Theor Appl Genet (1990) 79:785–792

THEORETICAL AND APPLIED GENETICS

Characterization of soybean vegetative storage proteins and genes *

W. D. Rapp **, G. G. Lilley** *** **and N. C. Nielsen** ****

USDA/Agricultural Research Service and the Department of Agronomy, Purdue University, West Lafayette, IN 47907, USA

Received September 6, 1989; Accepted January 22, 1990
Communicated by J. Beckmann

Summary. Soybean vegetative storage proteins (VSPs) were purified and characterized. Anion exchange HPLC resolved partially purified VSPs into fractions containing 27-kD/27-kD and 29-kD/29-kD homodimers and 27-kD/29-kD heterodimers. Reversed-phase HPLC resolved partially purified VSPs into three fractions. One fraction contained only 27-kD VSP and the other two contained 29-kD VSP. The two 29-kD VSP fractions differed with respect to their cyanogen bromide cleavage patterns, an observation that indicated the 29-kD VSPs were heterogeneous. Genomic clones that contained 29-kD VSP genes were also isolated and characterized. One genomic clone contained a complete 29-kD VSP gene and was sequenced. The coding region in the clone contained two introns whose borders had regulatory sequences typical of other eukaryotic genes. Putative polyadenlyation signals were present in the 3′-flanking region of the gene, while putative TATA, CAAT, and enhancer core sequences were found in the 5′-flanking regions. A second genomic clone that was studied contained the 5′ regions of two partial 29-kD VSP genes in an inverted linkage. Genomic DNA gel blots showed that the two genes were organized in the same arrangement in the soybean genome.

Key words: Vegetative storage protein – Nitrogen – Soybean – *Glycine max*

Introduction

Soybeans synthesize massive amounts of protein during seed development. This activity requires a large supply of nitrogen, carbon, and other nutrients, most of which are supplied by the vegetative tissues of the plant via the phloem transport system (Pate 1980; Simpson 1986). A significant amount of the nitrogen delivered to the seed comes from remobilization of proteins that have accumulated in the vegetative tissues (Thibodeau and Jaworski 1975; Sesay and Shibles 1980; Wittenbach et al. 1980). Although degradation of metabolic enzymes is responsible for supplying some of these compounds, other proteins have evolved that are temporary reservoirs for nitrogen and carbon (Wittenbach 1983 a,b). These vegetative storage proteins (VSPs) accumulate prior to the beginning of seed development, and can account for up to 15% of the total soluble protein in the leaves at the time of flowering (Wittenbach 1983 b; Staswick 1988). When the pods and seeds begin to develop, VSP levels decrease. As seed maturation nears completion, VSPs begin to accumulate again in the vegetative tissues (Wittenbach 1983 a,b; Staswick 1989).

The concentration of VSP in vegetative tissues can be manipulated artificially. Decreasing sink strength, either by depodding or interruption of phloem transport, causes VSPs to accumulate to high levels in the leaves. In plants that have been continually depodded for a number of weeks, VSPs account for over 50% of the total soluble leaf protein (Wittenbach 1983 b).

Partially purified VSPs have been characterized (Wittenbach 1983 b). They are glycoproteins with apparent molecular weights of 27-kD and 29-kD in SDS-polyacrylamide gels, and they behave as dimers on a gel filtration column. The amino acid sequences of these polypeptides have been derived from the nucleotide sequences of 27-

* Cooperative research between USDA/Agricultural Research Service and the Indiana Agricultural Experiment Station. Journal Paper No. 12, 192 from the Indiana Agricultural Experiment Station

** Present address: Boyce Thompson Institute, Cornell University, Ithaca, NY 14853, USA
*** Present address: CSIRO, Division of Biotechnology, 343 Royal Parade, Parkville, Victoria, 3052, Australia
**** To whom correspondence should be addressed

kD and 29-kD VSP cDNA clones, and a comparison of these sequences reveals that the polypeptides are homologous (Staswick 1988).

Genomic DNA gel blot hybridization experiments have revealed that there are approximately two genes for each size class of VSP (Staswick 1988). Expression of the 27-kD and 29-kD VSP genes is tightly regulated. Accumulation of VSP mRNAs parallel VSP accumulation in plants that are allowed to develop normally and in plants that have been manipulated to increase VSP levels (Staswick 1988, 1989). Jasmonic acid, a proposed phytohormone (Meyer et al. 1984), causes the concentration of the 29-kD VSP to increase in soybean suspension cultures. It also increases the content of 27-kD and 29-kD VSPs when sprayed on the leaves of soybean plants (Anderson et al. 1989). Consequently, it is thought that jasmonic acid is involved in the regulation of VSP accumulation.

A more thorough knowledge of the structure of the mature VSPs, as well as an understanding of how their synthesis and accumulation are regulated, could lead to a better understanding of the role that VSPs play in the metabolism of nitrogen in plants. In this paper, we report further purification and characterization of the 27-kD and 29-kD VSPs and the isolation and characterization of two 29-kD VSP genomic clones.

