The Plant Cell, Vol. 1, 141–150, January 1989, © 1989 American Society of Plant Physiologists

Multiple *cis* Regulatory Elements for Maximal Expression of the Cauliflower Mosaic Virus 35S Promoter in Transgenic Plants

Rong-Xiang Fang,[1] Ferenc Nagy,[2] Shanthi Sivasubramaniam,[3] and Nam-Hai Chua[4]

Laboratory of Plant Molecular Biology, The Rockefeller University, 1230 York Avenue, New York, New York 10021-6399

The 35S promoter is a major promoter of the cauliflower mosaic virus that infects crucifers. This promoter is still active when excised from cauliflower mosaic virus and integrated into the nuclear genome of transgenic tobacco. Previous work has shown that the −343 to −46 upstream fragment is responsible for the majority of the 35S promoter strength (Odell, J.T., Nagy, F., and Chua, N.-H. [1985]. Nature 313, 810-812). Here we show by 5′, 3′, and internal deletions that this upstream fragment can be subdivided into three functional regions, −343 to −208, −208 to −90, and −90 to −46. The first two regions can potentiate transcriptional activity when tested with the appropriate 35S promoter sequence. In contrast, the −90 to −46 region by itself has little activity but it plays an accessory role by increasing transcriptional activity of the two distal regions. Finally, we show that monomers and multimers of a 35S fragment (−209 to −46) can act as enhancers to potentiate transcription from a heterologous promoter.

INTRODUCTION

Cauliflower mosaic virus (CaMV) is a DNA virus that infects members of the *Cruciferae*. The virus is approximately 8 kb long and its complete nucleotide sequence has been determined (cf. Hohn et al., 1982). Transcript mapping experiments have identified two viral promoters, designated 19S and 35S. During the virus life cycle, the 35S promoter is transcribed from the viral DNA minus-strand to produce an 8-kb transcript referred to as the 35S RNA (Guilley et al., 1982). The 5′ and 3′ termini of this RNA have been determined by S1 nuclease analysis and shown to have an overlapping sequence of about 200 nucleotides (Covey et al., 1981; Guilley, et al., 1982).

In addition to serving as a template for translation, the 35S RNA, with its direct terminal repeat sequence, also functions as an intermediate for viral DNA synthesis through a reverse transcription process (cf. Pfeiffer and Hohn, 1983). Rapid viral replication in infected cells requires a copious supply of the 35S RNA. Therefore, it is not surprising that this RNA is a major transcript in infected cells (Guilley et al., 1982). The high level production of the 35S RNA is evidence of the strength of this promoter.

[1] Current address: Department of Virology, Institute of Microbiology, Academia Sinica, Beijing, People's Republic of China.
[2] Current address: Institute of Plant Physiology, Hungarian Academy of Sciences, P.O. Box 521, 6701 Szeged, Hungary.
[3] Current address: Institute of Molecular and Cell Biology, National University of Singapore, Kent Ridge, Republic of Singapore 0511.
[4] To whom correspondence should be addressed.

Information on the 35S transcriptional start site as well as the availability of cloned CaMV sequences have encouraged the use of the 35S promoter as a model system for investigation of plant gene expression (cf. Kuhlemeier et al., 1987a). Thus, CaMV fragments containing 400 to 1000 base pairs of 35S upstream sequences have been shown to be active when integrated into the nuclear genome of transgenic tobacco (Odell et al., 1985, 1987; Nagy et al., 1987; Kay et al., 1987; Jefferson et al., 1987) and petunia (Sanders et al., 1987). Moreover, the promoter can also be expressed transiently in protoplasts of several dicots and monocots (On-Lee et al., 1986; Fromm et al., 1985; Ow et al., 1987; Nagata et al., 1987; Odell et al., 1988). Quantitative measurements of relative transcript levels in transformed tobacco cells (Morelli et al., 1985; Nagy et al., 1985) or transgenic petunia plants (Sanders et al., 1987) showed that the 35S promoter is at least 30 times stronger than the nopaline synthase promoter. The strength of the 35S promoter accounts for its widespread use for high level expression of desirable traits in transgenic plants (e.g., Hemenway et al., 1988; Cuozzo et al., 1988).

We have used transgenic tobacco previously as an expression system to compare transcript levels of four 5′ deletion mutants of the 35S promoter (Odell et al., 1985; Nagy et al., 1985; Odell et al., 1987). We found that deletion of the −343 to −105 region leads to a 75% decrease in promoter strength. Extending this observation,

Kay et al. (1987) showed that a duplication of the −343 to −90 fragment can enhance transcription by 10-fold. DNA sequences for 35S promoter activity have also been studied by transient expression assays in protoplast cultures. Ow et al. (1987) placed the 5′ boundary of the 35S promoter at −148 and showed that deletion of the distal region (−148 to −89) of the promoter decreases activity by 80%. Odell et al. (1988) reported that an upstream fragment, from −393 to −90, can increase the expression level of the nopaline synthase promoter by about threefold. However, increasing the copy number of this activating fragment does not increase the expression level further.

One problem encountered with the quantitative analysis of the 35S promoter is the wide variation in activity among independent transgenic plants (Odell et al., 1987; Sanders et al., 1987). This variation is presumably due to the varying effects of different chromosomal locations on the expression level of the 35S promoter. In this paper, we sought to minimize position effect by including an internal reference gene in our plasmid, as has been done for the analysis of the pea *rbcS-3A* gene (Kuhlemeier et al., 1987b, 1988). With this system we have analyzed the expression level of not only 5′ deletion mutants but also 3′ and internal deletion mutants. Our results show that at least three regions in the −343 to −46 upstream sequences are needed for maximal expression of the 35S promoter.

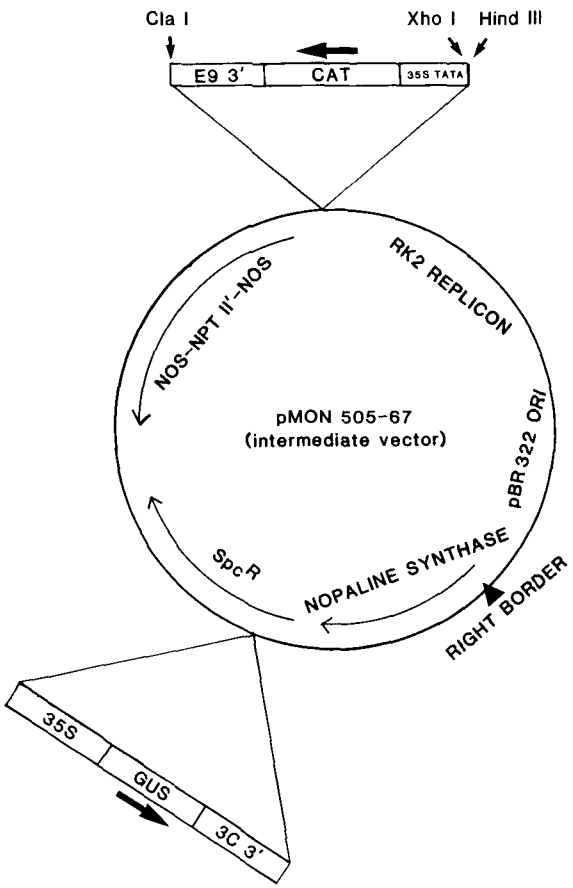

Figure 1. Binary Intermediate Vector pMON505–67 with Test and Reference Genes.

