The Plant Cell, Vol. 1, 609–621, June 1989

Cotyledon Nuclear Proteins Bind to DNA Fragments Harboring Regulatory Elements of Phytohemagglutinin Genes

C. Daniel Riggs, Toni A. Voelker,[1] and Maarten J. Chrispeels[2]

Department of Biology and Center for Molecular Genetics, University of California, San Diego, La Jolla, California 92093-0116

The effects of deleting DNA sequences upstream from the phytohemagglutinin-L gene of *Phaseolus vulgaris* have been examined with respect to the level of gene product produced in the seeds of transgenic tobacco. Our studies indicate that several upstream regions quantitatively modulate expression. Between −1000 and −675, a negative regulatory element reduces expression approximately threefold relative to shorter deletion mutants that do not contain this region. Positive regulatory elements lie between −550 and −125 and, compared with constructs containing only 125 base pairs of upstream sequences (−125), the presence of these two regions can be correlated with a 25-fold and a 200-fold enhancement of phytohemagglutinin-L levels. These experiments were complemented by gel retardation assays, which demonstrated that two of the three regions bind cotyledon nuclear proteins from mid-mature seeds. One of the binding sites maps near a DNA sequence that is highly homologous to protein binding domains located upstream from the soybean seed lectin and Kunitz trypsin inhibitor genes. Competition experiments demonstrated that the upstream regions of a bean β-phaseolin gene, the soybean seed lectin gene, and an oligonucleotide from the upstream region of the trypsin inhibitor gene can compete differentially for factor binding. We suggest that these legume genes may be regulated in part by evolutionarily conserved protein/DNA interactions.

INTRODUCTION

Embryogenesis in leguminous plants is accompanied by the transcriptional activation of the seed storage protein genes and the subsequent accumulation of the reserve proteins within vacuoles of the developing cotyledons (Gatehouse et al., 1986). In the common bean, the seed storage protein/lectin phytohemagglutinin (PHA) may account for up to 10% of the total protein at seed maturity. Mature PHA is a tetrameric glycoprotein composed of one or both of the subunits, PHA-E (erythroagglutinating) and/or PHA-L (leukoagglutinating). These subunits are encoded by two highly homologous and tandemly linked genes: *lec1* (formerly dlec1), which encodes PHA-E, and *lec2* (formerly dlec2), which encodes PHA-L (Hoffman and Donaldson, 1985). Transcriptional activation of the genes occurs early in seed ontogeny (approximately 10 days after fertilization), and mRNA levels quickly peak and fall thereafter as the cotyledons mature (approximately 30 days) and PHA accumulates in protein storage vacuoles known as protein bodies (Staswick and Chrispeels, 1984; Chappell and Chrispeels, 1986). In contrast, little or no PHA is found in the vegetative tissues of the plant (Borrebaeck, 1984). Upon germination, the reserve proteins are hydrolyzed to provide nutrients for the developing seedling. These characteristics make seed storage proteins like PHA excellent model systems with which to study tissue-specific and developmentally regulated gene expression.

Molecular analyses of the PHA genes of a high expression cultivar, Greensleeves, as well as from a lectin-deficient cultivar, Pinto UI111 (Voelker, Staswick, and Chrispeels, 1986), revealed that the four genes contain a substantial amount (over 300 bp of nearly identical sequences) of sequence homology 5′ to the coding region. In three of the four promoter regions, some of this homology is due to the presence of a 63-bp direct repeat, which could play a role in PHA gene regulation. Subsequent transformation studies, employing the reporter gene luciferase, demonstrated that the upstream flanking region of *lec2* directs proper temporal and spatial regulation of gene expression in transgenic tobacco (Riggs et al., 1989) and indicates that regulatory elements governing PHA gene expression lie within 1 kb of the coding region.

We have sought to identify sequences upstream from the *lec2* gene that modulate gene expression. Of particular interest was the 63-bp repeat structure, as this region

[1] Current address: Calgene, Inc., 1920 Fifth Street, Davis, CA 95616.

[2] To whom correspondence should be addressed.

contains multiple copies of the CATGCATG motif identified by Dickinson, Evans, and Nielsen (1988) as being present upstream from all legume seed storage protein genes sequenced to date. To facilitate such an analysis, we introduced 5′ deletion mutants of the *lec2* gene into the tobacco genome and monitored gene expression by immunoblotting. These experiments demonstrated that two regions up-regulate PHA gene expression and identified a region further upstream that attenuates expression. Gel retardation assays, using defined restriction fragments and synthetic oligonucleotides to compete for factor binding, were then employed to map the sites to which cotyledon nuclear proteins bind. In this report, we describe the DNA sequences that likely are responsible for modulating PHA expression in normal and lectin-deficient cultivars, and present evidence that these sequences may modulate the expression of related genes in bean and soybean.

RESULTS

We have shown previously that proper temporal and spatial control of PHA-L gene expression occurs in transgenic tobacco, and that the difference in the level of gene product produced in a normal and a lectin-deficient cultivar of bean is maintained in tobacco following transfer of the cognate genes (Voelker, Sturm, and Chrispeels, 1987). Additionally, PHA is correctly glycosylated, assembled into tetramers, and targeted to the protein bodies; this also parallels the situation found in beans (Sturm, Voelker, and Chrispeels, 1988). Thus, the cellular machinery of tobacco not only recognizes the nucleotide sequences that govern gene expression, but also distinguishes PHA as a storage protein for correct post-translational modifications and targeting to the proper subcellular compartment. These characteristics indicate that transgenic tobacco is an excellent system with which to conduct functional analyses of the promoters of PHA genes from normal and lectin-deficient cultivars.

Construction of Promoter Deletion Mutants of *lec2* and Analysis of Integrations into the Tobacco Genome

To investigate the potential involvement of upstream sequences in PHA gene regulation, we constructed a series of progressive deletions from the 5′ end of a cloned fragment [beginning at −1191 relative to the translation initiation codon (see Methods)], introduced these constructs into the tobacco genome, and monitored gene expression by immunoblotting with antibodies directed against PHA. A schematic of the wild-type *lec2* gene fragment and the locations of deletion endpoints are shown in Figure 1. The deletion mutants were cloned into the binary vector Bin19, and a number of transgenic plants

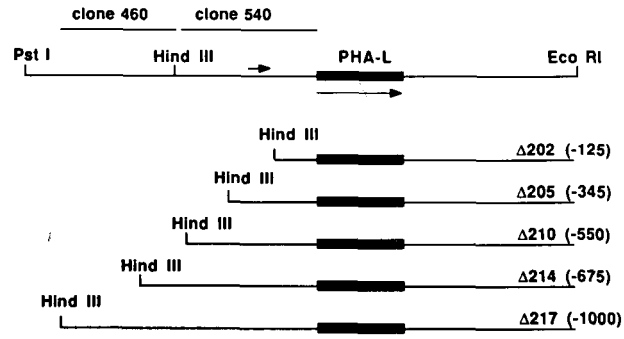

Figure 1. Structure of 5′ Deletion Mutants of the Phytohemagglutinin Gene *lec2*.

Construction of mutants is described in Methods. The top line represents the Greensleeves PHA-L gene and flanking regions. The direction of transcription and upstream direct repeats are indicated by the large and small arrows, respectively. The individual deletion mutants are shown below, and the construct designations and approximate 5′ endpoints (estimated by restriction enzyme mapping) are listed. All constructs have a common 3′ end, delimited by an EcoRI site located 1.5 kb downstream from the translation stop codon. The upstream regions defining clones 460 and 540, used in gel retardation assays, are indicated. The figure is not drawn to scale.

were produced for each construction. Genomic DNA from these plants was analyzed by DNA gel blotting to determine the gene dosage and fidelity of integrations, to show that the level of PHA produced by the plants is a function of promoter strength and not a copy number or "position" effect. Some of the plants did not contain intact PHA genes, and others contained a large number of copies and/or the genes were rearranged during integration. From this screening, we selected a number of plants of each deletion derivative, which were judged to contain at least one intact PHA gene.

