

Isolation and characterization of a genomic clone encoding the β -subunit of β -conglycinin

Mary L. Tierney, Elizabeth A. Bray*, Randy D. Allen, Yu Ma**, Roger F. Drong***, Jerry Slightom*** and Roger N. Beachy****

Department of Biology, Plant Biology Program, Washington University, St. Louis, MO 63130, USA

Abstract. β -Conglycinin, an abundant storage protein in soybean (*Glycine max* (L.) Merr.) seeds, is a trimeric protein consisting of various isomers containing the three subunits, α' , α and β . Accumulation of the β -subunit is unique because it appears to be regulated by a variety of developmental and environmental signals. In this paper we describe the isolation and characterization of a genomic clone encoding the β -subunit of β -conglycinin. The genomic clone was characterized by restriction-enzyme mapping and partial DNA sequence analysis, by immunoprecipitation of a hybrid-selected in-vitro translation product, and by RNA blot hybridization reactions. An mRNA of approx. 1700 nucleotides hybridized to an internal 2-kilobase (kb) region of this 4.4-kb cloned DNA restriction fragment and was translated to yield a polypeptide with an approximate molecular weight of 48 kilodalton. This polypeptide is immunoprecipitable by antibody against β -conglycinin and is of appropriate size to represent the precursor polypeptide of the β -subunit. When this sequence was used as a probe in RNA blot hybridization experiments, the β -gene transcript was first detected by stage K and accumulated through stage O during soybean seed development, coincident with appearance of the β -subunit. Partial DNA sequence analysis of the 5' end of the gene confirmed that the isolated gene encoded a β -subunit, based upon the

previously reported amino terminal sequence for this protein. Genomic DNA blot hybridization analyses indicate that multiple DNA restriction fragments are highly homologous to this cloned β -gene sequence.

Key words: β -Conglycinin gene sequence – *Glycine* (β -conglycinin) – Seed development – Storage protein (seed).

Introduction

Seed storage proteins in soybean are composed of two major proteins, β -conglycinin and glycinin. β -Conglycinin is a trimeric 7S molecule consisting of three subunits: α' (76 kilodalton, kDa), α (72 kDa) and β (53 kDa). The three subunits of β -conglycinin are all glycoproteins which undergo post-translational processing to give rise to their mature forms (Beachy et al. 1981). The genes encoding the α' and α subunits of β -conglycinin are highly homologous (Schuler et al. 1982b) and have been shown to have strong homology with both phaseolin and vicilin storage-protein genes (Doyle et al. 1986; Lycett et al. 1983).

The β -subunit of β -conglycinin is differentially affected by several developmental and environmental conditions during the development of soybean embryos. Accumulation of the β -subunit begins at a later stage of development than that of the α' - and α -subunits (Meinke et al. 1981). The synthesis and stability of the β -subunit is distinctly different from that of the α' - and α -subunits in the axis and cotyledon of the developing embryo (Ladin et al. 1987), and the levels of both the β -subunit transcript and protein are affected by abscisic acid (ABA; Bray and Beachy 1985) and me-

Present addresses:

* Department of Botany and Plant Sciences, University of California, Riverside, CA 92521, USA

** Agrigenetics Corp., Advanced Research Division, 5649 E. Buckeye Rd., Madison, WI 53716, USA

*** Upjohn Company, Division of Molecular Biology, Kalamazoo, MI 49007, USA

**** To whom correspondence should be addressed

Abbreviations: cDNA = copy DNA; kb = kilobase; kDa = kilodalton; SDS-PAGE = sodium dodecyl sulfate polyacrylamide gel electrophoresis

thionine treatments (Holowach et al. 1986) in cultured developing cotyledons.

In this paper we describe the isolation and characterization of a genomic clone that encodes a β -subunit of β -conglycinin. The clone was characterized using hybrid-selection of mRNA followed by immunoprecipitation of the in-vitro translation reaction products, and by RNA and DNA blot hybridization analyses. The restriction-enzyme map and partial DNA sequence of the coding region of the cloned fragment were determined and compared with DNA sequences previously determined for a β -conglycinin α -subunit gene. This analysis will now make it possible to dissect those regions of the β -subunit gene sequence which are responsible for its varied developmental expression in soybean seeds.

Material and methods

Genomic library construction. A genomic library of soybean (*Glycine max* (L.) Merr. cv. Corsoy, Agrigenetics Lubrizol Corp., Wickliffe, Oh., USA) DNA was constructed in lambda 1059 as previously described (Slightom et al. 1980). Soybean genomic DNA was isolated, partially digested with Mbo I and cloned into lambda 1059 using isolated arms which had been digested with Bam HI (Karn et al. 1980).

