The Plant Cell, Vol. 4, 785–798, July 1992 © 1992 American Society of Plant Physiologists

A Novel Light-Regulated Promoter Is Conserved in Cereal and Dicot Chloroplasts

David A. Christopher,[1] Minkyun Kim, and John E. Mullet

Department of Biochemistry and Biophysics, Texas A & M University, College Station, Texas 77843-2128

The chloroplast *psb*D-*psb*C genes encode D2 and cp43, a reaction center protein and chlorophyll-binding antenna protein of photosystem II, respectively. We have previously shown that differential accumulation of light-induced *psb*D-*psb*C mRNAs in barley chloroplasts is due to transcription from a blue light-responsive promoter (LRP). It is hypothesized that the light-induced mRNAs help to maintain levels of the D2 polypeptide, which is photodamaged and degraded in illuminated plants. To determine if light-induced accumulation of *psb*D-*psb*C mRNAs was a conserved phenomenon in chloroplasts, the expression of *psb*D-*psb*C operons from five cereals (barley, wheat, rice, maize, and sorghum) and three dicot (tobacco, spinach, and pea) species was examined. Cereal and dicot *psb*D-*psb*C operons differ due to several DNA rearrangements that moved *psb*K-*psb*I proximal to *psb*D-*psb*C, allowing cotranscription of these genes and production of several unique transcripts in cereals. Despite differences in the structure and expression of the cereal and dicot *psb*D-*psb*C operons, the accumulation of light-induced *psb*D-*psb*C mRNAs was conserved in all species studied. An unusual feature of the light-induced mRNAs was the occurrence of 5′ end microheterogeneity. The multiple 5′ termini were mapped to several consecutive nucleotides (8 to 25 bp) within a highly conserved (61%) DNA region that represents the transcription initiation site for the mRNAs in barley and tobacco. The novel LRP differs in sequence from typical plastid promoters that have prokaryotic "−10" and "−35" elements and is centered 570 bp (cereals), 900 bp (tobacco, spinach), or 1100 bp (pea) upstream from the *psb*D translational start codon. We propose that physiological and gene regulatory demands of the chloroplast act as constraints that preserved the linkage of the LRP with *psb*D despite DNA inversions involving the *psb*D upstream region.

INTRODUCTION

Angiosperm chloroplast genomes are, in general, similar in size and encode, with a few exceptions (for review, see Palmer, 1990; Gantt et al., 1991), the same genes. The chloroplast genes are compactly organized, often in complex operons, in a highly conserved order (Ohyama et al., 1986; Shinozaki et al., 1986; Hiratsuka et al., 1989). Compared with mitochondrial genomes, few recombination events have become fixed in chloroplast DNA (for review, see Newton, 1988; Palmer, 1990). The conserved architecture of the chloroplast genome suggests that genome evolution may be occurring under strong constraint (Palmer, 1990). A current question in chloroplast biology concerns the mechanisms responsible for preserving cohesive gene arrangements in chloroplast DNA. Moreover, little is known about the impact of gene rearrangements on chloroplast gene expression.

Chloroplast genes encode tRNAs, rRNAs, and proteins involved in transcription, translation, and photosynthesis. Genes encoding subunits of one protein complex, or proteins having related functions, are often clustered on the chloroplast genome. For example, the gene cluster *rpo*B-*rpo*C1-*rpo*C2 encodes subunits of the RNA polymerase, whereas *psb*I-*psb*K-*psb*D-*psb*C encode subunits of photosystem II (PSII) (for chloroplast gene nomenclature, see Hallick, 1989). The sequestration of related genes into operons facilitates stoichiometric production of subunits at proper levels. This is significant because the abundance of different plant protein complexes varies from 10^6 per plastid for PSII to $<10^4$ per plastid for RNA polymerase (Mullet et al., 1990). As expected, RNA levels from the *rpo*B-*rpo*C gene cluster are 200-fold lower than *psb*D-*psb*C mRNA levels (J.C. Rapp, B.J. Baumgartner, and J.E. Mullet, manuscript submitted).

Gene clusters may also be conserved due to special regulatory mechanisms that optimize gene expression. These mechanisms involve an ensemble of differentially regulated promoters, RNA processing/stability motifs, and translational regulatory signals (Gruissem and Zurawski, 1985; Klein and Mullet, 1987, 1990; Mullet and Klein, 1987; Stern et al., 1989; Eisermann et al., 1990; Haley and Bogorad, 1990; Danon and Mayfield, 1991; Kim et al., 1991; reviewed in Gruissem et al., 1988). The accumulation of various regulatory mechanisms in an operon may account for the complexity of chloroplast operon expression. The *psb*I-*psb*K-*psb*D-*psb*C-*orf*62 (*orf*, open reading frame) operon in the cereal barley, for example, accumulates 12 different mRNAs through the action of at least three different promoters, RNA processing, and transcription termination events (Berends-Sexton et al., 1990b). Most of these mRNAs accumulate in a light-independent manner dur-

[1] To whom correspondence should be addressed.

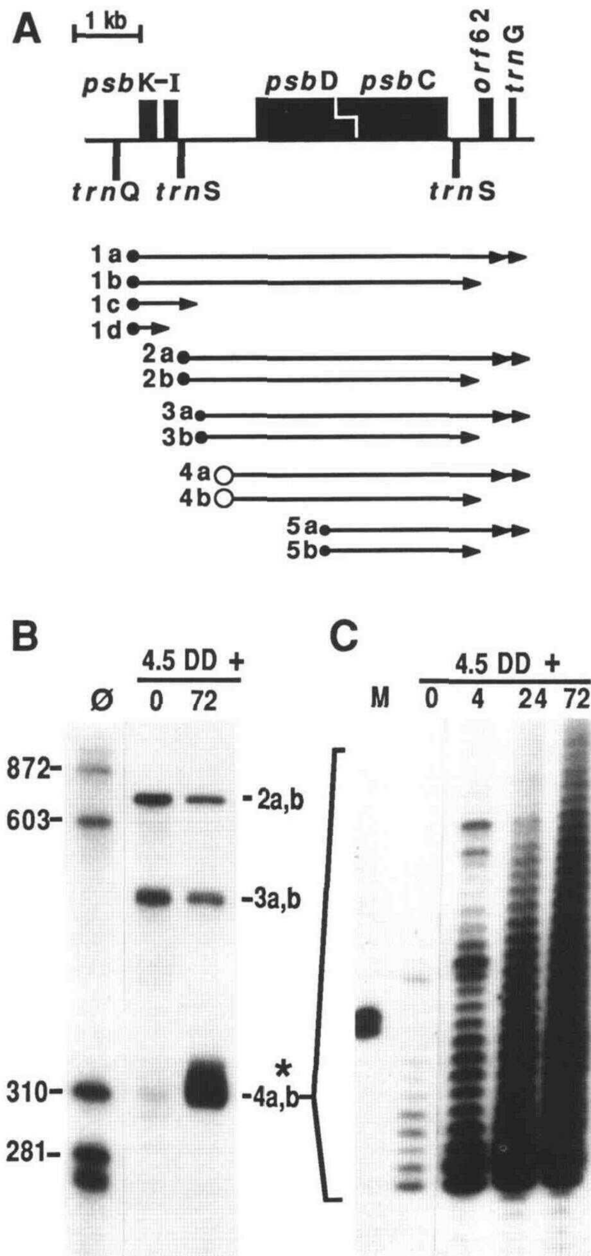

Figure 1. Transcripts and Gene Map of the Barley Chloroplast *psb*D-*psb*C operon.

(A) Operon encoding the loci 5'-*psbK-psbI-psbD-psbC-orf62-trnG* (GCC) for four thylakoid polypeptides of PSII, a conserved orf of 62 amino acids, and a glycyl tRNA (GCC). Genes are indicated by solid boxes and follow the described chloroplast gene nomenclature (Hallick, 1989). The 3' end of the protein coding region of the *psb*D gene overlaps 17 bp of the 5' end of the protein coding region of the *psb*C gene. Three tRNA genes, from left to right, Q(UUG), S(GCU), and S(UGA), are encoded on the opposite strand to that of the *psb*D-*psb*C operon. Mapped transcripts are designated by long arrows and labeled 1a to 5b (Berends-Sexton et al., 1990a, 1990b). Solid circles indicate 5' ends of mRNAs that accumulate in etioplasts of dark-grown barley. Open circles represent 5' ends of the light-induced mRNAs. Arrowheads represent 3' termini.

(B) Detection of light-induced *psb*D-*psb*C mRNAs (4a and 4b) by primer extension analysis. 4.5 DD (days, dark) refers to growth in complete darkness for 4.5 days. Numbers above each lane (0, 4, 24, and 72) refer to hours grown in the light after the 4.5-DD period. Hence, 4.5 DD + 0 indicates 4.5-day-old, dark-grown barley with no additional illumination. 4.5 DD + 72 refers to seedlings grown in 72 hr of light after the 4.5-DD period. Plastid RNAs were analyzed with primer 3 (described in Methods). RNA 5' ends detected in the experiments are labeled 2a to 4b corresponding to the transcripts in **(A)**. Light-induced RNAs are designated with an asterisk. Lane φ refers to some of the φX174/HaeIII restriction fragments used as markers, which are labeled in base pairs to the left.

(C) High resolution primer extension analysis of the 5' ends of the light-induced *psb*D-*psb*C mRNAs using primer 2. Barley seedlings were grown as described above in **(B)**, with two additional light treatments of 4 and 24 hr, after the 4.5-DD period. 72 bp refers to a ^{32}P-labeled 72-bp φX174/HaeIII restriction fragment. A black bracket connects bands for the light-induced RNA 5'-end signal in **(B)** with the multiple 5' end signals in **(C)**.

ing the early phase of chloroplast biogenesis, but two mRNAs, which have a common 5' end, accumulate differentially when plants are illuminated. The accumulation of these special *psb*D-*psb*C mRNAs in illuminated plants occurs by way of differential activation of one of the three *psb*D-*psb*C promoters (Berends-Sexton et al., 1990b) by blue light (Gamble and Mullet, 1989). We proposed that the physiological reason for blue light-induced accumulation of *psb*D mRNA is related to light-induced damage and turnover of the *psb*D gene product, D2. D2 is a chlorophyll protein that binds to a second chlorophyll protein, D1, to form the reaction center of PSII (for review, see Mattoo et al., 1989). During photochemistry, these proteins suffer damage that necessitates turnover and resynthesis to maintain PSII function (Mattoo et al., 1984; Fromm et al., 1985; Ohad et al., 1985; Schuster et al., 1988; Greenberg et al., 1989; Shipton and Barber, 1991). We proposed that activation of *psb*D transcription by blue light is necessary to maintain D2 synthesis and, consequently, PSII activity under these conditions (Berends-Sexton et al., 1990b).