Materials and methods

Plant material

Soybean plants (variety Century) grown in the greenhouse were depodded at weekly intervals beginning 1 week after flowering. Leaves from which protein was purified were collected from depodded plants 6–7 weeks after flowering and immediately frozen in liquid N_2. Leaves from which DNA was isolated (variety Forrest) were collected from plants grown both in the field and in the greenhouse. Leaves were stored at −80 °C.

Protein purification

A modification of the method of Wittenbach (1983 b) was used to obtain a fraction highly enriched in the 27-kD and 29-kD VSPs. Leaves from depodded plants (variety Century) were homogenized in 20 m*M* Tris-HCl (pH 7.6), 4 m*M* DTT, 1 m*M* EDTA, and 0.1 m*M* 1,10-phenanthroline. The extract was fractionated with ammonium sulfate. Protein that precipitated from the extract between 2.0 and 2.8 *M* ammonium sulfate was desalted and subjected to Con A-Sepharose column chromatography. The column buffer contained 20 m*M* triethanolamine-HCl (pH 7.6), 0.5 *M* NaCl, and 0.02% (w/v) sodium azide.

High performance liquid chromatography

Con A-Sepharose purified protein (Con A-2) was further fractionated by either reversed-phase HPLC or anion exchange HPLC. For reversed-phase HPLC, a Synchrom RP-P column (250 mm × 4.1 mm) with a mobile phase of 0.1% (v/v) trifluoroacetic acid (TFA) was used. Lyophilized Con A-2 was dissolved in 0.1% (v/v) TFA to give a concentration of 2 mg/ml, and then 0.5 ml was chromatographed at a buffer flow rate of 1 ml/min. After adjusting the acetonitrile concentration in the solvent to 36% during an initial 2-min period, a 36%–44% (v/v) acetonitrile gradient (0.5% increase per min) was used to elute VSPs.

For anion exchange HPLC, a Synchrom AX300 column (250 mm × 4.1 mm) with a mobile phase of 25 m*M* phosphate buffer (pH 6.5) was used. Con A-2 protein was concentrated by ultrafiltration and dialyzed against 25 m*M* phosphate buffer (pH 6.5). Of this protein solution 0.5 ml (2 mg protein) was fractionated chromatographically with a 0–0.2 *M* NaCl gradient (1% increase in concentration per min; flow rate 1 ml/min). Protein peaks as determined by A_{280} were collected manually.

Cross-linking

Proteins were cross-linked by a modification of the method of Siezen et al. (1980). Aliquots of anion exchange fractions containing 50 mg protein were brought to final volumes of 1 ml by the addition of 25 m*M* phosphate buffer (pH 6.5). After the addition of 0.4 ml of acetonitrile that contained 8 mg/ml of the cross-linking reagent dithiobis (succinimidyl propionate), the samples were incubated at room temperature for 30 min and then dialyzed overnight against 50 m*M* phosphate buffer (pH 7.0). The samples were dialyzed against distilled water, lyophilized, and analyzed by SDS-PAGE in the absence of 2-mercaptoethanol.

CNBr mapping

Protein was cleaved at methionyl residues by dissolving it in 1 ml of 70% (v/v) formic acid and adding a 100-fold molar excess (over methionine) of CNBr (Nute and Mahoney 1979). Digestion products were visualized on 20% SDS-polyacrylamide gels.

Construction and screening of genomic libraries

Genomic clone λ 20-2B was isolated from an EcoRI partial digestion library of genomic DNA from the soybean variety Forrest (Fischer and Goldberg 1982). About 3.9×10^5 recombinants wre screened using as probe the insert from pVSP25, a 27-kD VSP cDNA clone obtained from Staswick (1988). Genomic clone λ 2A2 was isolated from a library generated by digesting leaf DNA to completion with XbaI, size-fractionating the digestion produts in 0.6% agarose, pooling those fractions that contained DNA fragments in the 7- to 12-kb range, and cloning those fragments into vector LambdaGEM-12. About 1.4×10^5 recombinants were screened using as probe the insert from pVSP27, a 29-kD VSP cDNA clone obtained from Staswick (1988). Plaque hybridizations (Benton and Davis 1977) were in 50% formamide and 1 *M* Na^+ at 42 °C. Following hybridizations the filters were washed in 0.1 × SSC (0.3 *M* NaCl/0.034 sodium citrate, pH 7.0) and 0.1% (w/v) SDS at 53 °C, and exposed to film overnight. The 10.5-kb EcoRI fragment from λ20-2B was subcloned into pGEM-3blue to yield pLLPG5. The 10-kb XbaI fragment λ2A2 was inserted into the same vector to yield p2A2-1. *Escherichia coli* strains DH1 and NM539 were used as hosts for the screening of the EcoRI partial digestion library and the Xbal size-selected library, respectively. Plasmids were propagated in *E. coli* strains TB1 and DH1.

