Strategies for Development of Functionally Equivalent Promoters with Minimum Sequence Homology for Transgene Expression in Plants: cis-Elements in a Novel DNA Context versus Domain Swapping[1]

Simran Bhullar, Suma Chakravarthy[2,3], Sonia Advani[2], Sudipta Datta, Deepak Pental, and Pradeep Kumar Burma*

Department of Genetics, University of Delhi South Campus, Benito Juarez Road, New Delhi–110021, India

The cauliflower mosaic virus 35S (35S) promoter has been extensively used for the constitutive expression of transgenes in dicotyledonous plants. The repetitive use of the same promoter is known to induce transgene inactivation due to promoter homology. As a way to circumvent this problem, we tested two different strategies for the development of synthetic promoters that are functionally equivalent but have a minimum sequence homology. Such promoters can be generated by (a) introducing known cis-elements in a novel or synthetic stretch of DNA or (b) "domain swapping," wherein domains of one promoter can be replaced with functionally equivalent domains from other heterologous promoters. We evaluated the two strategies for promoter modifications using domain A (consisting of minimal promoter and subdomain A1) of the 35S promoter as a model. A set of modified 35S promoters were developed whose strength was compared with the 35S promoter per se using *β-glucuronidase* as the reporter gene. Analysis of the expression of the reporter gene in transient assay system showed that domain swapping led to a significant fall in promoter activity. In contrast, promoters developed by placing cis-elements in a novel DNA context showed levels of expression comparable with that of the 35S. Two promoter constructs Mod2A1T and Mod3A1T were then designed by placing the core sequences of minimal promoter and subdomain A1 in divergent DNA sequences. Transgenics developed in tobacco (*Nicotiana tabacum*) with the two constructs and with 35S as control were used to assess the promoter activity in different tissues of primary transformants. Mod2A1T and Mod3A1T were found to be active in all of the tissues tested, at levels comparable with that of 35S. Further, the expression of the Mod2A1T promoter in the seedlings of the T_1 generation was also similar to that of the 35S promoter. The present strategy opens up the possibility of creating a set of synthetic promoters with minimum sequence homology and with expression levels comparable with the wild-type prototype by modifying sequences present between cis-elements for transgene expression in plants.

Homology-based gene silencing (HBGS) has been reported to occur extensively in transgenic plants (Meyer and Saedler, 1996; Vaucheret and Fagard, 2001). HBGS is a generic term for both transcriptional gene silencing and posttranscriptional gene silencing. Such situations arise when (a) multiple copies of a gene cassette integrate into plants during transformation (van der Krol et al., 1990), (b) introduced transgenes have homologous promoters driving them (Mol et al., 1989), and (c) an introduced gene has homology with an endogenous gene in the coding region (Napoli et al., 1990). Several mechanisms have been suggested to explain the phenomena of HBGS (Matzke and Matzke, 1995, 1998; Meyer and Saedler, 1996; Fire, 1999; Hamilton and Baulcombe, 1999; Steimer et al., 2000), the common denominator being that homology triggers cellular recognition mechanisms that result in silencing of the repeated genes.

One approach to escape the silencing of the introduced DNA is to develop gene cassettes that contain very little sequence similarity either to endogenous sequences or to one another if a number of transgenes need to be introduced. This ability to develop gene cassettes that have no sequence homology or minimum sequence homology is of critical importance to fully exploit the transgenic technologies for crop improvement. In several cases, insertion of multiple genes coding for different traits or a metabolic pathway may be required (Ye et al., 2000) to introduce the desired phenotype in crops. Such transgenics can be developed by simultaneous or sequential transformation of multiple genes or by pyramiding different genes that are available in separate genetic stocks of a crop developed in independent laboratories.

With reference to promoters, homology can be avoided by either using diverse promoters isolated

[1] This work was supported by the Department of Biotechnology, Government of India and by Dow AgroSciences (Indianapolis). S.B., S.C., and S.D. were supported by Research Fellowships from Council of Scientific and Industrial Research, Government of India.

[2] These authors contributed equally to the paper.

[3] Present address: Boyce Thompson Institute for Plant Research, Ithaca, NY 14853.