Light-induced damage and turnover of D1 and, to a lesser extent, D2 are widespread phenomena in oxygenic photosynthetic organisms (Mattoo et al., 1984, 1989; Schuster et al., 1988; Bustos and Golden, 1992). However, at present, the light-induced transcription of *psb*D in higher plants has only been documented in barley (Berends-Sexton et al., 1990b). Therefore, one objective of this study was to determine if this response was common to other monocots and dicots. This is an interesting question because DNA inversions have altered the plastid genomes such that the *psb*D-*psb*C operons in dicots and monocot members of the cereal group are quite different (Palmer, 1990). In cereals, the operon has the gene order *psb*K-*psb*I-*psb*D-*psb*C-*orf*62. In contrast, the *psb*K-*psb*I loci in dicots are separated from *psb*D-*psb*C-*orf*62 by more than 23 kb of

DNA. Therefore, a second objective of this study was to determine the influence that these DNA rearrangements have on the expression of this operon. Despite rearrangements involving *psb*D-*psb*C loci, we found the novel light–responsive promoter (LRP) to be part of *psb*D regulation in all angiosperms studied. The conserved linkage of the LRP with *psb*D is an intriguing feature of the *psb*D-*psb*C operon and underlines a more widespread occurrence of light-regulated *psb*D-*psb*C expression in plants. We explain the possible evolutionary origin of the cereal operon and suggest that one of the molecular constraints that preserved the linkage of LRP with the *psb*D gene is the physiological need to replace D2 subunits in mature photosynthetically active chloroplasts.

RESULTS

Light-Induced *psbD-psbC* RNAs: Molecular Markers for the LRP in Barley Chloroplasts

The barley *psb*K-*psb*I and *psb*D-*psb*C-*orf*62 loci form a single polycistronic operon from which at least 12 overlapping mRNAs are produced, as illustrated in Figure 1A. We have previously shown that mRNAs 4a and 4b (Figure 1A) specifically accumulate in response to light as a result of light-induced transcription from a promoter that is relatively inactive in dark-grown plants (Berends-Sexton et al., 1990b). We showed that this accumulation is due to exposure to light and not due to barley seedling age (Gamble et al., 1988). This result can also be observed by comparing the third lane of Figure 1B (4.5 DD [days old, dark grown] + 72 hr light) with the first lane of Figure 2A (7.5 DD). In this comparison, RNA samples were analyzed from seedlings that are both 7.5 days old. Only the light-treated seedlings exhibit the induction of mRNAs 4a and 4b. After 72 hr of illumination of 4.5-day-old, dark-grown barley seedlings (Figure 1B), the light-induced mRNAs 4a and 4b have accumulated to become the predominant mRNAs encoding *psb*D-*psb*C in mature barley chloroplasts (Gamble et al., 1988).

As presented in Figure 1C, the 5′ ends of light-induced mRNAs 4a and 4b are heterogeneous, mapping within a 25-nucleotide region. Signals corresponding to mRNAs 4a and 4b are barely detected in 4.5-day-old, dark-grown seedlings. However, illumination with 4, 24, and 72 hr of light produced a significant increase in signal intensity observed as multiple bands, which, after 72 hr of illumination, spanned at least 25 nucleotides. Band intensity was not uniform over these 25 nucleotides. The largest proportion of the mRNA population possessed shorter 5′ ends, which yielded primer extension products with faster mobility than the 72-bp DNA marker, regardless of the length of light treatment. It is possible that this result was an experimental artifact or was due to RNA secondary structure. However, 5′ end microheterogeneity was observed using two different purified primers. We have not found any major RNA secondary structures in this region that could contribute to the 5′ end microheterogeneity. Furthermore,

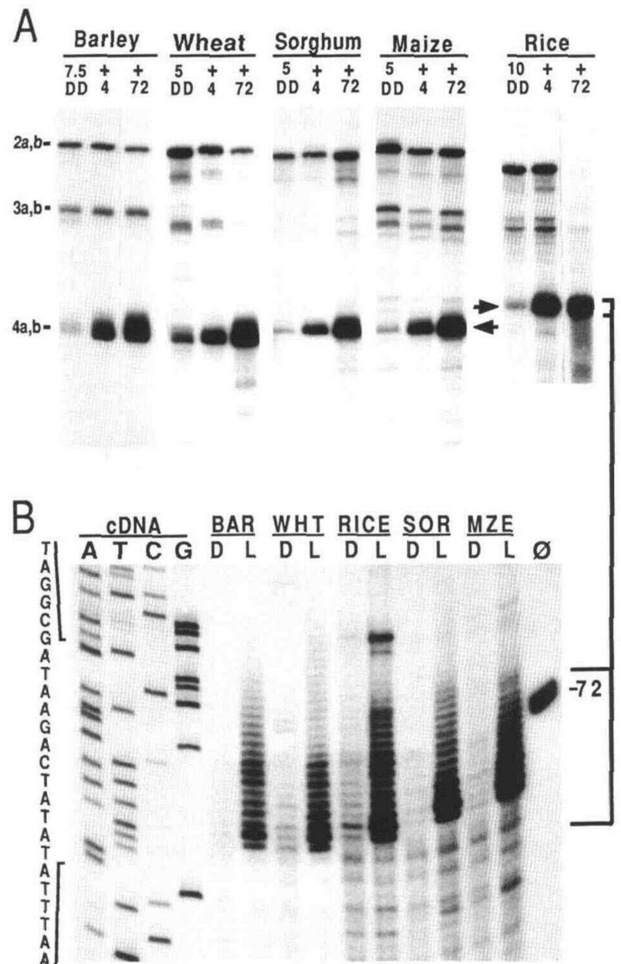

Figure 2. Accumulation of Light-Induced *psb*D-*psb*C Transcripts in Five Different Cereals during Greening of Etiolated Seedlings.

(A) Low resolution analysis of *psbD-psbC* mRNAs during chloroplast development. Primers 1 and 3 were utilized in the experiments. Lanes are labeled according to type of plant and corresponding dark and light treatment. Seedlings were grown in complete darkness for 7.5 days (7.5 DD; barley), 5 days (5 DD; wheat, sorghum, and maize), and 10 days (10 DD; rice). After the dark treatment, some seedlings were transferred to an additional 4 and 72 hr of light. 7.5 + 4 refers to 7.5 days of darkness followed by 4 hr of light. Bands corresponding to specific barley *psb*D-*psb*C RNAs in Figure 1 are labeled 2a to 4b (light induced) to the left of the barley lanes. Arrowheads designate the 5′-end signals for the light-induced RNAs.

(B) High resolution primer extension mapping of the 5′-ends of the light-induced RNAs from barley (BAR), wheat (WHT), rice, sorghum (SOR), and maize (MZE). D and L refer to the dark and 4-hr light treatments described in **(A)**. Lanes labeled cDNA (A, T, C, and G) refer to dideoxy nucleotide sequencing reactions conducted on plasmid pBE3-5 (Berends-Sexton et al., 1990b) using the same primer as for primer extension, which was primer 2. The cDNA sequence read from the lanes is printed to the left of the panel. φ refers to a ^{32}P-labeled, 72-bp φX174/HaeIII restriction fragment used as an additional marker. A black bracket connects signals for the same light-induced RNA 5′ ends in **(A)** with **(B)**.

this feature was not observed for the 5′ ends of *rbc*L mRNAs presented in Figure 3. Hence, a noteworthy feature of the light-induced *psb*D-*psb*C mRNAs is the occurrence of 5′ end microheterogeneity during all stages of light-induced chloroplast development. These light-induced mRNAs serve as molecular markers for the presence and activity of the LRP.

Light-Induced Accumulation of *psb*D-*psb*C RNAs Is Conserved in Cereals

The LRP could be a conserved feature of other cereal *psb*D-*psb*C operons. To understand more about the light response mechanism and to identify light-induced *psb*D-*psb*C mRNAs in other cereal chloroplasts, the expression of *psb*D-*psb*C operons from five different cereal genera was examined during light-induced chloroplast development. The results of the primer extension analysis of *psb*D-*psb*C RNAs from wheat, rice, maize, sorghum, and barley are shown in Figure 2A. Similar results are obtained when the period of growth in the dark is from 4 to 7.5 days prior to illumination. Results from 7.5-day-old (barley) and 5-day-old (wheat, maize, and sorghum), dark-grown seedlings are shown here. For barley, wheat, and rice, the top 4 cm of the primary leaf was used in all treatments. For maize and sorghum, tissue from the top 1 cm of the emerging primary leaf was harvested from 5-day-old, dark-grown (5 DD) seedlings and 5 DD seedlings illuminated with an additional 4 hr of light. In the 72-hr light treatment, the top 4 cm of primary leaf tissue was utilized. Rice required longer time (10 days) and warmer temperatures (28°C) in the dark for emergence and growth. In plastids of each plant species, illumination of dark-grown seedlings with an additional 4 or 72 hr of light caused the specific accumulation of RNAs that correspond to 4a and 4b in barley, as indicated by arrowheads in Figure 2A. After 72 hr of illumination, light-induced *psb*D-*psb*C RNAs accumulated to significant levels. This is direct evidence of evolutionary conservation of light-induced *psb*D-*psb*C RNAs among the cereals. This experiment also provided evidence supporting the idea that an LRP, analogous to the one in barley, is a conserved feature of other cereal *psb*D-*psb*C operons.