pMON505–67, which is derived from pMON505 (Horsch and Klee, 1986), contains a reference gene unit (35S promoter-GUS gene-*rbcS-3C* 3′ end) inserted at the HpaI site and a basic test gene unit (35S TATA sequence-CAT gene-*rbcS-E9* 3′ end) at the polylinker region between XbaI and ClaI. The 35S promoter of the reference gene is from −941 to +9, whereas the 35S TATA sequence of the test gene is from −46 to +9. HindIII and XhoI upstream of the 35S TATA sequence and ClaI downstream of the *rbcS-E9* 3′ end are unique sites used for insertion of various 35S 5′ upstream sequences.

RESULTS

Test and Reference Gene System

Previous studies have indicated that the activity of the 35S promoter depends on the chromosomal insertion site (Nagy et al., 1985; Sanders et al., 1987; and Odell et al., 1987), and variations up to 50 times among independent transgenic plants have been noted (Sanders et al., 1987; Odell et al., 1987). In the case of the rbcS gene, chromosomal position effect can be minimized by the use of a reference gene that serves as an internal control (Kuhlemeier et al., 1987a). Moreover, the estimation of the test gene activity relative to that of the reference gene also obviates the problem of differences in the transgene copy number in different transgenic plants. In the present work, we have inserted into the intermediate vector pMON505 (Horsch and Klee, 1986), a reference gene unit that is 5 kb away from the test gene. The reference gene is composed of the full-size 35S promoter (−941 to +9) (Odell et al., 1985), the *Escherichia coli* β-glucuronidase (GUS) coding region (Jefferson et al., 1987), and the *rbcS-3C* 3′ sequence. The test gene unit consists of the 35S TATA sequence, from −46 to +9, the chloramphenicol acetyltransferase (CAT) coding sequence, and the 3′ sequence of *rbcS-E9* (Figure 1). The partial sequence homology at the 3′ ends of the two rbcS genes allows us to use a single ^{32}P-labeled probe prepared from the 3′ end of *rbcS-E9* gene to discriminate between the protected signals derived from the test and reference transcripts by S1 nuclease digestion assays (Figure 2b) (Fluhr et al., 1986). The use of a single probe also has the added advantage of eliminating any possible errors due to different efficiencies of labeling and hybridization when two different probes are employed.

Figure 2a shows the results obtained from 10 independent transgenic plants in which the test gene contains the

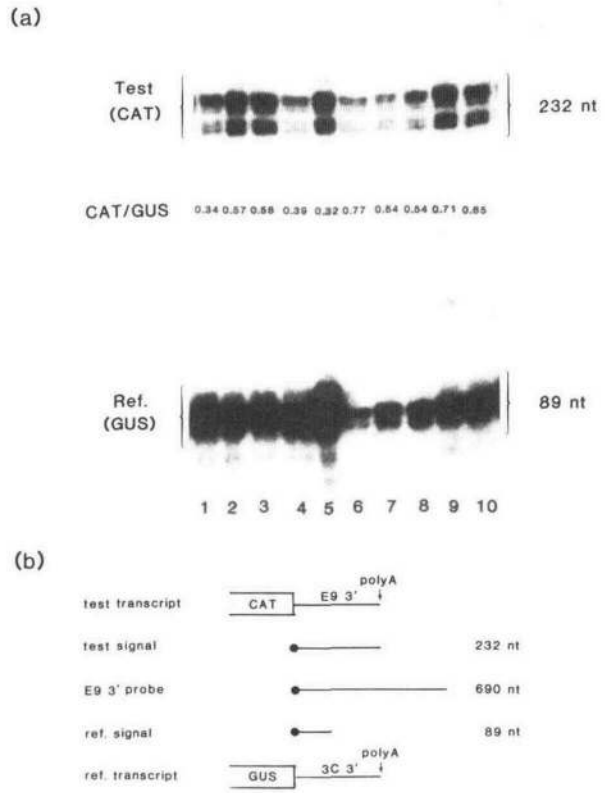

Figure 2. Variation of Transcript Levels among Independent Transgenic Plants.

(a) The test gene contains the −343 to −46 fragment fused to the 35S TATA sequence. Fifty μg of leaf RNA from 10 independent transgenic plants were assayed by 3′ S1 analysis, and the CAT/GUS ratio for each plant was determined as described in "Methods."

(b) Schematic diagram of 3′ S1 analysis. The cluster of bands around 232 bases is derived from the protection of the CAT-E9 chimeric transcripts, whereas bands around 89 bases are from the GUS-3C transcripts. nt, nucleotides.

−343 to −46 fragment fused to the 35S TATA sequence at −46. Whereas the strength of the CAT signal varies by at least five times among the transformants, the CAT/GUS ratios vary by only twofold. In spite of the use of this test-reference system, we found that approximately 25% of the transgenic plants express only one of the two signals either at a very low or undetectable level. This could be due to preferential suppression of either the test or the reference gene at a certain chromosomal location. Alternatively, one of the two genes could be lost selectively during transfer or integration. Therefore, these plants were excluded from our analysis. For each construct, we screened routinely eight to 10 independent transgenic plants by S1 nuclease protection assays. RNA samples from plants that give both signals were pooled and used for the estimation of CAT/GUS ratios.

Activities of 5′ and 3′ Deletion Mutants

To define functional elements of the 35S promoter other than the TATA box, we created 3′ deletions by digesting the −343 to +9 fragment with Bal31 nuclease, and mutants with breakpoints at −46, −78, −107, −127, −157, and −208 were selected (Figures 3 and 4b). The fragment, −343 to −46, was used to generate a series of 5′ deletion mutants with endpoints at −208, −168, −105, and −90 (Figures 3 and 4b). All 3′ and 5′ deletion mutants were inserted into the HindIII and XhoI site in both orientations upstream of the basic test gene unit in pMON505–67 (Figure 1) and transferred into tobacco. The nucleotide sequence of the −343 to +9 fragment and the endpoints of the deletion mutants are shown in Figure 3.