To determine the gene dosage and number of intact copies in each plant, we performed DNA gel blot hybridization experiments. Table 1 summarizes these studies. The gene dosage was determined by counting bands on autoradiograms following cleavage of the DNA with EcoRI and DNA gel blotting. This enzyme cleaves all deletion mutants only once at the 3′ end; thus, each hybridizing fragment must necessarily be delimited on the 5′ end by a unique, genomic EcoRI site, or another 3′ EcoRI site (in the case of tandem integrations). These blots established that many plants had only one insertion of the gene, but others received up to nine copies. To establish the number of intact copies, the DNA samples were restricted with enzymes (EcoRI and HindIII or PstI), which should release the entire PHA gene fragment from the tobacco genome, and, upon hybridization, an expected, stepwise increase in the apparent molecular weight was observed, indicating

that faithful insertion of most deletion constructs occurred without major rearrangements (see Table 1). Although the gene copy number among the individual transformants within a group does vary, comparisons between groups of plants representing the different deletion mutants show that the copy number of each group is similar. Thus, the level of expression exhibited by each group should reflect the presence or absence of sequences that control quantitative aspects of PHA-L gene expression.

Immunoblot Screening and Analyses of PHA-L in Transgenic Plants

Seeds from the plants described above were used as a source of protein to monitor PHA-L levels for the various groups of deletion mutants. Transgenic tobacco seed lots were collected at approximately 30 days post-anthesis, a time when the level of PHA is at a maximum in tobacco (Voelker, Sturm, and Chrispeels, 1987), and the seeds are mature and desiccated to provide a defined and comparable harvesting stage. Seed proteins soluble in a low-salt, Triton X-100 buffer (see Methods) were subjected to immunoblot analyses using anti-PHA serum. For individual plants containing the same deletion mutant, we found some variation among the samples (usually less than twofold). Thus, the gene dosage was not always indicative of the level of PHA expression (see below). To average any plant-to-plant variation, equal amounts of protein from seeds of plants transformed with the same deletion construct were pooled and 150 μg of total protein from each group was subjected to immunoblotting.

Table 1. Analysis of Integration Events

Plant	Gene Dosage	Intact Copies
Δ202-A	7	7
Δ202-B	1	1
Δ202-C	2	2
Δ202-D	2	2
Δ202-E	1	1
Δ205-A	1	1
Δ205-B	5	4
Δ205-C	3	1
Δ205-D	9	7
Δ205-F	2	1
Δ210-B	1	1
Δ210-D	1	1
Δ210-6	1	1
Δ210-8	1	1
Δ214-A	3	3
Δ214-B	5	5
Δ214-E	1	1
Δ214-J	1	1
Δ217-B	1	1
Δ217-D	1	1
Δ217-H	1	1
Δ217-J	2	2
Δ217-K	1	1

Individual transformants selected for expression studies are listed together with their gene dosage (number of bands on DNA gel blots) and number of intact copies (number of bands on DNA gel blots greater than or equal in size to that expected for the particular deletion mutant resected from the tobacco genome).

Figure 2A shows that, in addition to the authentic PHA-L [a doublet at 34 kD (Sturm, Voelker, and Chrispeels, 1988)], a substantial amount of the protein is cleaved into three to five discrete products that have molecular weights between 10 kD and 15 kD. These smaller polypeptides are processing products of PHA-L that are formed after the protein is sequestered in the protein storage vacuoles (Voelker, Herman, and Chrispeels, 1989). Anomalous processing of other storage proteins in the seeds of transgenic tobacco has also been reported by other laboratories (Beachy et al., 1985; Sengupta-Gopalan et al., 1985; Higgins et al., 1988).

The immunoblots showed marked differences in the amount of PHA produced by the different deletion mutants. The Δ217 pool had severalfold less PHA than did either the Δ214 or Δ210 pools. More striking differences were observed when these samples were compared with the Δ205 and Δ202 pools. Although the Δ202 pool gave rise to very little PHA, the presence of an additional 220 bp (Δ205, containing the direct repeat/CATGCATG motifs) resulted in a significant increase in the PHA levels. To quantitate PHA expression, we performed immunoblotting of dilution series of the pooled samples. Inspection of the blots revealed that the Δ217 sample indeed has about twofold to threefold less PHA than do the Δ214 and Δ210 samples (which are nearly equal). The Δ205 sample has eightfold less PHA than does Δ210, and Δ202 has about 25-fold less PHA than does Δ205. Compared with Δ210, the Δ202 pool has nearly 200-fold less PHA. A histogram of these data is presented in Figure 2B. The fact that PHA is produced by Δ202 indicates that this deletion mutant harbors the core sequences required for a basal level of expression. We conclude from these experiments that the region between −125 and −345 directs high-level expression of PHA-L in transgenic tobacco, and that the presence of an additional 205 bp (Δ210) further enhances PHA-L levels. The presence of regions further upstream has little effect (Δ214) or results in some attenuation (Δ217).

Cotyledon Nuclear Proteins Bind to the Upstream Region of the *lec2* Gene

Our deletion analysis suggested that there are two regions upstream from *lec2* that serve to enhance PHA expression, as well as one located further upstream that attenuates

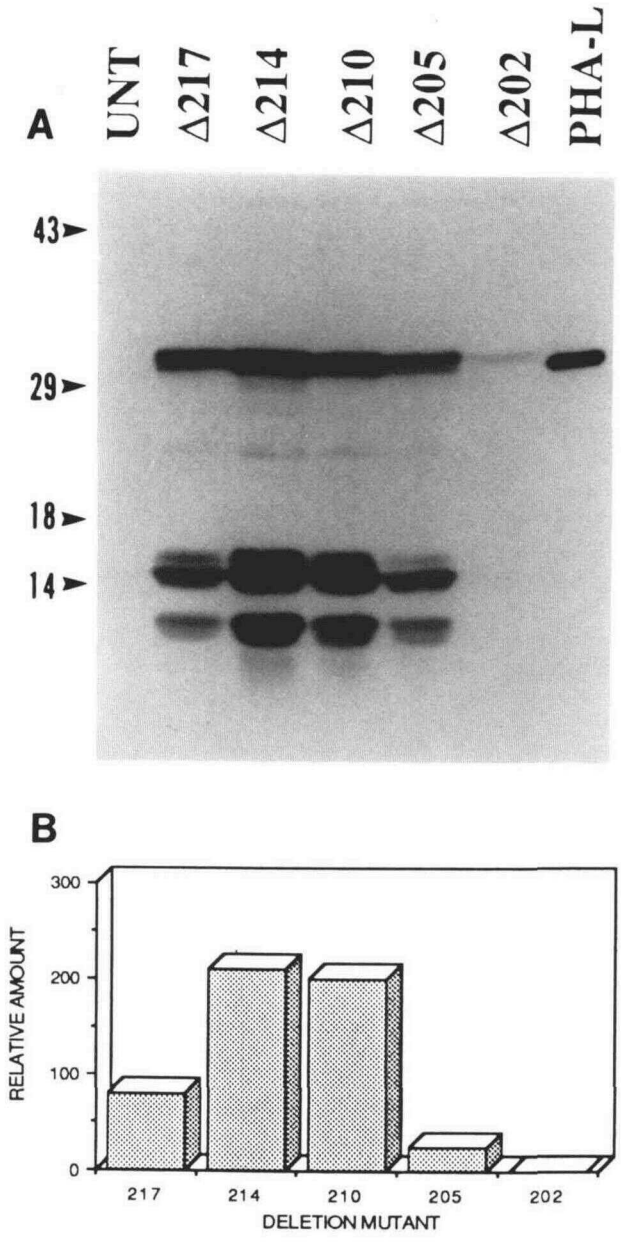

Figure 2. Immunoblot Analysis and Quantitation of PHA from Pooled Samples of Each Construct.

(A) Immunoblot analysis of PHA. Each lane contains 150 μg of pooled protein from the seeds of the plants listed in Table 1. Lane designations refer to the various deletion mutants described in the Figure 1 legend. UNT is seed protein from untransformed plants. The PHA lane contains 10 ng of purified PHA-L. Size markers are indicated in kilodaltons.

