The Plant Cell, Vol. 1, 427–435, April 1989 © 1989 American Society of Plant Physiologists

A Frameshift Mutation Prevents Kunitz Trypsin Inhibitor mRNA Accumulation in Soybean Embryos

K. Diane Jofuku,[1] Roberta D. Schipper, and Robert B. Goldberg[2]
Department of Biology, University of California, Los Angeles, Los Angeles, California 90024-1606

We investigated the molecular basis of a soybean Kunitz trypsin inhibitor (KTi) gene mutation that prevents the accumulation of Kunitz trypsin inhibitor protein during seed development. We found that mRNA encoding the major Kunitz trypsin inhibitor protein (KTi3 mRNA) is reduced at least 100-fold in null (KTi$^-$) embryos but that KTi3 gene transcriptional activity is similar in Kunitz trypsin inhibitor producing embryos (KTi$^+$) and in KTi$^-$ embryos. We sequenced the Kunitz trypsin inhibitor KTi3 gene from both KTi3$^+$ and KTi3$^-$ lines and found that these genes differ by only three nucleotides (+481, +486, and +487) within the KTi3 coding region. Alteration of these nucleotides results in a frameshift within the KTi3$^-$ gene that causes premature termination during translation. Our results suggest that the KTi3$^-$ frameshift mutation results in KTi3$^-$ mRNA destabilization and leads to a drastic reduction in KTi3 mRNA prevalence.

INTRODUCTION

Different soybean inbred lines have been identified that are defective in the expression of specific seed protein genes (see review by Goldberg, Barker, and Perez-Grau, 1989). This natural mutant collection includes varieties that either lack or underproduce Kunitz trypsin inhibitor (Orf and Hymowitz, 1979), seed lectin (Pull et al., 1978), β-amylase (Hildebrand and Hymowitz, 1980), lipoxygenase (Hildebrand and Hymowitz, 1981), glycinin (Scallon, Dickinson, and Nielsen, 1987; Cho et al., 1989), β-conglycinin (Kitamura and Kaizuma, 1981), and urease (Kloth, Polacco, and Hymowitz, 1987). In each inbred line, the defect is the result of a mutation that is inherited as a simple Mendelian recessive (e.g., Orf and Hymowitz, 1979). Several seed protein gene mutations have been studied and their molecular lesions have been characterized (Goldberg et al., 1989). These studies demonstrated that the mutations are caused by different processes and result in either transcriptional or posttranscriptional defects in seed protein gene expression. Soybean seed protein gene mutations include a large 5′ deletion (β-conglycinin cgy_1 gene; Ladin, Doyle, and Beachy, 1983), an insertion element within the gene (lectin *le 1* gene; Goldberg, Hoschek, and Vodkin, 1983), elimination of the translation start codon (glycinin gy_4 gene; Scallon, Dickinson, and Nielsen, 1987), and an inversion that separates the 5′ and 3′ gene regions (glycinin gy_3 gene; Cho et al., 1989).

[1] Current address: Laboratorium voor Genetica, Rijksuniversiteit Gent, B-9000 Gent, Belgium.
[2] To whom correspondence should be addressed.

In an attempt to understand the molecular processes controlling soybean seed protein gene expression, we investigated Kunitz trypsin inhibitor (KTi) gene structure and expression in a Kunitz trypsin inhibitor producing line (KTi$^+$) and in a Kunitz trypsin inhibitor null line (KTi$^-$). The KTi$^-$ genetic lesion results in a substantial reduction of Kunitz trypsin inhibitor protein and activity in soybean seeds (Orf and Hymowitz, 1979). We showed recently that KTi$^+$ soybean plants contain several Kunitz trypsin inhibitor genes and that these genes are differentially expressed during the soybean life cycle and in transformed tobacco plants (Jofuku, 1987; Jofuku and Goldberg, 1989). In this paper we show that only one member of the Kunitz trypsin inhibitor gene family is affected by the Kunitz trypsin inhibitor mutation. DNA sequence analysis indicates that this gene, designated as KTi3, encodes the major Kunitz trypsin inhibitor seed protein. Our results demonstrate that the mutant Kunitz trypsin inhibitor gene contains one base pair substitution and two base pair deletions that result in a translational frameshift. The inability to correctly translate the KTi3$^-$ mRNA probably causes the reduced Kunitz trypsin inhibitor mRNA and protein levels in KTi$^-$ embryos.

RESULTS

KTi$^-$ Embryos Contain Reduced Levels of KTi3 RNA

The soybean genome contains at least 10 distinct Kunitz trypsin inhibitor genes, several of which are expressed during the soybean life cycle (Jofuku, 1987; Jofuku and

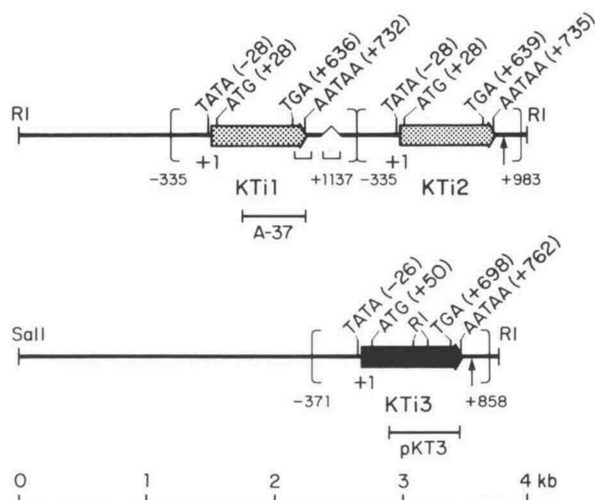


Figure 1. Schematic Representations of the KTi1, KTi2, and KTi3 Kunitz Trypsin Inhibitor Gene Regions.

Relevant consensus sequences and their locations are shown above the genes. RI and SalI represent EcoRI and SalI restriction endonucleases, respectively. Brackets surrounding genes enclose DNA sequences that are more than 75% similar between the KTi1/2 and KTi3 gene regions, and are more than 90% similar between the KTi1 and KTi2 gene regions. Thin brackets below the KTi1 gene and the KTi2 5′ region indicate KTi1 gene nucleotides +674 to +829 that are duplicated to form a portion of the 5′ flanking region of KTi2. Arrows point to KTi2 and KTi3 3′ regions that do not contain the duplicated KTi1 DNA segment. Bracketed lines designated as A-37 and pKT3 indicate the Kunitz trypsin inhibitor cDNA plasmids A-37 (Goldberg et al., 1981a) and pKT3 (Jofuku and Goldberg, 1989) that represent the KTi1/2 and KTi3 mRNAs, respectively.