To facilitate comparison, we have expressed the CAT/GUS rations of all the constructs relative to the CAT/GUS ratio of the −343 to −46, F construct which is used as the wild-type (WT) control. Figure 4 shows that 5′ deletion from −343 to −209 decreases the total activity by about 50% and further deletion to −168 reduces the activity by an additional 20%. The −105 mutant has a very low level of CAT transcripts, which makes estimation of the CAT/GUS ratio unreliable. However, direct comparison of the CAT transcript level reveals that this mutant has about 10% of WT activity (data not shown). The −90 and −46 deletion mutants show no detectable CAT transcript. Based on dilution experiments and by varying the exposure time of autoradiograms, we estimated that these mutants retain less than 5% of the WT activity.

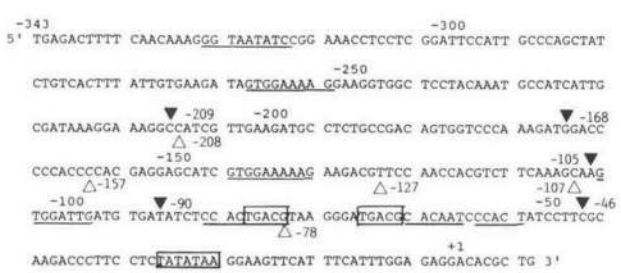

Figure 3. Nucleotide Sequence of the CaMV 35S Promoter and Upstream Region.

The nucleotide sequence from −343 to +9 is shown. The transcriptional start site is designated +1. The 5′ and 3′ deletions are indicated by solid and open arrowheads, respectively. The underlined sequences, GGTAATAC, GTGGAAAAG, GTGGAAAAAG, and GTGGATTG, resemble the enhancer core sequence of SV40 (Serfling et al., 1985). The putative CCAAT sequences are CCACT, CACAAT, and CCACT (Ow et al., 1987). The TATA box sequence and the TGACG repeats are boxed.

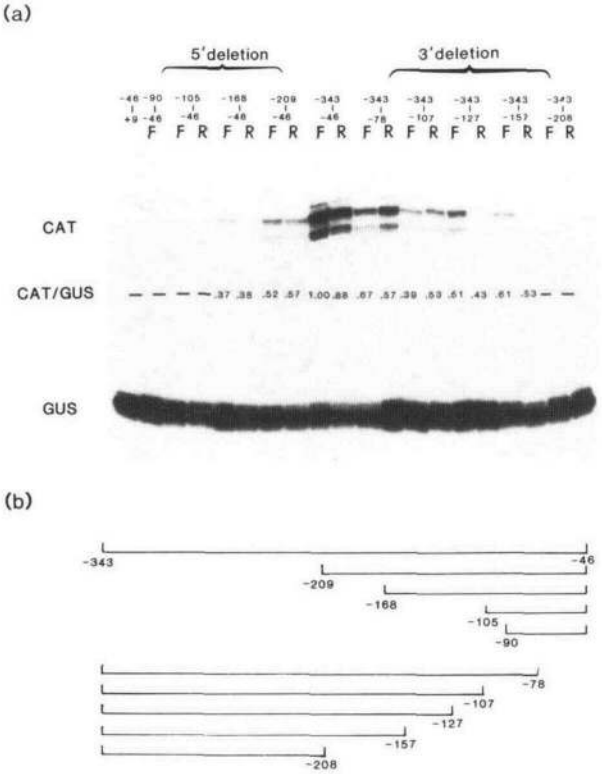

Figure 4. Activities of Various 5′ and 3′ Deletion Mutants of the 35S Promoter.

Each deletion mutant was inserted at position −46 of the test gene unit of pMON505–67 (Figure 1). F and R, forward and reverse orientation, respectively, relative to the TATA box. 3′ S1 analysis and quantitation of the S1 signals were as described in "Methods." To facilitate comparison of the expression levels, the CAT/GUS ratio of the −343 to −46 F construct (lane 9) was assigned a value of 1 and the CAT/GUS ratios of all other constructs were expressed relative to this value. These relative CAT/GUS values were given below the CAT signals in the autoradiogram. Number of independent transgenic plants used for analyses (from left to right) −46 to +9, 8; −90 to −46, 8; −105 to −46 (F), 8; −105 to −46 (R), 4; −168 to −46 (F), 5; −168 to −46 (R), 8; −209 to −46 (F), 8; −209 to −46 (R), 4; −343 to −46 (F), 7; −343 to −46 (R), 6; −343 to −78 (F), 5; −343 to −78 (R), 4; −343 to −107 (F), 8; −343 to −107 (R), 8; −343 to −127 (F), 8; −343 to −127 (R), 8; −343 to −157 (F), 8; −343 to −157 (R), 8; −343 to −208 (F), 4; and −343 to −208 (R), 7.

The 5′ deletion experiments indicate that DNA sequences between −343 and −105 are responsible for a majority of the 35S promoter activity. To investigate the role of sequences downstream of −105, we created a series of 3′ deletion fragments, all with 5′ end points at −343 (Figure 3). These deleted fragments were joined to the 35S TATA sequence at −46 (Figure 1), thus generating a series of internal deletions with 3′ break points at −46 but varying 5′ end points. Figure 4 shows that deletion of the sequence between −46 and −78 reduces the transcriptional activity to about 70% of control levels, and this value decreases slightly when the deletion extends to −157. However, an internal deletion mutant missing the sequence between −46 and −208 has no detectable activity (i.e., less than 5% of WT value).

All 5′ and 3′ deletion fragments give about the same levels of activity when tested in the reverse orientation. These results suggest strongly that the *cis*-regulatory elements contained therein have enhancer-like properties.

The −90 to −46 Region Plays an Accessory Role

Results in the previous section show that, although deletion of the −343 to −208 region reduces transcriptional activity to 50% (Figure 4), the same region by itself displays no activity when fused upstream of the 35S TATA box (Figure 5). This discrepancy can be resolved if the activity of the −343 to −208 region depends on sequences downstream of −208. In several animal genes the activity of certain enhancer elements has been shown to require the "CCAAT" box (Dierks et al., 1983; Bienz and Pelham, 1986). There are three CCAAT-like boxes in the 35S promoter: CCACT (−85 to −81), CACAAT (−64 to −59), and CCACT (−57 to −53) (Ow et al., 1987). Since all three boxes are contained within −90 to −46, we have examined the effect of this region on transcription. Figure 5a shows that the −90 to −46 region gives no detectable transcript when fused to the 35S TATA box (lane 6) and the −343 to −208 region is similarly inactive (lane 4). However, in combination, the two sequences are able to potentiate transcription from the same TATA box (Figure 5a, lane 5), suggesting that they act synergistically. By dilution experiments and by varying the exposure time of autoradiograms, we have estimated that the transcript level in construct 5 is at least 10 times higher than those in construct 4 and 6.