(B) Quantitation of PHA gene expression in transgenic tobacco. Dilutions of pooled samples from individual deletion mutants were subjected to immunoblotting, and the relative amounts of PHA in each sample were calculated by comparison to standards. A histogram of these data is presented with the level of PHA for groups of samples compared relative to that of Δ202.

expression. To determine whether these regions bind nuclear proteins that might modulate gene expression, we prepared nuclear extracts from bean cotyledons and performed gel retardation experiments using two probes: clone 540, harboring the region from −550 to −3 (containing a putative activator), and clone 460, which contains the region from −1000 to −550 (the putative silencer, see Figure 1). After binding conditions were established, a number of DNA competitors were included to compete for the nuclear factor(s) that promoted band retardation. Figure 3 illustrates such a gel retardation experiment. Incubation of the probe (clone 540) with nuclear proteins in the presence of the nonspecific competitor poly d(I)/d(C) (Figure 3A, lane PEC) permitted the identification of several bands with retarded mobility. Without this competitor, many proteins bind the probe nonspecifically, resulting in a high molecular weight smear (lane PE). The addition of a 10,000-fold molar excess of a 540-bp DdeI fragment of pUC19 (lane PECV) did not compete for the binding of the putative *trans*-acting factor(s), indicating that the probe contains a sequence(s) specifically recognized by a nuclear protein(s). That binding is authentic is demonstrated by titration of binding by increasing amounts of unlabeled probe (lanes 10, 100, and 1000).

To localize the binding site(s), a number of DNA competitors were then tested (Figures 3B and 3C). Titration of binding was attempted by including a 1000-fold molar excess of various restriction fragments derived from the clone 540, as well as the Δ205 deletion mutant, the synthetic oligonucleotide CATGCATGCATGCATG (annealed and ligated into multimers), and clone 460. Several conclusions can be drawn from these results. First, the direct repeats/CATGCATG motif probably does not interact with the abundant nuclear factors that give rise to the bound forms seen in the autoradiograms, as neither the Δ205 clone (lane 205), the oligonucleotide (lane OL), nor any of the proximal restriction fragments (lanes 1, 3, 5, and 7) competes for the bound forms. Second, the distal restriction fragments (lanes 2, 4, and 6) completely titrated the bound forms, indicating that this region contains the binding site(s). However, neither of the two DraI fragments (lanes 7 and 8) was an effective competitor, suggesting that the DraI site lies in or near the protein binding site.

Competition for Nuclear Factors Is Affected by DNA Fragments from Other Legume Genes

To ascertain whether the Pinto PHA genes contain binding sites for these nuclear proteins, we conducted additional gel shift experiments and used the appropriate restriction fragments to compete for factor binding. For these experiments, a 1000-fold molar excess of each fragment was used as a competitor. Figure 4 shows that the regions immediately upstream from both Pinto genes were strong competitors (lanes 2 and 6), and differential competition

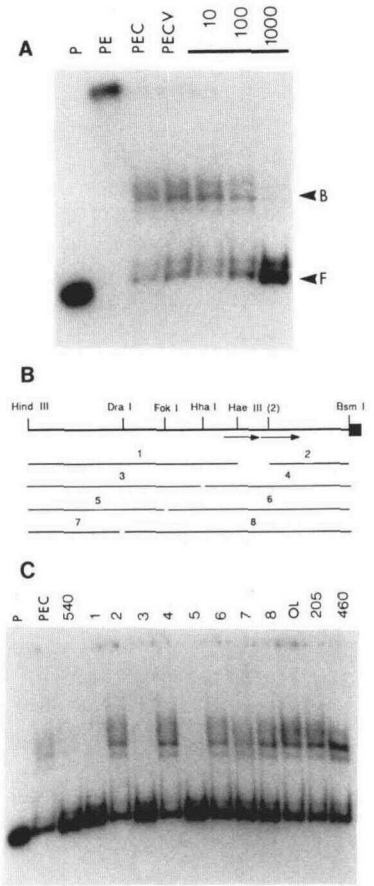

Figure 3. Binding of Cotyledon Nuclear Factors to Clone 540.

(A) Gel retardation assays were performed as described in Methods. Clone 540 (used as probe) contains the *lec2* upstream region from −550 (Δ210 endpoint) to −3 (BsmI site). Lane P, labeled probe only; lane PE, probe incubated with nuclear extract; lane PEC, probe incubated with extract and the synthetic competitor poly d(I)/d(C); lane PECV, a PEC reaction supplemented with a 10,000-fold molar excess of a 540-bp DdeI fragment of pUC19. The numbers 10, 100, and 1000 refer to the molar ratios of unlabeled to labeled probe (i.e., PEC reactions titrated with unlabeled probe). The arrows point to bound (B) and free (F) probe.
(B) Schematic representation of the pertinent features of clone 540. The top line includes restriction enzyme sites and the locations of the 63-bp direct repeats containing the CATGCATG motif (arrows), relative to the *lec2* structural gene. (The 5′ end of the gene is represented by the black box.) Various restriction fragments were used to compete for the binding of nuclear proteins, and these regions are numbered.
(C) Competition reactions. Reactions were performed as described above, using a 1000-fold molar excess of various competitors. Numbers refer to the addition of the restriction fragments specified in **(B)**. Other lanes are: 540, unlabeled probe; 205, the Δ205 deletion mutant containing sequences from −345 to −3; OL, a synthetic 16mer (CATGCATGCATGCATG) annealed and ligated into multimers of 2 to 50 in length; and 460, the region containing the sequences from −1000 to −550 (silencer-containing fragment).

was observed for several other fragments, including the proximal 3′ ends of both Pinto genes (lanes 1 and 7), the distal 3′ end of the *lec1-P* (which encodes PHA-E in the lectin-deficient cultivar Pinto) clone (lane 4), and the 5′-flanking region of a gene encoding β-phaseolin (lane β). These competitors affect the pattern of the bound form in a similar fashion; there is an apparent destabilization of the slowest migrating form and a concomitant increase in the amounts of the other two more rapidly migrating species. When titration was carried out with increasing amounts of the fragment containing the upstream region of a β-phaseolin gene, this destabilization/stabilization phenomenon was evident, as shown in Figure 5. The other fragments tested had little or no effect on the pattern of binding.

The availability of a soybean seed lectin gene clone and a cloned oligonucleotide homologous to a region upstream from a soybean Kunitz trypsin inhibitor gene (kindly provided by Dr. R.B. Goldberg) allowed us to test whether these clones, containing binding sites for a tissue-specific

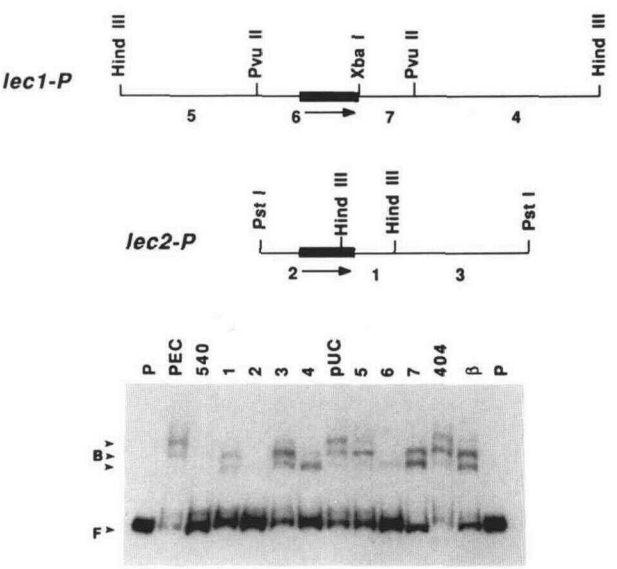

Figure 4. Competition Experiments with Pinto PHA Gene Fragments.

(Top) Schematic representation of partial restriction enzyme maps of the two Pinto PHA clones, *lec1-P* and *lec2-P*. The numbers below the maps indicate the restriction fragments used in the competition assays. Arrows indicate the direction of transcription.
(Bottom) Competition assays. Clone 540 was used as the probe, and competition reactions were carried out as described in the Figure 3 legend. Lane designations are described above, and other lanes include: pUC, linearized pUC19; 404, a 404-bp fragment that lies upstream from the *lec1* gene and contains organ-specific differentially methylated sites (Riggs and Chrispeels, 1988); β, an 850-bp fragment spanning the upstream region of a β-phaseolin gene (Hoffman et al., 1987).