DNA gel blot hybridization

Leaf DNA (variety Forrest) was digested with the appropriate restriction endonucleases, and 5 mg per lane separated electrophoretically in 0.6% agarose gels. After electrophoresis, the DNA was transferred to GeneScreen Plus as described in New England Nuclear catalog NEF-976. The membranes were probed either with the nick-translated pVSP27 insert or a nick-translated 3.3-kb ScaI fragment isolated from pLLPG5. Hybridization were at 53 °C in 50% formamide and 1 *M* Na^+.

Membranes were washed after hybridization in 0.1 × SSC and 0.1% SDS at 57 °C, and were exposed to film for several days with an intensifier screen.

DNA sequencing

Sequencing was performed according to the method of Sanger et al. (1977) using CsCl-purified plasmid DNA as template. The nucleotide sequence of both strands of the 29-kD VSP gene and its flanking regions was determined.

Results and discussion

Vegetative storage protein subunits are organized into homodimers and heterodimers

Vegetative storage proteins (VSPs) were purified from leaves of soybean plants that had been depodded at weekly intervals for 6–7 weeks after flowering. Because the VSPs were glycoproteins, they were purified by Con A-Sepharose chromatography (Wittenbach 1983 b). Elution of the VSP-enriched fraction from the Con A-Sepharose column was delayed relative to the breakthrough peak and was similar in yield to that in earlier reports. However, its elution from the column did not require washing with buffer containing mehtyl-α-D-mannopyranoside, as in previous reports (Wittenbach 1983 b).

The VSP-enriched fraction was further resolved by anion exchange HPLC (Fig. 1 A). Because these separations were done in aqueous buffers, the proteins presumably remained in their native conformations. Five major peaks were collected and denoted AEX 1 through AEX 5. The content of each fraction was examined by SDS-PAGE. Figure 1 B shows that AEX 1 contained only 35-kD proteins when reduced with 2-mercaptoethanol. The proteins from the AEX 1 fraction had hemagglutinating activity identical in specificity to that of Con A, and was due to lectins leached from the column during chromatography (Rapp 1989). When reduced with 2-mercaptoethanol, fraction AEX 2 contained 27-kD VSP plus small amounts of 35-kD protein from the tail of AEX 1. AEX 3 contained approximately equal amounts of the 27-kD and 29-kD VSPs, and AEX 4 contained mostly 29-kD VSPs. AEX 5 (not pictured in Fig. 1 B) contained a mixture of 27-kD, 29-kD, and 35-kD proteins.

Wittenbach (1983 b) has shown that when partially purified mixtures of the 27-kD and 29-kD VSPs are chromatographed on a gel filtration column, they migrate as dimers. With this in mind, the observation that the AEX 3 peak contained both 27-kD and 29-kD VSPs in apparently equal amounts, and that this peak eluted between the peaks containing only 27-kD VSPs (AEX 2) and only 29-kD VSPs (AEX 4), suggested that the proteins eluted from the anion exchange column as dimers. Heterodimers of 27- and 29-kD VSPs should exhibit charge characteristics intermediate between homodimers of 27- and 29-kD, respectively. This interpretation was tested by cross-linking the proteins in fractions AEX 2, AEX 3, and AEX 4 using dithiobis (succinimidyl propionate). Figure 2 shows that after the proteins