To determine whether the −90 to −46 region can also potentiate the activity of other regulatory sequences, we examined its effect on transcription driven by a longer upstream sequence (−343 to −107), which is active by itself. Figure 5 shows that the transcriptional activity of the −343 to −107 fragment can also be increased moderately by the −90 to −46 region. Note, however, that the elevated transcription seen in construct 2 is only 60% that of construct 1, implying a role for the −107 to −90 region that is missing from the former construct.

The −209 to −46 Region Can Function as an Enhancer Element

The results presented in Figure 4 show that a major part of the 35S promoter activity resides in the −209 to −46

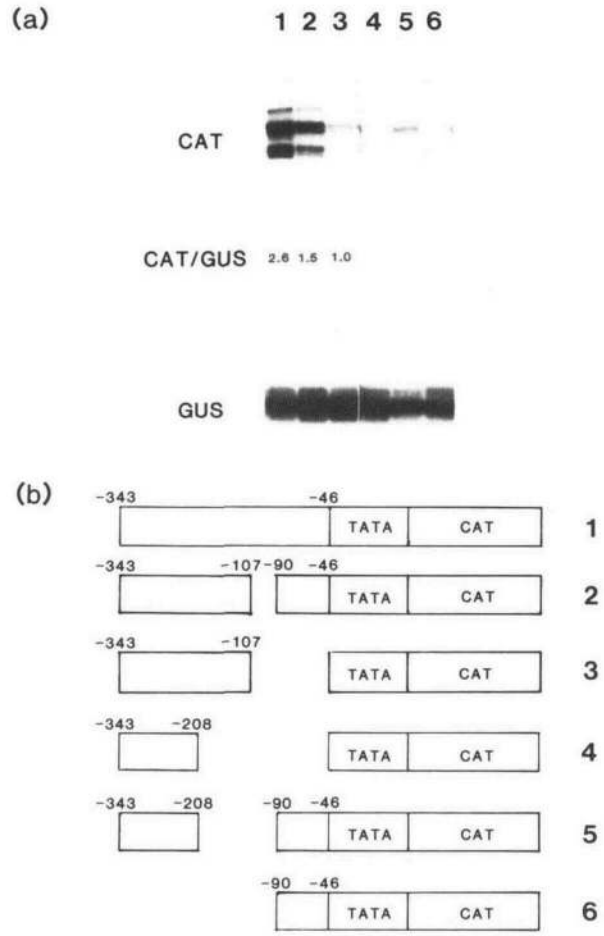

Figure 5. −90 to −46 Region of the 35S Promoter Potentiates Transcriptional Activity of Upstream Sequences.

Structures of the various constructs are shown in **(b)**. Fifty μg of leaf RNA were used for 3′ S1 analyses and the results quantitated as described in "Methods" and the legend to Figure 4. To facilitate comparison of expression levels, the CAT/GUS ratio of construct 3 was assigned a value of 1 and the CAT/GUS ratios of constructs 1 and 2 were expressed relative to this value. Number of independent transgenic plants used for analyses: construct 1, 7; construct 2, 8; construct 3, 8; construct 4, 4; construct 5, 5; and construct 6, 5.

fragment, which still retains its activity when placed in the reverse orientation. This bidirectionality suggests that the −209 to −46 fragment may function as a transcriptional enhancer. To investigate this point, we placed the fragment at the 3′ end of the transcription unit in both the forward and reverse orientation. Figure 6 shows that the fragment can potentiate transcription, albeit weakly, when placed at the 3′ end (lanes 1 and 2). Based on a fivefold longer exposure of the autoradiogram, we estimated that the activity of these constructs (lanes 5 and 6) is approximately 20% that of the controls (lanes 3 and 4).

Transcription driven by the −209 to −46 fragment is about one-half the level of that of the control fragment (−343 to −46) (Figure 4). To determine whether this activity can be increased over the control value, we investigated the effects of increasing the copy number of the −209 to −46 fragment. Figure 7 shows that the enhancing activity of the fragment increases linearly with copy number, and maximum stimulation is achieved with four head-to-tail copies in an orientation-independent manner. Eight copies of the same fragment do not give any further increase in activity.

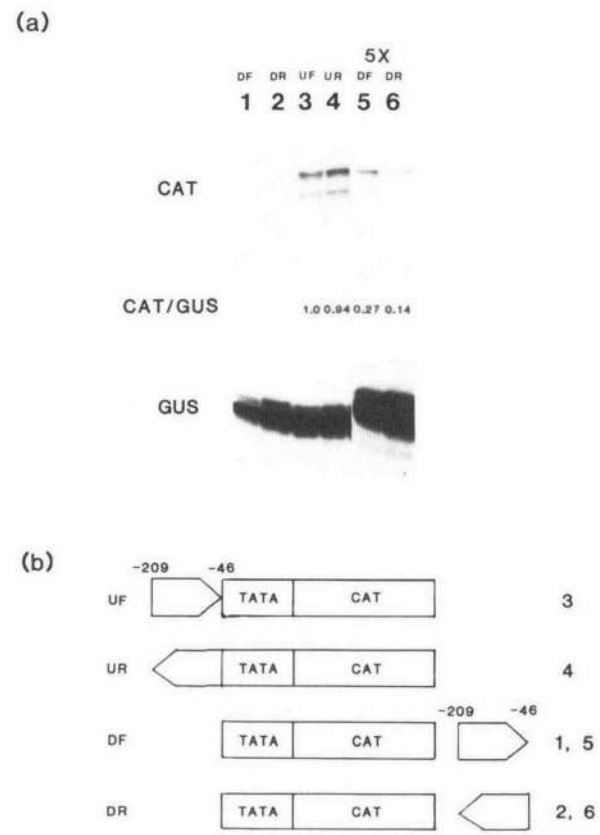

Figure 6. The −209 to −46 Region of the 35S Promoter Can Function as a Transcriptional Enhancer.

The −209 to −46 fragment was inserted in both orientations either upstream (UF or UR) or downstream (DF or DR) of the test gene unit as shown in the schematic diagram in **(b)**. Fifty μg of leaf RNA were used for each lane. Lanes 5 and 6 are lanes 1 and 2, but the autoradiograms were exposed five times longer. The CAT/GUS ratio of construct 3 was given a value of 1 and the CAT/GUS ratios of constructs 1 and 2 were expressed relative to this value. Number of independent transgenic plants used for analyses: construct 1, 8; construct 2, 4; construct 3, 8; and construct 4, 4.

