The Plant Cell, Vol. 3, 225–245, March 1991 © 1991 American Society of Plant Physiologists

Molecular Mechanisms Underlying the Differential Expression of Maize Pyruvate, Orthophosphate Dikinase Genes

Jen Sheen

Department of Genetics, Harvard Medical School, and Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114

I describe here the organization of maize C4 chloroplast and non-C4 cytosolic pyruvate, orthophosphate dikinase (PPDK) genes and the molecular mechanisms underlying their differential expression. The maize C4 chloroplast PPDK gene (*C4ppdkZm1*) appears to have been created by the addition of an exon encoding the chloroplast transit peptide at a site upstream of a cytosolic PPDK gene (*cyppdkZm1*). A splice acceptor sequence located in the first exon of *cyppdkZm1* allows the fusion of the transit peptide to the *cyppdkZm1* sequences. A second cyPPDK gene (*cyppdkZm2*) shares extensive homology with *cyppdkZm1* in the coding region and in the 5′ flanking region up to the TATA box. By a novel protoplast transient expression method, I show that the light-inducible expression of *C4ppdkZm1* is controlled by two expression programs mediated through separate upstream regulatory elements that are active in leaf, but inactive in root and stem. Light-mediated *C4ppdkZm1* expression in maize is apparently uncoupled from leaf development and partially associated with chloroplast development. For *cyppdkZm1* expression, distinct upstream elements and a specific TATA promoter element, located in the first intron of *C4ppdkZm1*, are required. The low expression of *cyppdkZm2* can be attributed to an absence of upstream positive elements and weak activity of the TATA promoter element.

INTRODUCTION

In an NADP-malic enzyme C4 plant such as maize, the genes encoding the C4 pathway enzymes—including pyruvate, orthophosphate dikinase (PPDK), phosphoenol pyruvate carboxylase (PEPC), NADP-malate dehydrogenase, and NADP-malic enzyme—have undergone evolutionary modification for C4 photosynthesis. The genes encoding these C4 enzymes are differentially expressed in two photosynthetic cell types, bundle sheath and mesophyll cells, unique to C4 plants (Sheen and Bogorad, 1987; Langdale et al., 1988; Nelson and Langdale, 1989). PPDK, PEPC, and NADP-malate dehydrogenase are required for the efficient fixation of CO_2 in mesophyll cells. The prefixed CO_2 is transferred to bundle sheath cells as malate and released by NADP-malic enzyme. The Calvin cycle pathway, active only in bundle sheath cells, then refixes the released CO_2 through ribulose bisphosphate carboxylase in a relatively high CO_2 environment that greatly reduces photorespiration and photooxidation (Hatch, 1976, 1987; Edwards and Walker, 1983; Foyer, 1984; Edwards and Ku, 1987; Nelson and Langdale, 1989). Under conditions of high photon flux, temperature, and aridity, the adaptation of C4 photosynthesis offers some advantages for plant survival (Widholm and Ogren, 1969; Björkman and Berry, 1973; Berry, 1975).

At present, the molecular mechanisms for the evolution and regulation of C4 photosynthetic genes are unknown. It has been postulated that either new C4 genes were created and/or the expression of existing genes was modified to facilitate C4 functions (Aoyagi and Bassham, 1983; Rothermel and Nelson, 1989). In a recent paper, Glackin and Grula (1990) showed that organ-specific transcripts of different size and abundance derive from the same PPDK gene in maize. I show here that there are three endogenous PPDK genes, one encoding the C4 chloroplast PPDK (C4PPDK) and two encoding cytosolic PPDK activities (cyPPDK1, cyPPDK2). The C4PPDK gene (*C4ppdkZm1*) overlaps the cyPPDK1 gene (*cyppdkZm1*) and has apparently been created from *cyppdkZm1* by genomic rearrangement. The cyPPDK2 gene (*cyppdkZm2*) is highly homologous to *cyppdkZm1* in coding and 5′ regions up to the TATA box. Each PPDK gene has a distinct expression pattern in root, stem, and leaf tissues and is differentially regulated by light in leaves. Using protoplasts isolated from maize seedlings and a transient expression method, I demonstrate that the differential expression of the maize PPDK genes is regulated by distinct promoters. The light-inducible expression of *C4ppdkZm1* in leaves is uncoupled from leaf development and partially associated with chloroplast development. Distinct upstream regulatory ele-

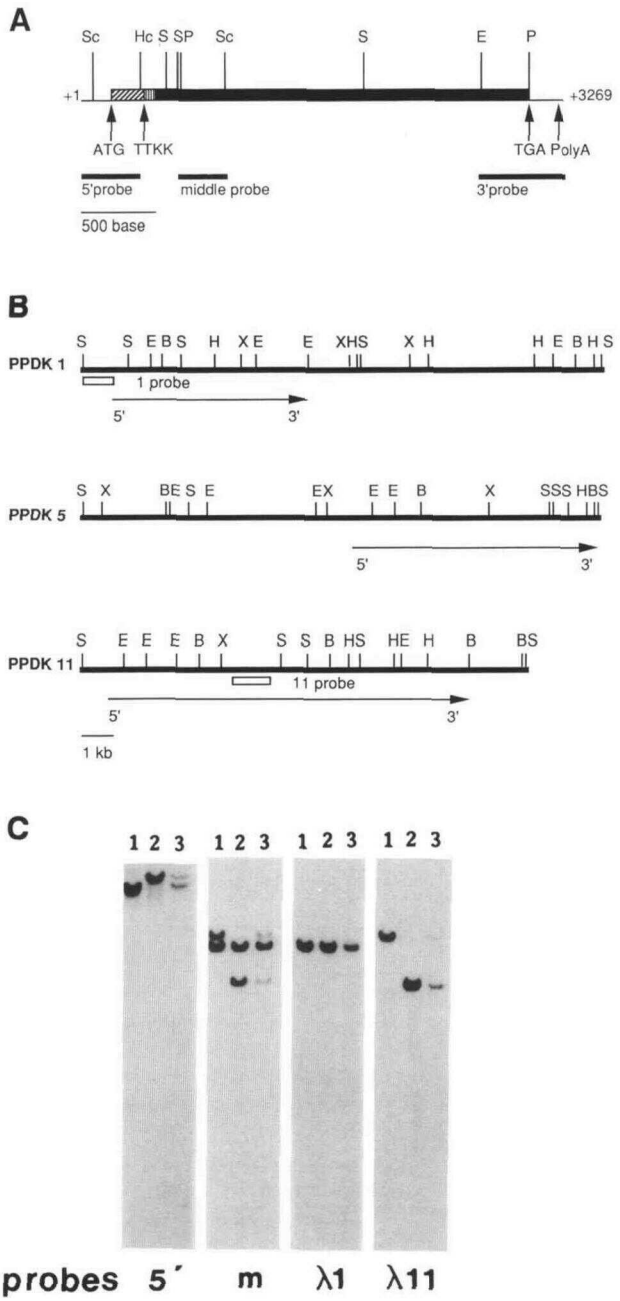


Figure 1. Maps and Genomic DNA Blot Analysis of the Maize PPDK Genes.

(A) Full-length cDNA map of the C4PPDK gene. The hatched box indicates the transit peptide coding region. The TTKK is the first 4 amino acids of the mature C4PPDK and is shown as the vertically striped box. The black box is the coding region of the mature C4PPDK. The short black lines beneath the map indicate the location of probes used for the analyses shown in **(B)** and **(C)**.

ments and TATA promoter elements are required for the expression of the PPDK genes in different tissues. Combinatorial control of the PPDK gene promoters is demonstrated. No transcriptional interference is observed for the promoter activities of the two overlapping PPDK genes.

RESULTS

Isolation of Maize PPDK cDNA and Genomic Clones

To isolate longer cDNA clones encoding C4PPDK, a maize leaf cDNA library was made with size-selected cDNA of 3 kb to 4 kb and screened with a partial C4PPDK cDNA clone identified previously (Sheen and Bogorad, 1987). In Figure 1A, the restriction map of the full-length C4PPDK cDNA is shown. The data presented in Figure 1A are consistent with the results presented by Matsuoka et al. (1988) and Glackin and Grula (1990).

To isolate genomic clones of PPDK genes, a maize genomic library was screened with the same partial C4PPDK cDNA probe. As shown in Figure 1B, three phage genomic clones, λppdk1, λppdk5, and λppdk11, were obtained, each containing a distinct insert.

The transcription orientation of each clone was determined by hybridizing restriction fragments of each with 5′, middle, and 3′ cDNA probes (Figures 1A, 1B, and data not shown). The λppdk1 clone showed strong homology with the middle and 3′ probes but not the 5′ probe that contained the transit peptide coding region. The λppdk5 clone showed strong homology with the 5′ and middle probes but not the 3′ probe. The λppdk11 clone showed strong homology with all three probes. The three genomic clones were highly homologous because hybridization signals were the same when washed at either high or low stringency (data not shown). The results suggest that λppdk1 probably encodes a cytosolic PPDK, whereas both λppdk5 and λppdk11 encode a chloroplast PPDK.

The Maize C4 Chloroplast PPDK Gene Is Highly Polymorphic

It remained unclear whether λppdk5 and λppdk11 were separate genes or polymorphic forms of a single gene. The

(B) Genomic maps. Arrowed lines represent transcripts. Open boxes represent the location of the λ1 and λ11 probes used for the analyses shown in **(C)**.

(C) Genomic DNA blot analysis. About 10 μg of maize genomic DNA isolated from the maize line FR9^{cms} (lanes 1), FR37 (lanes 2), and FR9^{cms} × RF37 (lanes 3) was digested with BamHI and electrophoresed in a 0.7% agarose gel. Probes are shown in **(A)** and **(B)**.

Restriction enzymes are B, BamHI; E, EcoRI; Hc, HincII; H, HindIII; P, PstI; S, SalI; Sc, SacI; X, XbaI.

latter was thought possible because the genomic library was constructed with DNA isolated from a hybrid maize line ($FR9^{cms} \times FR37$). To confirm the gene copy number of the maize PPDK genes, genomic DNA analysis was performed with DNA isolated from both the inbred lines ($FR9^{cms}$ or FR37) and the hybrid line. Four probes were used for the DNA blot analysis. They were derived from the 5′ and middle portions of the cDNA (Figure 1A), the 5′ flanking region of λppdk1, and the intron region of λppdk11 (Figure 1B). As shown in Figure 1C, the inbred lines yielded single bands of different sizes when the genomic DNA was hybridized with the 5′ and λ11 probes, whereas both bands appeared in genomic DNA from the hybrid line. The signals from the allelic genes in the hybrid line were much weaker. These results indicate that there is a single copy of the C4PPDK transit peptide sequence (homologous to the 5′ probe) in the maize genome and that λppdk and λppdk11 are the polymorphic forms of the C4PPDK gene derived from the maize lines $FR9^{cms}$ and FR37, respectively.

The middle probe, common to all PPDK genes, detected two bands in each inbred line and three bands in the hybrid line. The same band detected in both inbred and hybrid lines by the middle probe was also revealed by the λ1 probe. Thus, λppdk1 appears to encode a homologous cytosolic PPDK that does not show polymorphism between the two maize lines. Because λppdk1 contains only 1.5 kb of the 5′ upstream sequence, genomic DNA analysis was performed to ensure that the λppdk1 clone does not contain a chloroplast transit peptide sequence further upstream. BstXI, which cuts once in the middle of the λ1 fragment, was used to digest FR37 genomic DNA. Two BstXI genomic fragments of 8 kb and 10 kb hybridized with the λ1 probe, whereas a distinct band of larger size hybridized with the 5′ probe (data not shown). Therefore, there are two loci for PPDK genes in maize chromosomes. This conclusion agrees with the results of Helentjaris et al. (1988) and Glackin and Grula (1990). However, the locus lacking the transit peptide sequence shows no polymorphism (also examined with EcoRI, HindIII, and XbaI, data not shown), whereas the locus bearing the transit peptide sequence shows a high level of polymorphism (Figures 1B and 1C). In these two maize lines, the C4PEPC gene also shows higher polymorphism than other non-C4PEPC genes (A. Schäffner and J. Sheen, unpublished data). Although a high degree of polymorphism is always found with DNAs from unrelated maize lines (Helentjaris et al., 1988), the different levels of polymorphism between C4 genes and their non-C4 homologs in two related maize lines suggest that the C4 genes are located in dynamic regions of maize chromosomes, which might play a role in their evolution (Berry, 1975; McClintock, 1984; Walbot and Cullis, 1985).

C4 Chloroplast and Cytosolic PPDK Are Highly Homologous but Bear Distinct N Termini

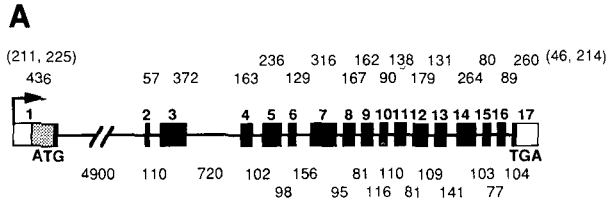
To understand the evolutionary relationship between the two homologous genes encoding the C4 chloroplast and non-C4 cytosolic PPDK, the structure of the C4PPDK gene (clone λppdk11) and 5′ portion of the cytosolic PPDK gene (clone λppdk1) were analyzed by DNA sequencing. Sequencing primers were made according to the sequence of a C4PPDK cDNA (Matsuoka et al., 1988) to determine the location and size of exons and introns of the C4PPDK gene. As shown in Figure 2A, the C4PPDK gene contains 17 exons and 16 introns. The first intron is unusually large (5.6 kb in λppdk5 and 4.9 kb in λppdk11) compared with other introns (77 bp to 720 bp) and separates the transit peptide coding region from the mature C4PPDK protein coding region. The sequence encoding the first 4 amino acids of the mature C4 chloroplast PPDK is in the first exon in agreement with the data presented by Matsuoka et al. (1988) and Glackin and Grula (1990).

As shown in Figure 2B, the homology between the 237-bp fragment of the two PPDK genes, encoding the N-terminal portions of the cytosolic PPDK and the mature C4PPDK, is 97% at the nucleotide level and 91% at the amino acid sequence level when the sequences encoding the first 11 residues of the cytosolic PPDK and the first 4 residues of the mature C4PPDK are ignored. The point of divergence between the two PPDK genes is located precisely at the 3′ end of the first intron of the C4PPDK gene (Figure 2B).