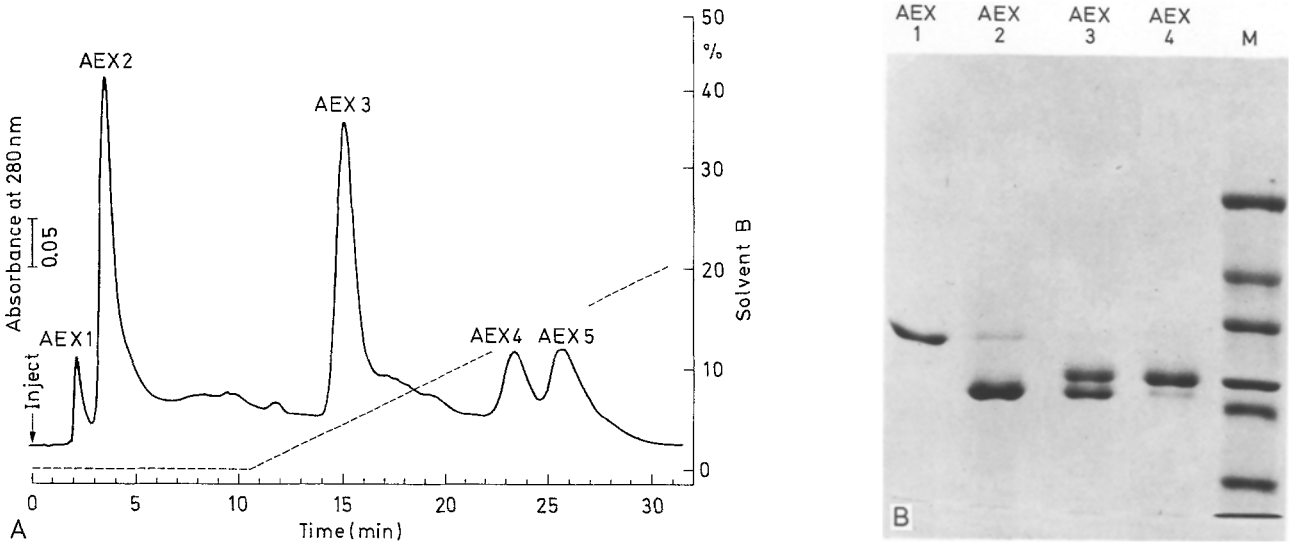

Fig. 1 A and B. Fractionation of native Con A-2 protein by anion exchange HPLC. **A** Anion-exchange HPLC elution profile. Solvent A was 25 m*M* phosphate buffer (pH 6.5) and solvent B was 50 m*M* phosphate buffer (pH 6.5) that contained 1 *M* NaCl. The flow rate was 1 ml/min. Fractions AEX 1 through AEX 5 were collected manually. **B** SDS-PAGE of fractions AEX 1 through AEX 4. Aliquots of each fraction were dialyzed in 8% formic acid and lyophilized. Protein was dissolved in loading buffer containing 2-mercaptoethanol and subjected to SDS-PAGE according to the method of Laemmli (1970). *Lane M* contains molecular-weight standards (66, 45, 36, 29, 24, and 20 kD)

in these fractions were cross-linked, they had apparent molecular weights between 52 kD and 55 kD in SDS-polyacrylamide gels, as would be expected for dimers. Therefore, we conclude that VSPs exist as 27-kD/27-kD and 29-kD/29-kD homodimers and as 27-kD/29-kD heterodimers. Because the electrophoretic mobilities of the 27- and 29-kD VSPs are similar in the presence and absence of 2-mercaptoethanol (data not shown), it appears that disulfide bonds are not involved in dimer formation.

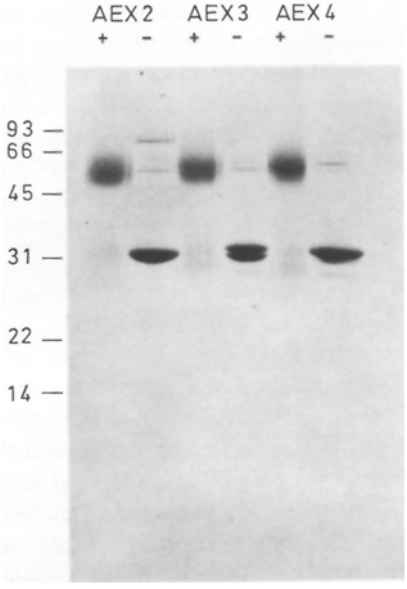

Fig. 2. Cross-linking of AEX fractions. AEX 2, AEX 3, and AEX 4 were allowed to react with the cross-linking reagent dithiobis (succinimidyl propionate) (lanes marked +). Lanes marked (−) contain AEX fractions incubated under the same conditions but without added cross-linker. Positions and sizes of molecular-weight standards are indicated

The 29-kD VSPs are heterogeneous

Figure 3 A shows that fraction Con A-2 was resolved into three fractions by reversed-phase HPLC. These separations were made after denaturation, which presumably eliminated higher order protein structure. Examination of these fractions by SDS-PAGE showed that RP-27 contained only 27-kD VSP and that RP-29A and RP-29B contained 29-kD VSP (Fig. 3 B). The 35-kD protein present in Con A-2 eluted as a shoulder on the back of the peak containing RP-29B. N-terminal amino acid sequence analyses of these fractions by Edman degradation (Rapp 1989) showed that the N-terminal sequence of the 27-kD protein was NH_2-Ala-Arg-Thr-Pro-Glu-Val-Lys. The N-terminus of the protein corresponded to the alanyl residue at position 35 of the initial translation product deduced from the cDNA nucleotide sequence reported by Staswick (1988) and Mason et al. (1988). The 29-kD proteins both had the N-terminal sequence NH_2-Glu-Arg-Ser-Glu-Val-Lys, the N-terminal residue corresponding to the glutamic acid residue at position 36 in its corresponding deduced sequence. The N-termini of the mature VSP proteins that we isolated from leaves of depodded plants were identical of the N-termini of VSPs

