysis in which the ^{32}P-labeled insert from pcGlb1S was used as probe (Fig. 5). As expected, a single band in the *Glb1-L/L* samples exhibited strong hybridization with the probe (Fig. 5, lanes 1 and 3). A faint band of slightly larger size was detected in the *Glb1-O/O* samples at both 24 and 27 DAP (Fig. 5, lanes 2 and 4). The apparent size of the transcript in *Glb1-L/L* embryos is 2.4 kb and that of the *Glb1-O/O* embryos is 2.5 kb. There was no detectable size difference between transcripts from embryos of *Glb1-L/L* and *Glb1-S/S* inbred lines, and only a single band was observed in RNA from *Glb1-L/S* embryos (data not shown).

Developmental Accumulation and Seed-Specificity of *Glb1* Transcripts

The accumulation of *Glb1* transcripts during embryo development is depicted in Figure 6. Total RNA from *Glb1-L/L* embryos at 18 to 42 DAP was subjected to Northern blot analysis in which the pcGlb1S insert was used as radiolabeled probe. The level of Glb1-specific transcripts increased from 18 to 24 DAP, then decreased by 30 DAP. The 30 DAP level was maintained through 42 DAP.

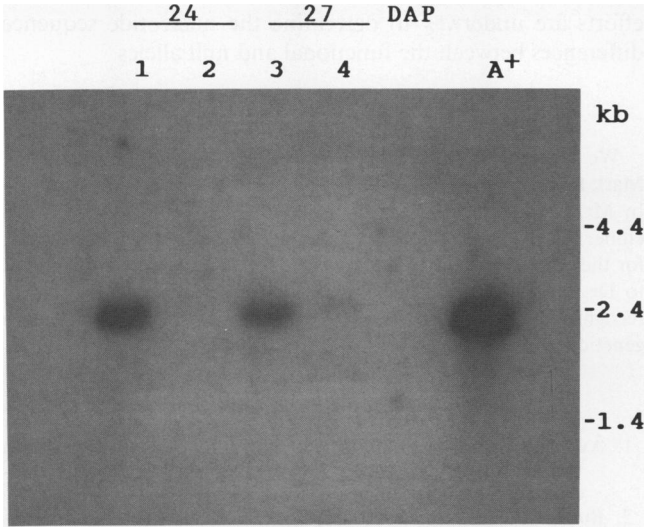

Figure 5. Northern blot analysis of RNA from embryos of different *Glb1* genotypes. Ten μg of total RNA from *Glb1-L/L* (W64A; lanes 1 and 3) or *Glb1-O/O* (lanes 2 and 4) embryos at 24 or 27 DAP, as indicated, or 1 μg of polyadenylated RNA from 27 DAP *Glb1-S/S* (Va26) embryos (A$^+$), was subjected to Northern blot analysis in which the pcGlb1S insert was used as probe. Positions of RNA size markers (BRL) are as indicated.

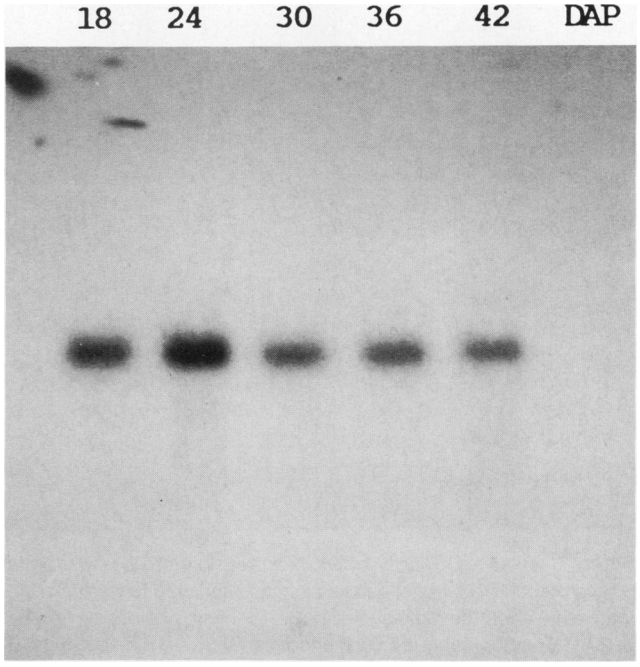

Figure 6. Developmental accumulation of *Glb1* transcripts. Total RNA (10 μg/lane) from developing *Glb1-L/L* embryos (18–42 DAP, as indicated) was subjected to Northern blot analysis in which the radiolabeled pcGlb1S cDNA insert was used as probe.