Goldberg, 1989). Figure 1 presents schematic representations of three Kunitz trypsin inhibitor genes, designated KTi1, KTi2, and KTi3, with their relevant consensus sequences and transcriptional orientations. Regions of the Kunitz trypsin inhibitor genes represented by cDNA plasmids A-37 and pKT3 are also shown in Figure 1. Plasmid A-37 contains sequences present in the KTi1 and KTi2 genes, whereas plasmid pKT3 contains sequences present in the KTi3 gene (Jofuku, 1987; Jofuku and Goldberg, 1989). We demonstrated that the KTi1, KTi2, and KTi3 genes are expressed at relatively high levels during embryogenesis and at lower levels in mature plant organ systems (Jofuku, 1987; Jofuku and Goldberg, 1989). DNA sequencing studies presented here and elsewhere (Jofuku, 1987; Jofuku and Goldberg, 1989) indicated that (1) the KTi1 and KTi2 genes encode minor Kunitz trypsin inhibitor proteins, (2) the KTi1 and KTi2 genes are nearly identical in nucleotide sequence, (3) the KTi3 gene encodes the predominant Kunitz trypsin inhibitor found in soybean seeds, and (4) the KTi1/2 and KTi3 genes have diverged 20% at the DNA sequence level.

We hybridized Kunitz trypsin inhibitor cDNA plasmids A-37 and pKT3 (Figure 1) with soybean midmaturation stage embryo polysomal mRNA gel blots to determine Kunitz trypsin inhibitor KTi1/2 and KTi3 mRNA concentrations in KTi⁺ and KTi⁻ plants. Figure 2A shows that A-37 hybridized with approximately equal intensity to a 0.9-kb mRNA in both KTi⁺ and KTi⁻ mRNA populations. This result indicated that the KTi1/2 mRNAs were present at similar levels in wild-type and mutant embryos. By contrast, Figure 2B shows that pKT3 hybridized with a 0.9-kb KTi⁻ embryo mRNA that was approximately 100-fold less prevalent than that present in KTi⁺ embryo mRNA. Similar

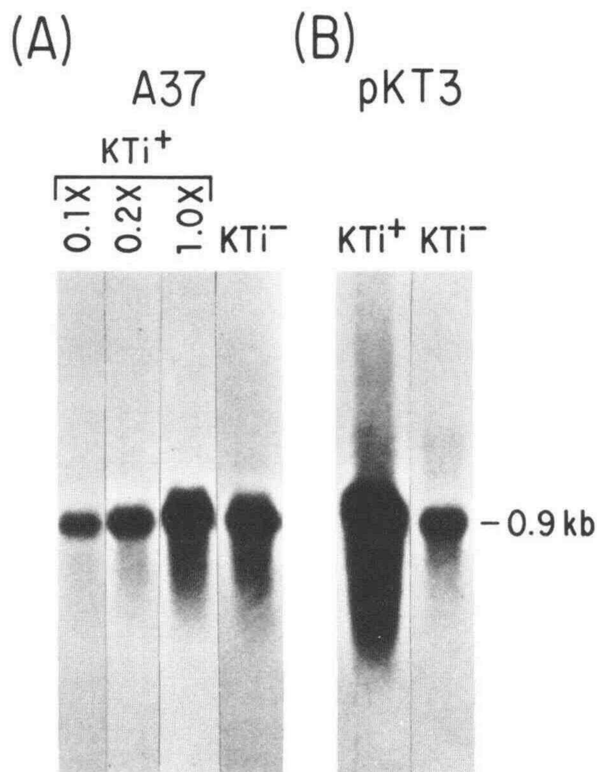


Figure 2. Representation of KTi1/2 and KTi3 Kunitz Trypsin Inhibitor mRNAs in KTi⁺ and KTi⁻ Soybean Lines.

Two micrograms of KTi⁺ and KTi⁻ midmaturation stage embryo polysomal poly(A) mRNAs (Goldberg et al., 1981b) were fractionated by electrophoresis on methylmercury hydroxide gels, transferred to nitrocellulose paper, and then hybridized with labeled A-37 and pKT3 plasmid DNAs (Figure 1) under moderately stringent reaction conditions (42°C; $\Delta T_m = -18°C$). We showed elsewhere that, at this criterion, there is roughly 20% cross-reaction between each plasmid and heterologous KTi1/KTi2 or KTi3 mRNAs (Jofuku and Goldberg, 1989). The 0.1X, 0.2X, and 1.0X lanes contained 0.2 μg, 0.4 μg, and 2 μg of KTi⁺ midmaturation stage mRNA, respectively. These reconstructions represented approximately 0.28%, 0.56%, and 2.8% of the embryo mRNA mass, respectively (Goldberg, et al., 1981a).

results were obtained with KTi⁻ total embryo RNA (data not shown), indicating that the concentration of the KTi3 mRNA was significantly reduced in KTi⁻ cells.

To determine whether the residual KTi3 mRNA detected in KTi⁻ embryos was derived from the KTi3 gene, we utilized two experimental approaches. First, we hybridized KTi⁺ and KTi⁻ embryo mRNAs with a synthetic oligonucleotide probe complementary to KTi1, KTi2, and KTi3 mRNAs, and then extended the primer with reverse transcriptase in the presence of only ^{32}P-dATP (Jofuku and Goldberg, 1989). As shown in Figure 3A, the KTi1/2 and KTi3 mRNAs differ by one nucleotide in the region 5′ to the site of oligonucleotide hybridization (5′-GUUU-3′ versus 5′-UUUU-3′). The extra uracil residue in the KTi3 mRNA permits extension of the oligonucleotide/KTi3 mRNA hybrid by one additional adenine in comparison to that driven by KTi1 or KTi2 mRNA oligonucleotide hybrids. Thus, the primer extension products produced by these reactions should specify the relative KTi1/2 and KTi3 mRNA levels in KTi⁺ and KTi⁻ embryos. Figure 3B shows that the 24-nucleotide KTi1/2 mRNA product was produced at approximately equal levels by both KTi⁺ and KTi⁻ embryo mRNAs. By contrast, the 25-nucleotide KTi3 mRNA product was produced at a higher level than the KTi1/2 mRNA product with KTi⁺ mRNA, and was greatly reduced in the KTi⁻ embryo mRNA population.