The C4PPDK Gene Overlaps a Cytosolic PPDK Gene

To understand the divergence between the C4PPDK and cyPPDK genes, the 3′ portion of the first intron of the C4PPDK gene was sequenced. In Figures 3A and 3B, the sequence shows that the first intron contains sequences highly similar to the nonallelic cytosolic PPDK gene in the 5′ coding and noncoding regions up to the TATA box. Thus, an intact gene, which is capable of encoding another cytosolic PPDK, overlaps the C4PPDK gene. The two overlapping genes, encoding the C4PPDK and one cytosolic PPDK (cyPPDK1), were designated *C4ppdkZm1* and *cyppdkZm1*. The independent PPDK gene found in clone λppdk1 encodes another cytosolic PPDK (cyPPDK2) and was designated *cyppdkZm2*.

The first intron of *C4ppdkZm1* is large enough (5.6 kb in λppdk5 and 4.9 kb in λppdk11) to encode another gene. RNA blot analysis was therefore carried out to explore that possibility. As shown in Figure 4A, several probes (from the λppdk5 clone) that covered the whole intron were used for the analysis. The 5′ and middle cDNA probes (Figure 1A) were used as positive controls. As shown in Figure 4B, no other transcripts, which might be regulated by the *C4ppdkZm1* promoter or an internal promoter, were re-



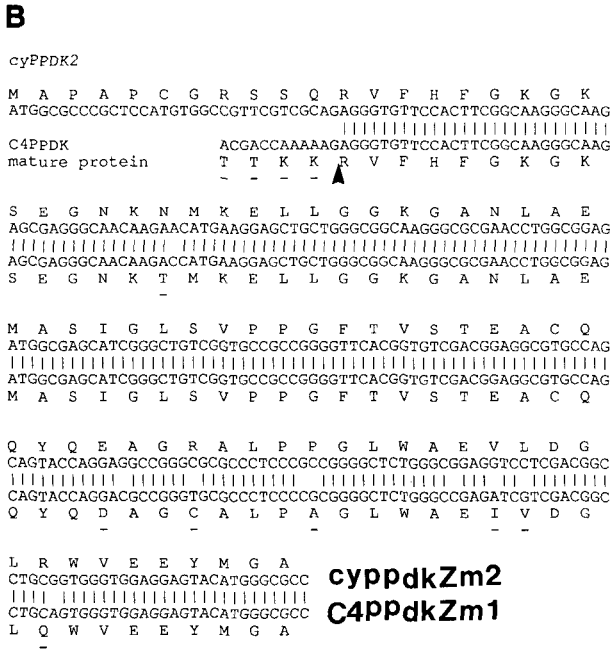


Figure 2. Structure and Sequence Comparison of PPDK Genes.

(A) A schematic representation of the C4PPDK gene. The arrow indicates the transcription initiation site. Open boxes are the 5′ and 3′ noncoding regions of the first and 17th exons. Stippled box is the transit peptide coding region in the first exon. Black boxes represent the coding regions of exons for the mature C4PPDK. Lines represent introns. The number of exons is indicated with boldface numbers. The sizes of exons and introns are shown. The cDNA sequence for the complete coding region of the C4PPDK has been published (Matsuoka et al., 1988).

(B) Sequence comparison of C4PPDK and cyPPDK2 genes. The N-terminal portions of the coding sequences for the mature C4PPDK and cyPPDK2 are shown. The upper DNA sequence encodes the cyPPDK2 and the lower DNA sequence encodes the mature C4PPDK. Matched nucleotide sequences are shown. The unique amino acids of the mature C4PPDK are underlined. The arrow indicates the divergent point of the two PPDK genes and the splice acceptor site of the first intron of the C4PPDK gene.

vealed even after long exposure of the filters. Genomic DNA blot analysis was also carried out with the same probes. The intron contains mostly repetitive sequences (data not shown). These repetitive sequences are present at about 50 to 100 copies per genome (data not shown) and might have played a role in bringing the duplicated *cyppdkZm1* to the vicinity of the regulatory and transit peptide sequences of the *C4ppdkZm1* by genomic rearrangement.

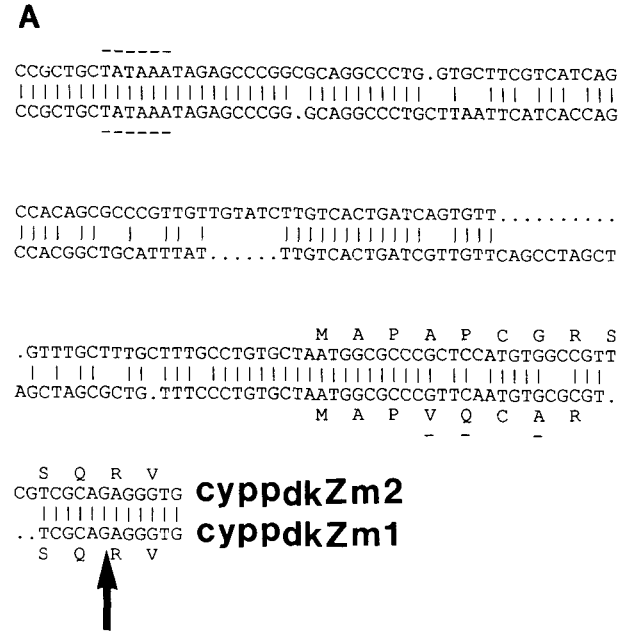


Figure 3. Sequence Comparison of *cyppdkZm1* and *cyppdkZm2*.

(A) Sequence comparison of the 5′ noncoding and coding regions of *cyppdkZm1* and *cyppdkZm2*. The 5′ coding and noncoding sequences of *cyppdkZm1* in the first intron of *C4ppdkZm1* are compared with the homologous regions of *cyppdkZm2*. The TATA box and the distinct amino acids of cyPPDK1 are indicated by dashed lines. The arrow indicates the splice acceptor site of the first intron of *C4ppdkZm1*. The coding sequence of *cyppdkZm1* after the splice acceptor site is identical to that of *C4ppdkZm1*, which has been compared with *cyppdkZm2* in Figure 2B.

(B) A schematic representation of the 5′ portions of *cyppdkZm2* and *C4ppdkZm1* in the context of *C4ppdkZm1*. The TATA boxes, putative promoters, first ATG sites, transit peptide, splicing sites, and the first intron are shown. The first 10 residues of cyPPDK1 and the first 11 residues of cyPPDK2 are presented. The dashed line indicates a deletion in *cyppdkZm1*.

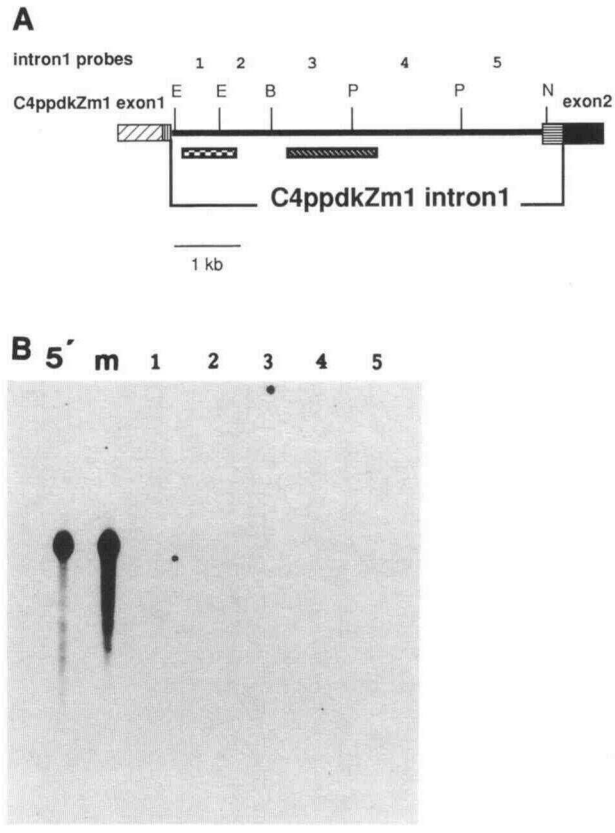


Figure 4. Analysis of the *C4ppdkZm1* Intron 1.

(A) Map of the *C4ppdkZm1* intron 1. The exon 1, 2 and the intron 1 of *C4ppdkZm1* are shown. The large hatched box in exon1 is the transit peptide coding region. The striped boxes in exon1 and intron1 are the unique N termini of C4PPDK and cyPPDK1, respectively. The black box is the exon 2 of *C4ppdkZm1*. The boxes under the map represent the location of two types of repetitive sequences (data not shown). Probes 1 to 5 used for analyses shown in **(B)** are indicated. Restriction enzmes are E, EcoRI; B, BamHI; P, PstI; N, Nar.

(B) RNA gel blot analysis with the *C4ppdkZm1* intron 1 probes. About 1 μg of polyA+ mRNA isolated from 24-hr greening leaves was loaded in each lane and electrophoresed in a 1% agarose gel. The autoradiogram was overexposed to detect potential transcripts. Probes 5′ and m (middle) are indicated in Figure 1A.

Transcription Initiation Sites of PPDK Genes

To determine the transcription initiation sites of the PPDK genes, primer extension was performed with gene-specific primers and reverse transcriptase. Because the 5′ regions of the two cyPPDK genes share extensive homology, it is difficult to obtain gene-specific results by nuclease protection methods based on hybridization of longer probes. To map the transcription initiation sites precisely, the same ^{32}P-labeled primers were used for primer extension and dideoxy DNA sequencing. The products of both reactions were loaded onto the same sequencing gel in parallel.

As shown in Figure 5A, the first 10 nucleotides of the *C4ppdkZm1* transcript are ACCCGTTCGC. This result is

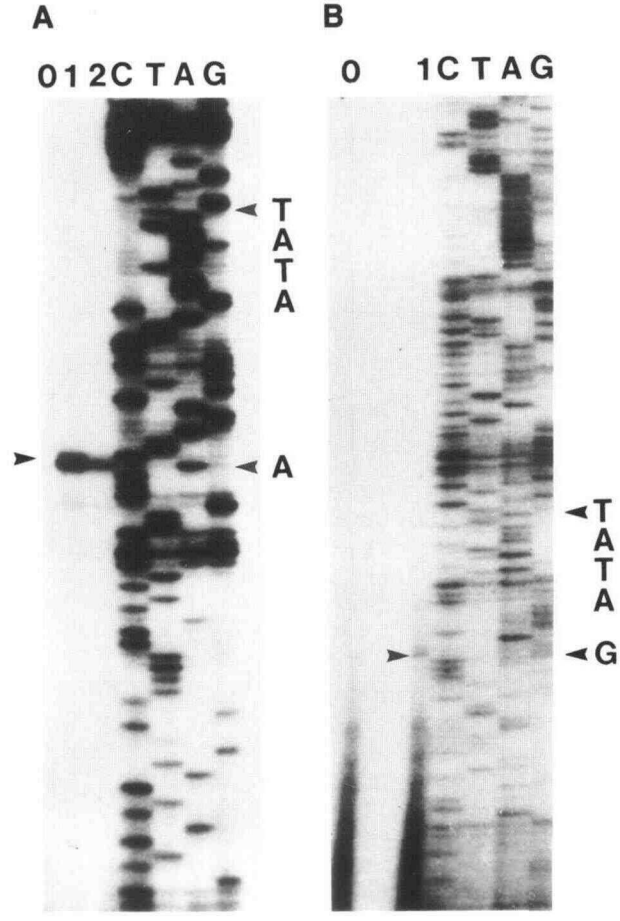


Figure 5. Transcription Initiation Sites of PPDK Genes.

(A) Transcription initiation site of *C4ppdkZm1*. About 1 μg of the polyA+ mRNA isolated from greening maize leaves (lanes 1) or green maize leaves (lanes 2) was used for primer extension with the *C4ppdkZm1* primer. Controls (lanes 0) were carried out with primers and 20 μg of tRNA. About 20% of the primer extension product was loaded on an 8 M urea, 6% polyacrylamide gel. DNA sequences generated with the same ^{32}P-labeled primers were loaded on the same gel in parallel. The autoradiogram was exposed for 2 days. The transcription initiation sites and TATA boxes are shown. The primer extension experiments were repeated twice with the same results.

(B) Transcription initiation site of *cyppdkZm1*. About 10 μg of polyA+ mRNA isolated from greening maize leaves was used for primer extension with the *cyppdkZm1* primer.

Arrows indicate the locations of transcription initiation sites in TATA boxes.

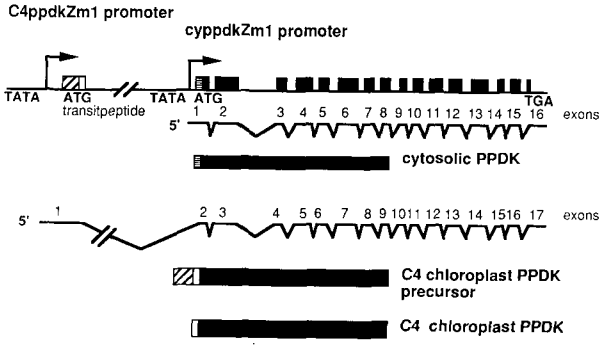


Figure 6. Summary of the Two Overlapping PPDK Genes.

Arrows indicate the transcription initiation sites. The hatched box is the transit peptide coding region unique to *C4ppdkZm1*. Open and striped boxes indicate the unique N termini of mature C4PPDK and cyPPDK1, respectively. Black boxes represent coding regions. The number of exons is shown. Black lines under the number of exons represent the pretranscripts of *C4ppdkZm1* and *cyppdkZm1*.

consistent with the 5′ mapping data of Glackin and Grula (1990). However, the A instead of the C is the predominant transcription initiation site (Figure 5A). As shown in Figure 5B, the 10 nucleotides surrounding the transcription initiation site of *cyppdkZmq* are GCAGGCCCTG and identical between *cyppdkZm1* and *cyppdkZm2* (Figures 3A and 3B). The *cyppdkZm1* transcript initiates predominantly at GCCCTG (Figure 5B). A gene-specific primer was also made from the unique sequence of the *cyppdkZm2*. However, the primer extension product is too short to be separated clearly from the background generated by the ^{32}P-labeled primer (data not shown).