We have shown previously that deletion of the upstream region of the *rbcS-3A* gene from −410 to −50 results in a severe drop of transcription (>20-fold), and little or no transcript can be detected with the −50 deletion mutant (Kuhlemeier et al., 1987a). This mutant serves as a convenient vehicle to test for the potential transcriptional activity of any DNA sequences placed upstream of it. To determine whether the −209 to −46 fragment of 35S can potentiate transcription from a heterologous TATA box, we cloned eight copies of this fragment at −50 of the *rbcS-3A* deletion mutant (Figure 8, construct 1). Figure 8 shows that the octamer increases the *rbcS-3A* transcript to a level even higher than that obtained with the *rbcS-3A* upstream region (cf. lanes 1 and 2).

DISCUSSION

The −343 to −46 Upstream Sequence Is Made Up of at Least Three Functional Regions

Our previous work has shown that the 5′ upstream sequence from −343 to −46 is responsible for the majority of the 35S transcription activity in transgenic tobacco plants (Odell et al., 1985; Nagy et al., 1985). In this paper we have focused on a detailed dissection of this upstream fragment in an attempt to identify sequences that contribute to the activity. For both the 5′ and 3′ deletion series, our results show that there is a gradual loss of activity with progressive deletion. Based on the deletion end points, we can divide the upstream sequence into three functional regions, all of which are needed for maximal 35S activity: (1) −343 to −208. As shown by 5′ deletion analysis, the −343 to −208 region is responsible for about 50% of the total 35S activity. Surprisingly, this 135-bp fragment does not give any detectable activity when fused to the 35S TATA box at −46; its activity can be demonstrated only in the presence of the −90 to −46 region

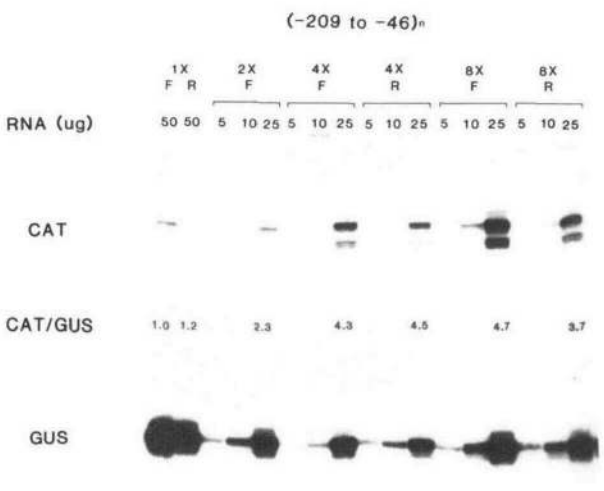

Figure 7. Activity of the −209 to −46 Fragment Increases with Copy Number.

Head-to-tail multimers of the −209 to −46 fragment were fused in the forward (F) or the reverse (R) orientation to −46 of the test gene unit as shown in the lower panel. 1×, 2×, 4×, and 8× are monomer, dimer, tetramer, and octamer, respectively. 3′ S1 analyses and quantitation of results were as described in "Methods." The CAT/GUS ratio of the first lane from the left was given a value of 1 and the CAT/GUS ratios of the other constructs were expressed relative to this value. Number of independent transgenic plants used for analyses: 1 × F, 8; 1 × R, 4; 2 × F, 8; 4 × F, 5; 4 × R, 7; 8 × F, 5; and 8 × R, 5.

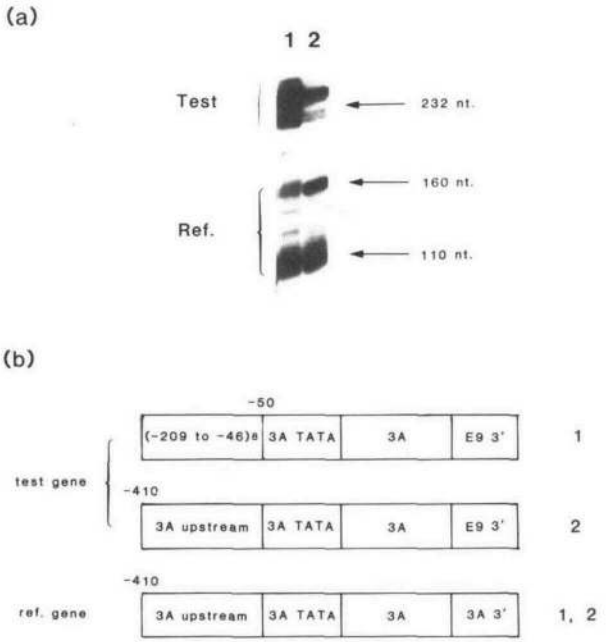

Figure 8. Eight Copies of the −209 to −46 Fragment Can Potentiate Transcription from a Heterologous Promoter.

(a) In lane 1, eight copies of the −209 to −46 fragment were fused to −50 of the truncated *rbcS-3A* test gene which contains the *rbcS-E9* 3′ sequence. An *rbcS-3A* gene with 410 bp of 5′ upstream sequence serves as the reference. Leaf RNAs from two independent transgenic plants were pooled and analyzed (Kuhlemeier et al., 1987). In lane 2, the test gene contains 410 bp of 5′ upstream sequence from the *rbcS-3A* gene. RNAs from five independent transgenic plants were pooled and analyzed. Twenty μg of RNA were used in 3′ S1 mapping using as a probe, a ^{32}P-labeled single-stranded DNA fragment derived from the *rbcS-E9* 3′ sequence (Morelli et al., 1985; Fluhr et al., 1986). Transcripts from the test gene gave protected fragments around 232 nucleotides, whereas transcripts from the reference gene produced signals at 160 and 110 nucleotides.
(b) Schematic diagram of the test and reference genes.

(Figure 5). (2) −208 to −90. The importance of this region has been shown by two independent series of experiments. First, 5′ deletion from −208 to −90 leads to a significant decrease in activity (Figure 4). Second, a severe drop in expression level is observed also when the sequences between −208 and −90 are removed from the WT construct, which has 343 bp of 5′ upstream sequence (Figure 5, cf. lanes 1 and 5). Based on our 5′ deletion results, we estimated that this 118-bp region probably contributes to approximately 40% of the total 35S activity. (3) −90 to −46. Because of the low expression level of the −90 5′ deletion mutant, it was not possible to assign any function to this region from 5′ deletion analysis. Its accessory role in transcription, however, can be uncovered when an upstream element is placed 5′ of it (Figure 5).

Examination of the nucleotide sequence in the −343 to −208 and −208 to −90 regions reveals several motifs similar to the SV40 enhancer core sequence (Figure 3). The role of these motifs, if any, in potentiating transcription of the 35S promoter remains to be established by mutagenesis experiments.