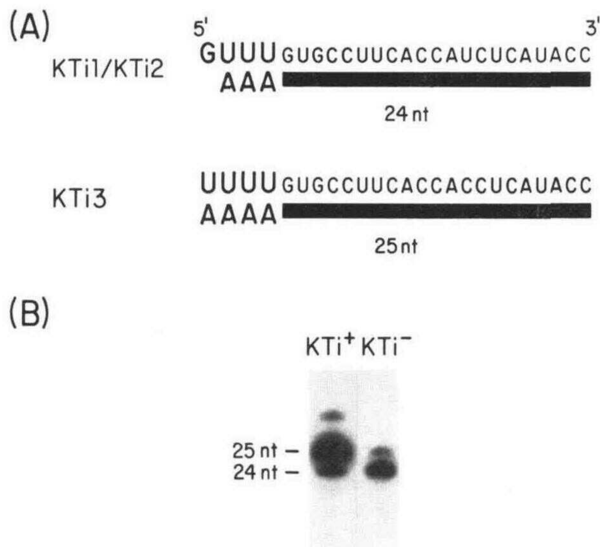


Figure 3. Primer Extension Analysis of KTi1/2 and KTi3 mRNAs in KTi3⁺ and KTi3⁻ Plants.

A 21-base primer was synthesized that is complementary to nucleotides +84 to +105 in the KTi3 Kunitz trypsin inhibitor gene (Figure 8) and to nucleotides +61 to +82 in the KTi1/2 Kunitz trypsin inhibitor genes (Jofuku and Goldberg, 1989). Nucleotide 13 of the primer is divergent between the KTi3 and KTi1/2 genes. An excess of labeled primer was hybridized separately with KTi3⁺ and KTi3⁻ midmaturation stage polysomal poly(A) mRNAs, and reverse transcriptase was used to extend the primer in the presence of only ^{32}P-dATP (Jofuku and Goldberg, 1989). The extended primers were sized on a DNA sequencing gel. The 25-nucleotide (nt) product represents the KTi3 mRNA, and the 24-nucleotide product represents the KTi1/2 mRNAs. The 26-nucleotide product in the KTi⁺ lane may represent extension of a related Kunitz trypsin inhibitor mRNA (Jofuku and Goldberg, 1989).

In the second approach, we hybridized KTi⁺ and KTi⁻ embryo mRNAs with the KTi3 gene probe shown in Figure 4A, and then measured the size of S1 nuclease-resistant DNA fragments. The stringency of both the hybridization criterion and the S1 nuclease protection assay permitted the detection of only KTi3 mRNA sequences. Figure 4B shows that KTi⁺ embryo mRNA protected the predicted DNA fragments (Figure 4B, lane KTi⁺). By contrast, KTi⁻ mRNA did not detectably protect any KTi3 DNA sequence (Figure 4B, lane KTi⁻). Together, these results indicate that the level of KTi3 mRNA is greatly reduced in KTi⁻ embryo cells, and that the concentrations of KTi1, KTi2, and related KTi embryo mRNAs are not detectably affected by the KTi3 gene mutation.

KTi3⁻ Gene Transcription Is Unaffected in Mutant Embryos

The studies presented in Figures 2 to 4 suggested that KTi3 mRNA is reduced in KTi⁻ embryos as a consequence of a defect at either the transcriptional or posttranscriptional levels. We synthesized ^{32}P-nuclear RNAs (nRNAs) in vitro using KTi⁺ and KTi⁻ embryo nuclei (Walling, Drews, and Goldberg, 1986; Cox and Goldberg, 1988), and then hybridized the ^{32}P-nRNAs with DNA gel blots containing KTi3 gene sequences to determine whether the mutant Kunitz trypsin inhibitor gene was transcriptionally active. Figure 5 shows that the KTi⁺ and KTi⁻ ^{32}P-nRNAs produced equivalent hybridization signals with the KTi3 gene (Figure 5, lanes KTi3). Similar results were obtained with genes unaffected by the KTi3⁻ gene mutation, although the hybridization signals were gene-specific (Figure 5, lanes *Le 1* and CG-4). These data indicate that the KTi3 Kunitz trypsin inhibitor gene is transcribed at the same relative rate in both KTi⁺ and KTi⁻ embryos, and that the reduction in KTi3⁻ embryo mRNA is due to posttranscriptional events.

KTi⁺ and KTi⁻ Plants Contain a Single KTi3 Kunitz Trypsin Inhibitor Gene

We hybridized Kunitz trypsin inhibitor cDNA plasmids (pKT3 and A-37) with DNA gel blots containing EcoRI-digested KTi⁺ and KTi⁻ leaf DNAs at a reduced criterion ($\Delta T_m = -30°C$) to determine the KTi3 gene copy number in these lines, and to compare the sizes of DNA fragments that contained Kunitz trypsin inhibitor gene sequences. We

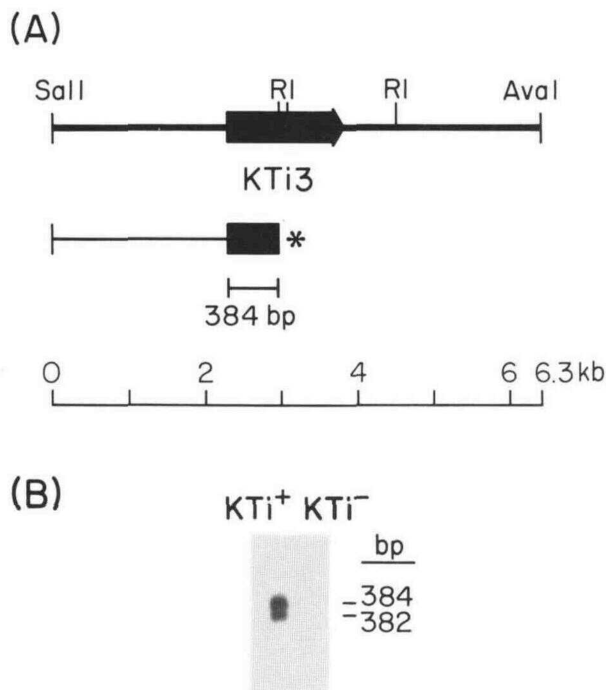


Figure 4. S1 Nuclease Analysis of KTi3 Gene Expression in KTi⁺ and KTi⁻ Embryos.

(A) Map of relevant restriction endonuclease sites within the KTi3 gene region. The line labeled with an asterisk indicates the 3.1-kb SalI/EcoRI DNA fragment from λClone 4A (Figure 7) that was used as a probe in the S1 nuclease mapping experiments shown in **(B)**. The predicted size of the S1 nuclease-resistant KTi3 mRNA/DNA hybrid is indicated. bp, base pairs.