Identification of β -conglycinin genomic clones. The soybean genomic library was first screened with a cloned cDNA (copy DNA) encoding the β -subunit of phaseolin, the analogous seed protein in *Phaseolus vulgaris* (Slightom et al. 1983). Positive clones were isolated and the DNA was purified and characterized by restriction-enzyme analysis. In order to identify a genomic clone encoding a soybean β -subunit gene DNA from these clones was nick-translated and used to probe RNA from immature soybean seeds by Northern blot analysis. Polyadenylated (Poly A⁺) RNA (10 μ g per lane) from developing soybean seeds was separated on denaturing 1.5% agarose gels containing 5 mM methyl mercury hydroxide (Bailey and Davidson 1976). The RNAs were transferred to activated diazobenzylmethyl (DBM) paper (Alwine et al. 1979) and the filters were pre-hybridized for 2 h at 42° C in 40% formamide, 0.9 M NaCl, 50 mM sodium phosphate, pH 7.5, 5 mM ethylenediaminetetraacetic acid (EDTA), 0.1% sodium dodecyl sulfate (SDS), 0.02% (w/v) each of bovine serum albumin (BSA), Ficoll, and polyvinylpyrrolidone (PVP) (1 \times Denhardt's solution), 250 μ g \cdot ml⁻¹ denatured salmon-sperm DNA and 1% (w/v) glycine. Hybridization with the nick-translated, ³²P-labeled genomic clones was carried out in the same buffer without glycine for 18 h at 42° C. The filters were washed first at room temperature in 0.36 M sodium chloride, 20 mM sodium phosphate, pH 7.0, 2 mM EDTA, 0.1% SDS (wash 1) followed by two washes in 18 mM sodium chloride, 1 mM sodium phosphate, pH 7, 0.1 mM EDTA, 0.1% SDS (wash 2) at 42° C (Tm-35). The filters were then washed at 57° C in wash 2 (Tm-20) and finally washed at 67° C in wash 2 (Tm-10). The filters were exposed to Kodak (Rochester, N.Y., USA) X-OMAT film at -80° C with two intensifying screens after each wash.

Restriction-enzyme analysis. Desoxyribonucleic acid isolated from the lambda clone 2.1.1.1 (which contained a β -subunit gene) was characterized by restriction-enzyme analysis using

the Southern Cross method (Potter and Dressler 1986). Restriction fragments which contained the coding region of the β -subunit gene were determined by hybridizing end-labeled soybean-seed mRNA to DNA from lambda clone 2.1.1.1 which had been cut with the indicated restriction enzymes and transferred to nitrocellulose filters (Maniatis et al. 1982). The region of genomic clone 2.1.1.1 which hybridized with ³²P-labeled soybean-seed mRNA was subcloned into the plasmid pSP64 (Promega Biotech, Madison, Wis., USA) and a more detailed restriction-enzyme map was deduced for this plasmid, designated pGmg 91.

Hybrid selection. Ribonucleic acid was isolated from immature soybean seeds that were greater than stage O, a stage at which β -subunit mRNA accumulates (Meinke et al. 1981; Ladin et al. 1987). The presence of the β -subunit transcript was confirmed by in-vitro translation of this RNA preparation (data not shown). Ribonucleic acid that was homologous to pGmg 91 was selected by hybridization to the plasmid DNA which had been denatured and bound to activated APT paper (Bunemann and Westhoff 1983). A 28- μ g sample of total RNA was used in the hybridization reaction. The filters were washed at Tm-10 five times, after which the bound RNA was eluted by holding the filters at 65° C for 5 min in 99% formamide, pH 7.8. The mRNA was then ethanol-precipitated using glycogen as a carrier. The precipitated RNA was washed well with 70% ethanol and translated in a cell-free wheat-germ extract (Anderson et al. 1983). The in-vitro translation reaction products were immunoprecipitated (Lingappa et al. 1978) with anti- β -conglycinin serum and size-separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE; Laemmli 1970). The labeled in-vitro translation products were detected by fluorography.

Blot analysis of RNA. Total RNA was isolated from soybean seeds at early and late stages in soybean development and 15 μ g of RNA from each stage was subjected to electrophoresis in 1.5% agarose gels containing formaldehyde (Maniatis et al. 1982). The RNA was transferred to nitrocellulose (Maniatis et al. 1982) and probed with nick-translated ³²P-labeled pGmg 91. The nitrocellulose was first prehybridized in 5 \times SSC (1 \times SSC = 0.15 M NaCl, 0.015 M sodium citrate, pH 7), 5 \times Denhardt's, 0.1% SDS, 0.1% salmon-sperm DNA and 50% formamide for 2 h at 42° C. The RNA bound to the filter was hybridized with ³²P-labeled pGmg 91 at 42° C overnight in the same solution and then washed four times at 42° C with 2 \times SSC, 0.5% SDS. The filter was then exposed to X-ray film at -80° C with two intensifying screens.

Genomic Southern blot analysis. Total DNA was isolated from soybean leaves (Dellaporta et al. 1983) and 10 μ g of DNA was digested with Hind III or Eco RI according to manufacturer's specifications (New England BioLabs, Inc., Beverly, Mass., USA). The DNA was size-separated on 0.8% agarose gels, transferred to nitrocellulose and prehybridized in 5 \times SSC, 5 \times Denhardt's for 2 h at 60° C. The DNA was hybridized with ³²P-labeled pGmg 91 in 5 \times SSC, 5 \times Denhardt's, 10% dextran sulfate, 0.1% SDS and 0.1 mg \cdot ml⁻¹ salmon-sperm DNA at 60° C overnight. The filter was then washed sequentially at 60° C in 1) 6 \times SSC, 5 \times Denhardt's, 0.2% SDS; 2) 6 \times SSC, 0.2% SDS; 3) 2 \times SSC, 0.2% SDS; 4) 0.1 \times SSC, 0.2% SDS and exposed to X-ray film at -80° C with two intensifying screens.