Recently, Ow et al. (1987) reported a detailed 5′ deletion analysis of the 35S promoter using a protoplast transient expression system. They found little change in activity when the upstream sequence is deleted from −1600 to −148 and, accordingly, they fixed the 5′ boundary of the 35S promoter at −148. In contrast, we show here that deletion of the −343 to −208 region reduces transcription activity by about 50% (Figure 4). Moreover, the same region can potentiate transcription when joined to the 35S promoter at −90 (Figure 5). The latter observation is particularly compelling since it demonstrates clearly a positive function for the −343 to −208 region. The discrepancy between our results and those of Ow et al. (1987) is due likely to the differences in the assay systems employed, i.e., transgenic plants versus protoplasts. One possibility is that, in protoplasts, *trans*-acting factors that interact with the −343 to −208 region may have become inactivated. Ebert et al. (1987) have also detected differences between results from transient expression assays and stable expression assays with the same constructs of the nopaline synthase promoter.

The −90 to −46 Region Plays an Accessory Role in Transcription

An interesting point that emerges from our analysis is the role of the −90 to −46 region. Ow et al. (1987) reported that, in carrot protoplasts, the −89 5′ deletion mutant has 23% WT activity, whereas the −68 5′ deletion mutant has only 0.8%. Odell et al. (1988) found that deletion of the sequence between −90 and −55 reduces transient expression levels of the −392 to −90 fragment by twofold. We have confirmed and extended these results in transgenic tobacco plants. We show here that the −90 to −46 region can potentiate the activity of two upstream fragments: −343 to −209 and −343 to −107 (Figure 5). Moreover, a similar region (−105 to −46) of the 35S promoter is needed for expression of a leaf-specific enhancer of the *Nicotiana plumbaginifolia rbcS-8B* gene that is located far upstream (−1038 to −589) (Poulsen and Chua, 1988). Together, these results suggest strongly that the −90 to −46 region performs an accessory role in increasing the transcriptional activity of upstream enhancers. Since the three upstream fragments, −343 to −46, −209 to −46, and −168 to −46 are active equally in either the forward or the reverse orientation, this 44-bp element (−90 to −46) appears to function independent of its orientation with respect to the 35S TATA box (Figure 4).

In animal systems, there is evidence that some enhancers can only function in conjunction with a CAAT box (Dierks et al., 1983; Bienz and Pelham, 1986). Ow et al. (1987) have pointed out three CAAT boxes in this region, CCACT (−85 to −81), CACAAT (−64 to −59), and CCACT (−57 to −53) (Figure 3). In addition to these motifs, there is also a pentanucleotide (TGACG) repeat located between −82 to −78 and −70 to −66 (Figure 3). This pentanucleotide is highly homologous to the hexamer motif identified upstream of histone genes by Iwabuchi and co-workers (Mikami et al., 1987). The relative importance of these sequence motifs to the 35S activity is being evaluated. It should be pointed out that there is no evidence yet that the CAAT box is important for plant gene expression. On the contrary, in the pea *rbcS-E9* gene, the putative CAAT box could be deleted without a negative effect on transcript levels (Morelli et al., 1985).

The −208 to −46 Fragment Can Function as an Enhancer

We have shown here that the 35S upstream fragment, containing sequences between −208 and −46, fulfills the principal criteria for a eukaryotic transcription enhancer (Serfling et al., 1985). This 162-bp fragment can function in an orientation-independent manner when located either upstream or downstream of the 35S transcription unit (Figure 6), and it can activate transcription from a heterologous TATA box (Figure 8). Moreover, there is a near linear increase in transcription activity with multimerization of the enhancer up to four copies (Figure 7). Therefore, this DNA fragment and its multimers can be used to enhance gene expression in transgenic plants. Our results here are similar to those of Kay et al. (1987), who claimed that duplication of the −343 to −90 region results in a 10-fold increase in transcription.

The 35S upstream sequences can activate heterologous promoters not only in transgenic plants but also in protoplasts. Whereas Ow et al. (1985) reported an increase in activity with multimerization of the distal regions (−148 to −90), Odell et al. (1988) obtained the same activity with either one or two copies of their activating fragment (−338 to −55).

Conclusions

Our investigations of the 35S upstream fragment (−343 to −46) indicate that maximal expression requires the cooperation of multiple *cis*-regulatory elements. These results are reminiscent of those obtained with simian virus 40 (Zenke et al., 1986; Schaffner et al., 1988). It has been shown that the SV40 enhancer is composed of several regulatory elements, each with a distinct cell specificity (Schirm et al., 1987; Ondek et al., 1987). Currently, we are examining whether this is also the case with the CaMV 35S promoter.

METHODS

Deletion Mutants of the CaMV 35S Promoter

We used the CaMV 35S 5′ regions from −343 to +9 subcloned in pUC13 as a ClaI/HindIII fragment to generate a series of 3′ deletion mutants (Odell et al., 1985). The plasmid was linearized with HindIII at +9, digested with Bal31 exonuclease, and ligated to SalI or XhoI linkers. The 3′ break points were identified by dideoxy sequencing, and deletion mutants ending at −46, −78, −107, −127, −187, and −208 were selected. The fragment extending from −343 to −46 was chosen to produce 5′ deletion mutants, either by cutting at convenient restriction sites (HaeIII at −209, XmnI at −130, EcoRV at −90) or by exchanging the −343 to −90 region with −168 to −90 and −105 to −90, both of which were isolated previously (Odell et al., 1985). Internal deletions from −208 to −90 or −107 to −90 within the −343 to −46 region were constructed by attaching fragments −343 to −208 or −343 to −107 to −90 to −46 through filling in a SalI site at −208 or −107 and ligation to an EcoRV site at −90.

Multimerization of the −209 to −46 Region

The HaeIII-SalI fragment (−209 to −46) from the 35S 5′ upstream region was subcloned between the SmaI and SalI sites in pEMBL12$^+$, a derivative of the pEMBL plasmid (Dente et al., 1983) with the polylinker from pUC12. The SacI site in the polylinker next to −209 was changed to an XhoI site by SacI digestion followed by T4-DNA polymerase treatment to remove the 3′ overhanging bases and ligation to XhoI linkers. The resulting plasmid pXS1 contained the 35S 5′ upstream region (−209 to −46) as an XhoI-SalI fragment with an EcoRI site located 5′ to the XhoI site. The "head to tail" dimer of the −209 to −46 region was made by inserting the EcoRI-SalI fragment between the EcoRI and XhoI sites of pXS1. A similar strategy was used to generate the tetramer and octamer.