Glb1 expression was also investigated in plant tissues other than the embryo. Total RNA samples were obtained from *Glb1-L/L* embryos, endosperm, unfertilized ears, immature tassels, and 7-d old seedlings. *Glb1*-Specific transcripts were detected in the endosperm but at a lower level than that found in the embryo (Fig. 7, lanes 1 and 2). *Glb1* transcripts were not detected in any of the other tissues examined. Identical results were obtained when similar tissues from the *Glb1-S/S* inbred line were subjected to the same analysis (data not shown).

DISCUSSION

To further characterize the maize *Glb1* gene, which encodes an abundant protein in the embryo, we have isolated and determined the nucleotide sequence of a cDNA clone, pcGlb1S, corresponding to this locus. Nucleotide sequence analysis of this 2009 bp clone reveals an open reading frame corresponding to 573 amino acids. The deduced amino acid sequence and composition are consistent with information obtained for GLB1-S (8). The nucleotide sequence contains consensus sequences for the initiator methionine and a 3′ polyadenylation sequence. The pcGlb1S sequence has a high G+C content of 66%; similar high G+C contents have been observed for genes encoding other seed proteins (20).

Sequence analysis indicates that the protein product of the *Glb1* gene appears to contain a signal sequence which suggests it may be transported through a membrane. Many seed storage proteins are localized in protein bodies which require

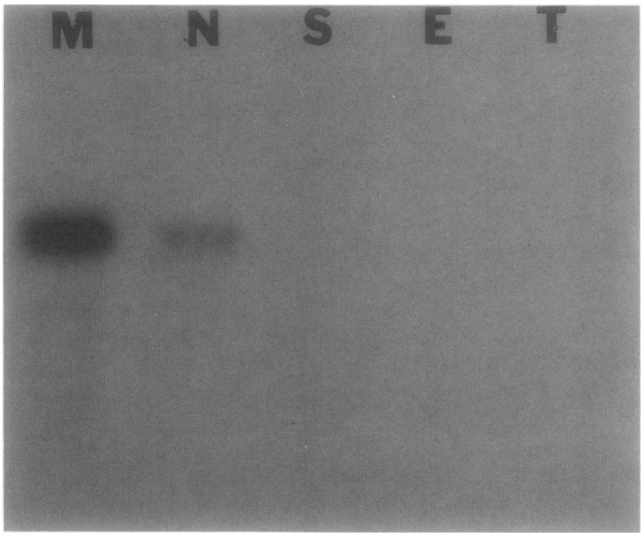

Figure 7. Tissue specificity of *Glb1* expression in *Glb1-L/L* plants. Ten μg of total RNA from each tissue was subjected to Northern blot analysis in which the radiolabeled pcGlb1S insert was used as probe. M, 27 DAP embryos; N, 27 DAP endosperm; S, 7-d old seedlings; E, unfertilized ears; T, immature tassels.

passage through the endoplasmic reticulum (24). The cellular location of the *Glb1* proteins remains to be determined.

Previous studies provided evidence for at least three co- and/or posttranslational protein processing steps in the production of GLB1 from the primary translation product of the *Glb1* gene (10). The final processing step in GLB1 synthesis is proteolytic cleavage of GLB1′. The nature of the two earlier processing steps has not been determined. The observation that the *in vitro* translation product preproGLB1′ is of a lower mol wt than proGLB1′, the first processing intermediate detected *in vivo* (10), indicates the involvement of protein modification processes. There is a potential N-linked glycosylation site of Asn-X-Ser/Thr (6) in the deduced amino acid sequence (Asn-Ile-Thr at nucleotide positions +1045 to +1053), although no direct evidence for such glycosylation has been obtained for GLB1 (8).