Western blot analysis. Proteins were extracted from developmentally staged soybean embryos (Meinke et al. 1981) and equal amounts of protein were separated by SDS-PAGE

(Laemmli 1970). The proteins were then transferred to nitrocellulose by electroblotting (Burnett 1981) and incubated with anti- β -conglycinin serum followed by ^{125}I -labeled goat-anti-rabbit antibody as described by Symington et al. (1981). After extensive washing to remove non-specifically bound ^{125}I , the filters were exposed to X-ray film at -80°C .

Analysis of DNA sequence. Restriction fragments that correspond to the 5' end of the β -subunit gene were subcloned into m13-mp19 and sequenced by di-deoxy chain termination procedures (Sanger 1977). The sequence obtained was compared, using the University of Wisconsin (Madison, USA) Genetics Group software package, to corresponding sequences of an α' -subunit gene (Doyle et al. 1986) and a β -subunit gene sequence isolated from soybean, cv. Forrest (S.J. Barker, J. Harada and R.B. Goldberg, University of California, Los Angeles, USA, personal communication).

Sources. Desoxyribonuclease I was purchased from Boehringer Mannheim Biochemicals, Indianapolis, Ind., USA. All other restriction and DNA-modifying enzymes were purchased from New England BioLabs, Beverly, Mass., USA. Nitrocellulose was purchased from Schleicher and Schuell, Keene, N.H., USA. Formamide for the hybrid-selection experiments was purchased from Bethesda Research Laboratories, Gaithersburg, Md., USA. All other chemicals and supplies were reagent grade.

Results

The objectives of this work were to isolate and characterize a genomic clone representing a gene encoding the β -subunit of β -conglycinin. The clone that was isolated was characterized by restriction-enzyme mapping and DNA sequence analysis, hybrid-selection of mRNA followed by in-vitro translation and immunoprecipitation, and RNA blot and genomic DNA blot hybridization analyses.

Isolation of a β -subunit gene. A lambda 1059 soybean genomic library (cv. Corsoy) was screened with a copy DNA (cDNA) encoding a *Phaseolus vulgaris* storage protein, phaseolin (Slightom et al. 1983). Phaseolin had previously been shown to have sequence homology with the β -conglycinin family of storage proteins in soybean (Doyle et al. 1986). Four classes of genomic clones, based on their restriction-enzyme patterns (data not shown), hybridized to the phaseolin cDNA in this screen. These clones represented candidates for genes encoding the α' -, α - and β -subunits of β -conglycinin.

The β -subunit gene can be distinguished from α' - and α -subunit genes by the size of its mRNA transcript. The β -subunit mRNA is approx. 1.7 kb (kilobase) in length while the α' and α transcripts are approx. 2.3 kb (Schuler et al. 1982b). This observation was used to determine which of the four classes of genomic clones encoded a β -subunit gene. The DNA from a representative clone from

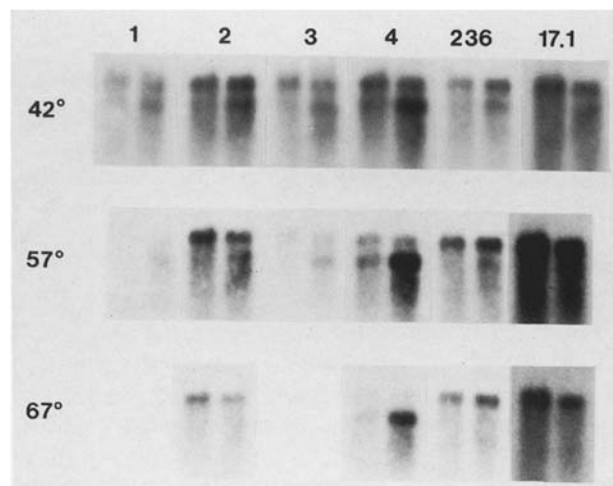


Fig. 1. RNA blot hybridization analysis of the four classes of putative β -conglycinin clones from soybean. The four classes of lambda genomic clones (1–4), selected because of their hybridization to phaseolin cDNA sequences, were hybridized to RNA blots of soybean-seed mRNA (lanes 1) and RNA from cotyledons cultured in the presence of 10 μM abscisic acid (lanes 2) (ABA increases the level of β -subunit mRNA under such conditions, Bray and Beachy 1985). These RNA samples were also probed with an α -subunit cDNA (pGmc 236) and an α -subunit genomic clone (pGmg 17.1). The RNA blots were washed successively at 42°C , 57°C and 67°C , and the patterns of hybridization were determined after each wash

each of the four classes was nick-translated and used to probe mRNA from developing soybean seeds. When the filters were washed at low stringency, each of the four classes of genomic clones hybridized to the α' , α and β transcripts (Fig. 1). However, when the stringency of the wash was increased, clone 2.1.1.1, which belonged to class 4, hybridized preferentially to the 1.7-kb mRNA transcript. On this basis, clone 2.1.1.1 was classified as a candidate for a β -subunit genomic clone.