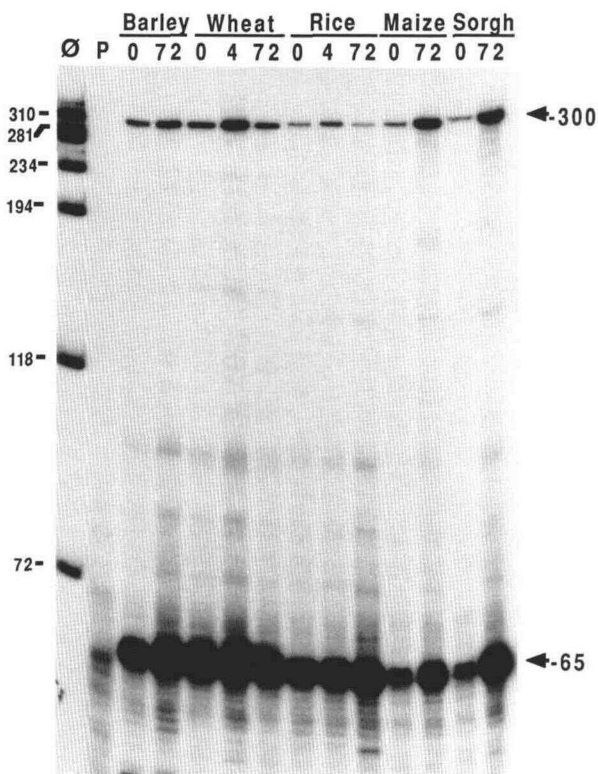

Figure 3. Primer Extension Analysis of *rbc*L mRNAs from the Five Cereals during Greening of Etiolated Seedlings.

Primer 6, which is described in Methods, was used. Treatments were as described in Figure 2A and are designated at the top of each lane as 0 (dark) and 4 or 72 hr of light. Arrowheads labeled −300 and −65 refer to specific *rbc*L RNA 5′ ends. Lane P designates a reaction with ^{32}P-labeled primer only, no RNA. φ refers to the φX174/HaeIII markers, which are labeled in base pairs at the left.

Some minor differences in the light response were also observed (Figure 2A). For instance, in rice the accumulation of the light-induced RNAs was rapid and after 4 hr of light almost equaled the signal intensity after 72 hr, whereas for the other species, the signal after 4 hr of light was weaker than after 72 hr. The higher temperatures and longer growing time required for rice could have allowed for a more rapid light response, or the response could be innately more rapid in rice. Very low levels of the light-induced RNAs do occur in etioplasts and may be a result of low basal levels of promoter activity in the dark (for example, see wheat in Figure 2).

The 5′ ends of barley *psb*D-*psb*C mRNAs 2a,b and 3a,b, which are major *psb*D-*psb*C RNAs present in etioplasts of dark-grown seedlings, were also detected using the same primer (Figure 2A). Similarly, at least one major dark-predominant *psb*D-*psb*C mRNA species was detected in plastid RNAs from dark-grown wheat, sorghum, maize, and rice. They were observed as bands with slower migration than the light-induced signals (Figure 2A) and appeared to correspond to the 5′ ends of barley mRNAs 2a,b. After 72 hr of illumination, the changes in the levels of these RNAs can be divided into the following three responses in the five cereals: (1) decreased slightly for barley and wheat, (2) increased slightly in maize and sorghum, and (3) disappeared in rice. These three molecular differences in RNA levels could reflect phylogenetic differences among the five species of cereals, which also fall into three groups. Barley and wheat are closely related and form one group, as do maize and sorghum, whereas rice is more distant from the four other species (Dobley et al., 1990). Additional *psb*D-*psb*C mRNA 5′ ends detected in dark-grown plants included the 5′ ends of barley mRNAs 3a,b. Bands with similar mobility were present in wheat, maize, and rice, but absent in sorghum. The sorghum nucleotide sequence for this region is not available for comparison with the transcript data. In any case, mRNAs

analogous to barley mRNAs 3a,b may not be present in the *psbD-psbC* mRNA population from etioplasts of all cereals.

Light-Induced *psbD-psbC* RNAs Possess 5′ End Microheterogeneity in Cereals

Mapping the 5′ termini of the light-induced RNAs to specific nucleotide sequences was a prerequisite to further studies on the promoter and light response mechanism. The 5′ ends of the light-induced RNAs were mapped in primer extension analysis experiments using a primer that hybridized 72 bp downstream from the 5′ ends of the RNAs. These results are presented in Figure 2B. To facilitate accurate mapping, a ladder of the barley cDNA sequence, which was generated using the same primer as for primer extension, was electrophoresed in parallel with the primer extension reactions. In each species studied, the primer appeared to be equidistant (within 1 to 2 nucleotides) from the light-induced mRNA 5′ ends. The primer extension signals for the 5′ ends of the light-induced RNAs consisted of multiple bands. Hence, mRNA 5′ end microheterogeneity appears to be a conserved feature of the light-induced *psbD-psbC* RNAs. Light-induced 5′ ends were scattered over 14 to 25 nucleotides that were located approximately 570 bp upstream from the translational initiator ATG codon of the *psbD* gene. The barley DNA sequence of the RNA-like strand for this region is CTCTATCCGCTATTCTGATATATAA (positions 5904 to 5928 in Berends-Sexton et al., 1990a). The greatest concentration of 5′ end signals was positioned within the sequence GATATATAAA. From nucleotide sequence comparisons (described below), this sequence and immediate flanking sequences were conserved among all cereals studied here.

As controls, the 5′ ends of the *rbcL* RNAs, transcribed from the gene that encodes the large subunit of ribulose bisphosphate carboxylase, were also assayed using primer extension analysis, as presented in Figure 3. Both previously mapped 5′ ends (Mullet et al., 1985) for the primary transcript, "−300," and the processed product, "−65," were detected. In contrast to *psbD-psbC*, no new *rbcL* mRNA 5′ ends were observed between the treatments. Unlike the broad series of multiple bands for the 5′ ends of the light-induced *psbD-psbC* RNAs, the major signals of the 5′ ends of the *rbcL* RNAs map to 1 to 2 nucleotides. Small increases in signal intensity between dark and light treatments were observed in barley, wheat, and rice. However, significant increases in the levels of both the "−300" and "−65" RNA 5′ ends were observed upon transfer of dark-grown maize and sorghum seedlings to 72 hr of light. The levels of *rbcL* mRNAs detected in dark-grown maize and sorghum were much lower than for dark-grown barley, wheat, and rice. This low level of *rbcL* mRNAs in dark-grown maize and sorghum could reflect the less developed state of the primary leaf in etiolated cereal seedlings that use the C_4 dicarboxylic pathway (maize and sorghum) versus cereals that use the C_3 photosynthetic pathway (barley, wheat, and rice). In these experiments, the primary leaf was used and was still partially retained in the coleoptilar sheath in 5-day-old, dark-grown and 5-day-old plus 4 hr, light-treated C_4 cereals, maize, and sorghum. After 72 hr of illumination, leaf mesophyll cell and chloroplast development are greatly stimulated (Mullet, 1988; Schrubar et al., 1991). This is accompanied by an increase in total plastid transcription (Schrubar et al., 1991), which could explain the light-induced increase in *rbcL* (Figure 3). The levels of *psbD-psbC* mRNAs already present in dark-grown maize and sorghum increased slightly after 72 hr of light, but less than the increase observed for *rbcL* mRNAs. The transcription of *psbD-psbC* and *rbcL* genes appears to follow a different pattern in dark-grown maize and sorghum. Moreover, the light-induced increase in *psbC-psbD* mRNAs 4a and 4b is selective and occurs in addition to a general stimulation of chloroplast transcription.

New Light-Induced *psbD-psbC* mRNA 5′ Ends Detected in Dicot Chloroplasts

In previous experiments (Berends-Sexton et al., 1990b), the barley LRP was transcribed in vitro using plastid extracts from light-grown pea, yet no light-induced *psbD-psbC* mRNAs have been detected in dicots. Therefore, we studied the expression of the *psbD-psbC* operons from three dicots (tobacco, pea, and spinach) to determine if they also had the light response. Tobacco and spinach were chosen because the DNA sequences for the *psbD-psbC* genes and upstream regions were available (Offerman, 1988; Yao et al., 1989). Pea was also chosen because the genes had previously been mapped, their coding regions sequenced (Rasmussen et al., 1984), and their expression analyzed using RNA dot blots (Woodbury et al., 1989). Furthermore, the pea chloroplast genome is unique in that it lacks the inverted repeat and has undergone additional recombination events relative to the other dicot chloroplast DNAs (Palmer, 1990).

The treatments used and results of the dicot gene analysis are presented in Figure 4B, whereas the positions of some of the 5′ ends with respect to the *psbD* coding locus are presented in Figure 4A. Primer extension experiments utilized a single primer that hybridized to a conserved sequence of the 5′ end of the protein coding region of *psbD* . The light-induced RNAs from the cereal barley are presented for reference. When 5-day-old, dark-grown spinach and pea seedlings were transferred to light for 96 and 72 hr, respectively, light-induced RNAs, previously unidentified for pea and spinach, were detected as bands at position 900 (spinach) and 1100 (pea) relative to the *psbD* translational initiator ATG codon (Figures 4A and 4B). The light induction is also observed in both spinach cotyledons and primary leaf tissue (data not shown). Data from primary leaf tissue are presented in Figure 4. Dark-adapted plants were utilized for tobacco and refer to 15-day-old, light-grown plants that were transferred to complete darkness for 3 days, followed by an additional 48 hr of illumination. In tobacco, the intensity of the band at estimated position 900 increased in the analysis of RNAs from primary

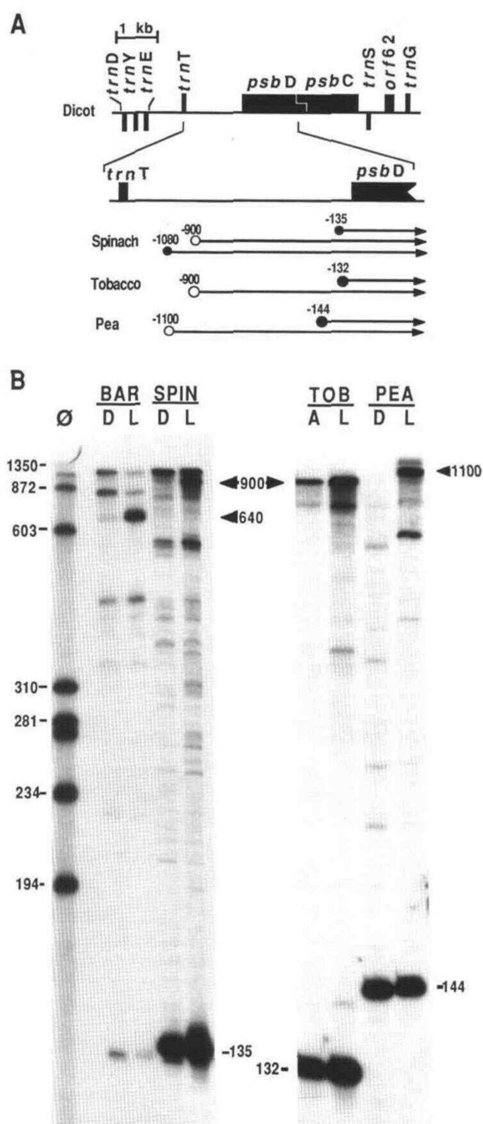

Figure 4. Detection and Primer Extension 5′ End Mapping of Light-Induced *psb*D-*psb*C RNAs from Three Dicots.