(B) S1 nuclease protection of KTi3 mRNA sequences in KTi⁺ and KTi⁻ embryos. An excess of labeled DNA fragment containing the 5′ end of the KTi3 gene was hybridized separately with 2 μg each of KTi⁺ and KTi⁻ embryo mRNA. The resulting hybrids were digested with S1 nuclease, and the S1 nuclease-resistant hybrids were fractionated by electrophoresis on a DNA sequencing gel. DNA fragment sizes (bp) were determined relative to a known DNA sequence ladder and are shown to the right.

showed elsewhere (Jofuku, 1987; Jofuku and Goldberg, 1989) that, under similar conditions, pKT3 and A-37 hybridized with the 5.5-kb and 0.6-kb KTi3 EcoRI DNA fragments, the 4.0-kb KTi1/KTi2 EcoRI DNA fragment (Figure 1), and other EcoRI DNA fragments containing related Kunitz trypsin inhibitor gene sequences. As seen in Figure 6, the hybridization patterns obtained with the KTi⁺ and KTi⁻ DNAs were identical with the exception of a KTi1/2 EcoRI DNA fragment polymorphism in one KTi⁺ line (compare Figure 6, lanes KTi_F^+ and KTi_D^+). The 5.5- and 0.6-kb KTi3 gene fragments were indistinguishable in size and copy number in the KTi⁺ and KTI⁻ DNAs (Figure 6, lanes KTi⁺, KTI⁻). We conclude that both the KTi⁺ and KTI⁻ genomes contain one KTi3 Kunitz trypsin inhibitor gene, and that no gross sequence alterations have occurred within the mutant KTi3 gene region.

KTi3⁺ and KTi3⁻ Genes Possess Similar Structures

We described elsewhere the isolation and characterization of the KTi3⁺ Kunitz trypsin inhibitor genomic clone, λClone 4A (Jofuku, 1987; Jofuku and Goldberg, 1989). To compare the KTi3⁺ and KTi3⁻ gene regions, we selected Kunitz trypsin inhibitor genomic clones from a KTi⁻ genomic DNA library. Figure 7 shows the restriction maps, gene locations, and transcriptional orientations of the KTi3⁺ genomic clone (λClone 4A) and the KTi3⁻ genomic clone (λClone TIM6). As seen in Figure 7, no significant differences in restriction endonuclease sites were observed between the two gene regions. In addition, at least 40 other restriction endonuclease sites within and flanking the KTi3⁻ gene were identical with those contiguous to the KTi3⁺ gene (data not shown). Heteroduplex analysis also showed that KTi3⁺ and KTi3⁻ gene regions were indistinguishable at the level of the electron microscope (data not shown).

We reacted Kunitz trypsin inhibitor phage DNAs with soybean KTi⁺ midmaturation stage embryo mRNA under conditions that form R-loops (Fischer and Goldberg, 1982) to visualize and compare KTi3⁺ and KTi3⁻ gene structures.

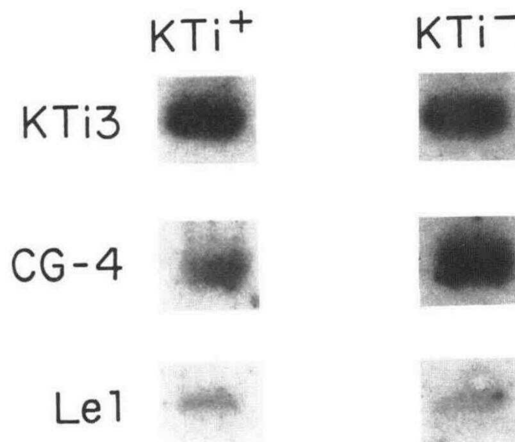


Figure 5. Hybridization of Seed Protein Gene Sequences with ³²P-Nuclear RNAs.

DNA fragments containing the KTi3 Kunitz trypsin inhibitor gene, the *Le 1* lectin gene (Goldberg et al., 1983), and the CG-4 β-conglycinin gene (Barker et al., 1988; Harada et al., 1989) were fractionated by electrophoresis, transferred to nitrocellulose paper, and hybridized with ³²P-nuclear RNAs synthesized in vitro from either KTi⁺ or KTi⁻ midmaturation stage embryo nuclei. Hybridization conditions used are described in the legend to Figure 2.

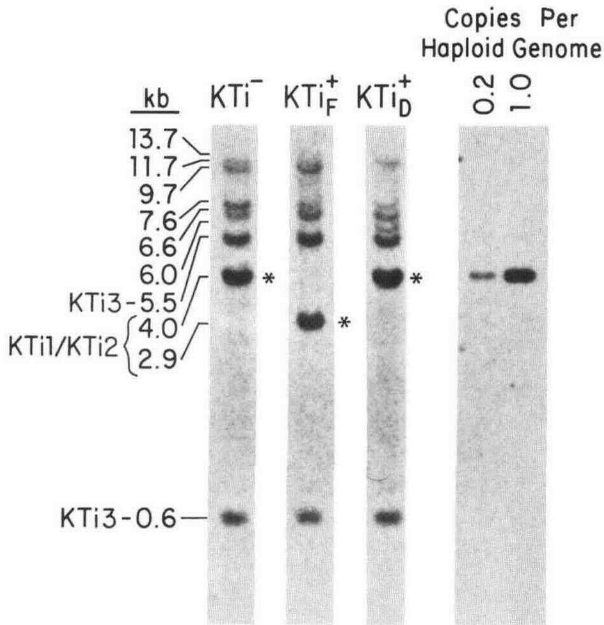


Figure 6. Kunitz Trypsin Inhibitor Gene Representation in KTi$^+$ and KTi$^-$ Genomes.

Soybean leaf nuclear DNAs were digested with EcoRI, fractionated by electrophoresis, transferred to nitrocellulose paper, and hybridized with a mixed probe containing plasmids A-37 and pKT3 (Figure 1) under a reduced hybridization criterion (30°C, $\Delta T_m = -30°$). KTi$^-$, KTi^+_F and KTi^+_D refer to DNAs from P.I. 157440, Forrest, and Dare soybean varieties, respectively. Sizes of EcoRI DNA fragments that reacted with the A-37/pKT3 probe are shown to the left. KTi1/KTi2 and KTi3 refer to the EcoRI DNA fragments that contain the KTi1/KTi2 and KTi3 gene sequences, respectively. The KTi1/2 DNA fragment that is polymorphic with respect to an EcoRI site is designated with an asterisk. The reconstruction lanes (copies per haploid genome) contained 0.2 and 1.0 copy equivalents of EcoRI-digested λClone 4A phage DNA (Figure 7). λClone 4A contains the KTi3$^+$ Kunitz trypsin inhibitor gene and is schematically shown in Figure 7.