We have examined *Glb1* expression in various plant tissues by Northern blot analysis. *Glb1*-Specific transcripts, detected only in seed tissues, were present at high levels in the embryo and at much lower levels in the endosperm. *Glb1*-Encoded proteins have also been detected in endosperm tissue by immunoblot analysis (JL Puckett, AL Kriz, unpublished data). In a developmental study, the amount of *Glb1*-specific transcripts in the embryo increased from 18 to 24 DAP and declined slightly to reach a steady level which was maintained from 30 to 42 DAP. *Glb1* transcripts are present at reduced, but significant, levels in embryos of dry mature seeds (our unpublished data), and it remains to be determined if this is due to continued transcription of the gene or to mRNA stability. This pattern of expression is quite different from that observed for genes encoding prolamin storage proteins in maize endosperm (9) and globulin storage proteins in legume cotyledons (5), where the corresponding transcripts decline to extremely low levels during the later stages of seed development. Expression patterns similar to those described here for *Glb1* transcripts have been observed for transcripts encoding a 7S globulin of wheat embryos (29).

Glb1 is apparently related to a gene encoding the 7S globulin storage protein of wheat embryos. GLB1-Specific antibodies cross-react with a protein of approximate mol wt 50,000 in mature wheat seeds (data not shown). We presume this wheat protein to be the 7S globulin described by Quatrano *et al.* (19). Computer comparisons of pcGlb1S and a genomic clone corresponding to the wheat 7S globulin (RS Quatrano, personal communication) indicate the presence of substantial similarities in both the DNA and amino acid sequences. The two clones exhibit similarity (75% nucleotide homology, 65% amino acid identity) in two regions, corresponding to nucleotide positions +320 to +720 and +900 to +1110 in the pcGlb1S sequence. Limited but significant amino acid similarity was also detected between the pcGlb1S sequence and those of two other seed globulin storage proteins, pea vicilin (15) and French bean phaseolin (25). Pea vicilin exhibits 28%, and phaseolin 25%, amino acid identity in the two pcGlb1S regions defined above. No cross-reactivity with GLB1-specific antiserum was observed in immunoblot analysis of pea or French bean seed protein extracts (data not shown). Detailed sequence comparisons of Glb1 and genes encoding seed globulins in other plants will be presented in a subsequent report.

Southern blot analysis of DNA from plants homozygous for the *Glb1-O* allele reveals that the null phenotype is not due to deletion of the *Glb1* structural gene. From Northern blot analysis it is evident that transcripts from the null allele do not accumulate to the same high levels as those of the functional alleles and that the *Glb1-O* transcripts are slightly larger than *Glb1* transcripts. Additional analysis of the null allele may provide insight into the mechanisms involved in transcriptional and/or posttranscriptional regulation of gene expression in developing seeds. We have obtained genomic clones corresponding to *Glb1-S, Glb1-L,* and *Glb1-O,* and efforts are underway to determine the nucleotide sequence differences between the functional and null alleles.

ACKNOWLEDGMENTS

We appreciate the excellent technical assistance provided by Mr. Mark S. Wallace in RNA isolations and Northern blot analysis and by Mr. Lee D. Stromberg in the isolation of the full-length pcGlb1S clone. We thank Drs. D. C. Bennett, J. W. Dudley, and P. D. Shaw for their critical reading of this manuscript. This paper is dedicated to Dr. Drew Schwartz in honor of his seventieth birthday and in recognition of his continuing contributions to the study of maize genetics.

LITERATURE CITED

1. **Aviv H, Leder P** (1972) Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc Natl Acad Sci USA **69:** 1408–1412
2. **Bunce NAC, White RP, Shewry PR** (1985) Variation in estimates of molecular weights of cereal prolamins by SDS-PAGE. J Cereal Sci **3:** 131–142
3. **Cox RA** (1968) The use of guanidinium chloride in the isolation of nucleic acids. Methods Enzymol **12B:** 120–129
4. **Gubler U, Hoffman BJ** (1983) A simple and very efficient method for generating cDNA libraries. Gene **25:** 263–269
5. **Higgins TJV** (1984) Synthesis and regulation of major proteins in seeds. Annu Rev Plant Physiol **35:** 191–221