(A) Dicot *psb*D-*psb*C operon, flanking genes, and positions of the 5′ ends of the *psb*D-*psb*C RNAs. Gene symbols are as given in Figure 1. The intergenic region between *trn*T and *psb*D is expanded below the gene map. Arrows designate transcripts (open circles, light-induced; closed circles, RNAs in dark-grown plants) mapped in the experiments presented in **(B)**.

(B) Analysis of *psb*D-*psb*C transcripts from dark- and light-grown dicots. Lanes contain barley (BAR), spinach (SPIN), tobacco (TOB), and pea. D, L, and A refer to plants grown in dark, light, and "dark-adapted" conditions, respectively, as described in Methods. Primer 1 (Methods) was utilized in the experiments. Arrowheads designate primer extension signals for light-induced RNA 5′ ends, which are labeled in base pairs to the right and left. Bands labeled 132, 135, and 144 refer to the size of fragments for abundant 5′ end signals present in both dark- and light-grown plants. φ refers to the φX174/HaeIII markers, which are given in base pairs at the left.

leaves of light-grown seedlings relative to dark-adapted seedlings. This mRNA 5′ end corresponds to the mRNA 5′ end previously localized at position −905 in tobacco (Yao et al., 1989). The light-induced *psb*D-*psb*C mRNAs appear to comprise a smaller proportion of the total *psb*D-*psb*C mRNA population in dicots than in cereals. No light-specific *rbc*L mRNAs were observed in a parallel analysis of the RNAs from these dicot plants (data not shown).

For each dicot species, additional 5′ ends for *psb*D-*psb*C RNAs were detected as strong bands at positions 132 (tobacco), 1080 and 135 (spinach), and 144 (pea) in both dark- and light-grown seedlings. A weak, previously unmapped signal was observed at position −135 for barley (Figure 4B). These *psb*D proximal mRNA 5′ ends had been previously mapped in tobacco (Yao et al., 1989) and spinach (Offerman, 1988). Hence, from these experiments, two classes of dicot *psb*D-*psb*C mRNAs were identified, as illustrated in Figure 4A. One class consisted of new, previously undescribed light-induced mRNAs. The second class of mRNAs was found in both dark- and light-grown plants and was primarily represented as major bands at positions 132 to 144 (Figure 4B). The weak primer extension signal at position −135 in barley could be an artifact, but weak RNA 5′-end signals were also detected at this position in the other cereals (data not shown). The difference between the intensity and ratio of barley and spinach signals at position 135 indicates a significant variation in abundance of these mRNAs.

Cereal and Dicot Light-Induced mRNA 5′ Ends Map to a Conserved Chloroplast DNA Region

To identify the DNA region mediating the accumulation of the light-induced *psb*D-*psb*C RNAs in dicots, we mapped the 5′ ends of these RNAs in each species by the method of primer extension analysis. Results from the analysis of tobacco and spinach RNAs are presented in Figure 5. Light-induced RNAs in tobacco were detected as multiple bands positioned over 6 to 8 nucleotides comigrating with the 72-bp DNA marker (72 to 78, Figure 5). These 5′ ends were mapped to the nucleotides 5′-TCTGATATTAAA within the larger sequence 5′-ATATCCGCTATTCTGATATTAAA of the RNA-like strand of the tobacco chloroplast DNA (positions 33,537 to 33,559; Shinozaki et al., 1986). Upon illumination of dark-grown spinach seedlings, a strong light-induced signal was detected at position 92 and a weaker one at position 106 relative to the primer annealing site. Although not as extensive as in cereals, the light-induced *psb*D-*psb*C RNAs of dicots also have some 5′ end microheterogeneity. These 5′ ends map to the nucleotides 5′-TTCTGATATTCAAA within the larger sequence 5′-ATACCCGCTATTCTGATATTCAAA of the RNA-like strand of the spinach chloroplast DNA (positions 10,908 to 10,933; Offerman, 1988). The upstream DNA sequences near the 5′ ends of the tobacco, spinach, and pea light-induced RNAs are highly similar to each other and to the DNA sequences near the 5′ ends of the cereal light-induced RNAs. In each case, the RNA with

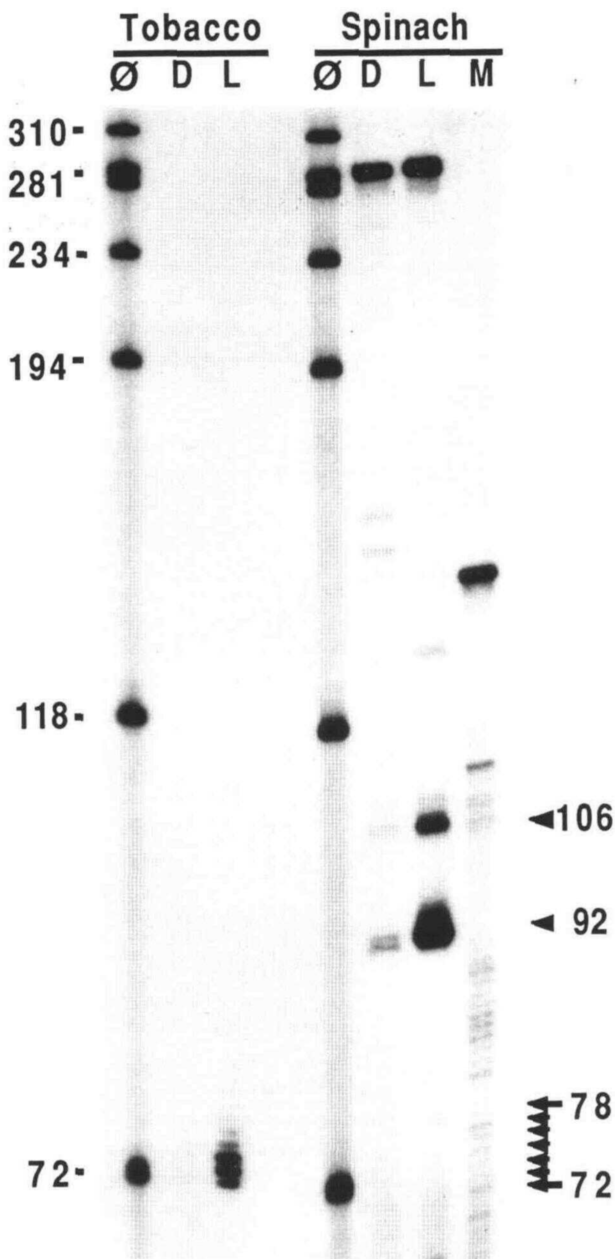

Figure 5. Fine Mapping the 5′ Ends of Light-Induced *psb*D-*psb*C RNAs in Two Dicots.

D and L refer to dark- and light-grown plants. Primers 4 (for spinach) and 5 (for tobacco) were utilized. Lane M designates a 143-bp restriction fragment from pHvLRP1 that was partially digested with DNaseI to yield smaller marker fragments differing by one nucleotide each. Numbered arrowheads on the right designate bands corresponding to the 5′ ends of light-induced RNAs and their molecular size as estimated from the markers. φ refers to the φX174/HaeIII markers, which are given in base pairs at the left.

the shortest 5′ end mapped exactly to the second adenosine residue within three conserved adenosine residues at the 3′ end of the sequence 5′-TTCTGATATTCAAA.

A sequence similarity comparison was conducted for the DNA sequences of the RNA-like strand of the *psb*D coding locus and 700 nucleotides (cereals) or 1100 nucleotides (dicots) of the upstream untranslated regions from spinach (Offerman, 1988), wheat (Howe et al., 1988), barley (Efimov et al., 1988; Berends-Sexton et al., 1990a), rice (Hiratsuka et al., 1989), tobacco (Shinozaki et al., 1986), and partial sequences from pea (Rasmussen et al., 1984), maize, and sorghum. The sequences of the LRP region from pea, maize, and sorghum were determined in this study. Sequence similarity and a consensus sequence were determined using computer-assisted sequence similarity search and alignment programs. The sequences could be classified into two levels of similarities that followed the two major cereal and dicot groupings. Generally, the sequences upstream from *psb*D were highly conserved among cereals (data not shown), but poorly conserved between cereals and dicots.

The results of this analysis are presented in Figures 6A and 6B. As expected, regions of nucleotide sequence conservation were confined to functional protein coding sequences. For example, a consensus sequence consisting of the first 70 nucleotides of the protein coding sequence of *psb*D and 60 nucleotides of the immediate 5′-untranslated region, which contains the ribosome binding site, was 89% conserved between cereals (barley, wheat, rice, and maize) and dicots (tobacco, spinach, and pea). Nucleotide sequence similarity drops to 32% in the adjacent 70-bp untranslated region (hatched box, Figure 6A). This region was located between −60 and −130 upstream from the *psb*D translational initiator ATG codon. A consensus sequence for the next 440 bp of cereal upstream region has virtually no sequence similarity (9%) with any segment of the 670-bp dicot intergenic region (Figure 6A). The 9% reflects a consensus sequence derived from an alignment of the five available sequences for the intergenic region from barley, wheat, rice, spinach, and tobacco. The complete sequences from pea, sorghum, and maize were not available for this region (9%) and were not included in the analysis. The lack of sequence identity in this large region between rice and tobacco has been previously noted (Hiratsuka et al., 1989).