Figure 7 shows that R-loops obtained with the KTi3$^+$ gene (λClone 4A) were simple, and that the average R-loop size correlated well with the size of the Kunitz trypsin inhibitor mRNA (data not shown). This finding suggested that, like the KTi1 and KTi2 genes (Jofuku, 1987; Jofuku and Goldberg, 1989), the KTi3$^+$ gene lacks detectable introns. Similarly, Figure 7 shows that R-loops formed between the KTi3$^-$ gene (λClone TIM6) and KTi$^+$ embryo mRNA were identical in structure and size to those obtained with the KTi3$^+$ Kunitz trypsin inhibitor gene. Using these experimental criteria, our results indicated that the KTi3$^+$ and KTi3$^-$ trypsin inhibitor genes were indistinguishable from each other, and that no detectable rearrangements, insertions, or deletions occurred within the KTi3$^-$ gene or contiguous DNA regions.

The KTi3$^-$ Kunitz Trypsin Inhibitor Gene Contains Mutations That Result in a Translational Frameshift

We sequenced the KTi3$^+$ and KTi3$^-$ genes to determine whether small sequence alterations may have occurred within the KTi3$^-$ gene region. Figure 8 presents the KTi3$^+$ gene sequence. As predicted from the R-loops shown in Figure 7, the KTi3 gene lacks introns. Translation of the KTi3 gene sequence indicates that the KTi3 gene encodes the major Kunitz trypsin inhibitor protein found in soybean seeds (Kim et al., 1985).

We compared the KTi3$^+$ and KTi3$^-$ gene sequences and found that they were identical from nucleotides −148 to +887, with the exception of nucleotides +481, +486, and +487 (Figure 8). As shown in Figure 9A, two deletions and one G → T transversion occurred within the KTi3$^-$ coding region. Figure 9B shows that these mutations result in a translational frameshift that causes four stop codons to be inserted into the KTi3$^-$ mRNA reading frame. We conclude that these mutations result in premature termination of KTi3$^-$ mRNA translation, and are responsible for the reduction in KTi3 mRNA (Figures 2 to 4) and Kunitz trypsin inhibitor protein in KTi$^-$ embryos.

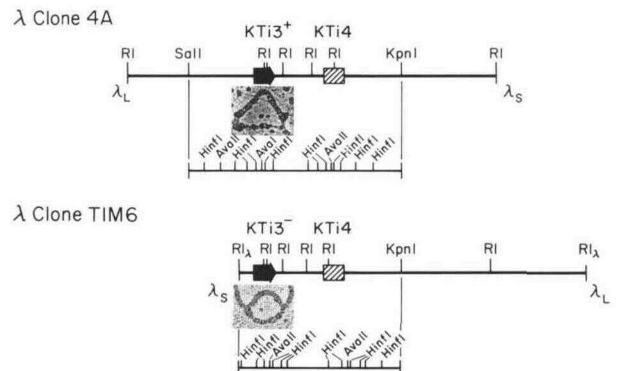


Figure 7. R-Loop and Restriction Endonuclease Site Analyses of KTi3 Kunitz Trypsin Inhibitor Gene Regions in KTi$^+$ and KTi$^-$ Plants.

Maps of restriction endonuclease sites that are present within KTi3 Kunitz trypsin inhibitor gene regions are shown. Insert sizes of the Kunitz trypsin inhibitor phages were 16.6 kb (λClone 4A) and 16 kb (λClone TIM 6), respectively. Boxes and arrows indicate gene locations and transcriptional orientations, respectively. λ_S and λ_L refer to the long and short λ phage arms, respectively. KTi4 box refers to the KTi4 Kunitz trypsin inhibitor gene (Jofuku and Goldberg, 1989) and represents maximum gene length. KTi4 gene structure has not yet been determined, nor is it known whether KTi4 is a functional gene. R-loops were formed by hybridizing Kunitz trypsin inhibitor phage DNAs with an excess of KTi$^+$ midmaturation stage embryo mRNA (Fischer and Goldberg, 1982; Jofuku and Goldberg, 1988). KTi3$^+$ and KTi3$^-$ R-loops are shown below their respective genes.

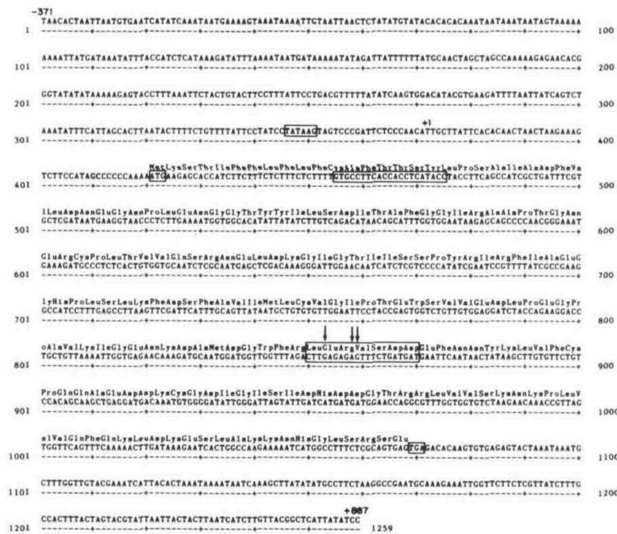


Figure 8. KTi3+ Kunitz Trypsin Inhibitor Gene Sequence.

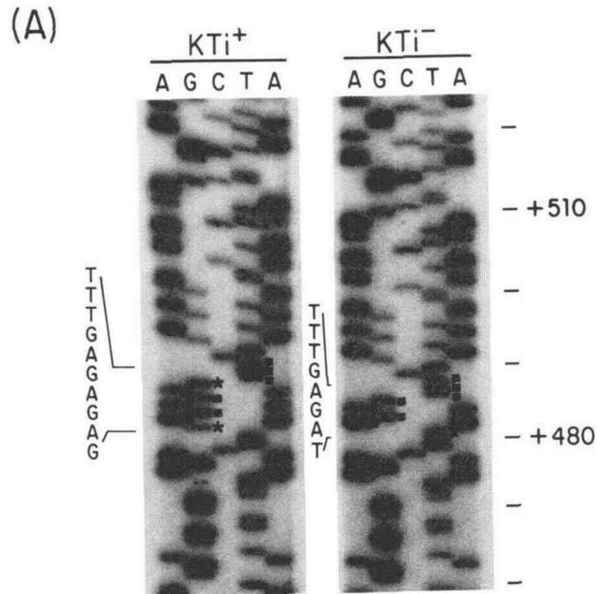
Nucleotides are numbered from the transcription start site (+1). The amino acid sequence derived from the nucleotide sequence is also shown. The sequence of the KTi3− gene is identical to the KTi3+ sequence shown with the exception of nucleotides +481, +486, and +487 (indicated by arrows). Boxes specify a TATA-like sequence, the translation start codon, the gene region that is mutated in the KTi3− genome, and the sequence complementary to the 21 nucleotide primer used for the primer extension experiment shown in Figure 3.