Restriction-enzyme mapping. A partial restriction map of genomic clone 2.1.1.1 was determined using the Southern Cross method (Fig. 2). To determine which of these restriction fragments contained the coding region for the β -subunit gene, ^{32}P -labeled mRNA from developing soybean seeds was hybridized to DNA from clone 2.1.1.1 that had been cut with various restriction enzymes. The fragment of genomic clone 2.1.1.1 that hybridized with the end-labeled mRNA was subcloned into the plasmid pSP64 and designated pGmg 91. Restriction fragments of pGmg 91 were also hybridized to end-labeled mRNA from soybean seeds to confirm and more precisely locate the coding region of this clone (Fig. 2).

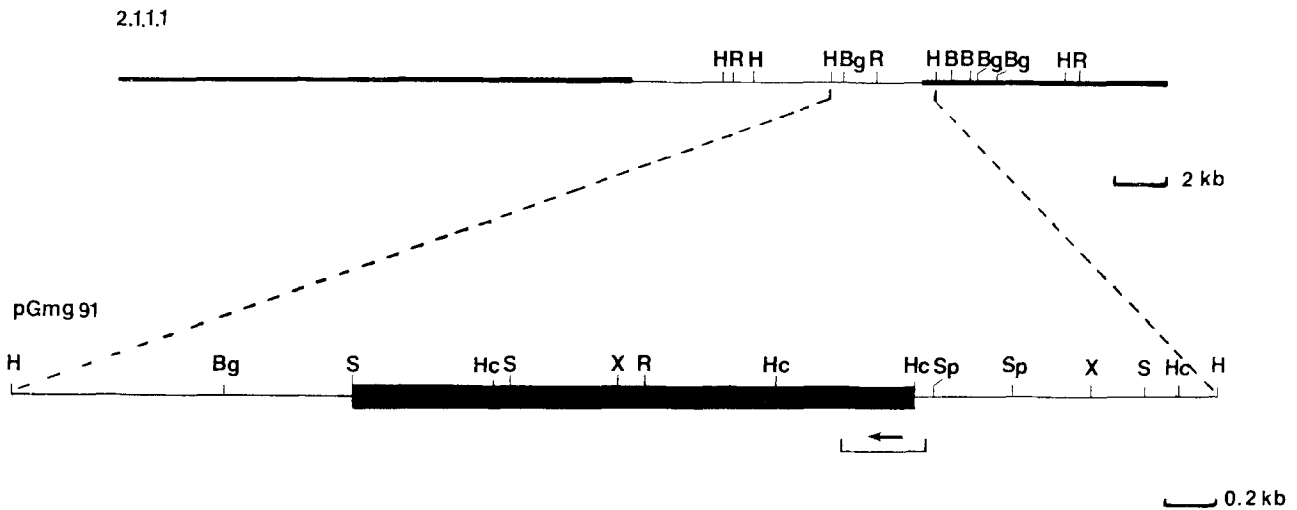


Fig. 2. Restriction-enzyme map of a genomic clone for the β -subunit of β -conglycinin from soybean. The lambda clone 2.1.1.1 represents a genomic soybean clone which contains the coding sequence for a β -subunit gene of β -conglycinin. The heavy lines represent the lambda 1059 arms and the internal thin line represents the cloned soybean DNA fragment. The 4.4 kb Hind III fragment of 2.1.1.1 which hybridized to soybean seed mRNA was subcloned into the Hind III site of pSP64 and designated pGmg 91. The coding region of pGmg 91 was confined to an internal 2 kb Sal I-Hinc II fragment, designated by the heavy line. The arrow indicates the direction of transcription of the β -subunit gene and the brackets beneath pGmg 91 indicate the restriction fragments for which DNA sequence is presented (see Fig. 4). Restriction-enzyme abbreviations used are: B=Bam HI; Bg=Bgl II; H=Hind III; Hc=Hinc II; R=Eco RI; S=Sal I; Sp=Sph I; X=Xba I

Confirmation of the identity of pGmg 91 as a clone encoding the β -subunit. To confirm that pGmg 91 contained the coding sequence for a β -subunit gene in soybean, cotyledon mRNA was hybrid-selected and translated in vitro. Soybean-seed mRNA from stage O (Meinke et al. 1981) was isolated, translated in vitro and immunoprecipitated to demonstrate that it contained β -subunit mRNA (Fig. 3, lane T). Soybean mRNA from stage O was then hybridized to pGmg 91 DNA which had been immobilized on 2-aminophenylthioether (APT) paper. The mRNA was eluted and translated in vitro and the translation products were immunoprecipitated using anti- β -conglycinin serum (Fig. 3, lane S). The major product that immunoprecipitated was the expected size of the pre- β -subunit in-vitro translation product (48 kDa). The pattern of in-vitro translation products derived from pGmg 91 hybrid-selected mRNA was the same before and after immunoprecipitation with β -conglycinin antiserum (data not shown). In-vitro translation products were also immunoprecipitated in minor amounts that corresponded in size with the polypeptides representing pre- α' and pre- α translation products. This is a consequence of the DNA sequence homology that is shared between all of the β -conglycinin subunit genes.