The 90-bp DNA region encoding the 5′ ends of the light-induced *psb*D-*psb*C RNAs (stippled box, Figure 6A) is located in cereals at −573 bp, except in wheat (−604 bp), and in dicots at −900 bp (tobacco and spinach) and −1100 bp (pea) upstream from the *psb*D translational initiator ATG codon (Figure 6A). This DNA region is 61% conserved at the nucleotide sequence level, which is significant for a chloroplast intergenic region (Figure 6B). The light-induced RNA 5′ ends were mapped to sequences at positions 47 to 69 (Figure 6B). Sequences weakly resembling chloroplast promoter elements (Gruissem and Zurawski, 1985; Hanley-Bowdoin and Chua, 1987) were also present in this region. A "−10" TATA motif is located at positions 62 to 65 (cereals) and positions 45 to 48 (dicots), and a "−35" motif is located at positions 19 to 24 (Fig-

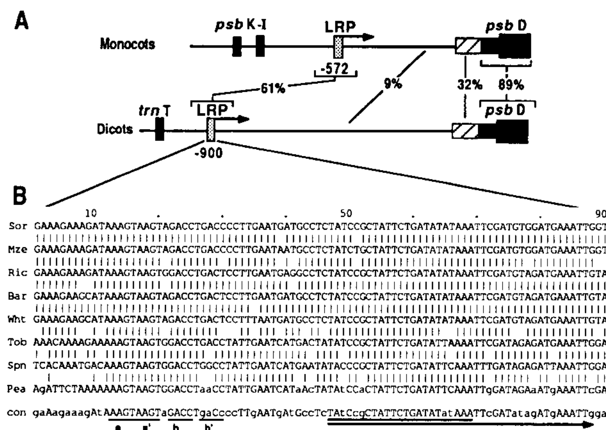

Figure 6. Comparison of the DNA Sequences of the mRNA-Like Strand in the *psb*D Upstream Region from Five Monocot Members of the Cereals (Grass Family) with Those from Three Dicots.

(A) The percentage of nucleotide sequence identity within specific subregions of the cereal and dicot *psb*D upstream regions is indicated. The sequence identity is based on a consensus sequence determined for the region by sequence alignment. The stippled box (LRP) and hatched box (from −60 to −130 upstream from *psb*D) highlight subregions of special sequence similarity. The numbers −572 and −900 refer to the number of nucleotides upstream from *psb*D.

(B) Alignment of the DNA sequences within a 90-bp region upstream of and spanning the 5′ ends of the light-induced *psb*D-*psb*C RNAs. Sequences from the following eight plant species are given: Sor, sorghum; Mze, maize; Ric, rice; Bar, barley; Wht, wheat; Tob, tobacco; Spn, spinach; and pea, pea. The consensus (Con) sequence is given at the bottom. Uppercase letters designate completely conserved nucleotides. Lowercase letters refer to a deviation of at least one nucleotide in the alignment. A lowercase, one-letter nucleotide abbreviation was selected based on the frequency of occurrence. Short, conserved direct repeats are single underlined and labeled a, a′ and b, b′. The multiple 5′ ends of the light-induced RNAs map to the double-underlined nucleotides. The arrow denotes the direction of transcription.

ure 6B). Direct repeats with the sequences of AAGT . . . AAGT and GACC . . . GACC were also observed (Figure 6B). The sequence GCCT (positions 42 to 45; Figure 6B) is repeated inversely as TCCG (positions 49 to 52; Figure 6B). A conserved GATA sequence is located at positions 60 to 63 (Figure 6B). No sequence resembling this region was found in the completely sequenced chloroplast genome of the bryophyte liverwort (Ohyama et al., 1986).

DISCUSSION

The LRP Mediates Light-Induced Transcription of *psbD-psbC*

The accumulation of light-induced *psb*D-*psb*C mRNAs was found to be conserved among the eight plant species studied. These RNAs arise from a conserved DNA region, presented in Figure 6B. Several lines of evidence support the idea that this conserved DNA region is a *psb*D LRP (Figure 6A) that is common to most plant chloroplast DNAs. We have previously determined by the method of in vitro RNA capping that the light-induced *psb*D-*psb*C mRNAs from barley are primary transcripts arising from a promoter distinct from other promoters in the *psb*D-*psb*C operon (Berends-Sexton et al., 1990b). Evidence that the light-induced mRNAs arise by light-induced transcription came from lysed chloroplast run-on transcription assays and from experiments in which the light-induced accumulation of RNAs was inhibited in vivo with tagetitoxin, a specific inhibitor of chloroplast transcription (Mathews and Durbin, 1990). In addition, a cloned DNA fragment containing the conserved DNA region (Figure 6A) initiated transcription in vitro using transcription extracts from light-grown pea (Berends-Sexton et al., 1990b). Likewise, the largest *psb*D-*psb*C mRNAs from tobacco were shown here to be light induced. Their 5′ ends mapped to the conserved DNA region at position −900 from the *psb*D translational start codon. The 5′ ends of these light-induced mRNAs were previously shown to be cappable primary transcripts (Yao et al., 1989). From these results, the 5′ ends of the barley and tobacco light-induced mRNAs represent the transcription initiation sites of these mRNAs. These transcription initiation sites reside within the conserved DNA region, as do the 5′ ends of the light-induced mRNAs from the other five plant species. It is possible that the promoter resides outside the conserved region. However, this would mean that the promoter resides in unconserved sequences between cereals and dicots, which seems a less likely alternative. Taken together, we interpret these results to support the existence of an LRP within the conserved DNA region.

A New Type of Chloroplast Promoter

The *psb*D LRP is exceptional in that: (1) It specifically responds to blue light (Gamble and Mullet, 1989). In contrast, genes such as *rbc*L, *atp*B (encodes β subunit of ATPase), and 16s rRNA are fully active in dark-grown barley (Klein and Mullet, 1990), as are other dark-active promoters within the barley *psb*D-*psb*C operon (Berends-Sexton et al., 1990b). Hence, the LRP has evolved a light-regulated specificity quite unlike other chloroplast promoters. (2) Despite genomic rearrangements and divergence of flanking DNA sequences in *psb*D-*psb*C operons of cereals relative to dicots, the nucleotide sequence of the LRP has been highly conserved, 61% at the nucleotide level among eight different sequences analyzed. A level of 61% nucleotide conservation is significant for a chloroplast DNA intergenic region considering that the cereal and dicot genes diverged approximately 100 to 200 million years ago (Wolfe et al., 1989; Doyle et al., 1992), and the *psb*D upstream regions in cereals relative to dicots do not share extensive regions of homology. Although the nucleotide sequence is highly conserved, typical prokaryotic-like "−10" and "−35" promoter elements of other known chloroplast promoters are poorly conserved or not in conventional positions in the LRP. For example,

the "−10" TATA motif is internal to the RNA 5′ end (Figure 6B). The plant sequences do not resemble the sequences identified upstream from the light-regulated cyanobacterial *psb*D gene (Bustos and Golden, 1992).

(3) Some of the conserved nucleotide sequences of the LRP resemble conserved sequences in light-responsive, transcriptionally activated nuclear genes (Gilmartin et al., 1990; Manzara et al., 1991; Piechulla et al., 1991). A conserved GATA sequence (positions 60 to 63; Figure 6B) and partially conserved sequence that is the inverse of GATA, CTATC (positions 46 to 50 and 53 to 58; Figure 6B), are found within the region near where the light-induced mRNA 5′ ends map. A GATA sequence is also found near the multiple 5′ ends of the light-predominant *pet*G (formerly *pet*E) RNAs from maize which encode polypeptide subunit V of the photosynthetic electron transport cytochrome b_6-f complex (Haley and Bogorad, 1990). Conserved GATA and CTATC (inverse of GATA) motifs are also found upstream of, and in some cases, exactly at the 5′ ends of light-regulated nuclear genes for chloroplast photosynthetic proteins (Piechulla et al., 1991). The relevance of these sequences in the chloroplast promoter is not known. However, it is tempting to speculate that specific sequences that confer light responsiveness in the chloroplast promoter may have been originally derived from a light-responsive nuclear gene. (4) One of the most unusual features of the light-induced *psb*D-*psb*C RNAs is the presence of 5′ end microheterogeneity. This feature could be due to reiterative transcription initiation as observed in *Escherichia coli* for the pBR322-derived *tet* promoter of the tetracycline resistance gene (Harley et al., 1990), loose initiation of transcription, or a 5′ to 3′ exonuclease that removes a few nucleotides from the 5′ end of the RNAs. In the latter case, some of the conserved nucleotides within the RNA terminus could form an RNA processing site.

Evolutionary Origin of the *psb*D-*psb*C Operon in Cereal Chloroplasts

The *psb*D-*psb*C operons of cereal and dicot chloroplast genomes are structurally distinct in their *psb*D upstream regions. In dicots, the LRP-*psb*D-*psb*C-*orf*62 loci are separated from *psb*K-*psb*I by over 23 kb of DNA. In contrast, the LRP-*psb*D-*psb*C-*orf*62 loci of cereals reside 1.2 kb downstream from, and in the same polarity as, the *psb*K-*psb*I loci. The basis for this structural distinction is outlined in a model that is depicted in Figure 7. This model applies to monocot members of the cereals only and not to monocots such as duckweed (DeHeij et al., 1983) and orchids (Chase and Palmer, 1989), which share the dicot gene arrangement. The model is used to explain the evolutionary origin of the cereal *psb*D-*psb*C operon, which came later in evolution (Doyle et al., 1992), from an ancestral dicot genome by way of a hypothetical intermediate genomic structure.

Included in Figure 7 are a series of three overlapping inversions (Howe et al., 1988; Hiratsuka et al., 1989; Palmer, 1990) that caused changes in *psb*D-*psb*C and *psb*K-*psb*I gene organization in cereals relative to dicots. For simplicity, each inversion has been assigned a number. The single large inversion 1 placed the LRP-*psb*D-*psb*C-*orf*62 loci adjacent to, but in the opposite polarity as, the *psb*K-*psb*I loci. Two smaller inversions, 2 and 3, placed LRP-*psb*D-*psb*C-*orf*62 and *trn*T in the same polarity as *psb*K-*psb*I. The LRP therefore became

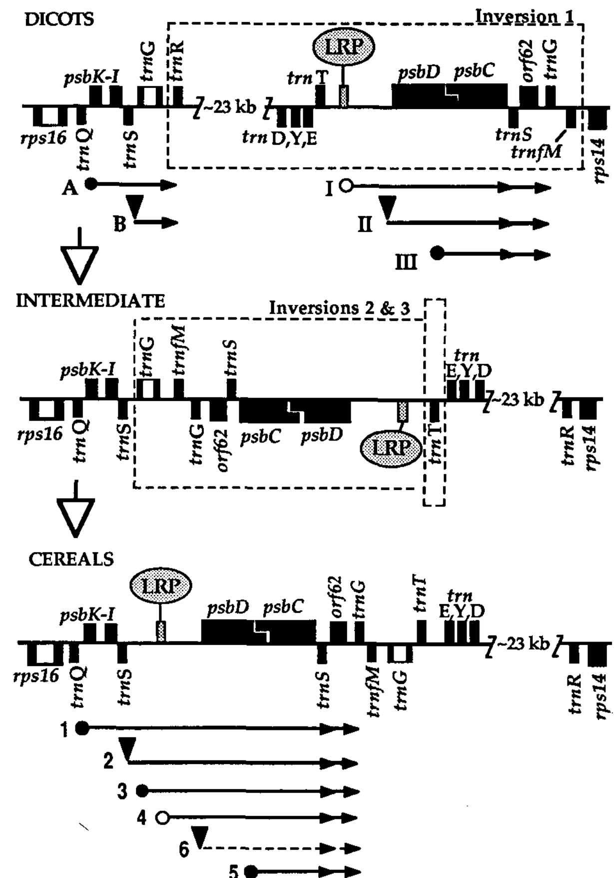

Figure 7. Model for the Evolutionary Origin of the Cereal Chloroplast *psb*D-*psb*C Operon from a Dicot-Like Ancestor.