DISCUSSION

The KTi3 Gene Encodes the Major Soybean Kunitz Trypsin Inhibitor Protein

Kunitz trypsin inhibitor is a prevalent protein that is found in soybean seeds (Ryan, 1981; Vodkin, 1981; Vodkin and Raikhel, 1986). This protein accumulates during seed development, is present in the embryonic axis and cotyledons (Horisberger and Tacchini-Vonlanthen, 1983), is packaged into protein bodies (Horisberger and Tacchini-Vonlanthen, 1983; Vodkin and Raikhel, 1986), and is not significantly degraded at germination (Freed and Ryan, 1978; Horisberger and Tacchini-Vonlanthen, 1983). In addition, Kunitz trypsin inhibitor protein (Vodkin and Raikhel, 1986) and Kunitz trypsin inhibitor mRNA (Jofuku and Goldberg, 1989) are present in lower amounts in mature plant organ systems. The physiological significance, if any, of Kunitz trypsin inhibitor protein is not known, although a nonrelated trypsin inhibitor has been shown to protect plants from insect predation (Hilder et al., 1987).

The soybean genome contains at least 10 distinct DNA sequences that are complementary to Kunitz trypsin inhibitor cDNA plasmids (Jofuku, 1987; Jofuku and Goldberg, 1989). Several of these DNA sequences (e.g., KTi1, KTi2, and KTi3) represent functional genes that are expressed in a developmental-specific manner during the soybean life cycle and in transformed tobacco plants (Jofuku and Goldberg, 1989). The KTi3 gene has a higher relative transcription rate than the KTi1 and KTi2 genes (Jofuku and Gold-



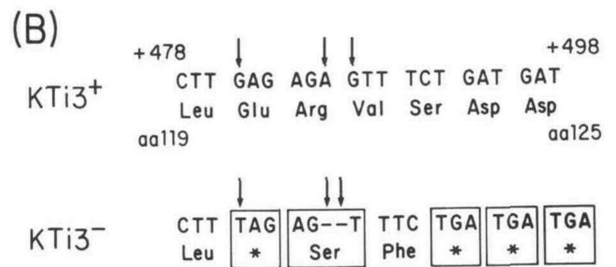


Figure 9. A Frameshift Mutation within the KTi3− Kunitz Trypsin Inhibitor Gene.

(A) Nucleotide sequence comparison of the KTi3+ and KTi3− genes. DNA sequence ladders represent nucleotides +460 to +520 of the KTi3+ and KTi3− gene regions. Nucleotides +481 to +490 of the KTi3+ gene sequence and the corresponding KTi3− sequence are shown to the left. Boxes correspond to nucleotides that are conserved between the two genes. The asterisks show the nucleotides that have been mutated in the KTi3− gene.

(B) Nucleotide and amino acid sequence comparisons of the KTi3+ and KTi3− Kunitz trypsin inhibitor genes and proteins. Only nucleotides +478 to +498 and translated amino acids 119 to 125 are shown. Arrows refer to the nucleotides that have been mutated in the KTi3− gene [shown as asterisks in **(A)**]. The asterisks designate translational stop codons that result from these three mutations.

berg, 1989), and produces a more prevalent Kunitz trypsin inhibitor mRNA (Figures 2 and 3). Comparison of the KTi3 gene translated protein sequence (Figure 8) with that of the Kunitz trypsin inhibitor Ti^a allelic form (Kim et al., 1985) indicates that they are identical. The translated protein sequence has 24 amino acids at the amino terminus and 11 amino acids at the carboxy terminus that are not present in the mature Kunitz trypsin inhibitor protein. We presume that the extra amino acids at the amino terminus represent a signal peptide that is processed during translation (Vodkin, 1981), and that those at the carboxy terminus are cleaved during seed maturation (Crouch et al., 1983). We conclude that the KTi3 gene is highly expressed in KTi^+ plants, and that this gene encodes the predominant Kunitz trypsin inhibitor protein stored in soybean seeds.

The KTi^- Phenotype Is Due to a Mutation in the KTi3 Structural Gene

KTi^- seeds fail to accumulate significant amounts of both Kunitz trypsin inhibitor protein and activity (Orf and Hymowitz, 1979), although there is cross-reacting material that interacts with Kunitz trypsin inhibitor antibodies (K. D. Jofuku and R. B. Goldberg, unpublished results). Presumably, the cross-reacting material represents the products of divergent Kunitz trypsin inhibitor genes (e.g., KTi1 and KTi2). The KTi^- phenotype is inherited as a simple Mendelian recessive trait (Orf and Hymowitz, 1979). Figures 8 and 9 show that three mutations have occurred in the $KTi3^-$ gene coding region. These mutations lead to a frameshift that causes premature termination of KTi3 mRNA translation and result in the production of a truncated protein (Figure 9B). Because there are no other detectable alterations in the $KTi3^-$ gene (Figure 8), or flanking DNA regions (Figures 6 and 7), we conclude that mutations within the $KTi3^-$ structural gene are responsible for the Kunitz trypsin inhibitorless phenotype.