Comparison of DNA sequences. The nucleotide sequence of the 5' end of Gmg 91 was determined

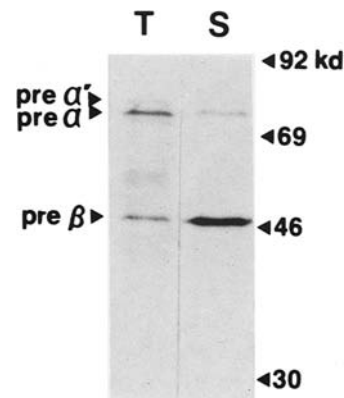


Fig. 3. Hybrid-selection of the β -subunit transcript with pGmg 91. RNA from immature soybean seeds (stage O) was hybridized with pGmg 91. The RNA was eluted, translated in vitro, and the ^3H -labeled in-vitro translation reaction products were immunoprecipitated with anti- β -conglycinin serum, and analyzed following SDS-PAGE and fluorography. T=total soybean RNA; S = hybrid-selected soybean RNA

to confirm the identity of this clone further, by comparing the amino-acid sequence derived from the DNA sequence of Gmg 91 with the NH_2 -terminal amino-acid sequence of the β -subunit (Coates et al. 1985). The sequence of the NH_2 -terminal amino acids deduced from DNA sequence analysis is identical to that determined for the mature protein (indicated by the solid line from position No. 128 through position No. 205 in Fig. 4) with

Gmg91 Beta																			
1	ATA	ACT	<u>ATA</u>	<u>AAT</u>	AGC	CCT	AAT	CTC	ACT	CCA	TGT	TT-	CAT	CGT	CCA	ATA	47		
	::	:::	:::	:::	:	:	:::	:::	:	:::	:::	:::	:::	:::	:::	:::			
1	ATG	ACT	<u>ATA</u>	<u>AAT</u>	ATC	TGC	AAT	CTC	GGC	CCA	AGT	TTT	CAT	CAT	C-A	AGA	47		
Gmg171 Alpha'																			
	ATA	T--	---	---	---	---	---	---	---	---	---	---	---	---	---	--A	TAT	ACT	58
	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:::	:::	
48	ACC	AGT	TCA	ATA	TCC	TAG	TAC	GCC	GTA	TTA	AAG	AAT	TTA	AGA	TAT	ACT		95	
	MET	MET	Arg	Val	Arg	Phe	Pro	Leu	Leu	Val	Leu	Leu	Gly	Thr	Val	Phe			
59	ATG	ATG	AGA	GTG	CGG	TTT	CCT	TTG	TTG	GTG	TTG	CTG	GGA	ACT	GTT	TTC		106	
	:::	:::	:::	:	:::	:::	:::	:::	:::	:::	:::	:::	:::	:	:::	:::			
96	ATG	ATG	AGA	GCG	CGG	TTC	CCA	TTA	CTG	---	TTG	CTG	GGA	GTT	GTT	TTC		141	
	MET	MET	Arg	Ala	Arg	Phe	Pro	Leu	Leu	---	Leu	Leu	Gly	Val	Val	Phe			
	Leu	Ala	Ser	Val	Cys	Val	Ser	Leu	Lys	Val	Arg	Glu	Asp	Glu	Asn	Asn			
107	CTG	GCA	TCA	GTT	TGT	GTC	TCA	TTA	AAG	GTG	ACA	GAG	GAT	GAG	AAT	AAC		154	
	::	:::	:::	:::	:	:::	:::	:::	:	:::	:	:	:	:	:::	:::			
142	CTA	GCA	TCA	GTT	TCT	GTC	TCA	TTT	GGC	ATT/AGA	CAT	AAG	AAT	AAG	AAC			726	
	Leu	Ala	Ser	Val	Ser	Val	Ser	Phe	Gly	Ile/Arg	His	Lys	Asn	Lys	Asn				
	<div style="text-align: center;"> Δ [537 base pair sequence in alpha' not in beta] </div>																		
	Pro	Phe	Tyr	[*] Phe	Arg	Ser	Ser	Asn	Ser	Phe	Gln	Thr	Leu	Phe	Glu	Asn			
155	CCT	TTC	TAC	TTT	AGA	AGC	TCT	AAC	AGC	TTC	CAA	ACT	CTC	TTT	GAG	AAC		202	
	:::	:::	:::	:	:	:	:	:	:	:::	:::	:::	:::	:::	:	:::			
727	CCT	TTT	CAC	TTC	AAC	TCT	AAA	AGG	---	TTC	CAA	ACT	CTC	TTC	AAA	AAC		771	
	Pro	Phe	His	Phe	Asn	Ser	Lys	Arg	---	Phe	Gln	Thr	Leu	Phe	Lys	Asn			
	Gln	Asn	Val	Arg	Ile	Arg	Leu	Leu	Gln	Arg	Phe	Asn	Lys	Arg	Ser	Gln			
203	CAA	AAC	GTT	CGC	ATT	CGT	CTC	CTC	CAG	AGA	TTC	AAC	AAA	CGC	TCC	CAA		251	
	:::	:	:	:	:::	:::	:::	:::	:::	:::	:::	:::	:::	:::	:::	:::			
772	CAA	TAT	GGC	CAC	GTT	CGC	GTC	CTC	CAG	AGG	TTC	AAC	AAA	CGC	TCC	CAA		819	
	Gln	Tyr	Gly	His	Val	Arg	Val	Leu	Gln	Arg	Phe	Asn	Lys	Arg	Ser	Gln			
	Gln	Leu	Glu	Asn	Leu	Arg	Asp	Tyr	Arg	Ile	Val	Gln	Phe	Gln	Ser	Lys			
252	CAA	CTT	GAG	AAC	CTT	CGA	GAC	TAC	CGG	ATT	GTC	CAG	TTT	CAG	TCA	AAA		299	
	::	:::	:::	:::	:::	:::	:::	:::	:::	:::	:	:::	:::	:	:::	:::			
820	CAG	CTT	CAG	AAT	CTC	CGA	GAC	TAC	CGC	ATT	TTG	GAG	TTC	AAC	TCC	AAA		867	
	Gln	Leu	Gln	Asn	Leu	Arg	Asp	Tyr	Arg	Ile	Leu	Glu	Phe	Asn	Ser	Lys			
	Pro	Asn	Thr	Ile	Leu	Leu	Pro	His	His	Ala	Asp	Ala	Asp	Phe	Leu	Leu			
300	CCC	AAC	ACA	ATC	CTT	CTC	CCC	CAC	CAT	GCT	GAC	GCC	GAT	TTC	CTC	CTC		347	
	:::	:::	:::	:	:::	:::	:::	:::	:::	:::	:::	:::	:::	:::	:::	:::			
868	CCC	AAC	ACC	CTT	CTT	CTC	CCC	CAC	CAT	GCT	GAC	GCT	GAT	TAC	CTC	ATC		915	
	Pro	Asn	Thr	Leu	Leu	Leu	Pro	His	His	Ala	Asp	Ala	Asp	Tyr	Leu	Ile			