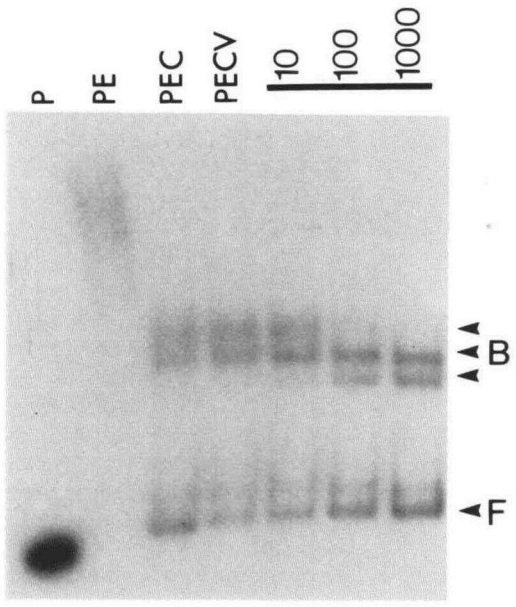

Figure 5. Competition for Nuclear Factor Binding by the Upstream Region of a β-Phaseolin Gene.

Increasing amounts of an 850-bp fragment containing a β-phaseolin promoter were used to titrate nuclear factor binding to labeled clone 540. Lane designations are as described in the Figure 3A legend.

and developmentally regulated protein (Jofuku, Okamuro, and Goldberg, 1987), are capable of competing for the binding of common bean nuclear factors to the *lec2* clone 540. The synthetic oligonucleotide (six copies of the 37-mer of the sequence: 5′-GTAAATTTATAATTAAATAA-TATGTATTACAAACACA-3′), spanning the region from −298 to −261, can compete for the binding of a developmentally regulated nuclear protein (Jofuku, Okamuro, and Goldberg, 1987) and DNase I footprinting demonstrated that this region and the lectin gene binding domain are very similar (C.R. Reeves and R.B. Goldberg, personal communication). The soybean upstream region (−550 to −30) is specifically recognized, albeit weakly, by at least one nuclear protein from bean (data not shown) and, as shown in Figure 6, it is capable of titrating the binding of factors to the upstream region of *lec2*. The oligonucleotide also affected competition, yet neither of the soybean sequences was as effective a competitor as were the PHA upstream fragments. The weak binding of the common bean nuclear proteins to the lectin promoter and the incomplete titration of *lec2* binding by equimolar amounts of the lectin gene fragment and the oligonucleotide suggest that, although a common protein is likely involved, the binding affinity of the bean protein(s) is lower for these heterologous gene fragments.

Surprisingly, the upstream region of the *lec1* gene did not compete for the binding of nuclear proteins to *lec2* (Figure 6, lanes *lec1*). For this experiment, we used the upstream region of the *lec1* gene from the cultivar Tendergreen (Hoffman and Donaldson, 1985), linked to clone 404 from Greensleeves. The rationale for this is as follows. DNA sequencing (Voelker, Staswick, and Chrispeels, 1986) has established that the PHA-L genes and flanking regions in the two cultivars are nearly identical at the nucleotide level, and mapping experiments (C.D. Riggs and M.J. Chrispeels, unpublished results) have shown that the Greensleeves genes are closely linked, just as they are in Tendergreen (the two genes are separated by about 4 kb; Hoffman and Donaldson, 1985). The Tendergreen clone (−380 to −171) was ligated to the most distal 404 bp (clone 404) of our Greensleeves *lec2* clone and did not effect competition (Figure 6). This clone effectively creates the *lec1* upstream region from −784 to −171. Neither the *lec1* upstream region (−380 to −3, data not shown), nor

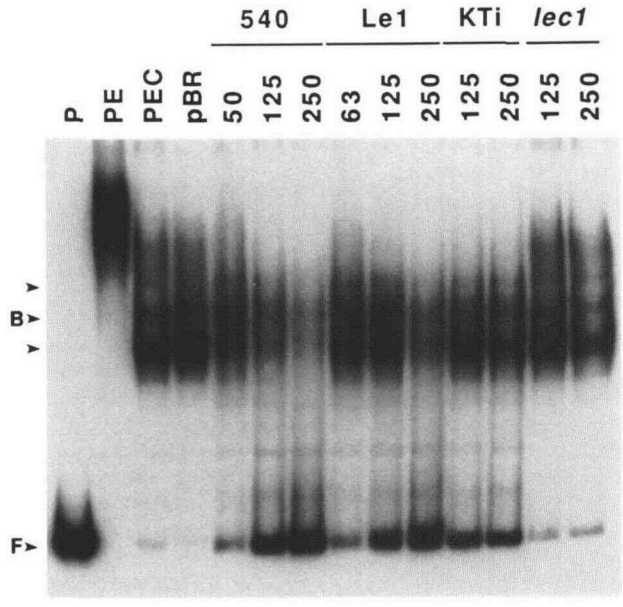

Figure 6. Gene Fragments of Other Genes Expressed in Legume Seeds Are Effective Competitors for Protein Binding.

A 240-bp fragment of *lec2* (−550 to −310) was used as a probe in competition experiments. Reactions were carried out essentially as described in the Figure 3 legend and Methods, except that the extract was that fraction of the total cotyledon nuclear proteins precipitated by addition of ammonium sulfate to 45% saturation. Lanes P, PE, and PEC are as previously described. The competitors are listed along with the molar ratio of unlabeled competitor to probe ratio. pBR, a 200-fold molar excess of unlabeled pBR322; 540, the upstream region of *lec2*; Le1, soybean lectin upstream region; KTi, Kunitz trypsin inhibitor oligonucleotide; *lec1*, the upstream region of the Tendergreen *lec1* gene linked to clone 404. Refer to the text for more details.

clone 404 (Figure 4), nor the linked fragments (Figure 6) were capable of competition. We also observed no shifted bands in gel retardation experiments in which this region was used as a probe (data not shown). If this DNA fragment does bind nuclear proteins, they must be different from those that effect the gel shift patterns seen for *lec2* and the soybean gene fragments.

To determine whether common sequence motifs exist that might be responsible for factors binding to *lec2* and the soybean clones, the DNA sequences of fragments that were effective competitors were compared by computer analysis. Figure 7 illustrates that there are similarities between the 5′-flanking regions of the PHA genes, the β-phaseolin gene, the soybean seed lectin gene, and the Kunitz trypsin inhibitor gene. These DNA fragments have a substantial amount of sequence homology in the area of −395 to −350 of *lec2*, and, outside the regions shown, there are no discernible similarities. Perhaps significantly, the regions of homology are centered around that region of *lec2* that contains the single DraI site. When clone 540 is restricted with DraI, neither fragment generated was an effective competitor, suggesting that this restriction site lies in or near the site of binding (see Figures 3B and 3C and the text).

A DNA Fragment Harboring the Silencer Element Interacts with Nuclear Proteins

The attenuator effect observed in the deletion analyses was also investigated by gel retardation in an attempt to localize a binding site for a putative silencer. For these studies, clone 460, containing the region from −1000 to −550, was used as a probe (see Figure 1), and restriction enzyme fragments from the Pinto genes were used to titrate binding. Figure 8 shows that this distal upstream region binds (albeit weakly in our standard reaction conditions) to perhaps two proteins (two bands with retarded mobility can be observed), and the binding can be titrated by unlabeled probe (lane 460). Competition experiments with restriction fragments of this region enabled us to localize the binding site within a 269-bp RsaI fragment (lane 269, coordinates from transcription start site −587 to −856). Within this region is an A/T-rich region that may form a stem/loop structure spanning 70 bp, and, as such, could be a motif recognized by a repressor protein. As with the activator elements, analogous restriction fragments of genes from both Pinto and Greensleeves, as well as the upstream region of the β-phaseolin gene, compete for the binding of nuclear proteins to this clone. Sequence analysis of the fragment containing the silencer revealed homology to several areas of the 5′-flanking region of the PHA genes, including part of the activator region. However, the silencer clone did not compete for binding of factors that bound to the activator clone (see Figure 3C), and, as the upstream regions of all of the genes are very A/T-rich, this partial homology may have no functional significance.