Solid boxes indicate gene coding sequences. Open boxes designate introns. Genes on top of the lined axis are transcribed from left to right. Genes of the opposite polarity are below the lined axis. The stippled box and oval signify the conserved LRP (Figure 6). DNA regions involved in three inversions are delineated with dashed boxes. Long horizontal and dashed arrows refer to transcripts, which are subdivided into classes for the *psb*K-*psb*I-*trn*G loci in dicot tobacco (class A and B; Meng et al., 1991), *psb*D-*psb*C loci in the dicot tobacco (I, II, and III; Yao et al., 1989), and cereal barley (classes 1 to 5 from Berends-Sexton et al., 1990b). Low abundance class 6 (dashed line) was determined in this study (Figure 4). The LRPs and promoters from mRNAs that accumulate in the dark are indicated with open and solid circles, respectively, that are located at the 5′ end of the transcripts. Horizontal arrowheads at the 5′ end of the transcripts (for example, class B and class II, dicots; class 2 and 6, cereals) indicate 5′ ends that arise by RNA processing. Open vertical arrowheads indicate the proposed direction of evolution from the dicot to cereals.

part of the intergenic region between *psbI* and *psbD*. A new order of juxtaposed genes encoding polypeptide components of PSII was thus formed and, intriguingly, a unique operon was created that has a series of overlapping polycistronic mRNAs even more complex than in dicots (Figure 7).

This notion is supported further when the positions of promoters and RNA processing sites for the polycistronic mRNAs are compared in dicots and cereals. Of the dicots, the *psbD-psbC* transcripts from tobacco chloroplasts are the most extensively studied (Yao et al., 1989; Meng et al., 1991) and will be used as the representative dicot for comparison with the cereals. It must be noted that from this study the map positions of spinach and pea *psbD-psbC* RNAs are similar to tobacco but spinach chloroplasts have an additional dark-predominant *psbD-psbC* mRNA not detected in tobacco. The barley *psbD-psbC* mRNAs are the best studied from cereals (Berends-Sexton et al., 1990a, 1990b) and will be used as the representative cereal in the comparison. Three major classes of transcripts (I, II, and III; Figure 7) arise from the dicot *psbD-psbC* operon. Two major classes of transcripts are shown from the dicot *psbK-psbI-trnG* operon (A and B; Figure 7). Additional transcripts only encoding *trnG* are not shown (Meng et al., 1991). Five major classes of transcripts and an additional low abundance class designated class 6 found in this work arise from the cereal *psbD-psbC* operon (Figure 7).

We suggest that cereal *psbD-psbC* mRNAs in class 4, class 6, and class 5 are homologous to the dicot *psbD-psbC* mRNA classes I, II, and III, respectively. Both class 4 and class I arise from the LRP. Class 6 and class II mRNAs arise by way of RNA processing (however, see Rogers et al., 1990), and class 5 and class III arise from a promoter internal to *psbD*. The entire ensemble consisting of the *psbD-psbC-orf62-trnG* coding loci, LRP, RNA processing site, and *psbC* promoter internal to *psbD* appears to have been retained in cereals. However, three main differences between cereals and dicots are evident. First, the LRP is ~400 bp closer to *psbD* in cereals than dicots, an indication that a ~400-bp deletion in cereals, or an insertion in dicots, has occurred between the LRP and *psbD* during evolution. Second, the light response was greater in cereals than dicots. Therefore, the light response has been further enhanced in cereals or diminished in dicots.

The levels of the light-induced mRNAs in dicots did not appear to comprise as large a portion of the total mRNA population as in monocots. However, the dicot light-induced mRNAs can indirectly contribute to the total *psbD-psbC* mRNA population by being processed to the smaller mRNAs corresponding to the 132 to 144-bp signals, which are not abundant in monocots (class 6 mRNAs). Third, the level of class 6 cereal RNAs relative to other *psbD-psbC* mRNAs was much less than the levels of the dicot counterpart class II RNAs. On the one hand, because the dicot RNA 5′ ends localized at 132, 135, and 144 are derived from RNA processing, this could be due to differences in RNA processing between cereals and dicots at this site or due to differences in RNA stability. On the other hand, the low abundance in cereals could be evidence that cereals have decreased use of this RNA 5′ end. The presence of new upstream promoters (classes 1 and 3) and a processing site (class 2) in cereals may obviate any selective advantage for retaining the class 6 mRNA processing site. There is one complication with the interpretation of the origin of the class II RNAs in dark-grown tobacco (Figure 4). In these seedlings, no dark-predominant transcript has been identified that would serve as the candidate precursor for the class II RNAs. However, evidence exists in spinach for a dark-predominant RNA that could serve in this role as observed comigrating with the 281-bp marker in Figure 5. Furthermore, additional promoters could exist immediately upstream from the dicot *psbD* gene as identified in vitro by Rogers et al. (1990). It still remains to be clarified whether the origin of the dicot class II RNAs occurs from a *psbD* proximal promoter followed by rapid processing or by processing of the larger RNAs, which have been observed in spinach and pea.

Cereal mRNAs in class 1, 2, and 3 are interpreted to have been added to *psbD-psbC* by recombination during cereal chloroplast genome evolution from dicots. They are specific to cereals, particularly barley and wheat. Class 1 mRNAs arise from a promoter upstream from *psbK* and have also been mapped in sorghum and maize (Schrubar et al., 1991). The cereal class 1 promoter and class 2 processing site are suggested to be homologous to the dicot class A promoter and class B processing site. The *psbK* promoter region, which is 65% conserved at the nucleotide sequence level between barley and tobacco (Meng et al., 1991), is interpreted to have been retained in spite of the DNA inversions that have occurred. This promoter would then be in position to promote transcription of class 1 cereal mRNAs, which span the entire *psbD-psbC* operon. Hence, what was formerly a promoter driving the ancestral tricistronic operon, *psbK-psbI-trnG*, now promotes synthesis of a unique pentacistronic transcript, *psbK-psbI-psbD-psbC-orf62*.

Although class 1 mRNAs were not searched for in maize and sorghum, class 2 RNAs were detected in this work. They arise by processing of the class 1 mRNAs (Berends-Sexton et al., 1990b). This class 2 processing event is presumed to result from moving a processing site located between *psbI* and *trnG* (possibly at the *trnS* locus in Figure 7) in dicots (Meng et al., 1991) to the *psbI-psbD* intergenic region of cereals (Berends-Sexton et al., 1990b). The abundance of class 3 mRNA is more variable among cereal *psbD-psbC* mRNA classes with low levels in maize, rice, and sorghum. These plant species may not have a strong promoter for class 3 RNAs. Class 3 cereal mRNAs could be analogous to the spinach mRNAs with 5′ ends that map to a promoter at −1080 from *psbD* translational ATG start site (Figure 4). RNAs analogous to the spinach −1080 mRNAs are not detected in pea or tobacco and are, therefore, also variable in dicots.

Proposed Function of Light-Induced Transcripts Encoding Photosynthetic Polypeptides

By their overlapping nature, the *psbD-psbC* transcripts are redundant in protein coding capacity. The redundancy of transcripts may reflect distinct roles for these mRNAs under

changing light conditions and specific chloroplast developmental states. In barley, the nonlight-responsive *psb*D-*psb*C mRNAs accumulate during early chloroplast biogenesis and then decline once high levels of PSII complexes are established in mature chloroplasts (Berends-Sexton et al., 1990b). These mRNAs provide templates for initial construction of PSII complexes during chloroplast biogenesis (for review, see Mullet, 1988). In dicots such as pea, spinach, and tobacco and cereals such as sorghum and maize, which exhibit light-stimulated leaf and chloroplast development, the accumulation of all of the *psb*D-*psb*C mRNAs, including those involved in PSII complex buildup, show some light modulation. This is because leaf and chloroplast biogenesis in these plants is modulated by red light acting through the photoreceptor phytochrome (Smith and Whitelam, 1990; Chory, 1991).

In contrast, activation of *psb*D-*psb*C mRNA accumulation occurs through high fluence blue light stimulation of the LRP for *psb*D-*psb*C. This activation is proposed to be based on the physiological need to replace D2 proteins that are damaged and turned over as a consequence of photochemistry, especially at high light intensity (Schuster et al., 1988; Shipton and Barber, 1991). We suggest that one of the molecular constraints that preserved the linkage of the LRP with the *psb*D gene is the physiological need to replace D2 subunits in mature photosynthetically active chloroplasts. In this view, light-induced transcription from the LRP of *psb*D-*psb*C is suggested to maintain *psb*D mRNA levels required for D2 synthesis in mature chloroplasts (Berends-Sexton et al., 1990b) and does not appear to be due to the general increase in plastid transcription that occurs early during chloroplast biogenesis (Baumgartner et al., 1989).

Despite genomic rearrangements involving the *psb*D-*psb*C loci, the LRP was found to be part of *psb*D regulation in all angiosperms studied. Cyanobacteria also have a light-regulated *psb*D gene, which gives rise to light-induced *psb*D mRNAs (Bustos and Golden, 1992). Therefore, the widespread characteristic of having a light-induced *psb*D gene would appear to be functionally relevant to be preserved in such evolutionarily diverse organisms. However, the LRP–*psb*D linkage may have been preserved simply due to the close physical proximity of the LRP upstream from *psb*D. No recombination events would have then separated them during evolution. Yet, this is not the case in liverwort chloroplast DNA, where the LRP is not found. Liverwort also lacks the ability to grow in high fluence light (Ohyama et al., 1983). One interpretation is that liverwort does not grow in light environments that would damage D2 and select for the LRP.