Fig. 4. Sequence comparison between pGmg 91 β -subunit (*top*) and an α -subunit gene (*bottom*) starting near the TATA sequences (underlined) of each gene. DNA sequence analysis of the 5' end of the β -gene (indicated by the brackets in Fig. 2) was determined by the di-deoxy method described by Sanger (1977). Two large gaps, as well as other small gaps, have been added to the sequences to facilitate alignment. The first large gap, beginning after β -subunit nucleotide No. 51, extends 37 bases. The second large gap is indicated by a 537 nucleotide insert in the α sequence at base No. 172. Homology between nucleotides is indicated by *colons*. Amino-acid sequence derived by translating the nucleotide sequences is also presented. The sequence that corresponds to the first 26 amino acids of the NH₂-terminus of the mature β -subunit (Coates et al. 1985) is indicated by a *solid line*. A probable transcription start site (Slightom et al. 1983) is marked with an *arrow*

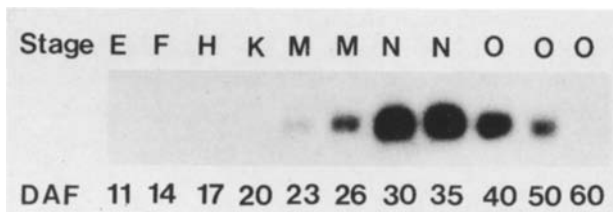


Fig. 5. RNA blot hybridization analysis of accumulation of the β -subunit transcript during soybean seed development. Total RNA was isolated from soybean seeds at different developmental stages (E–O) (Meinke et al. 1981) and separated on denaturing agarose gels. The RNA was transferred to nitrocellulose and probed with ^{32}P -labeled pGmg 91 as described in *Materials and methods*

one exception (indicated by * in Fig. 4), confirming that Gmg 91 codes for a β -subunit protein. The sequence determined for Gmg 91 was compared to the corresponding sequence of an independently isolated β -subunit genomic clone and was found to be identical (S.J. Barker, J.J. Harada, R.B. Goldberg, UCLA, personal communication).

The β -subunit gene sequence also shows strong homology with the corresponding sequence of an α' -subunit gene, pGmg 17.1 (Fig. 4). In this figure the α' - and β -subunit gene nucleotide sequences are aligned from their TATA boxes (underlined), through most of their first exons. Optimal alignment is achieved with the addition of two large gaps in the β -subunit sequence. The first gap is 37 nucleotides in length and is located in the 5' untranslated sequence. The second gap is much larger (537 nucleotides in length) and occurs in the first exon of the β -subunit gene (indicated by Δ in Fig. 4). Comparison of the derived amino-acid sequences of these genes following the second gap shows close homology, indicating that the addition of the gap has not altered the reading frame. If the 537-base α' insert is ignored, the α' and β se-

quences are approx. 75% identical at the nucleotide level within the coding sequences.

Analysis of β -mRNA accumulation in soybean seeds. During soybean seed development there is differential accumulation of the α' -, α - and β -subunits of β -conglycinin (Meinke et al. 1981; Ladin et al. 1987). While the α' - and α -subunits begin to accumulate 16–18 d after anthesis, the β -subunit does not accumulate for several additional days. To analyze the accumulation of the β -subunit mRNA directly, total RNA was isolated from developing soybean seeds at several stages of development (Meinke et al. 1981), transferred to a nitrocellulose filter, and hybridized with the ^{32}P -labeled β -genomic clone pGmg 91. Low levels of RNA encoding the β -subunit could be detected in embryos as early as stage K (data not shown), and continued to accumulate to high levels (Fig. 5). Accumulation of β -subunit protein was followed in developing soybean seeds using immunoblot analysis. As shown in Fig. 6, the β -subunit could be detected at stage L and continued to accumulate through stage R (400–450 mg/seed).