The primer extension data presented here should be reliable because all primers contain gene-specific sequences, are close to the transcription initiation sites, form extension products consistent with the location of the upstream TATA box (Figure 3A), and have generated reproducible results in three independent experiments (data not shown).

The primer extension results document the existence of the putative *cyppdkZm1* promoter located in the first intron of *C4ppdkZm1* (Figures 3A and 3B). The activity of this putative *cyppdkZm1* promoter is further demonstrated by a protoplast transient expression method described below.

Figure 6 summarizes the organization of the promoters, exons, introns, transcripts, and protein products of the two overlapping PPDK genes (*C4ppdkZm1* and *cyppdkZm1*).

Three PPDK Genes Have Distinct Promoters

In Figure 7, the 5′ noncoding and coding sequences of *C4ppdkZm1*, *cyppdkZm1*, and *cyppdkZm2* are shown. The 320-bp sequence upstream of the TATA box of the putative *C4ppdkZm1* promoter contains at least six inverted repeats and two 31-bp direct repeats, whereas the sequence between 321 bp to 500 bp upstream of the TATA box is AT-rich (about 80%) (Figure 7A). These sequences agree mostly with the 5′ noncoding sequences reported by Glackin and Grula (1990) except for five nucleotide differences, which may be attributed to polymorphism in different maize lines.

As shown by Glackin and Grula (1990), there are two potential start codons in the first exon. Translation initiated at the first ATG codon will add 24 more residues to the 71-residue transit peptide first reported by Matsuoka et al. (1988). The transit peptide shown here is translated from the second ATG codon because a TAG stop codon is present after the first ATG codon in the Golden Cross Bantam maize line used by Matsuoka et al. (1988) and the sequence surrounding the second ATG is closer to the eukaryotic translation initiation consensus (Kozak, 1987; Lutcke et al., 1987; Glackin and Grula, 1990). Moreover, as pointed out by Matsuoka et al. (1988), the first 6 residues of the chloroplast transit peptide starting from the second ATG more closely resemble the proposed consensus sequence of transit peptide (Karlin-Neumann and Tobin, 1986). However, further evidence will be required to assign the first ATG conclusively. The second, ninth, 21st, 49th, and 66th residues of the transit peptide are T, I, D, D, and D, respectively, and differ from those reported by Glackin and Grula (1990). DNA sequence comparison indicates that the second residue codon is ACG in the FR9 and FR37 maize lines used here and GCG in the other two maize lines (Matsuoka et al., 1988; Glackin and Grula, 1990).

As shown in Figures 7B and 7C, the putative promoters of *cyppdkZm1* and *cyppdkZm2* are identical only up to 7 bp upstream of the TATAAA sequence. There are AT-rich and AG-rich sequences scattered around the putative promoters of *cyppdkZm1* and *cyppdkZm2*.

Differential Expression and Splicing of PPDK Genes in Leaf, Root, and Stem Tissues of Maize

To determine the expression patterns of each of the PPDK genes, a rapid and sensitive polymerase chain reaction (PCR) method with cDNA templates and gene-specific primers was used (Wang et al., 1989). Because the three PPDK genes share identical or homologous sequences, it is difficult to obtain results of high resolution and specificity by nuclease protection or RNA blot analyses. For instance, using a 1.8-kb 3′ PPDK probe, Glackin and Grula (1990) detected a broad band on RNA blots with etiolated and green maize leaf RNA. As shown in Figures 8C and 8D, this RNA is actually a mix of three PPDK transcripts revealed by PCR using gene-specific primers. Analyzing cDNA with gene-specific primers by PCR, cloning, and

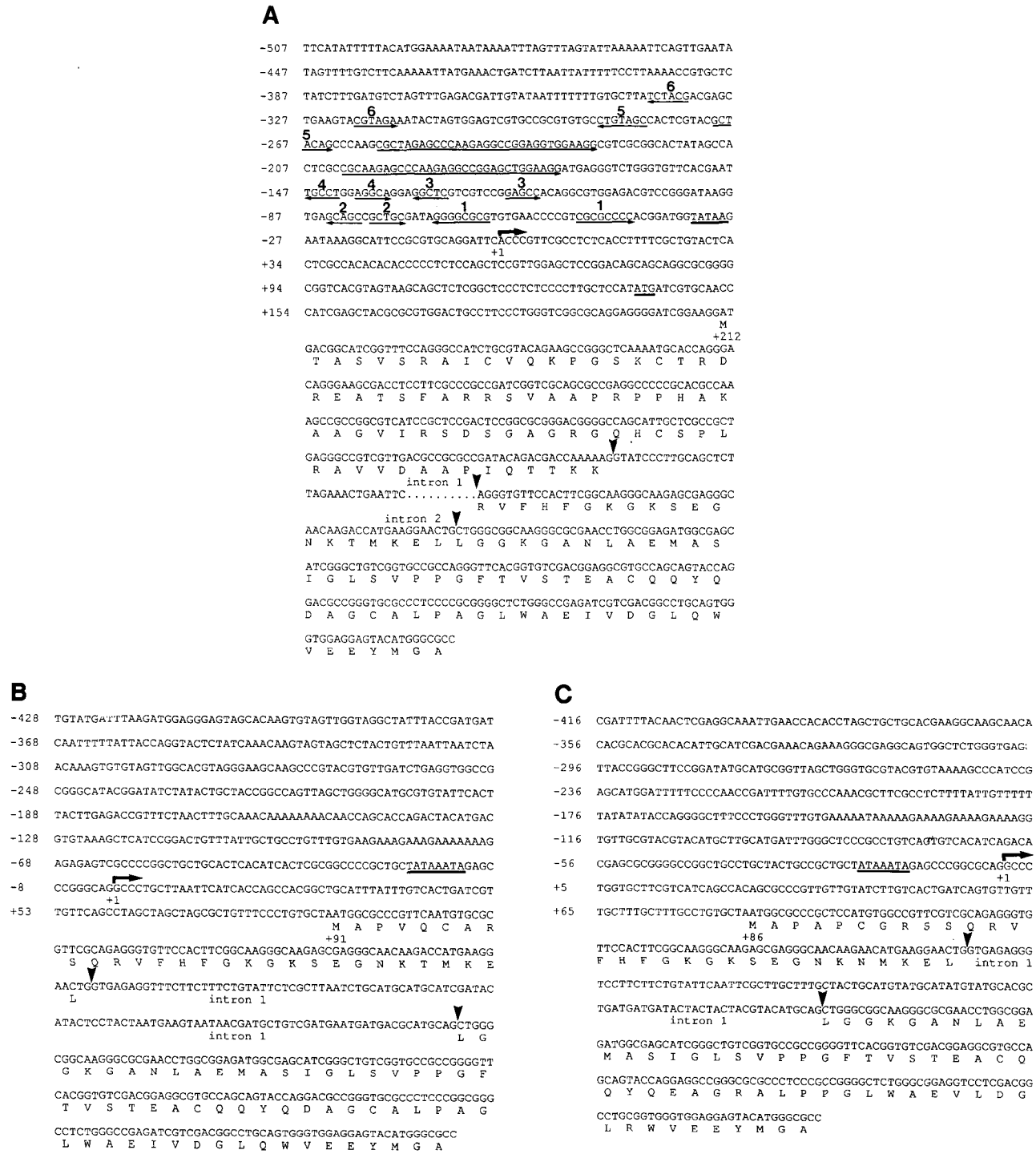


Figure 7. Sequence Analysis of the 5′ Regions of Three PPDK Genes.

(A) The 5′ noncoding and coding sequences of *C4ppdkZm1*. The TATA box, the transcription initiation site, the first ATG, and the location of introns are shown. Six inverted repeats are indicated with numbers and arrowed lines, and the two 31-bp direct repeats are indicated.
(B) The 5′ noncoding and coding sequences of the *cyppdkZm1*.
(C) The 5′ noncoding and coding sequences of the *cyppdkZm2*. The putative transcription initiation site is indicated based on the sequence similarity between the two cyPPDK genes.

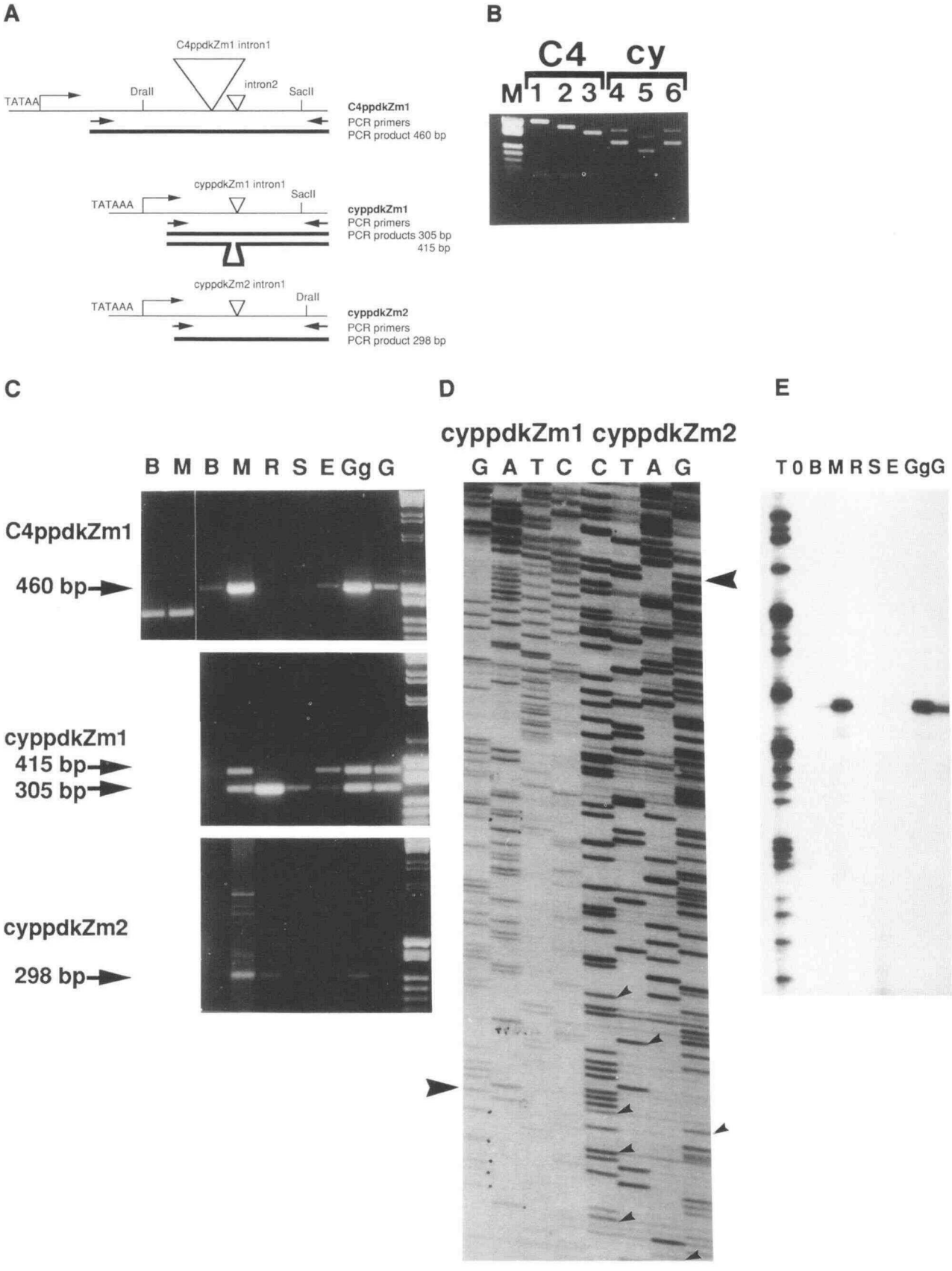


Figure 8. Differential Expression of Three PPDK Genes in Maize.

sequencing is a more reliable and definitive method to demonstrate the expression of a specific gene among genes sharing identical or homologous sequences.

The use of 5′ primers specific for each PPDK gene guaranteed the specificity of the PCR products and allowed the precise determination of intron splicing by DNA sequencing. For instance, the *C4ppdkZm1* PCR product was expected to be a unique fragment of 460 bp with the first two introns spliced, as shown in Figure 8A. The specific 5′ primer of the *cyppdkZm1* would give PCR products of 305 bp if the first intron of *cyppdkZm1* were completely spliced (Figure 8A). The specific 5′ primer of the *cyppdkZm2* would give a PCR product of 298 bp if its first intron were spliced (Figure 8A). The amount of cDNA templates and primers, number of cycles, and reaction temperature were calibrated first to ensure that all PCRs were carried out in a linear range and generated gene-specific products. The possibility of contamination by genomic DNA was eliminated because no DNA bands were observed when the PCR was performed with RNA samples directly without the cDNA synthesis step (data not shown). All PCR results were confirmed by digestion with enzymes recognizing characteristic sites (Figures 8A and 8B) and cDNA cloning and sequencing (Figure 8D and data not shown).

As shown in Figure 8C, the *C4ppdkZm1* transcript is detected only in leaves but not in either roots or stems, even when 20 times more cDNA is used. The accumulation of *C4ppdkZm1* transcript in greening leaves is about 50-fold higher than in etiolated leaves and fivefold higher than in green leaves (estimated by serial dilutions; data not shown). The *cyppdkZm1* transcript is expressed in leaves, roots, and stems, and the *cyppdkZm2* transcript is found at low levels in all tissues examined. Both cyPPDK transcripts are about fivefold higher in greening leaves than in etiolated leaves (Figure 8C). The majority of PPDK transcripts in leaves are located in mesophyll cells (Figure 8C). The coexpression of *C4ppdkZm1* and *cyppdkZm1* in etiolated, greening, and green mesophyll cells indicates that the two promoters of the overlapping PPDK genes do not interfere with each other (Figure 8C).