DISCUSSION

The identification of regulatory sequences and DNA binding factors related to the expression of storage protein

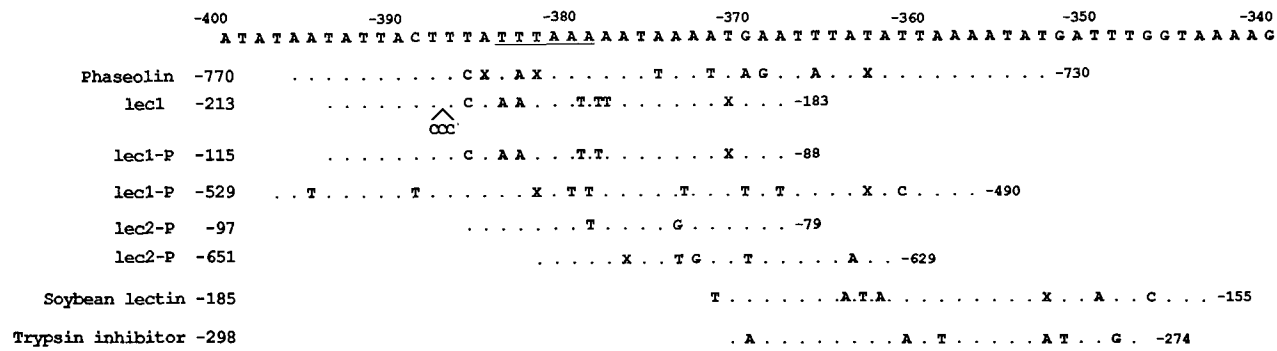

Figure 7. Comparison of the Homologous Regions Upstream from the PHA, β-Phaseolin, Soybean Seed Lectin, and Kunitz Trypsin Inhibitor Genes.

The top line represents part of clone 540 (upstream region of the Greensleeves *lec2* gene) and coordinates from the transcription start site (−11 relative to translation start) are shown. The underlined sequence defines the DraI site (see text). Gaps were introduced to maximize homology, and these are represented by the X. Additional bases are inserted between the register dots where appropriate. The sequences presented for the PHA alleles were determined in our laboratory (Voelker, Staswick, and Chrispeels, 1986); other sequence data are *lec1* (Hoffman and Donaldson, 1985), phaseolin (Doyle et al., 1986), soybean seed lectin (Jofuku, Okamuro, and Goldberg, 1987), and Kunitz trypsin inhibitor oligonucleotide (K.D. Jofuku and R.B. Goldberg, personal communication). The coordinates of each homologous region, relative to the transcription start site of the cognate gene, are shown before and after the individual sequences.

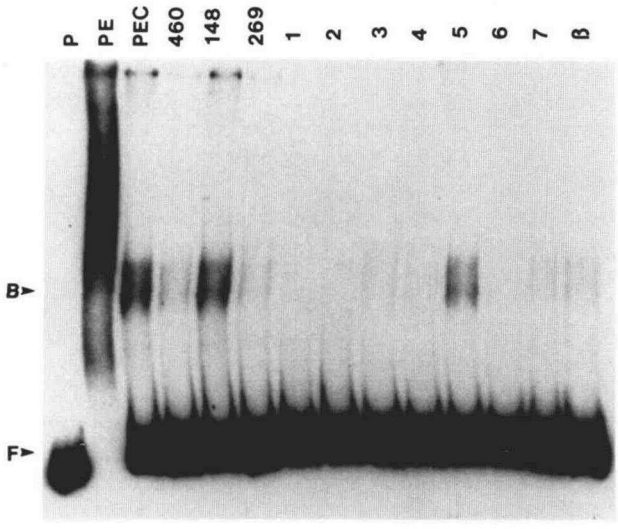

Figure 8. Competition of Pinto Restriction Fragments for the Putative Silencer Element.

Clone 460 was used as a probe, and gel retardation experiments were carried out as described in the Figure 3 legend. A 269-bp RsaI subfragment of clone 460 competes for binding, but the 148-bp RsaI subfragment does not. Numbers refer to the Pinto gene fragments shown in the Figure 4 legend.

genes is being investigated in a number of plant species. In soybean, the conglycinin genes (Beachy et al., 1985; Chen, Schuler, and Beachy, 1986; Allen, Lessard, and Beachy, 1988; Chen, Pan, and Beachy, 1988) and the seed lectin gene (Jofuku, Okamuro, and Goldberg, 1987) have received by far the most attention, and components involved in gene regulation have been identified. Similarly, *cis*-active sequences or protein binding sites have been identified upstream from genes encoding pea vicilin (Higgins et al., 1988), and legumin (Shirsat et al., 1989), maize zein (Maier et al., 1988), and wheat glutenin (Colot et al., 1987).

The experiments described here are a natural extension of our previous work centering on analyses of gene structure, mRNA levels, and PHA subunit levels in a high-expression cultivar (Greensleeves) and a lectin-deficient cultivar (Pinto) of the bean (Voelker, Staswick, and Chrispeels, 1986). At both the mRNA and protein levels, the expression of *lec2* is about 30-fold greater than that of *lec2-P* (which encodes PHA-L in Pinto). Furthermore, immunoblotting revealed that a similar ratio in expression levels (about 50-fold) is maintained following transfer of the two genes into the tobacco genome (Voelker, Sturm, and Chrispeels, 1987). These results are consistent with the interpretation that one or more *cis*-acting elements is involved in enhancing *lec2* expression and/or attenuating *lec2-P* expression. Sequence analysis of the four PHA genes revealed several interesting features in the 5′-flanking region. A schematic of all known and suspected regulatory elements upstream from the four alleles is shown in Figure 9. Included in this figure is the CATGCATG motif, identified by Dickinson, Evans, and Nielsen (1988) as being present upstream from all legume storage protein genes sequenced to date. Also included is the 63-bp direct repeat motif (double arrows), which is found upstream from all of the genes except *lec2-P*. We have postulated that the direct repeat/CATGCATG motif is involved in PHA gene regulation and may be responsible for the low expression exhibited by the *lec2-P* allele in both bean and transgenic tobacco (Voelker, Staswick, and Chrispeels, 1986; Voelker, Sturm, and Chrispeels, 1987).

To ascertain whether the direct repeat/CATGCATG motif is involved in modulating PHA gene expression, and to identify other *cis*-acting sequences governing expression, we monitored expression of promoter deletion mutants in transgenic tobacco. Our studies show that 125 bp of the upstream region is sufficient to direct a low level of PHA-L

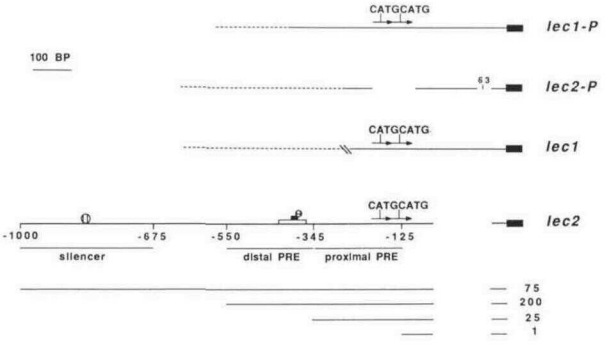

Figure 9. Summary of Known and Suspected Gene Regulatory Elements Governing PHA Gene Expression.

The 5′-flanking regions of the four PHA alleles are shown with the large black boxes representing the amino termini of the structural genes. The solid lines represent regions of high homology to *lec2* and the broken lines represent nonhomologous regions. The 63-bp direct repeats, each containing a CATGCATG motif (Dickinson, Evans, and Nielsen, 1988), are shown. Upstream from *lec2-P* there is a 63-bp insertion (labeled 63), as well as a 114-bp deletion. The broken bars in the *lec1* 5′ region indicate that this sequence is from clone 404 (see Figure 4 legend and the text), which lies upstream from *lec1* (approximate coordinates from the translation start site are −784 to −380). The *lec2* fragments containing the proximal and distal positive regulatory elements (PRE) are shown, as is that region containing the silencer element. The region containing homology to the soybean genes (see Figure 7) is represented by the open box. Protein factors that bind to the distal PRE and the silencer fragments are shown as differentially filled symbols. The lower part of the figure summarizes the deletion analysis data and includes a schematic of the deletion endpoints and the relative levels of PHA-L associated with these promoter deletion mutants.

expression in seeds. Several lines of evidence suggest that this "core" region contains a seed-specific, regulatory element. First, we could detect no PHA-L in non-seed organs (for example, leaves). If PHA-L were expressed in leaves, it could well be degraded in the protease-rich vacuoles. However, related experiments in which the PHA-L upstream region directs expression of the luciferase gene indicated that only background luciferase activity is present in non-seed tissues (Riggs et al., 1989), and, furthermore, luciferase activity is stable in leaves and other organs (Ow et al., 1986). Thus, the region between −125 and the transcription start site may contain a sequence controlling seed-specific gene expression, or perhaps the more distal upstream region (in the luciferase experiments) contains a silencer element that acts to suppress expression in all organ systems except seeds. Further studies are required to confirm the existence and location of such a regulatory element. In this respect, the approach used to identify a seed-specific enhancer immediately upstream from a soybean β-conglycinin gene (Chen, Pan, and Beachy, 1988) would provide a means to investigate this question.