Analysis of β -subunit genomic organization. Like many other storage-protein genes in plants, β -conglycinin is encoded by a multi-gene family. To estimate the number of β -subunit genes present in soybean, DNA was isolated from soybean leaves, digested with either Eco RI or Hind III, size-separated by agarose-gel electrophoresis, and transferred to nitrocellulose. The DNA was probed with the 4.2-kb Hind III DNA fragment insert of pGmg 91, encoding a β -subunit gene. The labeled probe hybridized with a number of restriction fragments when soybean DNA was cut with either enzyme (Fig. 7), indicating that there are between 8 and 13 β -subunit gene sequences per haploid genome in soybean.

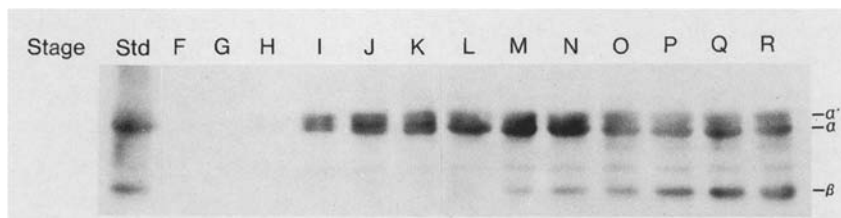


Fig. 6. Western blot analysis of β -conglycinin accumulation during soybean seed development. Protein was extracted from the seeds at different developmental stages (Meinke et al. 1981). The weights of the seeds at each of these stages were in mg/seed: F=5–10; G=10–25; H=25–45; I=45–65; J=65–85; K=85–120; L=120–160; M=160–200; N=200–250; O=250–300. Stages P–R represent later stages of seed development than those described by Meinke et al. (1981) in which the seed weights were, in mg/seed: P=300–350; Q=350–400; R=400–450. The proteins were separated by SDS-PAGE, transferred to nitrocellulose and incubated with β -conglycinin antibodies as described in *Materials and methods*. The standard lane (std) represents protein from dry seed

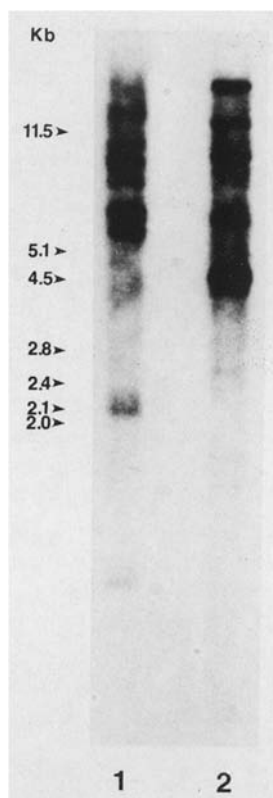


Fig. 7. Genomic DNA blot hybridization analysis of the β -subunit gene in soybean DNA. Genomic soybean DNA was isolated and digested with either Eco RI (1) or Hind III (2). The DNA was separated by electrophoresis on agarose gels, transferred to nitrocellulose, and probed with the ^{32}P -labeled Hind III insert of pGmg 91 containing the β -subunit, as described in *Materials and methods*

Discussion

In this paper we describe the isolation and partial characterization of a genomic clone (pGmg 91) that encodes the β -subunit of β -conglycinin in soybean. This clone shares sequence homology with the β -subunit of phaseolin, a related storage protein in *P. vulgaris*, as judged by the ability of phaseolin cDNAs to hybridize to genomic clones from a soybean library. It also shares homology with the α - and α' -subunits of β -conglycinin in soybean based on RNA hybridization reactions. The nucleotide sequence of the 5' end of pGmg 91 is perfectly homologous to an independently isolated β -subunit gene of Barker et al. (see Doyle et al. 1986). In contrast, the restriction map and DNA sequence of pGmg 91 shared fewer similarities with pGmg 17.1, an α' -subunit genomic clone (Schuler et al. 1982a).

Protein-sequence comparisons of β -conglycinin α' - and β -subunits demonstrate the high degree of

homology between these two proteins with the exception of a large insert near the amino terminus of the α' -subunit sequence (Doyle et al. 1986). If this large insert is omitted as in Fig. 4, the α' and β sequences are approx. 75% homologous. The close correlation between the amino-acid sequences derived from pGmg 91 and the amino-acid sequence of the first 26 NH_2 -terminal amino acids of the β -subunit (Coats et al. 1985) confirms the identity of this clone as a β -subunit gene of β -conglycinin.

The copy number of the β -gene sequence in the soybean genome was analyzed by hybridizing pGmg 91 to soybean DNA which had been digested with either Hind III or Eco RI. pGmg 91 hybridized to a number of restriction fragments of DNA from each of these reactions indicating that there are approx. 8–13 copies of the β -gene sequence per haploid genome.

Transgenic tobacco and petunia plants that contain and express pGmg 91 β -subunit mRNA have been prepared (Bray et al. 1987). Transcription products from this gene are detected in the developing seed, in a manner similar to that seen in soybean embryos. This analysis of pGmg 91 in transgenic plants indicates that the DNA sequences necessary for the developmental regulation of the β -subunit gene are intact. Further investigation of these sequences should provide important information about the DNA sequences and regulatory molecules that are necessary for the developmentally controlled expression of this interesting storage-protein gene.

This work was supported by research contract DE-ACO2-81ER19888 from the U.S. Department of Energy, Biological Energy Research to R.N.B. E.A.B. was supported by a fellowship from Monsanto Agricultural Products.