As predicted from DNA sequence data (Figures 3A, 3B, and 7B), the PCR product of *cyppdkZm1* cDNA should contain the splice acceptor site of the *C4ppdkZm1* first intron. This was confirmed by cloning and sequencing, as shown in Figure 8D. However, there are two sizes of PCR product derived from the *cyppdkZm1* transcripts (Figure 8C). The larger mRNA (415 bp) contains the first intron of *cyppdkZm1* (confirmed by cloning and DNA sequencing; data not shown). Both mRNAs are present at similar levels in greening and green leaves and in mesophyll cells of greening leaves. A higher amount of unspliced mRNAs was detected in etiolated leaves. In roots and stems, the spliced, small mRNA is dominant. In contrast, only the unspliced larger mRNA was detected in bundle sheath cells of greening leaves. Because the unspliced intron would terminate translation, differential splicing of the *cyppdkZm1* transcript might play a role in the cell-type-specific expression of *cyppdkZm1*.

The expression of *cyppdkZm2* was confirmed by cloning and sequencing of the cDNA amplified by PCR, as shown in Figure 8D. The activity of the *cyppdkZm2* promoter was also demonstrated by protoplast transient expression, as shown below. The DNA bands larger than 298 bp do not contain the first intron but consist of the 298-bp fragment fused to sequences of unknown origin (confirmed by cloning and sequencing; data not shown).

Figure 8. (continued).

(A) Schematic representation of the PCR analysis of PPDK transcripts.

(B) PCR products of *C4ppdkZm1* (C4) and *cyppdkZm1* (cy) transcripts. First-strand cDNA was synthesized from polyA+ mRNA isolated from 200 μg of total RNA of 24-hr greening leaves and diluted in water (100 μL final volume). For PCR analysis of the *C4ppdkZm1* and *cyppdkZm1* transcripts, 0.5 μL and 10 μL of cDNA were used in 100 μL of reaction, respectively. In lanes 1, 2, and 3, 5 μL of the PCR product was loaded, and in lanes 4, 5, and 6, 10 μL of the PCR product was loaded. Lanes 1 and 4 were the PCR product without enzyme digestion. Lanes 2 and 5 were the PCR product digested with SacII and lanes 3 and 6 with DraII.

(C) Tissue distribution of three PPDK transcripts. First-strand cDNAs were synthesized from polyA+ mRNA isolated from about 200 μg of total RNAs. About 0.5 μL, 10 μL, and 20 μL of cDNAs (from 100 μL total) were used to detect the *C4ppdkZm1*, *cyppdkZm1*, and *cyppdkZm2* transcripts, respectively, in various tissues by PCR, except 10 μL of cDNAs was used to detect the *C4ppdkZm1* transcript in roots and stems. In each lane, 5 μL, 10 μL, and 35 μL of the PCR product was loaded in a 2% agarose gel for the analysis of the *C4ppdkZm1*, *cyppdkZm1*, and *cyppdkZm2* transcripts, respectively. The first two lanes in the *C4ppdkZm1* panel were PCR products of *rbc* large subunit binding protein cDNA (J. Sheen, unpublished data) in bundle sheath cells (B) and mesophyll cells (M) as a control. R, roots; S, stems; E, etiolated leaves; Gg, 24-hr greening leaves; G, green leaves.

(D) An autorad showing the sequences of *cyppdkZm1* and *cyppdkZm2* leaf cDNA amplified by PCR. The large arrow in the *cyppdkZm1* panel indicates the splice acceptor site of the first intron of *C4ppdkZm1* located in the first exon of *cyppdkZm1* (shown in Figure 7B). The large arrow in the *cyppdkZm2* panel shows the fusion site of the first and second exon (indicated in Figure 7C). Small arrows indicate the unique bases in *cyppdkZm2*.

(E) Analysis of the *C4ppdkZm1* expression by primer extension. The experiment was performed with the same polyA+ mRNA samples as for the synthesis of cDNA for the PCR analysis in **(C)**. Lane T, T sequencing reaction product as marker; lane 0, tRNA control.

Because the specific *cyppdkZm1* 5′ primer is located in the first intron of *C4ppdkZm1*, it could detect the transcript initiated from the *cyppdkZm1* promoter or the *C4ppdkZm1* pretranscript. The bands shown in Figure 8C represent the *cyppdkZm1* transcript but not the *C4ppdkZm1* pretranscript for the following reasons. First, when the *C4ppdkZm1* intron was used as probe for RNA blot analysis, only probe 5 detected the 3-kb *cyppdkZm1* transcript, and no transcript larger than 3.3 kb was detected after exposure for 20 times longer than needed to show the *C4ppdkZm1* transcript (Figure 4B and data not shown). Second, Glackin and Grula (1990) detected PPDK transcripts of smaller size and low abundance in root and etiolated leaves. Their 3-kb transcript should be the equivalent of the *cyppdkZm1* and *cyppdkZm2* transcripts. No larger transcript of similar abundance, which would account for the *C4ppdkZm1* pretranscript, was detected in their RNA blots. Third, the DNA sequence data and the mapping of the transcription initiation site support the location and activity of the putative *cyppdkZm1* promoter (Figures 3A, 3B, 5B, and 7B). Finally, the activity of the *cyppdkZm1* promoter was directly demonstrated by protoplast transient expression shown below.

To confirm that the PCR method is reliable for the quantitation of mRNA, primer extension was also carried out with the same RNA samples for PCR analysis and the *C4ppdkZm1* primer. As shown in Figure 8E, the expression pattern of *C4ppdkZm1* revealed by the primer extension method is very similar to that revealed by the PCR method (Figure 8C).

Transient Expression in Maize Root, Stem, and Leaf Protoplasts

To study the molecular mechanisms underlying the differential regulation of maize PPDK genes, a novel protoplast transient expression method was developed. Protoplasts were isolated from roots, stems, and leaves of maize seedlings, as shown in Figure 9. To examine the transcriptional and translational activities of these protoplasts, a plasmid containing the cauliflower mosaic virus 35S promoter (35S), the chloramphenicol acetyltransferase (CAT) gene, and the 3′ sequence of the nopaline synthase gene (*nos*) was electroporated into protoplasts under optimized conditions. As shown in Table 1, high CAT activity was detected in cell extracts of all protoplasts 20 hr after DNA transfection. A construct without promoter gave background expression (data not shown) as reported previously (Sheen, 1990).

Tissue-Specific Regulation of ppdkCAT Chimeric Genes in Maize Root, Stem, and Leaf Protoplasts

To determine whether the putative PPDK promoters (Figure 7) mediate tissue-specific gene regulation, three ppdkCAT chimeric genes were constructed by fusing the 5′ regions of the three PPDK genes to the CAT coding region and the 3′ *nos* sequences, as shown in Figure 10. Plasmids containing these ppdkCAT chimeric genes were electroporated into protoplasts isolated from roots, stems, and leaves of maize seedlings. In all experiments, a plasmid containing the 35S promoter, the β-glucuronidase (GUS) gene, and the 3′ sequence of *nos* was included as an internal control. 35SCAT was also routinely included as a parallel control.

In the C4ppdkCAT construct, 1.3 kb of the 5′ flanking and untranslated region of *C4ppdkZm1* was fused at +72 (relative to the transcription initiation site) to the CAT gene. A high level of CAT activity was detected in mesophyll protoplasts but not in root and stem protoplasts when C4ppdkCAT was electroporated (Figure 10). To generate cyppdk1CAT, 1.1 kb of the 5′ flanking and untranslated region of *cyppdkZm1* was joined at +96 (with respect to the transcription initiation site) to the CAT gene. The CAT activity induced by cyppdk1CAT was detected in root, stem, and leaf protoplasts (Figure 10). The 5′ flanking and untranslated region of *cyppdkZm2* (1.1 kb) was placed upstream of the CAT gene at +91 (relative to a putative cap site) to create cyppdk2CAT. Low levels of CAT activity were observed in all protoplasts transfected with cyppdk2CAT (Figure 10). The data presented were the expression levels after subtracting the background. A construct containing no promoter gave only background expression (data not shown), as described previously (Sheen, 1990).

The tissue-specific expression patterns of the three ppdkCAT chimeric genes in protoplasts are similar to those of the endogenous PPDK genes in plants (Figures 8C, 8D, and 8E). Although regulation by other regions has not been ruled out yet, the results suggest that the 5′ regions of the three PPDK genes direct their expression in various tissues. CAT activities stimulated by cyppdk1CAT and cyppdk2CAT in mesophyll protoplasts were roughly proportional to the amount of *cyppdkZm1* and *cyppdkZm2* transcripts detected in leaves of plants. However, the promoter strength of C4ppdkCAT was lower than that of cyppdk1CAT in mesophyll protoplasts, even though the C4PPDK transcript is found at a much higher level than the cyPPDK1 transcript in leaves (Figure 8C). There could be at least three explanations. First, sequences other than −1300 to +72 may be required for the enhanced expression of *C4ppdkZm1* in leaves. Second, the endogenous *cyppdkZm1* promoter may be suppressed by the upstream *C4ppdkZm1* promoter. Third, the stability of the *C4ppdkZm1* transcript may be much higher than that of the *cyppdkZm1* transcripts. The possibility of transcriptional interference has been eliminated by analysis of a construct containing both promoters, as shown below. The lower expression of cyppdk1CAT in root than in leaf protoplasts may be attributed to the lower transfection and expression efficiency or the requirement of other regulatory elements.

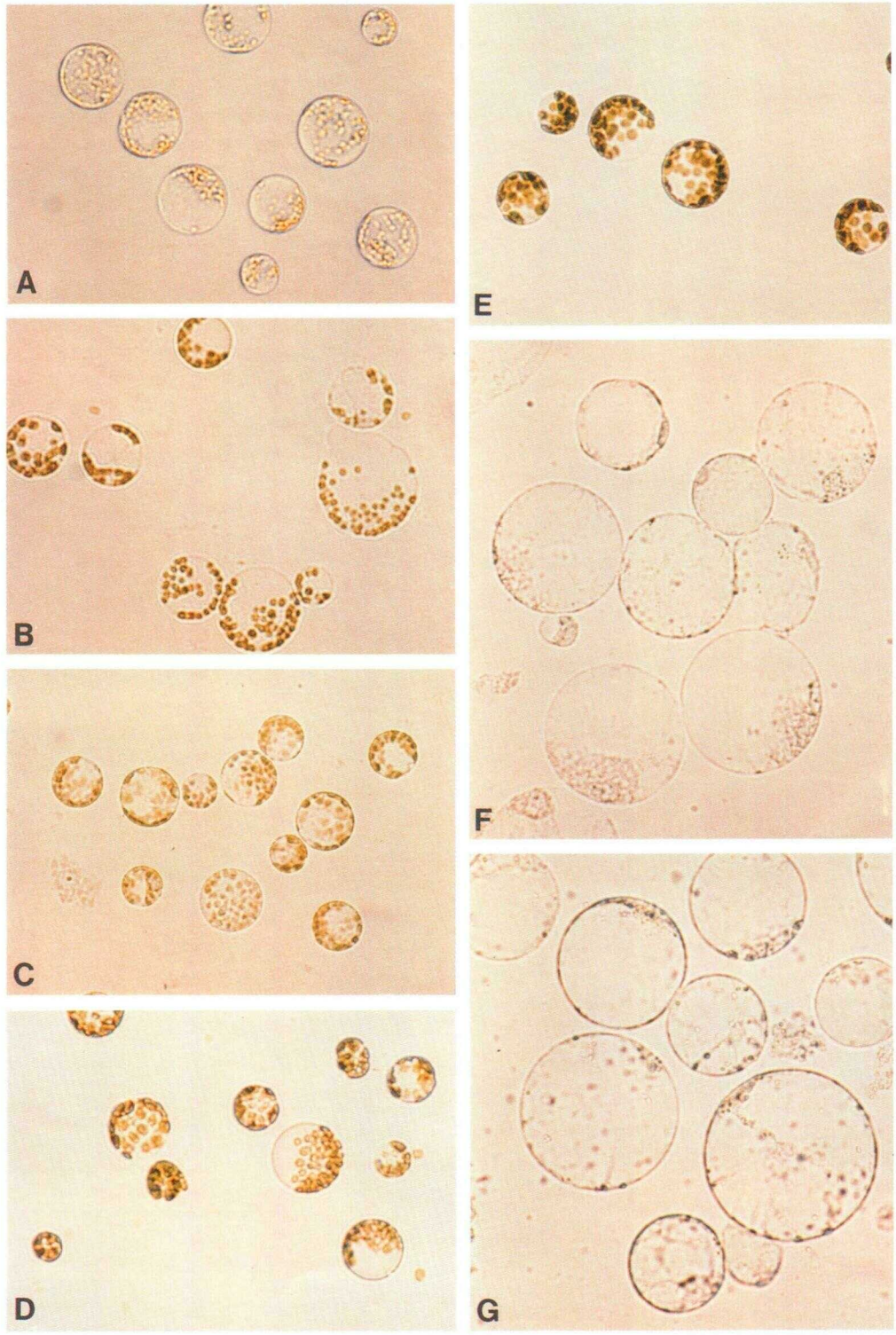


Figure 9. Maize Protoplasts Isolated from Roots, Stems, and Leaves.

(A) Protoplasts from etiolated leaves.
(B), **(C)**, and **(D)** Protoplasts from 12-hr, 24-hr, and 36-hr greening leaves.
(E) Protoplasts from green leaves.
(F) Protoplasts from roots.
(G) Protoplasts from stems.
Magnification ×280.

Table 1. Transient Expression of 35SCAT in Maize Root, Stem, and Leaf Protoplasts

35SCAT Samples	Relative CAT Activity		
	Root	*Stem*	*Leaf*
1	88	81	110
2	83	84	111
3	79	82	107

CAT assays were performed with cell extracts from 5×10^3 leaf protoplasts or 2.5×10^4 stem and root protoplasts for 90 min. The experiment was repeated twice with similar results.

As in transgenic dicot plants (Lam et al., 1989; Benfey and Chua, 1990), the 35S promoter is very active in maize root, stem, and leaf protoplasts. However, the expression of 35SGUS and 35SCAT is higher in leaf protoplasts than in root and stem protoplasts (Figure 10). These results are consistent with the observation that 35S promoter activity is higher in leaves than in roots of transgenic rice plants (Battraw and Hall, 1990). This, in turn, indicates that the 35S promoter activity is regulated differently in tissues of monocots and dicots.