The presence of more of the upstream region of *lec2* dramatically altered the levels of PHA-L in the seeds. Plants harboring the Δ205 mutant, containing the direct repeats/CATGCATG motif, gave rise to approximately 25-fold more PHA-L than did those plants harboring only the core promoter region. This difference is strikingly similar to the disparity in PHA-L levels exhibited by the two bean cultivars (Voelker, Staswick, and Chrispeels, 1986), and also parallels that observed for the two genes in transgenic tobacco (Voelker, Sturm, and Chrispeels, 1987). Although other DNA sequences are contained within this deletion mutant, we suggest that deleted direct repeats and/or the CATGCATG motif are responsible for the low expression exhibited by the *lec2-P* gene in both bean and tobacco. Additional transformation studies, in which this region is linked to the *lec2-P* gene fragment (or the chimeric upstream region is linked to a reporter gene), will be required to provide direct evidence for its involvement in gene regulation. Thus, although the role of the direct repeat/CATGCATG motif in enhancing the level of PHA-L is not known, the region between −345 and −125 contains a positive regulatory element (PRE), which we refer to as the proximal PRE (Figure 9).

Surprisingly, we could not detect binding of any nuclear factors to the proximal PRE in the gel retardation assays (tested for a number of extracts and under a variety of conditions). It is plausible that the region does contain binding sites but that the protein factors are unstable or the assay conditions were inappropriate to demonstrate binding. The evolutionary conservation of the CATGCATG motif (Dickinson, Evans, and Nielsen, 1988) is intriguing, and one might suspect that a conserved *trans*-acting factor(s) would bind to this sequence. However, neither the region containing the proximal PRE nor a synthetic 16-mer oligonucleotide (CATGCATGCATGCATG) was an effective competitor for the binding of nuclear factors to clone 540. We hypothesize that, as an alternative to regulation by soluble proteins, the CATGCATG motif may form a Z-DNA structure in vivo. One possibility is that, following recognition and binding by a putative activator protein(s) to the distal PRE (see below), the CATGCATG motif may adopt an altered conformation that enhances the recognition for or passage of transcriptional complexes.

The distal PRE (located between −550 and −345; see Figure 9) is responsible for an additional eightfold increase in the level of PHA-L. Compared with the core promoter, the presence of both the proximal and distal PREs increases the level of PHA-L by 200-fold. This region contains at least one binding site for cotyledon nuclear proteins. The additional bands seen in the gel retardation experiments may be due to different factors binding to other sites, protein/protein interactions at the primary binding site, or both. The experiments in which small restriction fragments of the *lec2* upstream region effected complete titration of the bound forms (see Figure 3) would suggest that the binding sites (if more than one) must lie in very close proximity to one another. The differential titration of individual bands by various restriction fragments of the *lec* genes of Pinto (Figure 4) and the 5′-region of a β-phaseolin gene (Figure 5) indicates that it is possible to separate components and maintain stable binding. The competition affected by the upstream region of the soybean seed lectin gene, and, perhaps more significantly, the Kunitz trypsin inhibitor oligonucleotide, indicates that there is probably one binding site. This site is specifically recognized by a nuclear protein(s) from both soybean (Jofuku, Okamuro, and Goldberg, 1987; C.R. Reeves and R.B. Goldberg, personal communication) and bean, and implies that the DNA binding domain of these proteins is conserved. The fact that a greater molar excess of the soybean fragments is required to titrate binding completely would suggest that, although the binding sites/nuclear factors are similar, the binding constants are different (at least under our standard assay conditions). Taken together, these studies point to common features of legume gene expression, yet the elucidation of the role of the protein that binds to the conserved sites must await the cloning of the cognate gene or the production of antibodies against the putative *trans*-acting factor.

The deletion analysis also revealed the presence of a negative regulatory element located between −1000 and −675 (see Figure 9) that is responsible for a threefold reduction of PHA-L levels. The level of PHA-L expression by individual transformants of this group (Δ217) was very similar, as was that of the Δ210 group. (In both groups most transformants contain one copy of the gene; see Table 1.) Thus, although some variation exists between the copy number of transformants in other groups, we believe that the different levels of PHA-L observed for these plants are a consequence of regulation by both

positive and negative regulatory elements, and that the upstream silencer can partly override the effect exerted by the more proximal enhancers. The silencer phenomenon has also been reported in several other plant systems. In transient assays, the existence of a negative regulatory element upstream from a chalcone synthase gene was elegantly demonstrated by co-electroporation of a chimeric chalcone synthase/chloramphenicol acetyltransferase gene together with and without a fragment containing the silencer region, and a dramatic increase in expression was found when the silencer fragment was included (Lawton et al., 1988). These authors suggest that the addition of silencer-containing fragments competitively bound a *trans*-acting repressor(s), and titration of the repressor was reflected by an increase in expression of the reporter gene. Similarly, Castresana et al. (1988) demonstrated that both positive and negative regulatory elements lie within 1 kb of a photoregulated *Cab* gene in tobacco, and found that the negative regulatory region is very A/T-rich. Like the *Cab* silencer, the putative PHA-L silencer is also very A/T-rich, and the three silencers are also similar in that positive regulatory elements lie between them and the genes they effect. We have shown that the region containing the PHA-L silencer binds at least one cotyledon nuclear protein. Binding was localized within a 269-bp RsaI fragment, which contains a 70-nucleotide region with some potential to form a stem/loop structure.

The gel retardation assays revealed two potentially significant features of legume storage protein gene regulation. First, we were unable to demonstrate that the 5′-flanking region of the *lec1* gene binds cotyledon nuclear proteins (data not shown), nor does this DNA fragment compete for binding of factors to the upstream region of *lec2*. The *lec1* gene fragment used was from another high-expression cultivar, Tendergreen (Hoffman and Donaldson, 1985), and, although hybridization experiments indicate that the two PHA gene loci (in Greensleeves and Tendergreen) are very similar, we cannot discount mutations that may interfere with binding. In this respect, sequence analysis revealed that the *lec1* clone contains the region of shared homology, but several insertions (in particular, a CCC insertion in an otherwise A/T-rich region; see Figure 7) could change the spatial arrangement of motifs important for recognition and/or protein binding. The *lec1* clone of Tendergreen is short relative to the other *lec* alleles we have tested. However, as the two genes are separated by about 4 kb, the 3′ end of our Greensleeves *lec2* clone (clone 404) would be located at coordinates −380 to −784 relative to the linked *lec1* gene (Figure 9). No competition was effected by this clone alone (see Figure 4) nor by a DNA fragment in which clone 404 was linked to the *lec1* upstream region (see Figure 6). Thus, the second block of conserved homology (present in the Pinto clones that do compete for binding) may be important for protein binding, but, in the case of the *lec1* gene, this region apparently does not exist. Conceptually, a single regulatory region could govern both tandemly linked genes, but this seems unlikely in that the putative *trans*-acting factors bind upstream from *lec2* (*lec1* lies approximately 4 kb downstream of *lec2*), yet the *lec1* gene product (PHA-E) begins to accumulate several days before that of *lec2*. One possibility is that a 170-bp region upstream from *lec1* (as well as the two *lec* genes of Pinto; see Figure 9), but absent from *lec2*, is responsible for modulating promoter strength. Indeed, the onset of PHA accumulation in seeds of transgenic tobacco is similar for the three alleles containing this upstream region but delayed for plants harboring *lec2* (T.A. Voelker and M.J. Chrispeels, unpublished results).