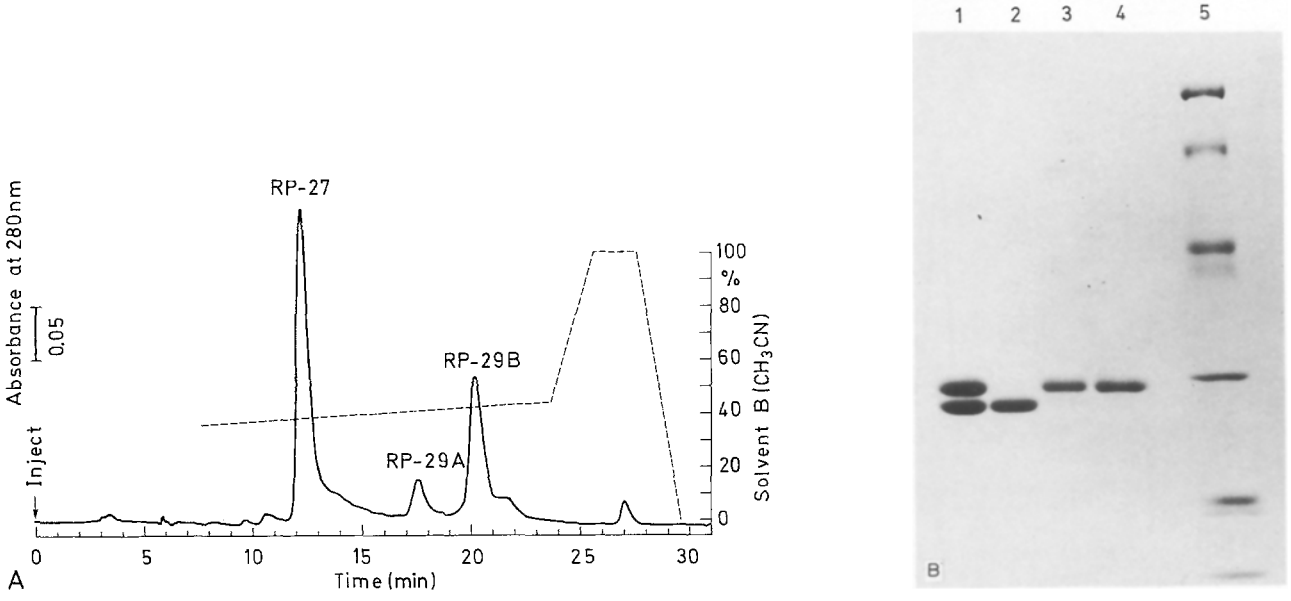

Fig. 3 A and B. Fractionation of denatured Con A-2 protein by reversed-phase HPLC. **A** Reversed-phase HPLC elution profile. Solvent A was 0.1% (v/v) TFA and solvent B was acetonitrile containing 0.1% (v/v) TFA. The flow rate was 1 ml/min. RP-27, RP-29 A, and RP-29B were collected manually. **B** SDS-PAGE of fractions RP-27 (*lane* 2), RP-29A (*lane* 3), and RP-29B (*lane* 4). Lyophilized protein was treated as described in Fig. 1. *Lane* 1 contains fraction Con A-2. *Lane M* contains molecular-weight standards (93, 66, 45, 31, 22, and 14 kD)

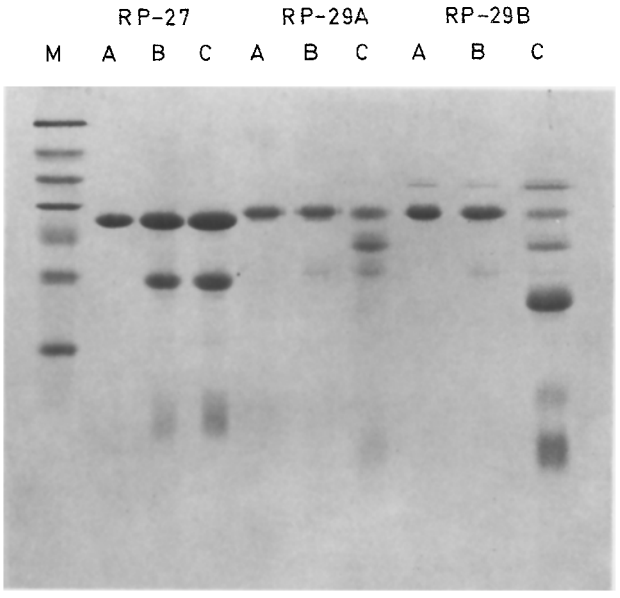

Fig. 4. Cyanogen bromide peptides from of RP-27, RP-29A, and RP-29B. Proteins were cleaved by CNBr in 70% formic acid (lanes marked *C*), exposed to 70% formic acid only (lanes marked *B*), or were not treated with either formic acid or CNBr (lanes marked *C*). The 35-kD protein present in the RP-29B fraction is not cleaved by CNBr (data not shown). Sizes of molecular-weight standards (*lane M*) are indicated in Fig. 1

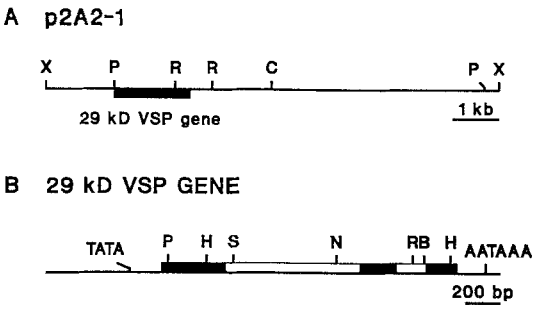

Fig. 5 A and B. Restriction maps of 29-kD VSP gene and flanking regions. **A** Clone p2A2-1 contains a complete 29-kD VSP gene on a 10.5-kb insert (introns and exons not shown). **B** The 29-kD VSP gene has three exons (*dark boxes*) and two introns (*light boxes*). A putative TATA box and polyadenylation (AATAA) signal were identified in the 5′-and 3′-flanking DNA, respectively

isolated from stems of soybean seedlings by Mason et al. (1988).