Light Regulation of ppdkCAT Chimeric Genes in Maize Leaf Protoplasts

To determine whether the 5′ regions of the PPDK genes are important for light regulation, the three ppdkCAT chimeric genes were electroporated into mesophyll protoplasts isolated from etiolated, greening, and green maize leaves. Well-differentiated etiolated leaves (grown in the dark) were used to examine the relationship between gene expression and leaf development in the absence of light. Greening leaves are defined to be etiolated leaves illuminated with continuous white light for 12 hr, 24 hr, or 36 hr to induce chloroplast development. Protoplasts were cultured with or without light after electroporation to test whether continuous illumination was required for transcriptional activity.

As shown in Figure 11, the C4ppdkCAT construct directed a very low level of CAT activity in etiolated mesophyll protoplasts cultured in the dark after electroporation, whereas the activity was sixfold higher in etiolated mesophyll protoplasts maintained under continuous illumination. In 12-hr and 24-hr greening mesophyll protoplasts cultured in the dark, the expression of the C4ppdkCAT chimeric gene was much higher than that in etiolated mesophyll protoplasts maintained in the dark. Culturing 12-hr and 24-hr greening mesophyll protoplasts under continuous illumination further boosted CAT activity about eightfold (Figure 11). In 36-hr greening and green mesophyll protoplasts cultured without light, the CAT activity directed by the C4ppdkCAT fusion gene was eightfold higher than that found in etiolated mesophyll protoplasts cultured in the dark. However, no further induction was observed when these protoplasts were cultured under continuous illumination (Figure 11). These data suggest that the dramatic increase of endogenous *C4ppdkZm1* transcript from etiolated to greening maize leaves (Figure 8C) (Hague et al., 1983; Hudspeth et al., 1986; Sheen and Bogorad, 1987) is mediated by the 5′ region of *C4ppdkZm1* and requires continuous illumination. Leaf development is uncoupled from, but chloroplast development is associated with, light-mediated *C4ppdkZm1* expression.

Light-mediated expression patterns were not observed with 35SCAT as a parallel control or 35SGUS as an internal control (Figure 11). CAT activity stimulated by 35SCAT fluctuated less than twofold between dark and light samples and between etiolated and greening samples (Figure 11). The pattern of GUS activity resembled that of CAT activity (Figure 11). As also shown in Figure 11, the cyppdk1CAT construct directed a high level of CAT activity

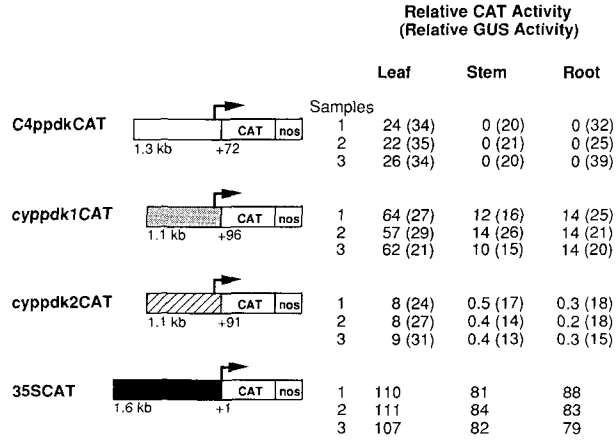


Figure 10. Tissue-Specific Expression of ppdkCAT Chimeric Genes in Maize Protoplasts.

Plasmids containing various promoter/CAT constructs (about 40 μg to 50 μg, in equal molar ratio) were electroporated into 1×10^5 protoplasts isolated from 24-hr greening leaves, root, and stems. A plasmid carrying the 35SGUS (10 μg) was coelectroporated as an internal control. Leaf protoplasts were cultured under light, whereas root and stem protoplasts were cultured in the dark. CAT assays were performed with cell extracts prepared from 5×10^4 protoplasts for 90 min as described (Seed and Sheen, 1988) for most samples except 35SCAT. The expression of 35SCAT and 35SGUS was assayed with cell extracts prepared from 5×10^3 leaf protoplasts and 2.5×10^4 stem and root protoplasts for 90 min. Three replicas (samples 1 to 3) were performed for each construct using the same batch of protoplasts to show the consistency among repeated samples. The experiment was repeated twice without internal controls with similar results.

	Relative CAT Activity (Relative GUS Avtivity)									
	E		G12		G24		G36		G	
	D	L	D	L	D	L	D	L	D	L
C4ppdkCAT	0.5 (26)	3 (16)	3 (34)	26 (46)	4 (27)	30 (41)	5 (45)	7 (53)	4 (19)	7 (25)
cyppdk1CAT	20 (14)	16 (22)	81 (32)	66 (43)	82 (20)	82 (52)	84 (25)	43 (23)	98 (16)	30 (16)
cyppdk2CAT	2 (14)	1 (20)	9 (39)	12 (42)	8 (30)	12 (53)	6 (35)	7 (24)	2 (19)	2 (18)
35SCAT	38	55	55	95	65	116	56	90	43	59

Figure 11. Light Regulation of ppdkCAT Chimeric Genes in Maize Mesophyll Protoplasts.

Plasmids of various promoter/CAT chimeric genes (about 40 μg to 50 μg in equal molar ratio) were electroporated into 2×10^5 mesophyll protoplasts isolated from etiolated (E), 12-hr (G12), 24-hr (G24), and 36-hr (G36) greening, and green leaves (G). A plasmid containing the 35SGUS (20 μg) was coelectroporated as an internal control. Each electroporated sample was divided into two parts (10^5 protoplasts each) and cultured either under light (L) or in the dark (D). CAT and GUS assays were performed as described in the legend to Figure 10. The experiment was repeated twice without internal controls with similar results.

that was independent of continuous illumination except in 36-hr greening and green mesophyll protoplasts, where light seemed to reduce CAT activity. The CAT activity was lower in etiolated mesophyll protoplasts than in greening and green mesophyll protoplasts. In maize, a lower amount of the *cyppdkZm1* transcript was also found in etiolated leaves than in greening and green leaves (Figure 8C). In all samples, the expression of 35SGUS internal control showed less than twofold variation.

The CAT activity stimulated by the cyppdk2CAT construct was low in etiolated and green mesophyll protoplasts, but was higher in greening protoplasts (Figure 11). The expression pattern of cyppdk2CAT was also similar to that of *cyppdkZm2* in plants (Figure 8D). The expression of cyppdk2CAT was not affected by continuous illumination of protoplasts.

Two Expression Programs of *C4ppdkZm1* Are Controlled by Separate Upstream Regulatory Elements

To determine the *cis*-acting elements important for the regulation of the light-inducible expression of *C4ppdkZm1* in leaves, seven chimeric genes were constructed with deleted, mutated, and hybrid promoters. As shown in Figure 12, C4del1CAT showed similar CAT activity to C4ppdkCAT in all mesophyll protoplasts. This result indicates that the sequences important for *C4ppdkZm1* expression are located between −347 and +72. Analysis of other deletions (between −1300 and −347) supports this conclusion (J. Sheen, data not shown). Deletion of the sequences between −347 to −109 (C4del2CAT) abolished the CAT activity stimulated by continuous illumination in etiolated and early greening (12-hr and 24-hr) mesophyll protoplasts. Further deletion of the sequences between −108 to −52 (C4del3CAT) eliminated the CAT activity in greening and green mesophyll protoplasts (Figure 12). The analysis of deletion mutants revealed separate upstream regulatory elements that control the two expression programs of *C4ppdkZm1* during chloroplast development mediated by light.

In the C4mutCAT construct, the mutation of the TATA sequence to a GCCC sequence abolished the activity of *C4ppdkZm1* promoter (Figure 12). This result demonstrates an intimate relationship between upstream and TATA elements.

The function of the upstream regulatory elements of *C4ppdkZm1* was also shown by fusion to a heterologous promoter containing a TATA box. As the activity of the C435SCAT construct shows, the sequences between −347 to −44 of the *C4ppdkZm1* promoter confer light regulation on the 35S promoter sequences (−45 to +1) in mesophyll protoplasts (Figure 12).

A Specific TATA Promoter Element and Distinct Upstream Regulatory Elements Are Important for the Expression of *cyppdkZm1* in Different Tissues

The regulatory elements important for the expression of *cyppdkZm1* and *cyppdkZm2* were also analyzed with de-

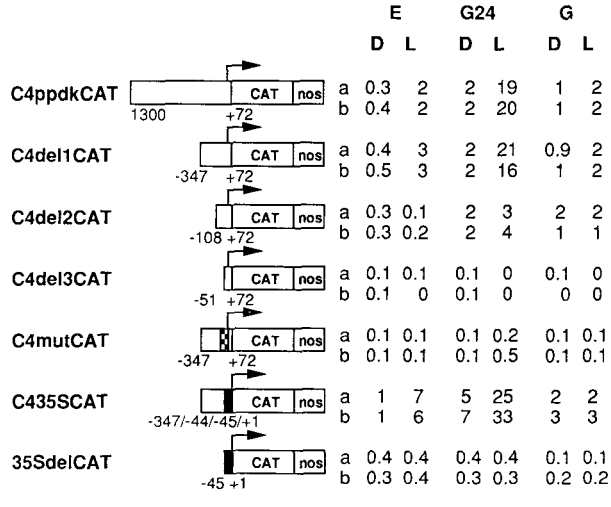
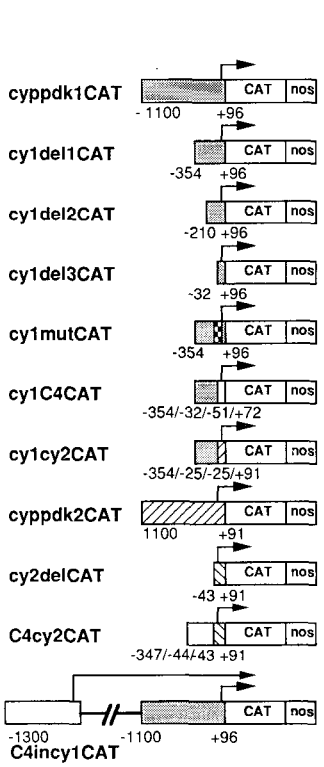


Figure 12. Analysis of the Deleted, Mutated, and Hybrid Promoters of *C4ppdkZm1*.

Electroporation and CAT assay were performed as described in the legends to Figures 10 and 11. Data from two independent experiments (a and b) are shown.

leted, mutated, and hybrid promoters. As shown in Figure 13, cy1del1CAT stimulated the same amount of CAT activity as cyppdk1CAT in protoplasts. This shows that the sequences important for expression of cyppdk1CAT in protoplasts of various tissues are located between −354 and +96. Analysis of other deletions between −1100 and −354 supports this conclusion (J. Sheen, data not shown). The deletion of the sequences between −354 to −211 (cy1del2CAT) reduced CAT activity about sixfold in root and stem protoplasts but not in leaf protoplasts (Figure 13). Further deletion of the sequences between −210 to −33 (cy1del3CAT) decreased CAT activity in all protoplasts examined (Figure 13). However, the truncated *cyppdkZm1* promoter bearing sequences from −32 to +96 (cy1del3CAT) still directed a substantial amount of CAT activity in leaf protoplasts and low levels of CAT activity in root and stem protoplasts (Figure 13). Analysis of promoter mutants indicated that the activity of the truncated *cyppdkZm1* promoter depends on the sequence between −32 to −15 including the TATAAA sequence (J. Sheen, unpublished data). The mutation of the TATAA sequence to a TCTAG sequence in the cy1mutCAT construct eliminated CAT activity (Figure 13).



		G24 D	G24 L	S D	R D
cyppdk1CAT	a	108	72	12	16
	b	99	70	16	25
cy1del1CAT	a	107	88	12	14
	b	83	62	12	17
cy1del2CAT	a	84	77	2	2
	b	81	61	2	3
cy1del3CAT	a	18	20	0.3	0.8
	b	16	22	0.4	0.5
cy1mutCAT	a	0.5	0.4	0.1	0.1
	b	0.4	0.4	0.1	0.1
cy1C4CAT	a	105	75	3	3
	b	90	30	2	2
cy1cy2CAT	a	104	98	8	10
	b	103	94	8	13
cyppdk2CAT	a	8	10	0.4	0.2
	b	3	7	0.3	0.4
cy2delCAT	a	5	6	0.2	0.2
	b	5	5	0.4	0.2
C4cy2CAT	a	4	24	3	0.2
	b	4	36	5	0.7
C4incy1CAT	a	111	107	14	17
	b	109	99	15	25

Figure 13. Analysis of the Deleted, Mutated, and Hybrid Promoters of *cyppdkZm1* and *cyppdkZm2*.

Electroporation and CAT assay were performed as described in the legends to Figures 10 and 11. Data from two independent experiments (a and b) are shown.

These experiments show that both upstream regulatory and the TATA promoter elements are required for the expression of cyppdk1CAT in protoplasts isolated from various tissues. However, the upstream regulatory element from −354 to −211 is more important for the expression in root and stem than the expression in leaf. The TATA promoter element is more important for leaf expression.

The function of the upstream regulatory elements of *cyppdkZm1* was also examined by fusions with heterologous promoters. The cy1C4CAT construct was created by fusing the 5′ region of *cyppdkZm1* (−354 to −32) to the *C4ppdkZm1* promoter at −51. The cy1cy2CAT construct was generated by fusing the 5′ region of *cyppdkZm1* (−354 to −25) to the *cyppdkZm2* promoter at −25. The analysis of hybrid promoters showed that the upstream regulatory elements of *cyppdkZm1* were essential for a high level of expression in all tissues examined. However, the CAT activity directed by cy1C4CAT was lower than that stimulated by cyppdk1CAT and cy1cy2CAT in root and stem protoplasts (Figure 13) and in etiolated and green mesophyll protoplasts (data not shown). These results imply that the TATA promoter elements of *cyppdkZm1* and *cyppdkZm2* interact better than that of *C4ppdkZm1* with the upstream regulatory elements of *cyppdkZm1*.

Deletion of sequences more than 8 bp upstream of the TATA box did not change the activity of the *cyppdkZm2* promoter (Figure 13). Analysis of other deletions between −1100 to −43 supports the same conclusion (J. Sheen, data not shown). The TATA promoter element of *cyppdkZm2* showed lower activity than that of *cyppdkZm1* in leaf protoplasts (Figure 13). This result indicates that the low expression level of *cyppdkZm2* in all tissues examined is due to an absence of positive upstream elements and to the low activity of the TATA box promoter element. The presence of the *C4ppdkZm1* upstream regulatory elements in C4cy2CAT did not reduce the activity of the *cyppdkZm2* promoter in root and stem protoplasts and conferred light regulation on the *cyppdkZm2* promoter in leaf protoplasts (Figure 13).