6. **Hubbard SC, Ivatt RJ** (1981) Synthesis and processing of asparagine-linked oligosaccharides. Annu Rev Biochem **50:** 555–583
7. **Huynh TY, Young RA, Davis RW** (1985) Constructing and screening cDNA libraries in λgt10 and λgt11. *In* DM Glover, ed, DNA Cloning: A Practical Approach, Vol 1. IRL Press, Oxford, pp 245–300
8. **Kriz AL** (1989) Characterization of embryo globulins encoded by the maize *Glb* genes. Biochem Genet **27:** 239–251
9. **Kriz AL, Boston RS, Larkins BA** (1987) Structural and transcriptional analysis of DNA sequences flanking genes that encode 19 kilodalton zeins. Mol Gen Genet **207:** 90–98
10. **Kriz AL, Schwartz D** (1986) Synthesis of globulins in maize embryos. Plant Physiol **82:** 1069–1075
11. **Kozak M** (1984) Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res **12:** 857–872
12. **Kyte J, Doolittle RF** (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol **157:** 105–132
13. **Lending CR, Kriz AL, Larkins BA, Bracker CE** (1988) Structure of maize protein bodies and immunocytochemical localization of zeins. Protoplasma **143:** 51–62
14. **Lutcke HA, Chow KC, Mickel FS, Moss KA, Kern BF, Scheele GA** (1987) Selection of AUG initiation codons differs in plants and animals. EMBO J **6:** 43–48
15. **Lycett GW, Delauney AJ, Gatehouse JA, Gilroy J, Croy RRD, Boulter D** (1983) The vicilin gene family of pea *(Pisum sativum* L.): A complete cDNA coding sequence for preprovicilin. Nucleic Acids Res **11:** 2367–2380
16. **Oshima RG** (1988) Rapid identification of M13 phage deletions for sequence analysis. BioTechniques **6:** 510–511
17. **Osterman JC** (1988) An allele of the *Prot* locus in maize is a variant for the site of protein processing. Biochem Genet **26:** 463–474
18. **Proudfoot NJ, Brownlee GG** (1976) 3′ noncoding region sequences in eukaryotic messenger RNA. Nature **263:** 211–214
19. **Quatrano RS, Litts J, Colwell G, Chakerian R, Hopkins R** (1986) Regulation of gene expression in wheat embryos by ABA: Characterization of cDNA clones for the EM and putative globulin proteins and localization of the lectin wheat germ agglutinin. *In* LM Shannon, MJ Chrispeels, eds, Molecular Biology of Seed Storage Proteins and Lectins. American Society of Plant Physiologists, Rockville, MD, pp 127–136
20. **Rogers JC, Dean D, Heck GR** (1985) Aleuraine: A barley thiol protease closely related to mammalian cathepsin H. Proc Natl Acad Sci USA **82:** 6512–6516
21. **Sanger F, Nicklen S, Coulson AR** (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA **74:** 5463–5467
22. **Schwartz D** (1979) Analysis of the size alleles of the *Pro* gene in maize—evidence for a mutant protein processor. Mol Gen Genet **174:** 233–240
23. **Selden RF** (1987) Analysis of RNA by northern hybridization. *In* FE Ausubel, R Brent, RE Kingston, DD Moore, JG Seidman, JA Smith, K Struhl, eds, Current Protocols in Molecular Biology. John Wiley and Sons, New York, unit 4.9
24. **Shotwell MA, Larkins BA** (1989) 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, pp 297–345
25. **Slightom JL, Sun SM, Hall TC** (1983) Complete nucleotide sequence of a French bean storage protein gene: Phaseolin. Proc Natl Acad Sci USA **80:** 1897–1901
26. **von Heijne G** (1986) A new method for predicting signal sequence cleavage sites. Nucleic Acids Res **14:** 4683–4690
27. **von Heijne G** (1987) Sequence Analysis in Molecular Biology. Academic Press, Inc, San Diego, CA
28. **Watson MEE** (1984) Compilation of published signal sequences. Nucleic Acids Res **12:** 5145–5164
29. **Williamson JD, Quatrano RS** (1988) ABA-regulation of two classes of embryo-specific sequences in mature wheat embryos. Plant Physiol **86:** 208–215
30. **Yanisch-Perron C, Vieira J, Messing J** (1985) Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:** 103–119