Second, the binding of the putative activator protein(s) may not be sufficient to direct high-level expression of PHA. Although the 5′-flanking region of the *lec2-P* gene is a good competitor for the binding of nuclear factors to clone 540, the level of *lec2-P* mRNA is about 30-fold less than that produced by *lec2* (Voelker, Staswick, and Chrispeels, 1986). Moreover, the expression of *lec2-P* in transgenic tobacco results in the accumulation of about 50 times less PHA-L than that exhibited by *lec2*-transformed plants (Voelker, Sturm, and Chrispeels, 1987). This enigma may be explained by either poor binding of the activator protein(s) in vivo, the lack of the proximal PRE, or both. Again, only transformation experiments in which the upstream from *lec2-P* is dissected will allow us to determine what effect the insertion of a proximal PRE has on PHA-L levels, and how different spatial arrangements of the upstream DNA sequences act in concert to modulate the level of gene product.

METHODS

Biological and Genetic Materials

Phaseolus vulgaris cv Greensleeves was grown in a greenhouse, and mid-mature cotyledons were used to prepare genomic DNA and nuclear proteins. *Nicotiana tabacum* cv Xanthi, both transformed and untransformed, was grown in a greenhouse, and various tissues were used as a source of DNA and proteins. Some of the plasmids used in these experiments have been described in other publications. The plasmids used include pTV794 (Voelker, Herman, and Chrispeels, 1989), an 850-bp BglII fragment containing the upstream region of a β-phaseolin gene (from p8.8pro, Hoffman et al., 1987); pLeHB1050, a 1050-bp HpaI fragment containing the upstream region of the soybean seed lectin gene (Jofuku, Okamuro, and Goldberg, 1987); dlec1A, a 6.0-kb PstI fragment containing the PHA-E gene and flanking regions from *P. vulgaris* cv Tendergreen (Hoffman and Donaldson, 1985); pKTiOligo 6 (a gift of Dr. R.B. Goldberg), a synthetic 37-mer of the sequence 5′-GTAAATTTATAATTAAATAATATGTATTACAAACACA-3′, annealed and ligated to yield six copies of this sequence, which was subsequently inserted into a cloning vector. Subclones from the PHA-L genomic clone pTV781 are described in the text.

Construction of Promoter Deletion Mutants of *lec2* and Transformation of Tobacco

The PHA-L gene, *lec2*, was isolated from the cultivar Greensleeves as described (Voelker, Sturm, and Chrispeels, 1987), and a 3.5-kb PstI/SalI fragment was subcloned into the vector Bluescribe (Stratagene) to yield pTV781. This clone, containing 1.2 kb of 5′-flanking DNA, the gene for PHA-L, and 1.5 kb of 3′-flanking DNA, was used to produce a series of progressive deletions from the unique PstI site (located 1.2 kb upstream from the gene). After linearization with PstI, nuclease Bal31 was employed as described (Maniatis, Fritsch, and Sambrook, 1982) and, after various times, aliquots of DNA were recovered and synthetic PstI linkers were added. Truncated inserts were released by digestion with PstI and EcoRI (which cuts in the polylinker region 3′ to the insert of pTV781) and were recloned into pUC12, such that they could be recovered with HindIII and EcoRI. [An internal HindIII site lies at position −550, and this subcloning would allow partial digestion and isolation of deletions longer than −550. Furthermore, the binary vector Bin19 (Bevan, 1984) has multiple PstI sites, circumventing the insertion of the primary deletion mutants.] Selected deletions were recloned into Bin19, and transgenic tobacco was produced by co-cultivation of tobacco leaf discs (Voelker, Sturm, and Chrispeels, 1987) with *Agrobacterium tumefaciens* harboring the various constructs. Eight to 10 plants of each construction were grown in a greenhouse.

DNA Isolation and Analysis of Integrations into the Tobacco Genome

Genomic DNA was purified from young leaves essentially as described by Dellaporta, Wood, and Hicks (1985). The DNA was cleaved with restriction enzymes under conditions suggested by the manufacturers (Bethesda Research Laboratories, New England BioLabs, Pharmacia), and aliquots were subjected to DNA gel blotting, using an alkaline transfer procedure (Chomczynski and Qasba, 1984) and Hybond-N (Amersham). PHA probes were prepared by oligolabeling (Feinberg and Vogelstein, 1983) the coding region of the *lec2* gene (an 820-bp BsmI/XbaI fragment). Filters were processed by standard techniques, and were washed at final stringencies of 0.1 × SSC at 65°C.

Isolation and Analysis of Seed Proteins of Transgenic Tobacco

Tobacco seeds from self-pollinated plants were collected at approximately 30 days post-anthesis, and aliquots were ground in ice-cold seed extraction buffer (50 mM Tris, pH 7.5, 50 mM NaCl, 0.1% Triton X-100, 1% β-mercaptoethanol, and 0.5 mM PMSF). Following clarification in a microcentrifuge, the supernatant was collected and proteins were precipitated with trichloroacetic acid. The pellets were washed with acetone, dried, and resuspended in denaturation buffer (20 mM Tris, pH 8.6, 1% SDS, 0.33% β-mercaptoethanol, and 10% glycerol). Aliquots of each sample were precipitated and the protein concentration was determined by the method of Lowry et al. (1951). Aliquots of protein from each transformant were subjected to SDS-PAGE, followed by immunoblot analysis. The primary antibody directed against PHA (Voelker, Staswick, and Chrispeels, 1986) was raised in a rabbit, and detection of PHA in the samples was facilitated using a secondary antibody obtained from Bio-Rad (goat anti-rabbit IgG conjugated to horseradish peroxidase). Titration of the primary antiserum was carried out with dilutions of purified PHA-L, and for all subsequent experiments it was diluted 500-fold. The secondary antibody was used at a 1:1000 dilution and detection was carried out as described in the Bio-Rad instruction manual.

Preparation of Bean Nuclear Extracts and DNA/Protein Binding Assays

The preparation of crude nuclear extracts was carried out essentially as described by Maier et al. (1988), with minor modifications. Cotyledons from mid-mature seeds (13 days to 18 days post-anthesis) were separated from seed coats on ice, and then powdered in liquid nitrogen in a mortar. The powder was transferred to an Omnimixer cup containing 10 volumes of Honda buffer, pH 8.5 (Luthe and Quatrano, 1980), and homogenized for 30 sec on medium power. All subsequent steps were performed at 4°C. The liquid was filtered through three layers of Miracloth, and a crude nuclei/starch pellet was recovered after centrifugation at 4000 rpm (Sorvall SS-34 rotor) for 5 min. The pellet was resuspended and washed twice in 10 mL of 25 mM Tris, pH 8.5, 5 mM $MgCl_2$, 25% glycerol, and 10 mM β-mercaptoethanol. Following centrifugation, 1 mL of high salt buffer (20 mM Hepes, pH 7.5, 25% glycerol, 0.42 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM PMSF, 0.5 mM DTT, and 0.5 mM trypsin inhibitor) was added, and the mixture was left on ice for 40 min with occasional vortexing. The nuclear extract was separated from starch grains by centrifugation at 6000 rpm for 10 min, and was then dialyzed overnight against 20 mM Tris, pH 7.5, 0.1 M KCl, 20% glycerol, 0.2 mM EDTA, 0.5 mM PMSF, and 0.5 mM DTT. Protein concentration was determined by the Bradford procedure (Bradford, 1976).

For gel retardation studies, binding reactions were carried out at room temperature for 30 min in a total volume of 20 μL. After addition of nuclear extract (which was added last), a typical reaction contained 20 mM Tris, pH 7.5, 1 mM EDTA, 0.5 mM DTT, 100 mM NaCl, 5% glycerol, 0.5 mM PMSF, and included 4 μg of poly d(I)/d(C) (Pharmacia), 5 μg of nuclear extract, and approximately 0.5 ng of labeled probe. Competition reactions were conducted with various competitors added in the molar ratios described in the text (based on the molecular weight of the restriction fragments or oligonucleotides relative to that of the labeled probe). Following incubation, reactions were loaded on 4% or 5% polyacrylamide gels (37.5:1 cross-linking ratio, in 0.2 × TBE), which had been prerun at 12 V/cm for 90 min. Electrophoresis was performed either at 4°C or at room temperature at 12 V/cm. Gels were then dried and exposed to x-ray film.