To compare RP-27, RP-29A, and RP-29B with respect to the number and position of methionyl residues, the products resulting from CNBr cleavage were compared on SDS-polyacrylamide gels. Figure 4 shows that RP-27 was not cleaved by CNBr, an observation that indicated methionyl residues were absent from this subunit. This result was expected because the deduced amino acid sequence of the mature 27-kD VSP contained no methionyl residues (Staswick 1988). Two bands were present in the digestion products of RP-29B that were absent from those of RP-29A. This indicated that RP-29A and RP-29B differed with respect to the position of at least one methionine residue, and that RP-29A and RP-29B were encoded by two different gene sequences. Because the VSPs were isolated from leaves that were collected from a population of Century soybean plants, and because allelic heterogenity fixed in populations of soybean varieties can occur using conventional breeding approaches, our data does not indicate whether the sequences originated from different genes or from alleles of the same gene in different plants.

A 29-kD VSP gene contains three exons and two introns

Based on the results of genomic DNA gel blot hybridizations done in this laboratory (data not shown) and by Staswick (1988), and on the analysis of a genomic clone that contained two 29-kD VSP coding sequences (see below), we determined that the complete 29-kD VSP gene should be contained on a 10-kb XbaI genomic DNA fragment. A Forrest genomic DNA library of size-fractionated XbaI fragments (7–12 kb) was constructed and screened, using as probe the insert from pVSP27. The insert encoded a 29-kD VSP cDNA obtained from Staswick (1988). One clone, λ2A2, was isolated and studied extensively.

Clone λ2A2 contained inserted XbaI fragments of approximately 7 and 10kb. The presence of two XbaI fragments was presumed due to a multiple insertion event. DNA gel blots of XbaI-digested λ2A2 DNA showed that only the 10-kb XbaI fragment hybridized to the 29-kD VSP cDNA probe (data not shown). This 10-kb fragment was subcloned into pGEM-3blue, and the resulting clone was called p2A2-1. The 7-kb insert was not studied further. Figure 5 shows the restriction map of p2A2-1, and Fig. 6 contains the complete nucleotide sequence of a 29-kD VSP gene and its flanking regions.

The 29-kD VSP gene from p2A2-1 is composed of three exons of 394, 210, and 182 bp, and two introns of 786 and 163 bp. Because the nucleotide sequence of a cDNA encoding a 29-kD VSP with its 5′ and 3′ transcribed noncoding regions matched exactly the sequence of the 29-kD VSP gene that was isolated and sequenced, we considered it likely that this gene was transcribed in the plant. Examination of the gene sequence revealed that putative CAAT and TATA boxes are located 142 and 83 bp upstream form the initiation codon, respectively. The 5′-flanking region also contained a repeated sequence motif with homology to the consensus enhancer core sequence TGTGG(AA/TT)AG from animal genes (Sassone-Corsi and Borrelli 1986). Variations of the consensus sequence TGTTGT(A/T)(G/T) are repeated six

Fig. 6. DNA sequence and derived amino acid sequence of a 29-kD VSP gene. Putative CAAT and TATA boxes (*) in the 5′-flanking region, and a polyadenylation signal (+) in the 3′-flanking region, are overmarked. The first amino acid of the mature 29-kD VSP is indicated (^). Regions similar in sequence to a consensus enhancer core sequence are *underlined*. A repeated 10 bp sequence in the 5′-flanking region is shown in *bold*

times within the 620-bp region immediately upstream from the initiation codon of the 29-kD VSP gene (Fig. 6). Two of these sequences were part of a larger 10-bp direct repeat, ATTTCTGTTG. This repeat occured twice with the first base of each repeat separated by 27 base pairs. Because sequences displaying homology to animal enhancer core sequences have been observed in the 5′-flanking region of the pea *rbs-3A* gene, and because these have been shown to function as *cis*-acting regulatory elements (Kuhlemeier et al. 1987, 1988), it is possible that the repeats in the 5′-flanking region of the 29-kD VSP gene have a similar function.