* Corresponding author; e-mail pburma@hotmail.com; fax 91–11–26885270.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.103.020602.

from different plant and viral genomes or by designing synthetic promoters. Synthetic promoters can be designed in two ways: (a) the identified cis-elements of a promoter can be placed in a synthetic stretch of DNA different from its own native DNA context to create a functionally similar promoter with a "novel" DNA sequence; and (b) "domain swapping," wherein cis-elements of a promoter can be replaced with functionally equivalent regions from heterologous promoters. For example, the *as-1* element in subdomain A1 of cauliflower mosaic virus 35S (35S) promoter, is a direct TGAGC repeat spaced by seven nucleotides (TGAGC*TAAGGGA*TGAGC), to which transcription factor ASF-1 binds (Lam et al., 1989). The *as-1* element in subdomain A1 is responsible for expression in the root (Benfey et al., 1989; Lam et al., 1989; Benfey and Chua, 1990) and has been reported to act synergistically with other subdomains (B1–B5) to confer expression in different tissues of the plant (Benfey and Chua, 1990; Lam, 1994). Elements like *as-1* are present in other plant, bacterial, and viral promoters. The pentameric direct repeat TGACG and a 7-bp spacer organization is conserved among different T-DNA gene promoters such as *ags* and *mas* (Bouchez et al., 1989) and viral (cauliflower mosaic virus, figwort mosaic virus, and mirabilis mosaic virus) promoters (Sanger et al., 1990; Dey and Maiti, 1999). Many of these elements have been shown to bind common protein factors isolated from tobacco (*Nicotiana tabacum*; An et al., 1986; Mikami et al., 1987; Katagiri et al., 1989). Thus the *as-1* element of the 35S promoter can be replaced with similar elements of other promoters. This strategy will introduce defined functional elements along with their own sequence context, which have been selected during evolution, whereas in the first strategy, cis-elements are placed in a novel (synthetic) DNA context.

In this paper, we present an evaluation of the two strategies for promoter modification of the 35S promoter using domain A of the promoter as a model because the cis-elements of the same have been characterized. The 35S promoter has been selected because it expresses at a high level in most plant tissues and organs (Odell et al., 1985) and is widely used for the development of transgenics particularly in the dicot species. The domain A of 35S has a minimal promoter (mp; −46 to +1) encompassing a TATA box and subdomain A1 (−90 to −46; Fig. 1a) with an *as-1* element. We show that domain swapping may not be as efficient a strategy as creating promoters with cis-elements in a synthetic context.

RESULTS

Generation of Synthetic Domains and Experimental Design

A program for the generation of random DNA sequences, DNASEQ (http://www.cbs.dtu.dk/ramneek/ds.zip), was developed and used to create synthetic stretches of DNA of the desired length and given GC content. The sequences generated lacked methylation-prone CG and CNG sites. Several "candidate" stretches of DNA were designed with the aim of replacing mp and subdomain A1 of the 35S promoter. The distribution of GC content in these sequences was examined in windows of 10 bp and compared with that of the corresponding 35S regions, using program DNASIS. The sequences that closely resembled the 35S promoter in GC content but were divergent in sequence were chosen. The cis-elements of domain A of 35S, i.e. either *as-1* or TATA box or both, were introduced in the synthetic stretches at distances equivalent to those observed in the 35S promoter.

To test the strength of different modified 35S promoters (described in Table I), these were cloned upstream to the *β-glucuronidase* (*gus*) reporter gene (Fig. 2). The binary vector developed also contained the gene for *chloramphenicol acetyl transferase* (*cat*) under the control of the 35S promoter (used as reference gene for transient expression assays) as well as a *nopaline synthase* promoter (*Pnos*)-driven *neomycin phosphotransferase* gene (*npt II*; used both as a selection marker for plant transformation and as the reference gene in transgenic plants).

The modified promoter constructs were tested in two phases. In the first phase of experiments, modified promoters (Table I) were tested by transient expression in tobacco cv Xanthi leaf mesophyll protoplasts. The strength of 35S and modified promoters was measured by GUS assays (Jefferson, 1987). In the second phase of experiments, based on the results with transient assays, two promoter constructs modified in domain A (Mod2A1T and Mod3A1T) were synthesized (Table I) and their activity was checked in different tissues of primary transformants (T_0) as well as in seedlings of the T_1 generation.