Posttranscriptional Events Reduce KTi3 mRNA Levels in $KTi3^-$ Embryos

The representation of KTi3 gene transcripts in $KTi3^-$ embryo polysomal RNA (Figures 2 to 4), and total RNA (R. D. Schipper and R. B. Goldberg, unpublished results), is at least 100-fold lower than that found at the same developmental stage in $KTi3^+$ embryos. Thus, the reduced KTi3 mRNA prevalence in KTi^- embryo cells is due to an actual reduction in KTi3 mRNA molecules rather than simply a release of KTi3 mRNA from $KTi3^-$ embryo polysomes. By contrast, the relative transcription rate of the KTi3 gene is the same in $KTi3^+$ and $KTi3^-$ embryos (Figure 5). This result indicates that the $KTi3^-$ gene is transcriptionally active, and that transcription is not affected detectably by the $KTi3^-$ frameshift mutation. We conclude that the reduction in $KTi3^-$ mRNA is due to posttranscriptional events such as increased turnover of $KTi3^-$ gene transcripts within the nucleus, altered nuclear export efficiency, reduced $KTi3^-$ mRNA cytoplasmic stability, or a combination of these events. The reduction of Kunitz trypsin inhibitor protein in KTi^- seeds is a direct consequence of the lower KTi3 mRNA levels, although the production of truncated polypeptides may contribute to the reduction as well.

Frameshift mutations in other plant and animal genes also result in lower cytoplasmic mRNA levels. For example, a frameshift that causes stop codons to be inserted into the bean (*Phaseolus vulgaris*) Pdlec1 lectin gene results in a 600-fold reduction of Pdlec1 mRNA (Voelker, Staswick, and Chrispeels, 1986). Similarly, nonsense mutations and translational frameshifts that produce stop codons in the human β-globin gene reading frame also result in a drastic lowering of β-globin cytoplasmic mRNA levels (Orkin and Kazazian, 1984). These studies, and the results presented here, strongly suggest that uninterrupted translational processes are required in order to establish correct cytoplasmic levels of individual mRNAs.

The precise mechanism by which the KTi3 mRNA level is reduced in KTi^- embryos is not yet known. Because KTi3 gene transcripts do not undergo splicing reactions in the nucleus (Figures 7 and 8), the simplest explanation for our results is that the half-life of KTi3 mRNA is lowered in KTi^- embryos. Premature uncoupling of $KTi3^-$ mRNA from embryo polysomes by translation termination could result in greater susceptibility of $KTi3^-$ mRNA molecules to ribonucleases that are responsible for normal mRNA turnover in embryo cells. Decreased cytoplasmic stability of eukaryotic mRNAs containing nonsense mutations has been directly demonstrated by pulse-chase studies in animal cells (Kinniburgh et al., 1982; Ross and Pizarro, 1983; Maquat, Chilcote, and Ryan, 1985). Although we have no evidence to suggest that this also occurs in KTi^- embryos, the direct correlation between polysome loss and seed protein mRNA degradation during seed dehydration (Goldberg et al., 1981a, 1981b; Walling et al., 1986; Barker, Harada, and Goldberg, 1988) suggests that the presence of seed protein mRNAs in polysome structures contributes to their stability during seed development.

Posttranscriptional processes have been shown to be important in regulating seed protein gene expression (Walling et al., 1986; Goldberg et al., 1989; Harada, Barker, and Goldberg, 1989). The precise details by which posttranscriptional events establish seed protein mRNA levels remain to be determined.

METHODS

Growth of Plants and Developmental Staging

KTi^+ soybean lines Dare and Forrest were obtained from the USDA soybean germplasm collection in Stoneville, MS. The Kunitz

trypsin inhibitor KTi⁻ line, P.I. 157740, was obtained from Dr. T. Hymowitz, Agronomy Department, University of Illinois. Plants were grown in the greenhouse and embryos were harvested and staged as described previously (Goldberg et al., 1981b).

Isolation of Kunitz Trypsin Inhibitor Phages and cDNA Plasmids

The characteristics of the $KTi3^+$ Kunitz trypsin inhibitor recombinant phage, λClone 4A, and cDNA plasmids A-37 and pKT3 were described by Jofuku and Goldberg (1989).

DNA Isolation and Labeling

Soybean, plasmid, and phage DNAs were isolated as described by Jofuku and Goldberg (1989). DNAs were labeled to specific activities of $>10^8$ cpm/μg by nick translation as specified by Bethesda Research Laboratories.

Polysomal mRNA Isolation

Soybean polysomal poly(A) mRNAs were isolated as outlined by Cox and Goldberg (1988). Total embryo cell RNAs were isolated according to the guanidinium thioisocyanate procedure of Chirgwin et al. (1979).

Construction of a KTi^- Genomic DNA Library

A λ Charon 35 partial Sau3A library of soybean KTi^- nuclear DNA was constructed as described by Jofuku and Goldberg (1988, 1989). This library contained over 10^6 phages of which >99% were recombinants.

Gel Blot Studies

DNA and mRNA gel blot experiments were carried out according to procedures as outlined in Jofuku and Goldberg (1988, 1989).

Primer Extension and S1 Nuclease Studies

Extension of DNA/RNA hybrids with reverse transcriptase was performed according to the procedure outlined by Jofuku and Goldberg (1989). S1 nuclease protection experiments were carried out according to the procedures of Berk and Sharp (1977) and Rosbash et al. (1979).

R-Loop Analysis

R-loops were formed between Kunitz trypsin inhibitor phage DNAs and soybean KTi^+ mid-maturation stage embryo mRNAs as described by Fischer and Goldberg (1982) and Rosbash et al. (1979).

^{32}P-Nuclear RNA Synthesis

Synthesis of labeled RNA from isolated nuclei was carried out as described by Walling et al. (1986) and Cox and Goldberg (1988).

ACKNOWLEDGMENTS

We are deeply indebted to Drs. John Harada and Lila Vodkin for their aid and advice in the experiments reported here. We thank Dr. Allan Caplan for critical reading of this manuscript. This research was supported by a United States Department of Agriculture grant to R.B.G. K.D.J. was supported by an ARCO Cell Research Institute Predoctoral Fellowship and R.D.S. was supported by a McKnight Foundation Predoctoral Fellowship.

Received February 22, 1989.

REFERENCES

Barker, S.J., Harada, J.J., and Goldberg, R.B. (1988). Cellular localization of soybean storage protein mRNA in transformed tobacco seeds. Proc. Natl. Acad. Sci. USA **85,** 458–462.

Berk, A.J. and Sharp, P.A. (1977). Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. Cell **12,** 721–732.

Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J., and Rutter, W.J. (1979). Isolation of biologically active RNA from sources enriched in ribonuclease. Biochemistry **18,** 5294–5299.

Cho, T.J., Davies, C.S., Fischer, R.L., Turner, N.E., Goldberg, R.B., and Nielsen, N.C. (1989). Molecular characterization of an aberrant allele for the Gy_3 glycinin gene: A chromosomal rearrangement. Plant Cell **1,** 339–350.

Cox, K.H., and Goldberg, R.B. (1988). Analysis of plant gene expression. In Plant Molecular Biology: A Practical Approach. C.H. Shaw, ed. (Oxford: IRL Press), pp. 1–34.