In maize, the endogenous *C4ppdkZm1* and *cyppdkZm1* overlap. To test whether transcriptional read-through from the upstream *C4ppdkZm1* promoter would turn off the downstream *cyppdkZm1* promoter, a long DNA fragment containing both the *C4ppdkZm1* and *cyppdkZm1* promoters was fused to the CAT and *nos* sequences to create C4incy1CAT. In mesophyll protoplasts, transcription should start from both promoters. Because the first intron cannot be spliced from the transcript initiated from the *C4ppdkZm1* promoter (the splice acceptor site is absent from the construct), CAT activity should only be derived from the transcript initiated from the *cyppdkZm1* promoter. The CAT activity stimulated by C4incy1CAT was similar

to that stimulated by cyppdk1CAT in all protoplasts (Figure 13). There was no obvious inhibition of the downstream *cyppdkZm1* promoter activity even in 24-hr greening leaf protoplasts when the upstream *C4ppdkZm1* promoter was most active (Figures 10 and 11).

DISCUSSION

The C4 Chloroplast PPDK Gene Is Derived from a Cytosolic PPDK Gene

Glackin and Grula (1990) have reported that organ-specific transcripts of different size and abundance derive from the same PPDK locus in maize. However, the gene organization, precise differences between the two transcripts, and the mechanisms for their differential regulation were unclear. In this work, I demonstrate that the two transcripts are derived from the same genetic locus but are the products of two overlapping genes, which are divergent at their 5′ ends because of differential transcriptional initiation and splicing. The expression of a second locus encoding a similar cytosolic PPDK gene is also shown. Each of the PPDK genes has a distinct expression pattern regulated by distinct promoters.

The designation of two overlapping genes, *C4ppdkZm1* and *cyppdkZm1*, instead of one gene with alternative promoters and splicing sites is preferred for the following reasons. First, alternative promoters are often employed when the same protein product from one gene is expressed in different tissues or at different developmental stages, e.g., the *Adh* gene in *Drosophila* (Corbin and Maniatis, 1989). The *C4ppdkZm1* and *cyppdkZm1* encode different protein products, C4 chloroplast and non-C4 cytosolic PPDK, with distinct compartmentation and function. Second, alternative promoters of the same gene are usually separated by less than a few hundred base pairs, whereas the *C4ppdkZm1* and *cyppdkZm1* promoters are separated by more than 5 kb. Third, the *cyppdkZm1* is an intact, independent, and active gene that is similar to *cyppdkZm2* in coding and 5′ regions up to the TATA box.

The discovery of the overlapping PPDK genes demonstrates an unusual and economical way of using the same coding sequence for divergent functions in plants. Several lines of evidence indicate that the C4PPDK gene was created from the ancestral cyPPDK gene but not vice versa. First, PPDK exists in microorganisms (both photosynthetic and nonphotosynthetic) and non-C4 plants (Milner et al., 1975; Aoyagi and Bassham, 1983, 1984; Matsuoka and Yamamoto, 1989). This indicates that PPDK is encoded by an "old" gene and has a general function in carbohydrate metabolism. Second, PPDK had not been implicated directly in photosynthesis until the recent evolution of C4 plants. It is widely believed that C4 photosynthesis occurred recently in evolution because both C4 and C3 species are found in the same family or even the same genus of higher plants (Hatch, 1976; Moore, 1982; Edwards and Ku, 1987). For instance, the upper bound for the maize (C4 plant) and wheat (C3 plant) divergence is less than 70 million years ago (Meagher et al., 1989; Wolfe et al., 1989). Third, the overlapping cyPPDK gene is an active, independent gene and is similar in structure and coding sequence to another cyPPDK gene. Fourth, the overlapping cyPPDK gene is expressed in all tissues examined, which implies that it is likely a housekeeping gene and should have existed before the specialized C4 chloroplast PPDK gene. Finally, the presence of repetitive sequences between distinct domains of the two overlapping PPDK genes suggests a convenient point for recombination without destroying the ancestral cyPPDK gene (Flavell, 1980).

However, the generation of the chloroplast PPDK gene might or might not correlate with the evolution of C4 photosynthesis in maize because the chloroplast PPDK gene is also found in C3 plants, although expressed at much lower levels for both mRNA and protein (Aoyagi and Bassham, 1983, 1984; Matsuoka and Yamamoto, 1989). Until the structure of a C3 chloroplast PPDK gene is determined, it remains unanswered whether the creation of the maize C4 chloroplast PPDK gene is unique to C4 plants.

Nuclear-Encoded Chloroplast Proteins Have Two Origins

In plants, enzymes involved in carbohydrate metabolism have at least two isoforms, the cytosolic enzymes and the chloroplast enzymes. The study of the genes encoding chloroplast and cytosolic glyceraldehyde-3-phosphate dehydrogenase in tobacco shows that the chloroplast and cytosolic glyceraldehyde-3-phosphate dehydrogenase genes evolved from different prokaryotic ancestral genes (Shih et al., 1986). This result supports the endosymbiotic theory of chloroplast origin and, hence, the endosymbiotic origin of nuclear-encoded chloroplast proteins (Shih et al., 1986).

In this study, I show that the gene encoding the C4 chloroplast PPDK is highly related to one cytosolic PPDK gene (*cyppdkZm2*) and overlaps another cytosolic PPDK gene (*cyppdkZm1*). The presence of PPDK in nonphotosynthetic prokaryotes and plants, which bypasses the need of mitochondria for the generation of phosphoenol pyruvate, indicates that the ancestral PPDK gene might exist before the divergence of prokaryotes and eukaryotes (Lehninger, 1975; Milner et al., 1975; Shih et al., 1986). Therefore, it is likely that an alternative evolutionary pathway existed for the generation of nuclear-encoded chloroplast isozymes from nonendosymbiotic origin. The existence of the genes encoding homologous chloroplast and cytosolic isozymes of glutamine synthetase (Tingey et al.,

1988) and malic enzymes (Rothermel and Nelson, 1989), which are also found in nonphotosynthetic organisms, further supports the nonendosymbiotic origin of some nuclear-encoded chloroplast proteins in higher plants.

It is intriguing that the structure of some plant nuclear genes, whose first exons encode mostly chloroplast transit peptides, is similar to that of *C4ppdkZm1*. However, these genes only encode chloroplast proteins such as ribulose bisphosphate carboxylase small subunit and chloroplast *a*/*b*-binding proteins, have no cytosolic counterparts, are only found in photosynthetic organisms, and predate the evolution of C4 photosynthesis (Weeden, 1981; Moore, 1982; Karlin-Neumann and Tobin, 1986; Meagher et al., 1989; Glackin and Grula, 1990). Unlike *C4ppdkZm1*, these genes likely evolved from endosymbiotic origin (Weeden, 1981). Perhaps the chloroplast transit peptide exon and associated upstream elements of *C4ppdkZm1* are derived from one of these genes.

Transient Expression in Protoplasts Isolated from Fresh Tissues

Despite intensive efforts, little is known about the differential gene regulation of monocots, the major crop plants. The principal obstacles have been the difficulties of producing transgenic monocot plants and of isolating high-quality protoplasts from differentiated monocot plant tissues (Dekeyser et al., 1990; Fromm et al., 1990; Gordon-Kamm et al., 1990; Nelson, 1990). So far, transient expression has only been applied to the protoplasts isolated from undifferentiated suspension cultures of monocots (Callis et al., 1987; Walker et al., 1987; Marcotte et al., 1988; Gallie and Walbot, 1990). The lack of activity of most monocot gene promoters in dicots has also limited the use of transgenic dicot plants for the study of monocot gene regulation (Keith and Chua, 1986; J. Sheen, unpublished data).

It has only recently become possible to study differential gene regulation in monocot plants by tissue transient expression methods (Sanford, 1988; Dekeyser et al., 1990). A high-velocity microprojectile gun can be used to stably or transiently transform various tissues of monocot plants (Sanford, 1988). The introduced genes show tissue-specific and light regulation (Bruce et al., 1989; Klein et al., 1989; Goff et al., 1990). However, the method requires expensive equipment. The other newly developed tissue transient expression method introduces DNA into fresh tissue by electroporation (Dekeyser et al., 1990). This method is less expensive and can be applied to many plant species. The introduced genes also show tissue-specific expression (Dekeyser et al., 1990). However, the method has only been extensively demonstrated for leaf base tissues and is subject to problems with nucleases in root explants (Dekeyser et al., 1990). In addition, tissue transient expression is less sensitive and less amenable to quantitative analysis than protoplast transient expression and does not lend itself well to the standardized analysis of multiple constructs introduced into identical samples.

I have developed a convenient and reproducible protoplast transient expression system that allows the study of differential gene regulation in root, stem, and leaf tissues. Using a rapid procedure for the isolation of protoplasts, a simple medium for the electroporation and incubation of protoplasts, and a sensitive CAT assay (Seed and Sheen, 1988), less than 100 mesophyll protoplasts are needed to detect an expression of 35SCAT that is 10-fold higher than background. Using the same method, transient expression can also be carried out with mesophyll protoplasts isolated from barley, wheat, rice, tobacco, pea, and *Arabidopsis* (J. Sheen, unpublished data). A complex medium (Paszkowski and Saul, 1986; Hauptmann et al., 1987; Junker et al., 1987; Prols et al., 1988) is apparently unnecessary for transient expression in protoplasts isolated from fresh tissues of both monocot and dicot plants. Moreover, the use of glucose or sucrose as osmotica or carbon sources in mesophyll protoplast cultures causes a feedback transcriptional repression of photosynthetic genes (Sheen, 1990). The finding of sugar repression in protoplasts is apparently physiologically significant because sugar accumulation induced by the overexpression of a yeast invertase in transgenic tobacco plants also results in the inhibition of photosynthesis and the repression of photosynthetic gene expression (Von Schaewen et al., 1990).

Light-Mediated *C4ppdkZm1* Expression Is Uncoupled from Leaf Development but Partially Associated with Chloroplast Development

In most dicot plants, primary leaf development is tightly linked to chloroplast development and photosynthetic gene expression and requires light. Chory et al. (1989) isolated *det* mutants in *Arabidopsis* that show leaf and chloroplast development and photosynthetic gene expression in the absence of light. Based on the study of these mutants, it was proposed that the primary role of light in gene expression is mediated by the activation of leaf development. In monocot plants, however, leaf development proceeds in the dark and is uncoupled from chloroplast development. Using well-differentiated etiolated and greening maize leaves of the same age and a protoplast transient expression method, I show here that the light-inducible expression of *C4ppdkZm1* in maize is apparently uncoupled from leaf development and is only partially associated with chloroplast development. The light-inducible expression of maize *rbcS* and *cab* also show similar properties (H. Huang, A. Schäffner, and J. Sheen, unpublished data).

Analysis of deletion mutants shows that *C4ppdkZm1* is regulated by two expression programs mediated by separate upstream regulatory elements. The distal upstream

regulatory element (−347 to −109) is essential for the expression of *C4ppdkZm1* in etiolated and early (12-hr and 24-hr) greening leaves. Continuous illumination but not the presence of chloroplasts is essential for this expression program. The requirement for continuous illumination for a high level of *C4ppdkZm1* promoter activity in etiolated leaves is unprecedented. In many other plants, the light-inducible expression of photosynthetic genes in etiolated leaves is regulated by phytochrome and requires only red light pulses (Tobin and Silverthorne, 1985; Lissemore and Quail, 1988; Mosinger et al., 1988). However, it has also been shown that phytochrome-mediated light regulation is limited to etiolated immature primary leaves. In mature leaves of dark-adapted pea and transgenic petunia plants, continuous illumination with white or blue light, but not a red light pulse, is effective for *rbcS* expression (Fluhr and Chua, 1986). Preliminary results indicate that a red light pulse does not increase the accumulation of *C4ppdkZm1* transcripts in etiolated mature maize leaves, and blue light instead of red light is essential for the light induction of C4ppdkCAT expression in mesophyll protoplasts (J. Sheen, unpublished data). Phytochrome-mediated light regulation is likely bypassed in differentiated maize leaves grown in the dark.

The proximal upstream regulatory element (−108 to −52) is important for *C4ppdkZm1* expression in greening and green leaves. This facet of the expression program of *C4ppdkZm1* does not require continuous illumination and seems to correlate with the development of chloroplasts. Although the signal transduction pathway for nucleus-plastid interaction is mostly unknown, the intimate relationship between chloroplasts and the expression of nuclear-encoded photosynthetic genes is well documented (Simpson et al., 1986b; Taylor, 1989).

Preliminary data indicate that the inverted and direct repeats shown in Figure 7A share no consensus with known *cis*-acting regulatory elements of photosynthetic genes (Dean et al., 1989; Gilmartin et al., 1990; Glackin and Grula, 1990; Schindler and Cashmore, 1990), but interact synergistically for *C4ppdkZm1* expression (J. Sheen, unpublished data). In addition, the light-inducible activity of the upstream regulatory elements of *C4ppdkZm1* absolutely requires the TATA promoter element (C4mutCAT in Figure 12). However, the interaction is rather nonspecific because the *C4ppdkZm1* TATA promoter elements can be functionally replaced by similar elements from the 35S promoter and *cyppdkZm2* promoter.

Combinatorial Control of *cyppdkZm1* Expression in Different Tissues

Unlike the activity of the *C4ppdkZm1* promoter, the activities of the *cyppdkZm1* and *cyppdkZm2* promoters are detectable in all tissues examined. The result of deletion and mutational analysis indicates that distinct upstream regulatory and TATA elements (−354 to −211, −210 to −33, and −32 to −15) are important for the activity of the *cyppdkZm1* promoter in different tissues (Figure 13). For *cyppdkZm1* expression in roots and stems, the upstream elements (−354 to −211 and −210 to −33) are most important. For *cyppdkZm1* expression in leaves, the proximal upstream and TATA promoter elements (−210 to −33 and −32 to −15) are essential. In transgenic tobacco plants, different subdomains of the 35S promoter also confer tissue-specific expression (Benfey and Chua, 1990). The absence of upstream regulatory elements and a low TATA element activity apparently account for the lower activity of the *cyppdkZm2* promoter in all tissues examined.