Expression Analysis of the 35S Promoter with Modified mp

The mp (−46 to +1; Fig. 1a) of 35S was replaced with mp regions of heterologous plant promoters (Table I). No changes were made in the rest of the domains of 35S promoter. The mp regions of *rbcs-3A* promoter of pea (*Pisum sativum*; Fluhr et al., 1986; Kuhlemeir et al., 1988), *ats-1A* promoter of Arabidopsis (Krebbers et al., 1988; Donald and Cashmore, 1990), and *cab22R* promoter of petunia (*Petunia hybrida*; Gidoni et al., 1988, 1989) were domain swapped with the mp of 35S to create promoters 3AT, 1AT, and 22RT, respectively. The mp of the genes *rbcs3A* from pea and *ats1A* from Arabidopsis were selected because they possessed consensus plant TATA sequence (TATATATA; Joshi, 1987; Sawant et al., 1999). The *cab22R* gene promoter was chosen because its TATA box (TAAATAAA) is slightly different from plant consensus but closer to viral consensus

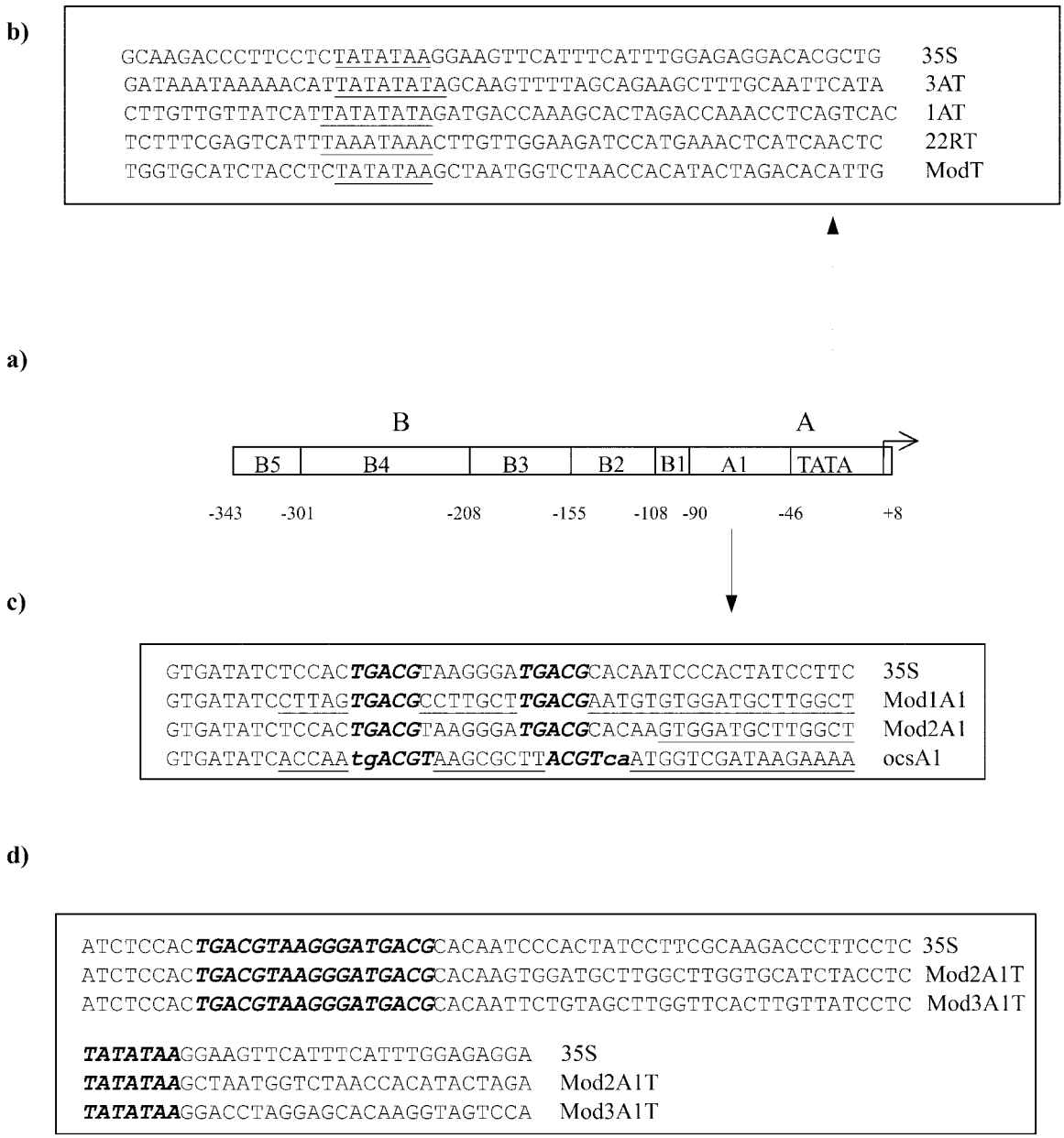

Figure 1. a, Modular organization of the 35S promoter as described by Benfey and Chua (1990). The 35S promoter consists of two domains, A and B. Domain A is further divided into the mp and subdomain A1. Domain B has five subdomains (B1–B5). b, Sequence comparison (from −46 to +8) of 35S and modified mps 3AT, 1AT, 22RT, and ModT. The TATA box in these promoters is underlined. In all of these promoters, the TATA box is located 25 bp upstream to the transcription start site of the original promoters. c, Sequence comparison of the subdomain A1 of the 35S promoter (−90 to −46) containing the *as-1* element shown along with the modified regions in the Mod1A1, Mod2A1, and ocsA1 promoters. Nucleotides that have been changed from the 35S sequence are underlined. The *ocs* element used in ocsA1 is a modification of the native sequence (Kim et al., 1994). It had nucleotide substitutions (lowercase) that made it identical to the consensus *as-1* element of the 35S promoter. The direct repeat of the *as-1* element (in 35S, Mod1A1, and Mod2A1) and the inverted repeat of *ocs* element (ocsA1) are shown in bold italics. d, Sequence comparison of domain A (−90 to +1) of the 35S promoter with that of the Mod2A1T and Mod3A1T promoters. The *as-1* element and TATA box are shown in bold italics. Domain A of Mod2A1T and Mod3A1T show sequence similarity of 35% and 41%, respectively with that of the 35S promoter in the regions flanking the identified cis-elements.