References

- Alwine, J.C., Kemp, D.J., Parker, B.A., Reiser, J., Renart, J., Stark, G.R., Wahl, G.M. (1979) Detection of specific RNAs or specific fragments of DNA by fractionation in gels and transfer to diazobenzoxymethyl paper. *Methods Enzymol.* **68**, 220–242
- Anderson, C.W., Strauss, J.W., Dudock, B.S. (1983) Preparation of a cell-free protein-synthesizing system from wheat germ. *Methods Enzymol.* **101**, 635–644
- Bailey, J.M., Davidson, N. (1976) Methylmercury as a reversible denaturing agent for agarose gel electrophoresis. *Anal. Biochem.* **70**, 75–85
- Beachy, R.N., Jarvis, N.P., Barton, K.A. (1981) Biosynthesis of subunits of the soybean 7S protein. *J. Mol. Appl. Genet.* **1**, 19–27
- Bray, E.A., Beachy, R.N. (1985) Regulation by ABA of β -conglycinin expression in cultured developing soybean cotyledons. *Plant Physiol.* **79**, 746–750
- Bray, E.A., Naito, S., Pan, N-S., Anderson, E., Dubé, P.,

- Beachy, R.N. (1987) Expression of the β -subunit of β -conglycinin in seeds of transgenic plants. *Planta* **172**, 364–370
- Bunemann, H., Westhoff, P. (1983) Hybrid selection of specific RNAs using DNA covalently coupled to macroporous supports. *Methods Enzymol.* **100**, 400–407
- Burnett, W.N. (1981) “Western blotting”: Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radiiodinated Protein A. *Anal. Biochem.* **112**, 195–203
- Coates, J.B., Medeiros, J.S., Thanh, V.H., Nielson, N.C. (1985) Characterization of the subunits of β -conglycinin. *Arch. Biochem. Biophys.* **243**, 184–194
- Dellaporta, S.L., Wood, J., Hicks, J.B. (1983) A Plant DNA Miniprep: Version II. *Plant Molec. Biol. Repr.* **1**, 19–21
- Doyle, J.J., Schuler, M.A., Godette, W.D., Zenger, V., Beachy, R.N., Slightom, J.L. (1986) The glycosylated seed storage proteins of *Glycine max* and *Phaseolus vulgaris*: structural homologies of genes and proteins. *J. Biol. Chem.* **261**, 9228–9238
- Holowach, L.P., Madison, J.T., Thompson, J.F. (1986) Studies on the mechanism of regulation of the mRNA for a soybean storage protein subunit by exogenous L-methionine. *Plant Physiol.* **80**, 561–567
- Karn, J., Brenner, S., Barnett, L., Cesareni, G. (1980) Novel bacteriophage λ cloning vector. *Proc. Natl. Acad. Sci. USA* **77**, 5172–5176
- Ladin, B.F., Tierney, M.L., Meinke, D.W., Hosangadi, P., Veith, M., Beachy, R.N. (1987) Developmental regulation of β -conglycinin in soybean axes and cotyledons. *Plant Physiol.* **84**, 35–41
- Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685
- Lingappa, V.R., Lingappa, J.R., Prasad, R., Ebner, K.E., Blobel, G. (1978) Coupled cell free synthesis, segregation and core-glycosylation of a secretory protein. *Proc. Natl. Acad. Sci. USA* **75**, 2338–2342
- Lycett, G.W., Delauney, A.J., Gatehouse, J.A., Gilroy, J., Croy, R.R.D., 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
- Maniatis, T., Fritsch, E.F., Sambrook, J. (1982) *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory, N.Y., USA
- Meinke, D.W., Chen, J., Beachy, R.N. (1981) Expression of storage protein genes during soybean seed development. *Planta* **153**, 130–139
- Potter, H., Dressler, D. (1986) A “Southern Cross” method for analysis of genome organization and the localization of transcription units. *Gene* **48**, 229–239
- Sanger, F., Nicklen, S., Coulson, A.R. (1979) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467
- Schuler, M.A., Schmitt, E.S., Beachy, R.N. (1982a) Closely related families of genes code for the α and α' subunits of the soybean 7S storage protein complex. *Nucleic Acids Res.* **10**, 8225–8244
- Schuler, M.A., Ladin, B.F., Polacco, J.C., Freyer, G., Beachy, R.N. (1982b) Structural sequences are conserved in the genes coding for the α , α' , and β -subunits of the soybean 7S seed storage protein. *Nucleic Acids Res.* **10**, 8245–8261
- Slightom, J.L., Blechi, A.E., Smithies, O. (1980) Human Fetal $G\gamma$ - and $A\gamma$ -Globin genes: complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. *Cell* **21**, 627–638
- Slightom, J.L., Sun, S.M., Hall, T.C. (1983) Complete nucleotide sequence of a French bean storage protein gene: Phaseolin. *Proc. Natl. Acad. Sci. USA* **80**, 1897–1901
- Symington, J., Green, M., Brackman, K. (1981) Immunoradiographic detection of proteins after electrophoretic transfer from gels to diazo-paper: Analysis of adenovirus encoded proteins. *Proc. Natl. Acad. Sci. USA* **78**, 177–181

Received 19 February; accepted 4 June 1987