The analysis of hybrid promoters shows that, except in greening protoplasts, the upstream regulatory elements of *cyppdkZm1* enhance expression better when fused to their own or the *cyppdkZm2* TATA promoter elements than when fused to the *C4ppdkZm1* TATA promoter element. These results indicate that specific interactions between the upstream and TATA elements are required in most tissues for maximal activity of the *cyppdkZm1* promoter, which may reflect the difference in the availability and affinity of the corresponding *trans*-acting factors in different tissues (Fischer and Maniatis, 1988; Lam et al., 1989; Benfey and Chua, 1990). Although all TATA promoter elements contain the TATAA sequence, the sequences surrounding the TATA boxes are divergent and may contribute to the differences of their activities in different tissues. The analysis of hybrid promoters (Figure 13) also suggests that the restricted expression of *C4ppdkZm1* in leaves can be attributed to the absence of root-specific and stem-specific positive regulatory elements, rather than the presence of an upstream negative regulatory element such as that found in the pea *cab* promoter (Simpson et al., 1986a).

METHODS

Plant Material and Growth Conditions

To obtain etiolated and greening maize seedlings with leaf and stem morphology more similar to that of green seedlings grown under physiological conditions, maize seeds were soaked in water overnight and allowed to germinate in moist vermiculite and pealite mix (2:1) for 3 days under room light and temperature (20 $\mu E/m^2$ and 23°C). The plants were then grown in a dark growth chamber for about 8 days to 9 days at 25°C. The initial light treatment inhibited extensive stem elongation in the dark, but did not change the photosynthetic gene expression patterns of second leaves in etiolated and greening seedlings (Sheen and Bogorad, 1987; Figure 8; and data not shown). Only the middle section (4 cm to 10 cm from the tip) of second leaves was used for RNA and protoplast isolation in all experiments. To obtain greening plants,

etiolated seedlings were illuminated under room light and temperature for various numbers of hours. Green plants were grown either in a growth chamber at 25°C under a 10 hr/14 hr light/dark cycle for about 12 days, or under constant room light and temperature for about 13 days. All plants used for RNA and mesophyll protoplast isolation were about the same age. Maize seedlings used to isolate root and stem protoplasts were grown under sterile conditions in the dark for 4 days at 27°C. A maize hybrid line, $FR9^{cms} \times FR37$ (Illinois Foundation Seed, Champaign, IL), was used for all experiments.

Library Construction and Screening

The methods of maize RNA and DNA isolation were the same as described (Sheen and Bogorad, 1987). The construction and screening of the maize leaf cDNA library were performed as described (Aruffo and Seed, 1987; Sheen and Bogorad, 1987). To obtain the full-length PPDK cDNA, fractionated cDNA (3 kb to 4 kb) was used to construct the library and screened with a partial cDNA clone identified before (Sheen and Bogorad, 1987). The genomic library was constructed with the EMBL4 vector, and maize DNA was partially digested with Sau3A and fractionated through a 5% to 20% sucrose gradient. After ligation, the packaging step was performed as described by the manufacturer (Stratagene). About 10^7 plaques were screened to obtain 15 PPDK genomic clones.

DNA and RNA Blot Analysis

Genomic DNA blot analysis was performed as described (Bernatzky and Tanksley, 1986) except that Genescreen Plus membrane (Du Pont-New England Nuclear) was used, and "dry blot" took only 3 hr. The RNA blot analysis was performed as described (Sheen and Bogorad, 1987) with polyA+ mRNA isolated from 24-hr greening maize leaves. The nick-translated DNA probes were made with DNA fragments isolated from a 1.5% Nusieve gel (FMC Bioproducts, Rockland, ME) and a kit from Boehringer Mannheim.

DNA Sequencing

DNA sequencing was performed with alkaline denatured double-stranded plasmid DNA template and Sequenase as described (Tabor and Richardson, 1987). The sequences of the primers used to analyze the structure of *C4ppdkZm1* were derived from the published PPDK cDNA sequence (Matsuoka et al., 1988). The sequence of a 15-nucleotide primer was chosen from every 200 bp to 250 bp of the PPDK cDNA sequence in both orientations.

PCR and Primer Extension for RNA Analysis

Total RNA, polyA+ mRNA, and the first-strand cDNA were made as described (Sheen and Bogorad, 1987) from various tissues and cell types. The polyA+ mRNA isolated from about 200 μg of total RNA was used for cDNA synthesis in 10 μL at 42°C for 1 hr. The cDNA was heated at 95°C for 5 min and diluted to 100 μL with water. The amount of total RNA was used for standardization but not the amount of polyA+ mRNA because the polyA+ mRNA/cell ratio varies during leaf development but not the total RNA/cell ratio (data not shown). PCR was carried out for 30 cycles in 100 μL of 50 mM Tris-HCl, pH 8.6, 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin, 10% DMSO, 150 μg of primers (20 to 35-mers), 250 μM deoxynucleotide triphosphates, 0.5 units of AmpliTaq (U.S. Biologicals), and 0.5 μL to 20 μL of diluted cDNA depending on the abundance of the specific cDNA (Wang et al., 1989). Each cycle was at 92°C for 1 min, 60°C for 2 min, and 72°C for 3 min. About 5 μL to 35 μL of the PCR products were analyzed in a 2% agarose gel stained with ethidium bromide. Three 5′ primers used to quantitate RNA were 5′-ATGACGGCATCGGTTTCCAGGGCCATCTGCGT (32-mer) for *C4ppdkZm1*, 5′-GTTGTTCAGCCTAGCTAGCTAGCGCTG (27-mer) for *cyppdkZm1*, and 5′-AGTGTTGTTTGCTTTGCTTTGCCT (24-mer) for *cyppdkZm2*. The 3′ primer was 5′-CGCCCATGT-ACTCCTCCACCCACCGCAGGCCGTC (34-mer) for all three genes. The specificity of three PCR products was confirmed by the detection of their characteristic restriction sites and cloning and sequencing (Figures 8B, 8D, and data not shown).

Primer extension was carried out as described (Dunsmuir et al., 1988). About 1 μg and 10 μg of polyA+ mRNA isolated from greening maize leaves were used to map the transcription initiation sites of *C4ppdkZm1* and *cyppdkZm1*, respectively. The transcription initiation sites were determined precisely by comparing the products of primer extension and DNA sequencing performed with the same ^{32}P-labeled primers and loaded on the same sequencing gel. The two primers used for primer extension were 5′-GTGACCGCCCCGCGCCTGCTGCTGTCC (27-mer) for *C4ppdkZm1* and 5′-CAGCGCTAGCTAGCTAGGCTGAACAAC (27-mer) for *cyppdkZm1*. All primers were chosen from the distinct sequences of the 5′ untranslated regions of each gene to ensure specificity. For the quantitation of the *C4ppdkZm1* transcript by primer extension, the same RNA samples for the PCR analysis were used.

Isolation of Protoplasts from Leaf, Root, and Stem

To isolate mesophyll protoplasts, the middle portions of the second leaves (4 cm to 10 cm from the tip) were cut with fresh razor blades to 0.5-mm strips. Up to 60 leaves could be piled and cut together. Cutting leaves to narrow strips (<0.5 mm) without mincing the tissues was critical for high yield and cell integrity. The leaf strips were digested in 1% cellulase RS, 0.1% macerozyme R10 (both from Yakult Honsha, Japan, or Karplan Co., Torrance, CA), 0.6 M mannitol, 10 mM Mes (pH 5.7), 1 mM $CaCl_2$, 1 mM $MgCl_2$, 0.1% BSA, and 15 mM β-mercaptoethanol (β-ME) for 2 hr at 23°C with gentle shaking (40 rpm). The addition of β-ME and BSA in the digestion solution reduced cell damage. Mesophyll protoplasts were released by gentle shaking (80 rpm) and filtered through 70-μm nylon filters. Protoplasts were collected by centrifugation at 100*g* for 3 min. Protoplasts were washed once in cold 0.6 M mannitol and counted with a hemocytometer under a microscope. The yield was about 5×10^6/g fresh weight for etiolated and greening protoplasts and about 2×10^6/g fresh weight for green protoplasts. Green mesophyll protoplasts were more fragile than greening and etiolated protoplasts. The use of a less potent cellulase mix (cellulase R10) could increase the successful rate for both the isolation and electroporation of green mesophyll protoplasts. Etiolated protoplasts were isolated in the dark, kept on ice, and electroporated under dim green light. Root

and stem protoplasts were isolated by the same method except 0.7 M mannitol (Junker et al., 1987), 2% cellulase YC, and 0.1% pectolyase Y23 (Karplan Co.) were used. Stems and roots were sliced to about 0.5-mm thick discs with fresh razor blades before digestion. The yield of the root and stem protoplasts was about 5×10^5/g fresh weight. Microphotographs were taken using a Zeiss ICM 405 microscope and Kodak Ektachrome tungsten film 160 with a tungsten light source.

Electroporation and Protoplast Incubation

Electroporation was typically carried out with 1 to 2×10^5 protoplasts in 300 μL of electroporation solution and about 40 μg to 50 μg of plasmid DNA (adjusted to about equal molar ratio by both OD reading and ethidium bromide staining). In the experiments with internal controls, 10μg to 20 μg of a plasmid DNA containing 35SGUS was coelectroporated with CAT plasmids. The electroporation condition was 200 μF, 400 V/cm, 5 msec or 10 msec, and one pulse with a Promega *X*-450 apparatus. The electroporation solution was 0.6 M (leaf) to 0.7 M (root and stem) mannitol and 25 mM KCl (leaf) to 40 mM KCl (root and stem). For root and stem protoplasts, 4 mM Mes (pH 5.7), and 1 mM β-ME were also added. The use of $CaCl_2$ and phosphate-buffered saline was detrimental to maize protoplasts isolated from fresh tissues during electroporation (J. Sheen, unpublished data). After electroporation, protoplasts were cultured (10^5/mL) in solutions containing 0.6 M mannitol, 4 mM Mes (pH 5.7), and 0.1 mM β-ME for leaf protoplasts, and 0.7 M mannitol, 4 mM Mes (pH 5.7), 40 mM KCl, and 1 mM β-ME for root and stem protoplasts. Similar results were obtained with either linearized or circular plasmid DNA (J. Sheen, unpublished data).

CAT and GUS Assay

The phase-extraction CAT assay was used for its rapidity, broad linear range, low price, high sensitivity, and low background (Seed and Sheen, 1988; Sheen, 1990). Cell extracts from 2×10^3 to 5×10^4 protoplasts were used for CAT assays. Each sample contained 0.2 μCi of ^{3}H-chloramphenicol (Du Pont-New England Nuclear) in 100 μL of 100 mM Tris-Cl (pH 8), 5 mM butyryl-CoA, and cell extract. The reaction was incubated at 37°C for 90 min before extraction with xylenes. The relative CAT activity was expressed in counts per minute divided by 1000. The background of CAT assay was about 100 cpm. The fluorogenic GUS assay (Jefferson, 1987) was performed with cell extracts from 5×10^3 to 2.5×10^4 protoplasts in 100 μL of 10 mM Tris-Cl (pH 8), 2 mM $MgCl_2$, and 1 mM 4-methylumbelliferyl glucuronide for 90 min at 37°C. The reaction was stopped by the addition of 900 μL of 0.2 M $NaCO_3$. The fluorescence was measured in a final volume of 2 mL. The relative GUS activity was the direct fluorescence reading divided by 10. All data were collected in the linear range of both CAT and GUS assays. The data generated from the same batch of protoplasts are presented to minimize variation among samples. Each construct was usually assayed at least three times with a different batch of protoplasts to ensure the consistency (data not shown).

Promoter/CAT Fusions

C4ppdkCAT, cyppdk1CAT, and cyppdk2CAT were constructed by fusing the 5′ noncoding regions of the three PPDK genes (1.3 kb, 1.1 kb, and 1.1 kb, respectively) to the CAT and *nos* sequences in the pUC8CaMVCATδN originated from Walbot's laboratory (Fromm et al., 1986). The fusion sites were the SacI site (+72) in the 5′ untranslated region of *C4ppdkZm1* and the blunted NarI sites (+96 and +91) in the 5′ regions of *cyppdkZm1* and *cyppdkZm2*. (All ATG introduced are in-frame with the ATG of CAT.) The hybrid promoters were constructed by fusing the upstream elements of *C4ppdkZm1* and *cyppdkZm1* to the truncated versions of 35SCAT, C4ppdk1CAT, and cyppdk2CAT. The fusion sites were shown in Figures 12 and 13. All fusions were made by blunt-end ligation and verified by DNA sequencing. Bal31 digestion was used to generate promoter deletion mutants. Kunkel's method was used for site-directed mutagenesis to mutate the TATA sequence as described (Kunkel, 1985; Sheen, 1990). The C4incy1CAT construct was generated by the insertion of a 5.3-kb XbaI fragment upstream of cyppdk1CAT at the XbaI site. The orientation of the insertion was verified by restriction site analysis.

ACKNOWLEDGMENTS

I thank Fred Ausubel and Howard Goodman for continuous support and two anonymous reviewers for critical comments. This work was supported by a grant from Hoechst AG.

Received December 12, 1990; accepted January 24, 1991.

REFERENCES

Aoyagi, K., and Bassham, J.A. (1983). Pyruvate orthophosphate dikinase in wheat leaves. Plant Physiol. **73,** 853–854.

Aoyagi, K., and Bassham, J.A. (1984). Pyruvate orthophosphate dikinase mRNA organ specificity in wheat and maize. Plant Physiol. **76,** 278–280.

Aruffo, A., and Seed, B. (1987). Molecular cloning of a CD28 cDNA by a high-efficiency COS cell expression system. Proc. Natl. Acad. Sci. USA **84,** 8573–8577.

Battraw, M.J., and Hall, T.C. (1990). Histochemical analysis of CaMV 35S promoter-β-glucuronidase gene expression in transgenic rice plants. Plant Mol. Biol. **15,** 527–538.

Benfey, P.N., and Chua, N.-H. (1990). The cauliflower mosaic virus 35S promoter: Combinatorial regulation of transcription in plants. Science **250,** 959–966.

Bernatzky, R., and Tanksley, S. (1986). Methods for detection of single or low copy sequences in tomato on Southern blots. Plant Mol. Biol. Rep. **4,** 37–41.