(TATATAA). In a fourth modification (Mod T), a synthetic mp was created by placing TATA box of 35S promoter (TATATAA) in a synthetic stretch of DNA. A comparison of sequences of mp of 3AT, 1AT, 22RT, and ModT is represented in Figure 1b. The strength of these promoters was tested in transient expression system. It was observed that whereas ModT exhibited about 85% of 35S promoter activity,

Table I. *Modified promoters developed for analysis*

Modified Domain of 35S	Modification	Name of the Promoter
Minimal Promoter	TATA box was placed in a synthetic DNA context	ModT
	Minimal promoters from other well-characterized promoters used for domain swapping	
	−46 to + 8 of *rbcs3A* of pea	3AT
	−46 to + 8 of *ats1A* of Arabidopsis	1AT
	−46 to +8 of *cab22RT* of petunia	22RT
Subdomain A1	The direct repeat (TGACG) of *as-1* element was placed in synthetic DNA context with different flanking and spacer nucleotides	Mod1A1
	Same as Mod1A1 except that intervening 7 bp between TGACG repeat and flanking 5 bp was conserved	Mod2A1
	Functionally equivalent *ocs* element from octopine synthase gene promoter from *Agrobacterium* sp. was used for domain swapping	ocsA1
Domain A	TATA box and defined *as-1* element as tested in ModT and Mod2A1 were placed in a synthetic stretch	Mod2A1T
	TATA box and defined *as-1* element placed in a second synthetic stretch	Mod3A1T

domain-swapped promoters were weaker. Of the three domain-swapped promoters, 22RT was the strongest showing about 68% of 35S activity, whereas 3AT and 1AT exhibited less than 20% of the activity shown by 35S promoter (Fig. 3). The results thus clearly showed that mp with a viral TATA box drives higher expression levels than mp with a consensus plant TATA box inserted in context of 35S promoter. However, none of the tested promoters with domain-swapped mp functioned better than the promoter with synthetic mp. A synthetic mp can thus replace 35S mp without significantly affecting its promoter activity.