Construction of the Intermediate Vector with Test and Reference Genes

A derivative of the binary intermediate vector pMON505 (Horsch and Klee, 1986) containing the 35S TATA sequence from −46 to +9 fused to the CAT coding sequence and polyadenylation sequence of the pea *rbcS-E9* gene was constructed. The 35S-CAT chimeric gene was located in the polylinker region in the orientation shown in Figure 1. A unique HpaI site in this plasmid, 5 kb away from the 35S TATA box of the chimeric CAT gene, was used for insertion of the reference gene unit. We chose the *E. coli* β-glucuronidase gene (GUS) coding sequence as the coding sequence for the reference gene (Jefferson et al., 1987). Plasmid pRAJ260, which is a pEMBL9 derivative containing the entire GUS coding region (1.8 kb) in a PstI site, was cut with BamHI. A 300-bp fragment containing the polyadenylation sequence of the pea *rbcS-3C* gene (Fluhr et al., 1986) was inserted into this site by blunt-end ligation. The HindIII site just upstream of the GUS gene was used for the insertion of a CaMV 35S upstream fragment (−941 to +9). The HindIII site was filled in and the resulting 35S-GUS-3C 3′ cassette was inserted into the HpaI site by blunt-end ligation in the orientation shown in Figure 1. The resulting plasmid, pMON505–67, about 16 kb in size, was used as an acceptor for the insertion of various 35S promoter fragments and mutant derivatives. Insertions were at the 5′ (HindIII and XhoI) or 3′ end (ClaI) of the 35S TATA-CAT-E9 3′ test gene unit (Figure 1).

The 35S TATA-CAT-E9 3′ unit in pMON505–67 was replaced by the CaMV 35S promoter (−941 to +9) CAT-E9 3′ chimeric gene unit (Nagy et al., 1987) to give the control construct pMON505–70.

An octamer of −203 to −46 as an XhoI-SalI fragment was inserted at −50 of the *rbcS-3A* gene in a pMON200 derivative, which also contains an intact *rbcS-3A* gene as a reference (C. Kuhlemeier, unpublished data).

Production of Transgenic Plants

Intermediate vectors containing various chimeric test and reference genes were mobilized into a "disarmed" *Agrobacterium tumefaciens* GV3111SE by triparental mating (Fraley et al., 1985). Exconjugants were used to inoculate leaf discs of Nicotiana tabacum SR1, and regenerated shoots were selected on a medium containing 200 μg/ml kanamycin (Horsch et al., 1985). After rooting, transgenic plantlets were transferred to soil and grown in a greenhouse.

Analysis of CaMV 35S Promoter Activities

Activities of the various CaMV 35S promoter constructs were screened first by CAT assay and the transcript levels measured by quantitative 3′ S1 mapping. Protein was extracted from leaves with the GUS extraction buffer (Jefferson et al., 1987). One to 5 μg of protein were used for CAT assay by the TLC method (Gorman et al., 1982). Transgenic plants with six to eight leaves were used for transcript analyses. Total RNA was extracted from fully expanded leaves as described by Nagy et al. (1987). Ten to 50 mg of RNA were used in quantitative 3′ S1 mapping of transcript levels of test and reference genes by a method similar to that described previously (Fluhr et al., 1986). A 3′ ^{32}P-labeled single-stranded DNA fragment derived from a 690-bp HindIII-ClaI fragment containing the *rbcS-E9* 3′ end sequence (Morelli et al., 1985) was used as a probe. Hybridization was carried out in a 10-μl solution containing 50% formamide, 0.4 M NaCl, 20 mM Pipes (pH 6.8), and 2 mM EDTA for 12 hr at 37°C. After hybridi-

zation, 150 μl of S1 digestion solution containing 0.25 M NaCl, 25 mM NaOAc (pH 4.6), 1 mM $ZnSO_4$, 25 μg/ml denatured salmon sperm DNA, and 1000 units/ml S1 nuclease (Bethesda Research Laboratories) was added. The mixture was incubated at room temperature for 90 min and digestion was terminated by adding 20 μl of a solution containing 6.4 M NH_4 acetate, 0.1 M EDTA, and 0.1 mg/μl tRNA. Protected single-stranded DNA fragments were precipitated with ethanol and sized in a 6% sequencing gel. This method gives protected fragments of 230 bases for the CAT-E9 transcripts, 89 bases for the GUS-3C transcripts, and 160 bases for the *rbcS-3A* transcripts. Signals on the autoradiogram were quantitated by the method of Suissa (1983), and the signal strengths were expressed as the ratio of CAT (test transcript) to GUS (reference transcript).

ACKNOWLEDGMENTS

We thank Cathy Ward and Ling Ren for excellent technical assistance, Drs. Eric Lam and Phillip Benfey for pointing out the TGACG motif, Hugh Williams for graphics, and Wendy Roine for typing the manuscript. We also thank Drs. Philip Benfey, Eric Lam, and Philip Gilmartin for their comments on the manuscript. This work was supported by grants from the Monsanto Company and the Pew Trust.

Received October 26, 1988.

REFERENCES

Bienz, M., and Pelham, H.R.B. (1986). Heat shock regulatory elements function as an inducible enhancer in the Xenopus hsp70 gene and when linked to a heterologous promoter. Cell **45**, 753–760.

Covey, S.-N. (1985). Organization and expression of the cauliflower mosaic virus genome. In Molecular Plant Virology, Vol. 2, J. W. Davies, ed (Boca Raton, FL: CRC Press, Inc.), pp. 121–160.

Covey, S.-N., Lomonossoff, G.P., and Hull, R. (1981). Characterization of cauliflower mosaic virus DNA sequences which encode major polyadenylated transcripts. Nucleic Acids Res. **9**, 6735–6747.

Cuozzo, M., O'Connell, K.M., Kaniewski, W., Fang, R.-X., Chua, N.-H., and Tumer, N.E. (1988). Viral protection in transgenic tobacco plants expressing the cucumber mosaic virus coat protein or its antisense RNA. Bio-Technology (NY) **6**, 549–557.

Dente, L., Cesareni, G., and Cortese, R. (1983). pEMBL: A new family of single stranded plasmids. Nucleic Acids Res. **11**, 1645–1655.

Dierks, P., Van Ooyen, A., Cochran, M.D., Dobkin, C., Reiser, J., and Weissmann, C. (1983). Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit beta-globin gene in mouse 3T6 cells. Cell **32**, 695–706.

Ebert, P.R., Ha, S.B., and An, G. (1987). Identification of an essential upstream element in the nopaline synthase promoter by stable and transient assays. Proc. Natl. Acad. Sci. USA **84**, 5745–5749.