Berry, J.A. (1975). Adaptation of photosynthetic processes to stress. Science **188,** 644–650.

Björkman, O., and Berry, J. (1973). High-efficiency photosynthesis. Sci. Am. **229,** 80–93.

Bruce, W.B., Christensen, A.H., Klein, T., Fromm, M., and Quail, P.H. (1989). Photoregulation of a phytochrome gene promoter from oat transferred into rice by particle bombardment. Proc. Natl. Acad. Sci. USA **86,** 9692–9696.

Callis, J., Fromm, M., and Walbot, V. (1987). Introns increase gene expression in cultured maize cells. Genes Dev. **1,** 1183–1200.

Chory, J., Peto, C., Feinbaum, R., Pratt, L., and Ausubel F. (1989). *Arabidopsis thaliana* mutant that develops as light-grown plant in the absence of light. Cell **58,** 991–999.

Corbin, V., and Maniatis, T. (1989). Role of transcriptional interference in the *Drosophila melanogaster Adh* promoter switch. Nature **337,** 279–282.

Dean, C., Pichersky, E., and Dunsmuir, P. (1989). Structure, evolution and regulation of *RbcS* genes in higher plants. Annu. Rev. Plant Physiol. **40,** 415–439.

Dekeyser, R.A., Claes, B., De Rycke, R.M.U., Habets, M.E., Van Montagu, M.C., and Caplan, A.B. (1990). Transient gene expression in intact and organized rice tissues. Plant Cell **2,** 591–602.

Dunsmuir, P., Bond, D., Lee, K., Gidoni, D., and Townsend, J. (1988). Stability of introduced genes and stability in expression. In Plant Molecular Biology Manual, S.B. Gelvin and R.A. Schilperoort, eds (Dordrecht, The Netherlands: Kluwer Academic Publishers), pp. c1/5–7.

Edwards, G.E., and Walker, D.A. (1983). C3, C4: Mechanisms and Cellular and Environmental Regulation of Photosynthesis. (Oxford: Blackwell Scientific).

Edwards, G.E., and Ku, M.S.B. (1987). Biochemistry of C3-C4 intermediates. In The Biochemistry of Plants, Vol. 10, M.D. Hatch and N.K. Boardman, eds (New York: Academic Press), pp. 275–325.

Fischer, J.A., and Maniatis, T. (1988). *Drosophila Adh*: A promoter element expands the tissue specificity of an enhancer. Cell **53,** 451–461.

Flavell, R. (1980). The molecular characterization and organization of plant chromosome DNA sequences. Annu. Rev. Plant Physiol. **31,** 569–596.

Fluhr, R., and Chua, N.-H. (1986). Developmental regulation of two genes encoding ribulose-bisphosphate carboxylase small subunit in pea and transgenic petunia plants: Phytochrome response and blue-light induction. Proc. Natl. Acad. Sci. USA **83,** 2358–2362.

Foyer, C.H. (1984). Photosynthesis. (New York: Wiley Interscience Publications), pp. 151–173.

Fromm, M.E., Taylor, L.P., and Walbot, V. (1986). Stable transformation of maize after gene transfer by electroporation. Nature **319,** 791–793.

Fromm, M.E., Morrish, F., Armstrong, C., Williams, R., Thomas, J., and Klein, T.M. (1990). Inheritance and expression of chimeric genes in the progeny of transgenic maize plants. Bio/Technology **8,** 833–839.

Gallie, D.R., and Walbot, V. (1990). RNA pseudoknot domain of tobacco mosaic virus can functionally substitute for a poly(A) tail in plant and animal cells. Genes Dev. **4,** 1149–1157.

Gilmartin, P.M., Sarokin, L., Memelink, J., and Chua, N.-H. (1990). Molecular light switches for plant genes. Plant Cell **2,** 369–378.

Glackin, C.A., and Grula, J.W. (1990). Organ-specific transcripts of different size and abundance derive from the same pyruvate orthophosphate dikinase gene in maize. Proc. Natl. Acad. Sci. USA **87,** 3004–3008.

Goff, S.A., Klein, T.M., Roth, B.A., Fromm, M.E., Cone, K.C., Radicella, J.P., and Chandler, V.L. (1990). Transactivation of anthocyanin biosynthetic genes following transfer of B regulatory genes into maize tissues. EMBO J. **9,** 2517–2522.

Gordon-Kamm, W.J., Spencer, T.M., Mangano, M.L., Adams, T.R., Daines, R.J., Start, W.G., O'Brien, J.V., Chambers, S.A., Adams, W.R., Jr., Willetts, N.G., Rice, T.B., Mackey, C.J., Krueger, R.W., Kausch, A.P., and Lemaux, P.G. (1990). Transformation of maize cells and regeneration of fertile transgenic plants. Plant Cell **2,** 603–618.

Hague, D.R., Uhler, M., and Collins, P.D. (1983). Cloning of cDNA for pyruvate, Pi dikinase from maize leaves. Nucl. Acids Res. **11,** 4853–4865.

Hatch, M.D. (1976). Photosynthesis: The path of carbon. In Plant Biochemistry, J. Bonner and J.E. Varner, eds (New York: Academic Press), pp. 797–844.

Hatch, M.D. (1987). C4 photosynthesis: A unique blend of modified biochemistry, anatomy and ultrastructure. Biochim. Biophys. Acta **895,** 81–106.

Hauptmann, R.M., Ozias-Akins, P., Vasil, V., Tabaeizadeh, Z., Rogers, S.G., Horsch, R.B., Vasil, I.K., and Fraley, R.T. (1987). Transient expression of electroporated DNA in monocotyledonous and dicotyledonous species. Plant Cell Rep. **6,** 265–270.

Helentjaris, T., Weber, D., and Wright, S. (1988). Identification of the genomic locations of duplicate nucleotide sequences in maize by analysis of restriction fragment length polymorphisms. Genetics **118,** 353–363.

Hudspeth, R.L., Glackin, C.A., Bonner, J., and Grula, J.W. (1986). Genomic and cDNA clones for maize phosphoenolpyruvate carboxylase and pyruvate orthophosphate dikinase: Expression of different gene-family members in leaves and roots. Proc. Natl. Acad. Sci. USA **83,** 2884–2888.

Jefferson, R. A. (1987). Assaying chimeric genes in plants: The GUS gene fusion system. Plant Mol. Biol. Rep. **5,** 387–405.

Junker, B., Zimny, J., Lührs, R., and Lörz, H. (1987). Transient expression of chimeric genes in dividing and non-dividing cereal protoplasts after PEG-induced DNA uptake. Plant Cell Rep. **6,** 329–332.

Karlin-Neumann, G.A., and Tobin, E.M. (1986). Transit peptides of nuclear-encoded chloroplast proteins share a common amino acid framework. EMBO J. **5,** 9–13.

Keith, B., and Chua, N.-H. (1986). Monocot and dicot premRNAs are processed with different efficiencies in transgenic tobacco. EMBO J. **5,** 2419–2425.

Klein, T.M., Roth, B.A., and Fromm, M.E. (1989). Regulation of anthocyanin biosynthetic genes introduced into intact maize tissues by microprojectiles. Proc. Natl. Acad. Sci. USA **86,** 6681–6685.

Kozak, M. (1987). An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucl. Acids Res. **15,** 8125–8148.

Kunkel, T.A. (1985). Rapid and efficient site-specific mutagenesis without phenotype selection. Proc. Natl. Acad. Sci. USA **82,** 488–492.

Langdale, J.A., Rothermel, B.A., and Nelson, T. (1988). Cellular pattern of photosynthetic gene expression in developing maize leaves. Genes Dev. **2,** 106–115.

Lam, E., Benfey, P.N., Gilmartin, P.M., Fang, R.-X., and Chua, N.H. (1989). Site-specific mutations alter in vitro factor binding and change promoter expression pattern in transgenic plants. Proc. Natl. Acad. Sci. USA **86,** 7890–7894.

Lehninger, A.L. (1975). Biochemistry: The Molecular Basis of Cell Structure and Function. (New York: Worth Publishers, Inc.), pp. 626–627.

Lissemore, J.L., and Quail, P.H. (1988). Rapid transcriptional regulation by phytochrome of the genes for phytochrome and chlorophyll *a/b* binding protein in *Avena sativa*. Mol. Cell. Biol. **8,** 4840–4850.

Lutcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F., and Scheele, G.A. (1987). Selection of AUG initiation codons differs in plants and animals. EMBO J. **6,** 43–48.

Marcotte, W.R., Bayley, C.C., and Quatrano, R.S. (1988). Regulation of a wheat promoter by abscisic acid in rice protoplasts. Nature **335,** 454–457.

Matsuoka, M., and Yamamoto, N. (1989). Induction of mRNAs for phosphoenolpyruvate carboxylase and pyruvate, orthophosphate dikinase in leaves of a C3 plant exposed to light. Plant Cell Physiol. **30,** 479–486.

Matsuoka, M., Ozeki, Y., Yamamoto, N., Hirano, H., Kano-Murakami, Y., and Tanaka, Y. (1988). Primary structure of maize pyruvate orthophosphate dikinase as deduced from cDNA sequence. J. Biol. Chem. **263,** 11080–11083.

McClintock, B. (1984). The significance of responses of the genome to challenge. Science **226,** 792–801.

Meagher, R.B., Berry-Lowe, S., and Rice, K. (1989). Molecular evolution of the small subunit of ribulose bisphosphate carboxylase: Nucleotide substitution and gene conversion. Genetics **123,** 845–863.

Milner, Y., Michaels, G., and Wood, H.G. (1975). Pyruvate, orthophosphate dikinase of *Bacteroides symbiosus* and *Propionibacterium shermanii*. Methods Enzymol. **42,** 199–212.

Moore, P.D. (1982). Evolution of photosynthetic pathways in flowering plants. Nature **295,** 647–648.

Mosinger, E., Batschauer, A., Apel, K., Schafer, E., and Briggs, W. (1988). Phytochrome regulation of greening in barley. Plant Physiol. **86,** 706–710.

Nelson, T. (1990). New horses for monocot gene jockeys. Plant Cell **2,** 589.

Nelson, T., and Langdale, J.A. (1989). Patterns of leaf development in C4 plants. Plant Cell **1,** 3–13.

Paszkowski, J., and Saul, M.W. (1986). Direct gene transfer to plants. Methods Enzymol. **118,** 668–684.

Prols, M., Topfer, R., Schell, J., and Steinbiss, H.H. (1988). Transient gene expression in tobacco protoplasts: I. Time course of CAT appearance. Plant Cell Rep. **7,** 221–224.

Rothermel, B.A., and Nelson, T. (1989). Primary structure of the maize NADP-dependent malic enzyme. J. Biol. Chem. **264,** 19587–19592.

Sanford, J.C. (1988). The biolistic process. Trends Biotechnol. **6,** 299–302.

Schindler, U., and Cashmore, A.R. (1990). Photoregulated gene expression may involve ubiquitous DNA binding proteins. EMBO J. **9,** 3415–3427.

Seed, B., and Sheen, J.-Y. (1988). A simple phase-extraction assay for chloramphenicol acetyltransferase activity. Gene **67,** 271–277.

Sheen, J. (1990). Metabolic repression of transcription in higher plants. Plant Cell **2,** 1027–1038.

Sheen, J.-Y., and Bogorad, L. (1987). Differential expression of the C4 pathway genes in mesophyll and bundle sheath cells of greening maize leaves. J. Biol. Chem. **262,** 11726–11730.

Shih, M.-C., Lazar, G., and Goodman, H.M. (1986). Evidence in favor of the symbiotic origin of chloroplasts: Primary structure and evolution of tobacco glyceraldehyde-3-phosphate dehydrogenase. Cell **47,** 73–80.

Simpson, J., Schell, J., Van Montagu, M., and Herrera-Estrella, L. (1986a). Light-inducible and tissue-specific pea *lhcp* gene expression involves an upstream element combining enhancer- and silencer-like properties. Nature **323,** 551–554.

Simpson, J., Van Montagu, M., and Herrera-Estrella, L. (1986b). Photosynthesis-associated gene families: Differences in response to tissue-specific and environmental factors. Science **233,** 34–38.

Tabor, S., and Richardson, C.C. (1987). DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA **84,** 4767–4771.

Taylor, W.C. (1989). Regulatory interactions between nuclear and plastid genomes. Annu. Rev. Plant Physiol. **40,** 211–233.

Tingey, S.V., Tsai, F.-Y., Edwards, J.W., Walker, E.L., and Coruzzi, G.M. (1988). Chloroplast and cytosolic glutamine synthetase are encoded by homologous nuclear genes which are differentially expressed in vivo. J. Biol. Chem. **263,** 9651–9657.

Tobin, E.M., and Silverthorne, J. (1985). Light regulation of gene expression in higher plants. Annu. Rev. Plant Physiol. **36,** 569–593.

Von Schaewen, A., Stitt, M., Schmidt, R., Sonnewald, U., and Willmitzer, L. (1990). Expression of a yeast-derived invertase in the cell wall of tobacco and *Arabidopsis* plants leads to accumulation of carbohydrate and inhibition of photosynthesis and strongly influences growth and phenotype of transgenic tobacco plants. EMBO J. **9,** 3033–3044.

Walbot, V., and Cullis, C.A. (1985). Rapid genomic change in higher plants. Annu. Rev. Plant Physiol. **36,** 367–396.

Walker, J.C., Howard, E.A., Dennis, E.S., and Peacock, W.J. (1987). DNA sequences required for anaerobic expression of the maize alcohol dehydrogenase 1 gene. Proc. Natl. Acad. Sci. USA **84,** 6624–6628.

Wang, A.M., Doyle, M.V., and Mark, D.F. (1989). Quantitation of mRNA by the polymerase chain reaction. Proc. Natl. Acad. Sci. USA **86,** 9717–9721.

Weeden, N.F. (1981). Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. J. Mol. Evol. **17,** 133–139.

Widholm, J.M., and Ogren, W.L. (1969). Photorespiratory-induced senescence of plants under conditions of low carbon dioxide. Proc. Natl. Acad. Sci. USA **63,** 668–675.

Wolfe, K.H., Gouy, M., Yang, Y.-W., Sharp, P.M., and Li, W.-H. (1989). Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. Proc. Natl. Acad. Sci. USA **86,** 6201–6205.