Two genes that encode 29-kD VSPs are in an inverted linkage

A second genomic clone (λ20-2B) that contained 29-kD VSP coding sequences was isolated from a Forrest EcoRI

partial digestion genomic library. It was identified using the insert from pVSP25, a 27-kD VSP cDNA (Staswick 1988). The probe was used at a hybridization stringency that allowed coding sequences for the 27-kD and 29-kD VSPs to cross-hybridize. DNA gel blot hybridizations showed that λ20-2B contained a 10.5-kb EcoRI fragment that hybridized with the 27-kD VSP cDNA probe (data not shown). To facilitate further analysis, the 10.5-kb EcoRI fragment was subcloned into pGEM-3blue to yield pLLPG5. When duplicate DNA gel blots of pLLPG5 were hybridized with nick-translated 27-kD and 29-kD VSP cDNA probes at a higher stringency, only the 29-kD VSP probe hybridized to the insert (data not shown). Furthermore, the probe hybridized to the outermost 1.5 kb of each end but not to the central region of the insert.

A restriction map of the 10.5-kb EcoRI insert in pLLPG5 was prepared and, as indicated in Fig. 7, a symmetrical arrangement of restriction sites was observed. For every restriction site mapped within the outer 3.6-kb region of one end, an identical site was located the same distance in from the opposite end (the region between the two internal ScaI sites was not mapped). Taken together, these data indicated that clone λ20-2B contained a fragment on which at least parts of two 29-kD VSP genes were linked in the opposite orientation. Sequence analyses of region on each end of the insert (data not shown), together with a comparison of the restriction map of pLLPG5 to p2A2-1, permitted us to conclude that each end of clone λ20-2B contained the first two (out of three total) exons and one plus part of the second (out of two total) introns.

Genomic DNA gel blot hybridizations, performed in this laboratory (data not shown) and by Staswick (1988), showed that the 29-kD VSP cDNA probe hybridizes strongly to a 10.5-kb EcoRI genomic fragment. This result would be anticipated if the arrangement of the genes in pLLPG5 represented their arrangement in the soybean genome. To determine if the gene were linked in this manner in the genome, we hybridized DNA gel blots of restriction enzyme digested Forrest genomic DNA with a 3.3-kb genome, ScaI probe isolated from the middle of the pLLPG5 insert (Fig. 7). Based on the restriction map of pLLPG5, we expected this probe to hybridize with restriction fragments of the following sizes: EcoRI, 10.5 kb; NdeI, 9.6 kb; SphI, 8.5 kb; PstI, 7.5 kb; XbaI, 4.5 kb; ScaI, 3.3 kb. Figure 8 shows that the probe hybridized to fragments in the expected size range. We concluded that two 29-kD VSP genes were linked in the soybean genome in the same manner as in pLLPG5.

As indicated above, regions on each end of the pLLPG5 insert were sequenced, including parts of exons, introns, and flanking regions. The corresponding areas in each partial gene had identical sequences (data not shown). Moreover, the sequence of the corresponding

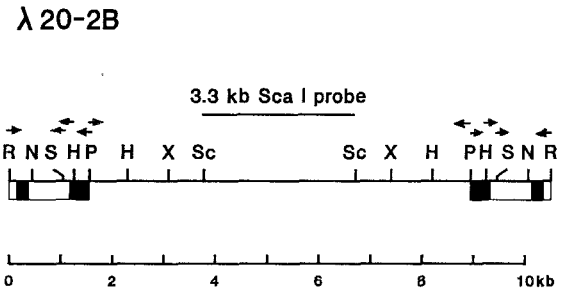

Fig. 7. Restriction map of the 10.5-kb insert from pLLPG5. Exon and intron regions from the 5′ portions of two 29-kD VSP genes are represented by *filled* and *open boxes*, respectively. Regions marked by *arrows* were sequenced

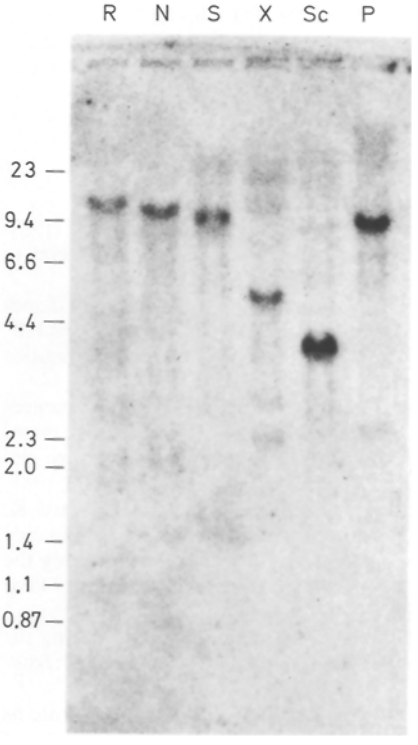

Fig. 8. DNA gel blot of restriction-enzyme-digested Forrest leaf DNA probed with a nick-translated 3.3-kb ScaI fragment isolated from pLLPG5 (Fig. 7). N, NdeI; P, PstI; R, EcoRI; S, SphI; Sc, ScaI; X, XbaI

regions of the complete gene in p2A2-1 matched exactly the nucleotide sequences of the partial genes. Because of the high degree of sequence conservation among the two inverted 29-kD VSP genes, it seems likely that the two genes are the product of a fairly recent gene duplication event. However, the functional significance, if any, of the unusual inverted linkage between the genes remains to be evaluated.