Fraley, R.T., Rogers, S.G., Horsch, R.B., Eichholtz, D.A., Flick, J., S., Fink, C.L., Hoffmann, N.L., and Sanders, P.R. (1985). The SEV system: A new disarmed Ti plasmid vector system for plant transformation. Bio-Technology (NY) **3**, 629–635.

Fluhr, R., Moses, P., Morelli, G., Coruzzi, G., and Chua, N.-H. (1986). Expression dynamics of the pea rbcS multigene family and organ distribution of the transcripts. EMBO J. **5**, 2065–2071.

Fromm, M., Taylor, L.P., and Walbot, V. (1985). Expression of genes transferred into monocot and dicot plant cells by electroporation. Proc. Natl. Acad. Sci. USA **82**, 5824–5828.

Gorman, C.M., Moffat, L.F., and Howard, B.H. (1982). Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell Biol. **2**, 1044–1051.

Guilley, H., Dudley, R.K., Jonand, G., Balazs, E., and Richards, K.E. (1982). Transcription of cauliflower mosaic virus DNA; detection of promoter sequences and characterization of transcripts. Cell **30**, 763–773.

Hemenway, C., Fang, R.-H., Kaniewski, W.K., Chua, N.-H., and Tumer, N.E. (1988). Analysis of the mechanism of protection in transgenic plants expressing the potato virus X coat protein or its antisense RNA. EMBO J. **7**, 1273–2380.

Hohn, T., Richards, K., and Lebeuvier, G. (1982). Cauliflower mosaic virus on its way to becoming a useful plant vector. Curr. Top. Microbiol. Immunol. **96**, 194–236.

Horsch, R.B., and Klee, H.J. (1986). Rapid assay of foreign gene expression in leaf discs transformed by Agrobacterium tumefaciens: role of T-DNA borders in the transfer process. Proc. Natl. Acad. Sci. USA **83**, 4428–4432.

Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G., Fraley, R.T. (1985). A simple and general method for transferring genes into plants. Science **227**, 1229–1231.

Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. (1987). GUS fusions: Beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. **6**, 3901–3907.

Kay, R., Chan, A., Daly, M., and McPherson, J. (1987). Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. Science **236**, 1299–1302.

Kuhlemeier, C., Green, P., and Chua, N.-H. (1987a). Regulation of gene expression in higher plants. Annu. Rev. Plant Physiol. **38**, 221–257.

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

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

Mikami, K., Tabata, T., Kawata, T., Nakayama, T., and Iwabuchi, M. (1987). Nuclear protein(s) binding to the conserved DNA

hexameric sequence postulated to regulate transcription of wheat histone genes. FEBS Lett. **223**, 273–278.

Morelli, G., Nagy, F., Fraley, R.T., Rogers, S.G., and Chua, N.-H. (1985). A short conserved sequence is involved in the light-inducibility of a gene encoding ribulose-1,5-bisphosphate carboxylase small subunit of pea. Nature **315**, 200–204.

Nagata, T., Okada, K., Kawazu, T., and Takebe, I. (1987). Cauliflower mosaic virus 35S promoter directs S phase specific expression in plant cells. Mol. Gen. Genet. **207**, 242–244.

Nagy, F., Odell, J.T., Morelli, G., and Chua, N.-H. (1985). In Biotechnology in Plant Science: Relevance to Agriculture in the Eighties, M. Zaitlin, P. Day, and A. Hollaender, eds (New York: Academic Press), pp. 227–236.

Nagy, F., Boutry, M., Hsu, M.-Y., Wong, M., and Chua, N.-H. (1987). 5′-proximal region of the wheat Cab-1 gene contains a 268-bp enhancer-like sequence for phytochrome response. EMBO J. **6**, 2537–2542.

Odell, J.T., Nagy, F., and Chua, N.-H. (1985). Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature **313**, 810–812.

Odell, J.T., Nagy, F., and Chua, N.-H. (1987). Variability in 35S promoter expression between independent transformants. In Plant Gene Systems and Their Biology, L. McIntosh and J. Key, eds (New York: Alan R. Liss, Inc.), **62**, pp. 329–331.

Odell, J.T.,, Knowlton, S., Lin, W., and Mauvais, C.J. (1988). Properties of an isolated transcription stimulating sequence derived from the cauliflower mosaic virus 35S promoter. Plant Mol. Biol. **10**, 263–273.

Ondek, B., Shepard, A., and Herr, W. (1987). Discrete elements within the SV40 enhancer region display different cell-specific enhancer activities. EMBO J. **6**, 1017–1025.

On-Lee, T.M., Turgeon, R., and Wu, R. (1986). Expression of a foreign gene linked to either a plant-virus or a Drosophila promoter, after electroporation of protoplasts of rice, wheat, and sorghum. Proc. Natl. Acad. Sci. USA **83**, 6815–6819.

Ow, D.W., Jacobs, J.D., and Howell, S.H. (1987). Functional regions of the cauliflower mosaic virus 35S RNA promoter determined by the firefly luciferase gene as a reporter of promoter activity. Proc. Natl. Acad. Sci. USA **84**, 4870–4874.

Pfeiffer, P., and Hohn, T. (1983). Involvement of reverse transcription in the replication of cauliflower mosaic virus: A detailed model and test of some aspects. Cell **33**, 781–789.

Poulsen, C., and Chua, N.-H. (1988). Dissection of 5′ upstream sequences for selective expression of *Nicotiana plumbaginifolia* rbcS-8B gene. Mol. Gen. Genet. **214**, 16–23.

Sanders, P.R., Winter, J.A., Zarnason, A.R., Rogers, S.G., and Fraley, R.T. (1987). Comparison of cauliflower mosaic virus 35S and nopaline synthase promoters in transgenic plants. Nucleic Acids Res. **15**, 1543–1558.

Schaffner, G., Schirma, S., Muller-Baden, B., Weber, F., and Schaffner, W. (1988). Redundancy of information in enhancers as a principle of mammalian transcription control. J. Mol. Biol. **201**, 81–90.

Schirm, S., Jirieny, J., and Schaffner, W. (1987). The SV40 enhancer can be dissected into multiple segments, each with a different cell type specificity. Genes Dev. **1**, 65–74.

Serfling, E., Jasin, M., and Schaffner, W. (1985). Enhancers and eukaryotic gene transcription. Trends Genet. **1**, 224–230.

Suissa, M. (1983). Spectrophotometric quantitation of silver grains eluted from autoradiograms. Anal. Biochem. **133**, 511–514.

Zenke, M., Grundstrom, T., Matthes, H., Wintzerith, M., Schatz, C., Wildeman, A., and Chambon, P. (1986). Multiple sequence motifs are involved in SV40 enhancer function. EMBO J. **5**, 387–397.