Crouch, M.L., Tenbarge, R.M., Simon, A.E., and Ferl, R. (1983). cDNA clones for *Brassica napus* seed storage proteins: Evidence from nucleotide sequence analysis that both subunits of napin are cleaved from a precursor polypeptide. J. Mol. Appl. Genet. **2,** 273–283.

Fischer, R.L., and Goldberg, R.B. (1982). Structure and flanking regions in soybean seed protein genes. Cell **29,** 651–660.

Freed, R.C., and Ryan, D.S. (1978). Changes in Kunitz trypsin inhibitor during germination of soybean: An immunological assay system. J. Food Sci. **43,** 1316–1319.

Goldberg, R.B., Hoschek, G., Ditta, G.S., and Breidenbach, R.W. (1981a). Developmental regulation of cloned superabundant embryo mRNAs in soybean. Dev. Biol. **83,** 201–217.

Goldberg, R.B., Hoschek, G., Tam, S.H., Ditta, G.S., and Breidenbach, R.W. (1981b). Abundance, diversity, and regulation of mRNA sequence sets in soybean embryogenesis. Dev. Biol. **83,** 201–217.

Goldberg, R.B., Hoschek, G., and Vodkin, L.O. (1983). An insertion sequence blocks the expression of a soybean lectin gene. Cell **34,** 465–475.

Goldberg, R.B., Barker, S.J., and Perez-Grau, L. (1989). Regulation of gene expression during plant embryogenesis. Cell **56,** 149–160.

Harada, J.J., Barker, S.J., and Goldberg, R.B. (1989). Soybean β-conglycinin genes are clustered in several DNA regions and are regulated at both transcriptional and posttranscriptional levels. Plant Cell **1,** 415–425.

Hildebrand, D.F. and Hymowitz, T. (1980). Inheritance of β-amylase nulls in soybean seed. Crop Sci. **20,** 727–730.

Hildebrand, D.F., and Hymowitz, T. (1981). Two soybean genotypes lacking lipoxygenase-1. J. Am. Oil Chem. Soc. **58,** 583–586.

Hilder, V.A., Gatehouse, A.M.R., Sheerman, S.E., Barker, R.F., and Boulter, D. (1987). A novel mechanism of insect resistance engineered into tobacco. Nature **330,** 160–163.

Horisberger, M., and Tacchini-Vonlanthen, M. (1983). Ultrastructural localization of Kunitz trypsin inhibitor on thin sections of soybean by the gold method. Histochemistry **77,** 37–50.

Jofuku, K.D. (1987). Developmental regulation of soybean seed protein gene expression. Ph.D. Dissertation, University of California, Los Angeles, California.

Jofuku, K.D., and Goldberg, R.B. (1988). Analysis of plant gene structure. In Plant Molecular Biology: A Practical Approach. C.H. Shaw, ed (Oxford: IRL Press), pp. 37–66.

Jofuku, K.D., and Goldberg, R.B. (1989). Kunitz trypsin inhibitor genes are differentially expressed during the soybean life cycle and in transformed tobacco plants. Plant Cell, in press.

Kim, S.-H., Hara, S., Hase, S., Ikenaka, T., Toda, H., Kitamura, K., and Kaizuma, N. (1985). Comparative study on amino acid sequences of Kunitz-type soybean trypsin inhibitors, Ti^a, Ti^b, and Ti^c. J. Biochem. **98,** 435–448.

Kinniburgh, A.J., Maquat, L.E., Schedl, T., Rachmilewitz, E., and Ross, J. (1982). mRNA-deficient β^o-thalassemia results from a single nucleotide deletion. Nucl. Acids Res. **10,** 5421–5427.

Kitamura, K., and Kaizuma, N. (1981). Mutant strains with low levels of subunits of 7S globulin in soybean seed. Japan J. Breed. **31,** 353–359.

Kloth, R.H., Placco, J.C., and Hymowitz, T. (1987). The inheritance of a urease null trait in soybean. Theor. Appl. Genet. **73,** 410–418.

Ladin, B.F., Doyle, J.J., and Beachy, R.N. (1983). Molecular characterization of a deletion mutation affecting the α'-subunit of β-conglycinin of soybean. J. Mol. Appl. Genet. **2,** 372–380.

Maquat, L.E., Chilcote, R., and Ryan, P.M. (1985). Human triosephosphate isomerase cDNA and protein structure: Studies of triosephosphate isomerase deficiency in man. J. Biol. Chem. **260,** 3748–3753.

Orf, J.H., and Hymowitz, T. (1979). Inheritance of the absence of the Kunitz trypsin inhibitor in seed protein of soybeans. Crop Sci. **19,** 107–109.

Orkin, S. A., and Kazazian, H.H. (1984). The mutation and polymorphism of the human β-globin gene and its surrounding DNA. Annu. Rev. Genet. **18,** 131–171.

Pull, S.P., Pueppke, S.G., Hymowitz, T., and Orf, J.H. (1978). Soybean lines lacking the 120,000-dalton seed lectin. Science **220,** 1277–1279.

Rosbash, M., Blank, D., Fahrner, K., Hereford, L., Ricciardi, R., Roberts, B., Ruby, S., and Woolford, J. (1979). R-looping and structural gene identification of recombinant DNA. Methods Enzymol. **68,** 454–469.

Ross, J., and Pizarro, A. (1983). Human beta and delta globin messenger RNAs turn over at different rates. J. Mol. Biol. **167,** 607–617.

Ryan, C.A. (1981). Proteinase inhibitors. In Biochemistry of Plants. Vol. 6, A. Marcus, ed. (New York: Academic Press), pp. 351–370.

Scallon, B.J., Dickinson, C.D., and Nielsen, N.C. (1987). Characterization of a null-allele for the Gy_4 glycinin gene from soybean. Mol. Gen. Genet. **209,** 197–113.

Vodkin, L.O. (1981). Isolation and characterization of messenger RNAs for seed lectin and Kunitz trypsin inhibitor in soybeans. Plant Physiol. **68,** 766–771.

Vodkin, L.O., and Raikhel, N.V. (1986). Soybean lectin and related proteins in seeds and roots of Le^+ and Le^- soybean varieties. Plant Physiol. **81,** 558–565.

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.

Walling, L., Drews, G.N., and Goldberg, R.B. (1986). Transcriptional and post-transcriptional regulation of soybean seed protein mRNA levels. Proc. Natl. Acad. Sci. USA **83,** 2123–2127.