METHODS

Plant Growth

Barley (*Hordeum vulgare* var Morex), wheat (*Triticum aestivum* var Tam108 seed, kindly provided by Dale Williams, Texas A & M University Foundation Seed Service), sorghum (*Sorghum bicolor* var Moench), maize (*Zea mays* var Conlee), pea (*Pisum sativum* var Little Marvel), and spinach (*Spinacea oleracea* var Bloomsdale Longstanding) were planted in vermiculite watered with half-strength Hoagland nutrient solution. Seedlings were grown at 23°C in a light-tight controlled environment chamber in a light-tight room. After 4.5, 5.0, and 7.5 days, the dark-grown seedlings were either harvested or transferred to a continuously illuminated chamber (fluorescent plus incandescent bulbs, light intensity 350 μE cm^{-2} sec^{-1}) for an additional 4, 24, 48, 72, or 96 hr before harvesting. Tobacco (*Nicotiana tabacum* var Xanthi) was planted as above, germinated, and grown under complete darkness or continuous illumination (350 μE cm^{-2} sec^{-1}) for 15 days and then transferred to a light-tight chamber for 3 days. Primary leaves from these previously illuminated, dark-adapted tobacco plants were then either harvested or returned to illumination for 48 hr before harvesting. Rice (*Oryza sativa* var Lamont) was planted according to Ayers et al. (1991). Seedlings were maintained at 28°C for 10 days in light-tight conditions and then either harvested or transferred to continuous illumination (350 μE cm^{-2} sec^{-1}) for 4 or 72 hr before harvesting. Manipulations of dark-grown plants were performed in complete darkness, or, when necessary, a green safelight was employed as described (Klein and Mullet, 1987).

Isolation of Plastids

Plastids were isolated from seedlings by centrifugation of cell lysates on percoll gradients (35%/75%) as previously described (Klein and Mullet, 1987). Intact plastids were quantitated (plastids per microliter) by counting a dilute suspension in a hemacytometer.

Plastid and Total Cell Nucleic Acid Isolation

Plastid nucleic acid was isolated on a per plastid basis by extraction of intact plastids with phenol, as described previously (Mullet et al., 1985). Total nucleic acid from leaf tissue was isolated according to Mullet and Klein (1987). DNA was removed from plastid and total cell nucleic acid samples by digestion with RNase-free DNase I (Promega).

Primer Extension Analysis

The following purified oligodeoxynucleotide primers (synthesized at Texas A & M University Oligonucleotide Synthesis Services) were utilized in primer extension analysis experiments: (1) 5′-GTCATAGTG-ATCCTCCTATTC, complementary to nucleotide positions −16 to +5 (Efimov et al., 1988) of the RNA-like strand of the *psb*D gene. This primer detects barley class 2, 3, and 4, and 6 RNAs 5′ ends (Figure 1), but is too distant to detect class 1 and is upstream from class 5 RNAs; (2) 5′-GCGTGGTCTAAGTCTAAGG, complementary to positions 5971 to 5989 of the RNA-like strand of the barley *psb*I-*psb*D intergenic region (Berends-Sexton et al., 1990a); (3) 5′-GAATTCAGCATTATT-CCAAG, complementary to positions 6226 to 6245 of the RNA-like strand of the barley *psb*I-*psb*D intergenic region (Berends-Sexton et al., 1990a); (4) 5′-GACCACAAGGATTTAATTGG, complementary to positions 10,818 to 10,837 of the RNA-like strand upstream from spinach *psb*D (Offerman, 1988); (5) 5′-CTTGTGGAATCCAAAAC, complementary to positions 33,613 to 33,629 of the RNA-like strand upstream of tobacco *psb*D (Shinozaki et al., 1986); (6) 5′-GACATAAGTCCCTCCCTAC, complementary to positions 57,573 to 57,591 of the RNA-like strand of tobacco *rbc*L (Shinozaki et al., 1986), which is a conserved site in several *rbc*L loci (Crossland et al., 1984); and (7) 5′-CAACAGATT-

CCTGCAGAGTTAG, complementary to positions 877 to 898 nucleotides upstream from the pea *psbD* translational initiator ATG codon (D.A. Christopher and J.E. Mullet, unpublished results). Primers (30 ng) were labeled at the 5′ end using T4 polynucleotide kinase and γ-^{32}P-ATP. A total of 10^5 dpm of ^{32}P-labeled primer was co-precipitated with either chloroplast RNA (2.05×10^6 plastids/reaction) or with total cellular RNA (12 to 14 μg). The annealing and primer extension reactions were conducted as described in Christopher and Hallick (1989), except final deoxynucleotide triphosphate concentration was 400 μM.

PCR Amplification of Light Response Regions, Molecular Cloning, and DNA Sequence Analysis of PCR Products

The pair of oligodeoxynucleotide primers used for polymerase chain reaction (PCR) with Taq DNA polymerase (Perkin-Elmer Cetus) were 5′-ATCGGATCCACCATAAAATTG (which is the RNA-like sequence, positions 5840 to 5860 [Berends-Sexton et al., 1990a], engineered BamHI site is underlined) and 5′-CGTGGTCTAAGCTTAAGG (which is the cDNA-like sequence, positions 5971 to 5988 [Berends-Sexton et al., 1990a], engineered HindIII site is underlined). When annealed to maize, sorghum, barley, wheat, or rice chloroplast DNA, the primers flank 148 bp of DNA sequence flanking the 5′ ends of the light-induced RNAs. The primers were thus designed to amplify the light response regions from maize and sorghum. PCRs contained 1 and 5 μg of chloroplast DNA (maize) or 2 and 10 μg of total cellular DNA (sorghum; generous gift of Dr. Sujata Pammi), 500 ng of each primer, and 300 μM each of deoxynucleotide triphosphate. PCR amplification consisted of 1.5 min denaturation at 95°C, 2.0 min annealing at 40°C, and 1 min elongation by Taq DNA polymerase at 72°C for 40 cycles. Taq DNA polymerase was added at 72°C during the first PCR cycle after the initial annealing step. Additional PCR amplifications were conducted as above except with an annealing temperature of 52°C. Control PCR amplifications contained previously sequenced plasmids of barley chloroplast DNA as templates (500 ng).

The BamHI- and HindIII-digested PCR products were purified by electroelution and were ligated into BamHI and HindIII sites of pBS+ (pBluescribe; Stratagene). A separate ligation reaction was conducted for DNA fragments derived from each of the different PCR conditions above. The resulting plasmids were designated pZmLRP1 (maize), pSbLRP1 (sorghum), and pHvLRP1 (barley). A minimum of 12 recombinant plasmid DNAs with inserts derived from three separate PCR reactions were sequenced by the dideoxynucleotide chain termination method (Sanger et al., 1977). Barley plasmid DNAs were sequenced as controls for assessing base misincorporation. The plasmid pP9 contains a 6.5-kb pea chloroplast DNA pstI fragment which encodes the *psbD* and its upstream region. The upstream region was sequenced as described above. Sequence data were analyzed on a VAX computer using the DNA analysis program from IntelliGenetics, Co. (Madison, WI).

ACKNOWLEDGMENTS

We thank Dr. Jeff Palmer for helpful suggestions and critical review of the manuscript; Sharyll Pressley for expert typing of references and revisions; and Jeff Rapp, Jay DuBell, Jung-Mook Kim, and Bing-Yuan Meng for fruitful discussions. This work was supported by Grant No. 5-R01 GM37987 from the National Institutes of Health (to J.E.M.).

Received February 25, 1992; accepted May 11, 1992.

REFERENCES

Ayers, N.M., Stansel, J.W., and Park, W.D. (1991). An improved method for transferring tissue culture plants to the greenhouse. Rice BioTechnol. Q. **7**, 5.

Baumgartner, B.J., Rapp, J.C., and Mullet, J.E. (1989). Plastid transcription activity and DNA copy number increase early in barley chloroplast development. Plant Physiol. **89**, 1011–1018.

Berends-Sexton, T., Jones, J.T., and Mullet, J.E. (1990a). Sequence and transcriptional analysis of the barley ctDNA region upstream of *psbD-psbC* encoding *trnK*(UUU), *rsp16*, *trnQ*(UUG), *psbK*, *psbI* and *trnS*(GCU). Curr. Genet. **17**, 445–454.

Berends-Sexton, T., Christopher, D.A., and Mullet, J.E. (1990b). Light-induced switch in barley *psbD-psbC* promoter utilization: A novel mechanism regulating chloroplast gene expression. EMBO J. **9**, 4485–4494.

Bustos, S.A., and Golden, S.S. (1992). Light-regulated expression of the *psbD* gene family in the *Synechococcus* sp. strain 7942: Evidence for the role of duplicated *psbD* genes in cyanobacteria. Mol. Gen. Genet. **232**, 221–230.

Chase, M.W., and Palmer, J.D. (1989). Chloroplast DNA systematics of liliod monocots: Resources, feasibility and an example from the orchidaceae. Am. J. Bot. **76**, 1720–1730.

Chory, J. (1991). Light signals in leaf and chloroplast development: Photoreceptors and downstream responses in search of a transduction pathway. New Biol. **3**, 538–548.

Christopher, D.A., and Hallick, R.B. (1989). *Euglena gracilis* chloroplast ribosomal protein operon: A new chloroplast gene for ribosomal protein L5 and description of a novel organelle intron category designated group III. Nucl. Acids Res. **17**, 7591–7608.

Crossland, L.D., Rodermal, S.R., and Bogorad, L. (1984). Single gene for the large subunit of ribulosebisphosphate carboxylase in maize yields two differentially regulated mRNAs. Proc. Natl. Acad. Sci. USA **81**, 4060–4064.

Danon, A., and Mayfield, S.P.Y. (1991). Light-regulated translational activators: Identification of chloroplast gene-specific mRNA binding proteins. EMBO J. **10**, 3993–4001.

DeHeij, H.T., Lustig, H., Moeskops, D.-J.M., Bovenberg, W.A., Bisanz, C., and Grott, G.S.P. (1983). Chloroplast DNAs of *Spinacia*, *Petunia*, and *Spirodela* have a similar gene organization. Curr. Genet. **7**, 1–6.