Expression Analysis of the 35S Promoters with Modified Subdomain A1

Three different modifications in subdomain A1 (Table I) were tested for their effect on the 35S promoter activity. In the first modification (Mod1A1), direct repeats of *as-1* element (TGACG; Lam et al., 1989) were placed in a synthetic stretch of DNA wherein the 7-bp sequences between TGACG repeat were also novel. Mod2A1 was similar to Mod1A1 except that 7 bp between the TGACG repeats and 5 bp on either side of the repeats were kept the same as that of 35S. In the third modification (ocsA1), a 44-bp region encompassing a modified *octopine synthase* (*ocs*) element from *ocs* promoter of *Agrobacterium* sp. (Kim et al., 1994) was used to replace subdomain A1 (−90 to −46). The *ocs* element is a hexameric inverted repeat with 8-bp spacer (16-bp palindrome sequence; Ellis et al., 1987) and is a binding site for transcription factor OCSTF (Singh et al., 1990). This element was chosen to replace the *as-1* element of the 35S promoter because it has also been shown to bind the transcription factor OCSTF (Bouchez et al., 1989). Transcription factor ASF-1 that binds to *as-1* element of 35S has also been reported to bind the *ocs* palindrome (Fromm et al., 1989). A comparison of the sequence of the modified subdomains A1 that were tested is shown in Figure 2c. Transient expression analysis showed that whereas modification made in Mod1A1 led to a 3-fold drop in promoter activity, Mod2A1 gave expression level similar to that of the 35S promoter. The promoter ocsA1, designed to test

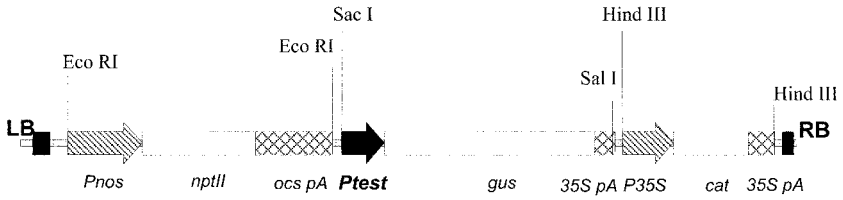

Figure 2. Binary vector developed for the transient expression analysis and transformation of tobacco. LB and RB are the left and right borders, respectively, of T-DNA in the binary vector. The selection marker gene *npt II* is driven by *Pnos* and has 3′ pA region of *ocs* gene. The *npt II* was used as a selectable marker and a reference gene in the transgenic lines. *Ptest* denotes the modified promoter driving *gus* reporter gene with 3′ pA region of 35S. The *cat* gene under the control of the 35S promoter (*P35*) was used as the reference gene in transient expression assays.

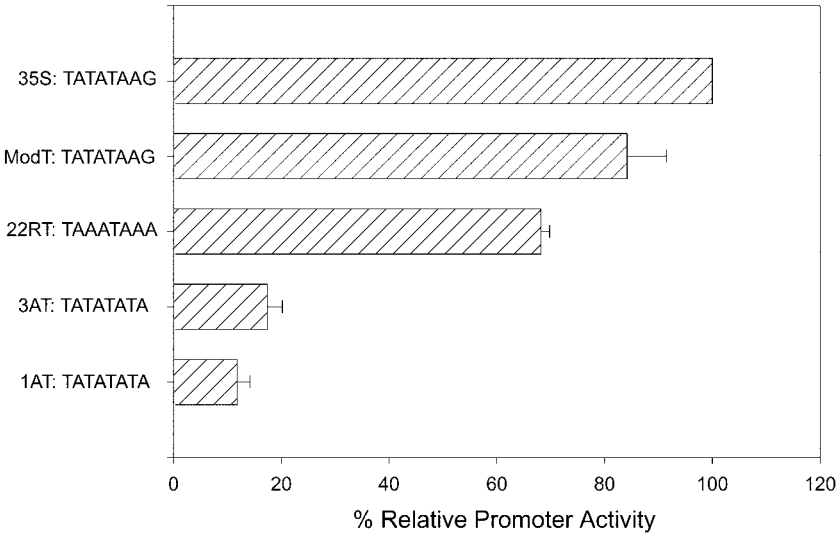

Figure 3. Relative activity of the modified promoters 3AT, 1AT, 22RT, and ModT in transient expression assays. The variations in transient expression were normalized using 35S-*cat* gene as a reference. Data of a minimum of three independent experiments were used to measure promoter activity of each construct. In addition, every experiment had internal replicates. The activity of the 35S promoter was measured as a control. Normalized GUS/CAT of the 35S promoter was assigned the value of 100%. Each bar represents the average relative promoter activity ± SE. *y* axis denotes the different modified promoters tested and the TATA box present therein.