ACKNOWLEDGMENTS

We thank Dr. Les Hoffman (Agrigenetics Corporation) for the β-phaseolin and Tendergreen *lec1* clones and Dr. Robert Goldberg (UCLA) for the soybean seed lectin clone, the Kunitz trypsin inhibitor oligonucleotide, and especially for sharing results prior to their publication. We also appreciate the advice and criticism given by other colleagues. This research was supported by a grant from

the United States Department of Agriculture (Competitive Research Grants Program/Genetic Mechanisms for Crop Improvement) (to M.J.C.) and a National Institutes of Health postdoctoral fellowship (to C.D.R.).

Received April 7, 1989.

REFERENCES

Allen R.D., Lessard, P.A., and Beachy, R.N. (1988). Interaction of nuclear proteins with β-conglycinin gene upstream sequences. J. Cell. Biochem. Suppl. **C12,** 223 (abstr. L709).

Beachy, R.N., Chen, Z.-L., Horsch, R.B., Rogers, S.G., Hoffman, N.J., and Fraley, R.T. (1985). Accumulation and assembly of soybean β-conglycinin in seed of transformed petunia plants. EMBO J. **4,** 3047–3053.

Bevan, M.W. (1984). Binary *Agrobacterium* vectors for plant transformation. Nucl. Acids Res. **12,** 8711–8712.

Borrebaeck, C.A.K. (1984). Detection and characterization of a lectin from non-seed tissue of *Phaseolus vulgaris*. Planta **161,** 223–228.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. Anal. Biochem. **72,** 248–254.

Castresana, C., Garcia-Luque, I., Alonso, E., Malik, V.S., and Cashmore, A.R. (1988). Both positive and negative regulatory elements mediate expression of a photoregulated CAB gene from *Nicotiana plumbaginifolia*. EMBO J. **7,** 1929–1936.

Chappell, J., and Chrispeels, M.J. (1986). Transcriptional and posttranscriptional control of phaseolin and phytohemagglutinin gene expression in developing cotyledons of *Phaseolus vulgaris*. Plant Physiol. **81,** 50–54.

Chen, Z.-L., Schuler, M.A., and Beachy, R.N. (1986). Functional analysis of regulatory elements in a plant embryo-specific gene. Proc. Natl. Acad. Sci. USA **83,** 8560–8564.

Chen, Z.-L., Pan, N.-S., and Beachy, R.N. (1988). A DNA sequence element that confers seed-specific enhancement to a constitutive promoter. EMBO J. **7,** 297–302.

Chomczynski, P., and Qasba, P.K. (1984). Transfer of DNA to plastic membrane. Biochem. Biophys. Res. Commun. **122,** 340–344.

Colot, V., Robert, L.S., Kavanagh, T.A., Bevan, M.W., and Thompson, R.D. (1987). Localization of sequences in wheat endosperm protein genes which confer tissue specific expression in tobacco. EMBO J. **6,** 3559–3564.

Dellaporta, S.L., Wood, J., and Hicks, J.B. (1985). Maize DNA miniprep. In Molecular Biology of Plants, R. Malmberg, J. Messing, and I. Sussex, eds (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory), pp 36–37.

Dickinson, C.D., Evans, R.P., and Nielsen, N.C. (1988). RY repeats are conserved in the 5′-flanking regions of legume seed-protein genes. Nucl. Acids Res. **16,** 371.

Doyle, J. J., Schuler, M.A., Godette, W.D., Zenger, V., Beachy, R.N., and 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.

Feinberg, A.P., and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132,** 6–13.

Gatehouse, J. A., Evans, I.M., Croy, R.R.D., and Boulter, D. (1986). Differential expression of genes during legume seed development. Philos. Trans. R. Soc. Lond. B. Biol. Sci. **314,** 367–384.

Higgins, T.J.V., Newbigin, E.J., Spencer, D., Llewellyn, D.J., and Craig, S. (1988). The sequence of a pea vicilin gene and its expression in transgenic tobacco seeds. Plant Mol. Biol. **11,** 683–695.

Hoffman, L.M., and Donaldson, D.D. (1985). Characterization of two *Phaseolus vulgaris* phytohemagglutinin genes closely linked on the chromosome. EMBO J. **4,** 883–889.

Hoffman, L.M., Donaldson, D.D., Bookland, R., Rashka, K., and Herman, E.M. (1987). Synthesis and protein body deposition of maize 15 kD zein in transgenic tobacco seeds. EMBO J. **6,** 3213–3221.

Jofuku, K.D., Okamuro, J.K., and Goldberg, R.B. (1987). An embryo DNA binding protein that interacts with a soybean lectin gene upstream region. Nature **328,** 734–737.

Lawton, M.A., Kragh, K., Jenkins, S.M., Dron, M., Clouse, S.D., Dixon, R.A., and Lamb, C.J. (1988). Nuclear protein binding to a cis-acting silencer in a defense gene promoter. J. Cell. Biochem. Suppl. **C12,** 222 (abstr. L707).

Lowry, O.H., Rosebrough, N.J., Farr, A.L., and Randall, R.J. (1951). Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193,** 265–275.

Luthe, D.S., and Quatrano, R.S. (1980). Transcription in isolated wheat nuclei. Plant Physiol. **65,** 305–308.

Maier, U.-G., Brown, J.J.W.S., Schmitz, L.M., Schwall, M., Dietrich, G., and Feix, G. (1988). Mapping of tissue-dependent and independent protein binding sites to the 5′ upstream region of a zein gene. Mol. Gen. Genet. **212,** 241–245.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory).

Ow, D.W., Wood, K.V., DeLuca, M., de Wet, J.R. Helinski, D.R., and Howell, S.H. (1986). Transient and stable expression of the firefly luciferase gene in plant cells and transgenic plants. Science **234,** 856–860.

Riggs, C.D., and Chrispeels, M.J. (1988). Phytohemagglutinin gene expression is correlated with tissue-specific DNA methylation patterns in *Phaseolus vulgaris*. Genome **30** (Suppl. 1), 95 (abstr. 31.23.4).

Riggs, C.D., Hunt, D.C., Lin, J., and Chrispeels, M.J. (1989). Utilization of luciferase fusion genes to monitor differential expression of phaseolin and phytohemagglutinin gene promoters in transgenic tobacco. Plant Sci., in press.

Sengupta-Gopalan, C., Reichert, N.A., Barker, R.F., and Hall, T.C. (1985). Developmentally regulated expression of the bean β-phaseolin gene in tobacco seed. Proc. Natl. Acad. Sci. USA **82,** 3320–3324.

Shirsat, A., Wilford, N., Croy, R., and Boulter, D. (1989). Sequences responsible for the tissue specific promoter activity of a pea legumin gene in tobacco. Mol. Gen. Genet. **215,** 326–331.

Staswick, P., and Chrispeels, M.J. (1984). Expression of lectin genes during seed development in normal and phytohemagglutinin deficient-cultivars of *Phaseolus vulgaris*. J. Mol. Appl. Genet. **2,** 525–535.

Sturm, A.J., Voelker, T.A., and Chrispeels, M.J. (1988). Correct glycosylation, Golgi-processing, and targeting to protein bodies of the vacuolar protein phytohemagglutinin in transgenic tobacco. Planta **175,** 170–183.

Voelker, T.A., Staswick, P., and Chrispeels, M.J. (1986). Molecular analysis of two phytohemagglutinin genes and their expression in *Phaseolus vulgaris* cv Pinto, a lectin-deficient cultivar of the bean. EMBO J. **5,** 3075–3082.

Voelker, T.A., Sturm, A.J., and Chrispeels, M.J. (1987). Differences in expression between two seed lectin alleles obtained from normal and lectin-deficient beans are maintained in transgenic tobacco. EMBO J. **6,** 3571–3577.

Voelker, T.A. Herman, E.M., and Chrispeels, M.J. (1989). In vitro mutated phytohemagglutinin genes expressed in tobacco seeds: Role of glycans in protein targeting and stability. Plant Cell **1,** 95–104.