Dobley, J., Durbin, M., Golenberg, E.M., Klegg, M.J., and Ma, D.P. (1990). Evolutionary analysis of the large subunit of carboxylase (*rbcL*) nucleotide sequence among the grasses (Gramineae). Evolution **44**, 1097–1108.

Doyle, J.J., Davis, J.I., Soreng, R.J., Garvin, D., and Anderson, M.J. (1992). Chloroplast DNA inversions and the origin of the grass family (Poaceae). Proc. Natl. Acad. Sci. USA, in press.

Efimov, V.A., Andreeva, A.V., Reverdatto, S.V., and Chakhmakhcheva, O.G. (1988). Nucleotide sequence of the barley chloroplast *psbD* gene for the D2 protein of photosystem II. Nucl. Acids Res. **16**, 5686.

Eisermann, A., Tiller, K., and Link, G. (1990). In vitro transcription and DNA binding characteristics of chloroplast and etioplast extracts

from mustard (*Sinapis alba*) indicate differential usage of the *psb*A promoter. EMBO J. **9**, 3981–3987.

Fromm, H., Devic, M., Fluhr, R., and Edelman, M. (1985). Control of *psb*A gene expression: In mature *Spirodela* chloroplasts light regulation of 32-kd protein synthesis is independent from transcript level. EMBO J. **4**, 291–295.

Gamble, P.E., and Mullet, J.E. (1989). Blue light regulates the accumulation of two *psb*D-*psb*C transcripts in barley chloroplasts. EMBO J. **8**, 2785–2794.

Gamble, P.E., Sexton, T.B., and Mullet, J.E. (1988). Light-dependent changes in *psb*D and *psb*C transcripts of barley chloroplasts: Accumulation of two transcripts maintains *psb*D and *psb*C translation capability in mature chloroplasts. EMBO J. **7**, 1289–1297.

Gantt, J.S., Baldauf, S.L., Calie, P.J., Weeder, N.F., and Palmer, J.D. (1991). Transfer of *rpl*22 to the nucleus greatly preceded its loss from the chloroplast and involved gain of an intron. EMBO J. **10**, 3073–3078.

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

Greenberg, B.M., Gaba, V., Canani, O., Malkin, S., Mattoo, A.K., and Edelman, M. (1989). Separate photosensitizers mediate degradation of the 32-kDa photosystem II reaction center protein in the visible and UV spectral regions. Proc. Natl. Acad. Sci. USA **86**, 6617–6620.

Gruissem, W., and Zurawski, G. (1985). Analysis of promoter regions of the spinach chloroplast *rbc*L, *atp*B and *psb*A genes. EMBO J. **4**, 3375–3383.

Gruissem, W., Barkan, A., Deng, X.-W., and Stern, D. (1988). Transcriptional and post-transcriptional control of plastid mRNA levels in higher plants. Trends Genet. **4**, 758–763.

Haley, J., and Bogorad, L. (1990). Alternative promoters are used for genes within maize chloroplast polycistronic transcription units. Plant Cell **2**, 323–333.

Hallick, R.B. (1989). Proposals for the naming of chloroplast genes. II. Update to the nomenclature of genes for thylakoid membrane polypeptides. Plant Mol. Biol. Reporter **7**, 266–275.

Hanley-Bowdoin, L., and Chua, N.-H. (1987). Chloroplast promoters. Trends Biochem. **12**, 67–70.

Harley, L.B., Lawrie, J., Boyer, H.W., and Hedgpeth, J. (1990). Reiterative copying by *E. coli* RNA polymerase during transcription initiation of mutant pBR322 *tet* promoters. Nucl. Acids Res. **18**, 547–552.

Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.-R., Meng, B.-Y., Li, Y.-Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K., and Sugiura, M. (1989). The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. Mol. Gen. Genet. **217**, 185–194.

Howe, C.J., Barker, R.F., Bowman, C.M., and Dyer, T.A. (1988). Common features of three inversions in wheat chloroplast DNA. Curr. Genet. **13**, 343–349.

Kim, J., Gamble-Klein, P., and Mullet, J.E. (1991). Ribosomes pause at specific sites during synthesis of membrane-bound chloroplast reaction center protein D1. J. Biol. Chem. **266**, 14931–14938.

Klein, R. R., and Mullet, J.E. (1987). Control of gene expression during higher plant chloroplast biogenesis: Protein synthesis and transcript levels of *psb*A, *psa*A-*psa*B, and *rbc*L in dark-grown and illuminated barley seedlings. J. Biol. Chem. **262**, 4341–4348.

Klein, R.R., and Mullet, J.E. (1990). Light-induced transcription of chloroplast genes: *psb*A transcription is differentially enhanced in illuminated barley. J. Biol. Chem. **5**, 1895–1902.

Manzara, T., Carrasco, P., and Gruissem, W. (1991). Developmental and organ-specific changes in promoter DNA-protein interactions in the tomato *rbc*S gene family. Plant Cell **3**, 1305–1316.

Mathews, D.E., and Durbin, R.D. (1990). Tagetitoxin inhibits RNA synthesis directed by RNA polymerases from chloroplasts and *Escherichia coli*. J. Biol. Chem. **265**, 493–498.

Mattoo, A.K., Hoffman-Falk, H., Marder, J.B., and Edelman, M. (1984). Regulation of protein metabolism: Coupling of photosynthetic electron transport to in vivo degradation of the rapidly metabolized 32-kilodalton protein of the chloroplast membrane. Proc. Natl. Acad. Sci. USA **81**, 1380–1384.

Mattoo, A.K., Marder, J.B., and Edelman, M. (1989). Dynamics of photosystem II reaction center. Cell **56**, 241–246.

Meng, B.-Y., Wakasugi, T., and Sugiura, M. (1991). Two promoters within the *psb*K-*psb*I-*trn*G gene cluster in tobacco chloroplast DNA. Curr. Genet. **20**, 259–264.

Mullet, J.E. (1988). Chloroplast development and gene expression. Annu. Rev. Plant Physiol. Plant Mol. Biol. **39**, 475–502.

Mullet, J.E., and Klein, R.R. (1987). Transcription and RNA stability are important determinants of higher plant chloroplast RNA levels. EMBO J. **6**, 1571–1579.

Mullet, J.E., Orozco, E.M., and Chua, N.-H. (1985). Multiple transcripts for higher plant *rbc*L and *atp*B genes and localization of the transcription initiation site of the *rbc*L gene. Plant Mol. Biol. **4**, 39–54.

Mullet, J.E., Gamble-Klein, P.E., and Klein, R.R. (1990). Chlorophyll regulates accumulation of the plastid-encoded chlorophyll apoproteins CP43 and D1 by increasing apoprotein stability. Proc. Natl. Acad. Sci. USA **87**, 4038–4042.

Newton, K.J. (1988). Plant mitochondrial genomes. Organization, expression and variation. Annu. Rev. Plant Physiol. Plant Mol. Biol. **39**, 503–532.

Offerman, K. (1988). Studien zu struktur und expression polycistronischer transkriptionseinheiten für plastomkodierte thylakoidmembranproteine. PhD Dissertation (Düsseldorf, Federal Republic of Germany: University of Düsseldorf), pp. 66–71.

Ohad, I., Kyle, D.J., and Hirschberg, J. (1985). Light-dependent degradation of the QB-protein in isolated pea thylakoids. EMBO J. **4**, 1655–1659.

Ohyama, K., Yamano, Y., Fukuzawa, H., Komano, T., Yamagishi, H., Fujimoto, S., and Sagiura, M. (1983). Physical mappings of chloroplast DNA from liverwort *Marchantia polymorpha* L. cell suspension cultures. Mol. Gen. Genet. **189**, 1–6.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., and Ozeki, H. (1986). Complete nucleotide sequence of liverwort, *Marchantia polymorpha* chloroplast DNA. Plant Mol. Biol. Rep. **4**, 150–175.

Palmer, J.D. (1990). Contrasting modes and tempos of genome evolution in land plant organelles. Trends Genet. **6**, 115–120.

Piechulla, B., Kellmann, J.-W., Pichersky, E., Schwartz, E., and Förster, H.-H. (1991). Determination of steady-state mRNA levels of individual chlorophyll *a*/*b* binding proteins of the tomato *cab* gene family. Mol. Gen. Genet. **230**, 413–422.

Rasmussen, O.F., Bookjans, G., Stummann, B.M., and Henningson, K.W. (1984). Localization and sequence of the gene for the mem-

brane polypeptide D2 from pea chloroplast DNAs. Plant Mol. Biol. **3**, 191–199.

Rogers, S.A., Chen, L.-J., and Orozco, E.M. (1990). Analysis of chloroplast promoters using bidirectional transcription vectors. Plant Mol. Biol. **15**, 421–435.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**, 5463–5467.

Schrubar, H., Warner, G., and Westhoff, P. (1991). Transcriptional control of plastid gene expression in greening sorghum seedlings. Planta **183**, 101–111.

Schuster, G., Timberg, R., and Ohad, I. (1988). Turnover of thylakoid photosystem II proteins during photoinhibition of *Chlamydomonas reinhardtii.* Eur. J. Biochem. **177**, 403–410.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B.Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., and Sugiura, M. (1986). The complete nucleotide sequence of the tobacco chloroplast genome. Plant Mol. Biol. Rep. **4**, 112–147.

Shipton, C.A., and Barber, J. (1991). Photoinduced degradation of the D1 polypeptide in isolated reaction centers of photosystem II: Evidence for an autoproteolytic process triggered by the oxidizing side of the photosystem. Proc. Natl. Acad. Sci. USA **88**, 6691–6695.

Smith, H., and Whitelam, G.C. (1990). Phytochrome, a family of photoreceptors with multiple physiological roles. Plant Cell Environ. **13**, 695–707.

Stern, D.B., Jones, H., and Gruissem, W. (1989). Function of plastid mRNA 3′ inverted repeats: RNA stabilization and gene-specific protein binding. J. Biol. Chem. **264**, 18742–18750.

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

Woodbury, N.W., Dobres, M., and Thompson, W.F. (1989). The identification and localization of 33 pea chloroplast transcription initiation sites. Curr. Genet. **16**, 433–445.

Yao, W.B., Meng, B.Y., Tanaka, M., and Sugiura, M. (1989). An additional promoter within the protein-coding region of the *psb*D-*psb*C gene cluster in tobacco chloroplast DNA. Nucl. Acids Res. **17**, 9583–9591.