In summary, our studies characterize several soybean VSPs and an example of a gene that encodes them. The results show that VSP subunits are apparently organized

into both homodimers and heterodimers, and that there is heterogeneity among 29-kD VSPs with regard to their primary protein structures. We also report the isolation of VSP genomic clones and the first complete nucleotide sequence of a VSP gene. Finally, we demonstrate that two genes in an inverted tandem repeat encode 29-kD VSPs. This information will provide a starting point to identify *cis*-acting regulatory elements that are associated with these genes, and to deduce the role of jasmonic acid in regulation of their expression. It will also be interesting to explore how VSP composition and accumulation affect seed development and quality.

Acknowledgements. We wish to thank Drs. M. Hermodson and S. Buckle for their advice with respect to protein purification and characterization. We also thank Dr. T.-J. Cho for his advice and help concerning library screening and other techniques, and Dr. G. Kohlhaw for comments made during preparation of this manuscript.

References

Anderson JM, Spilatro SR, Klauer SF, Franceschi VR (1989) Jasmonic acid-dependent increase in the level of vegetative storage proteins in soybean. Plant Sci 62:45–52

Benton WD, Davis RW (1977) Screening λgt recombinant clones by hybridization to single plaques in situ. Science 196:180

Fischer RL, Goldberg RB (1982) Structure and flanking regions of soybean seed protein genes. Cell 29:651–660

Kuhlemeier C, Fluhr R, Green PJ, Chua N-H (1987) Sequences in the pea *rbcS-3A* gene have homology to constitutive mammalian enhancers but function as negative regulatory elements. Genes Dev 1:247–255

Kuhlemeier C, Cuozzo M, Green PJ, Goyvaerts E, Ward K, Chua N-H (1988) Localization and conditional redundancy of regulatory elements in *rbcS-3A*, a pea gene encoding the small subunit of ribulose-biphosphate carboxylase. Proc Natl Acad Sci USA 85:4662–4666

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685

Mason HS, Guerrero FD, Boyer JS, Mullet JE (1988) Proteins homologous to leaf glycoproteins are abundant in stems of dark-grown soybean seedlings. Analysis of proteins and cDNAs. Plant Mol Biol 11:845–856

Meyer A, Miersch O, Buttner C, Dathe W, Sembdner G (1984) Occurrence of the plant growth regulator jasmonic acid in plants. J Plant Growth Reg 3:1–8

Nute PE, Mahoney WC (1979) Complete amino acid sequence of the λ chain from the major fetal hemaglobin of the pig-tailed macaque, *Macaca nemestrina*. Biochemistry 18:467–472

Pate JS (1980) Transport and partitioning of nitrogenous solutes. Annu Rev Plant Physiol 31:313–340

Rapp, WD (1989) Isolation and characterization of soybean vegetative storage proteins and their genes. PhD Thesis, Purdue University, West Lafayette/IN

Sanger F, Nicklen S, Coulsen AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Sassone-Corsi P, Borrelli E (1986) Transcriptional regulation by *trans*-acting factors. Trends Genet 2:215–219

Sesay A, Shibles R (1980) Mineral depletion and leaf senescence in soybean as influenced by foliar nutrient application during seed filling. Ann Bot 45:47–55

Siezen RJ, Bindels JG, Hoenders HJ (1980) The quaternary structure of bovine α-crystallin: chemical cross-linking with bifunctional imido esters. Eur J Biochem 107:243–249

Simpson RJ (1986) Translocation and metabolism of nitrogen: whole plant aspects. In: Lambers H, Neeteson JJ, Stulen I (eds) Fundamental, ecological, and agricultural aspects of nitrogen metabolism in higher plants. Nijhoff, Dordrecht, pp 71–96

Staswick PE (1988) Soybean vegetative storage protein structure and gene expression. Plant Physiol 87:250–254

Staswick PE (1989) Developmental regulation and the influence of plant sinks on vegetative storage protein gene expression in soybean leaves. Plant Physiol 89:309–315

Thibodeau PS, Jaworski EG (1975) Patterns of nitrogen utilization in the soybean. Planta 127:133–147

Wittenbach VA (1983 a) Effect of pod removal on leaf photosynthesis and soluble protein composition of field-grown soybeans. Plant Physiol 73:121–124

Wittenbach VA (1983 b) Purification and characterization of a soybean leaf storage protein glycoprotein. Plant Physiol 73:125–129

Wittenbach VA, Ackerson RC, Giaquinta RT, Hebert RR (1980) Changes in photosynthesis, ribulose bisphosphate carboxylase, proteolytic activity, and ultrastructure of soybean leaves during senescence. Crop Sci 20:225–231