the concept of domain swapping was only about 25% active as compared with 35S (Fig. 4). A comparison of promoter activity of Mod1A1 and Mod2A1 showed that the intervening 7 bp in *as-1* element is important, and homology cannot be broken in this region. When the intervening spacer sequence was changed (Mod1A1), there was a significant reduction in the expression of GUS. Furthermore, domain swapping of the *as-1* element with the *ocs* element in ocsA1 led to reduction in promoter activity, providing further evidence along with the data on mps that domain swapping may not be a viable strategy for breaking homology between promoters.

Expression Analysis of the 35S Promoters with Modified Domain A in Different Tissues of Primary Transformants (T_0) of Tobacco

On the basis of our observations on transient expression analysis of 35S promoter modified for the mp region and subdomain A1, we designed two promoters with a synthetic domain A: The TATA box (TATATAA) and the *as-1* element (TGACG*TAAGG-GA*TGACG) were placed in two different synthetic stretches of DNA. The two promoters thus designed, named Mod2A1T and Mod3A1T, had much reduced sequence homology among themselves and with the 35S promoter in the domain A. Twenty-two independent tobacco transgenic lines each with 35S-*gus* and Mod2A1T-*gus* constructs were developed and grown under the same conditions to test the promoter activity in the leaf tissue by GUS assays (Fig. 5). Twenty-seven independent transgenic lines were developed with Mod3A1T-*gus* construct and grown along with 11 35S-*gus* containing lines for analyzing the promoter activity in leaves (Fig. 6). Variation in the *gus* expression pattern among independent transgenic plants generated using 35S promoter and Mod2A1T promoter was similar (Fig. 5, a and b). Approximately 4.5-fold variation in expression between the highest and the lowest expresser was observed in case of both 35S-*gus* and Mod2A1T-*gus* containing lines. These variations could be due to multiple integrations in the genome and/or due to position effects (Peach and Velten, 1991; Iglesias et al., 1997). To normalize this variability, the expression levels of NPT II were used

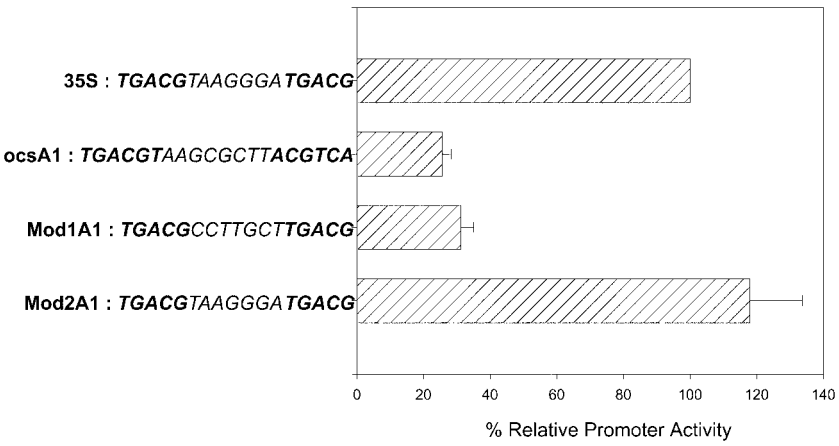

Figure 4. Relative activity of the modified promoters Mod1A1, Mod2A1, and ocsA1 in transient expression assays. The experiments were performed as detailed in Figure 3. The mean value of the normalized GUS/CAT of the 35S promoter was assigned the value of 100%. Each bar represents the average relative promoter activity ± SE. *y* axis represents the different modified promoters and *as-1* or *as-1* like elements present therein.

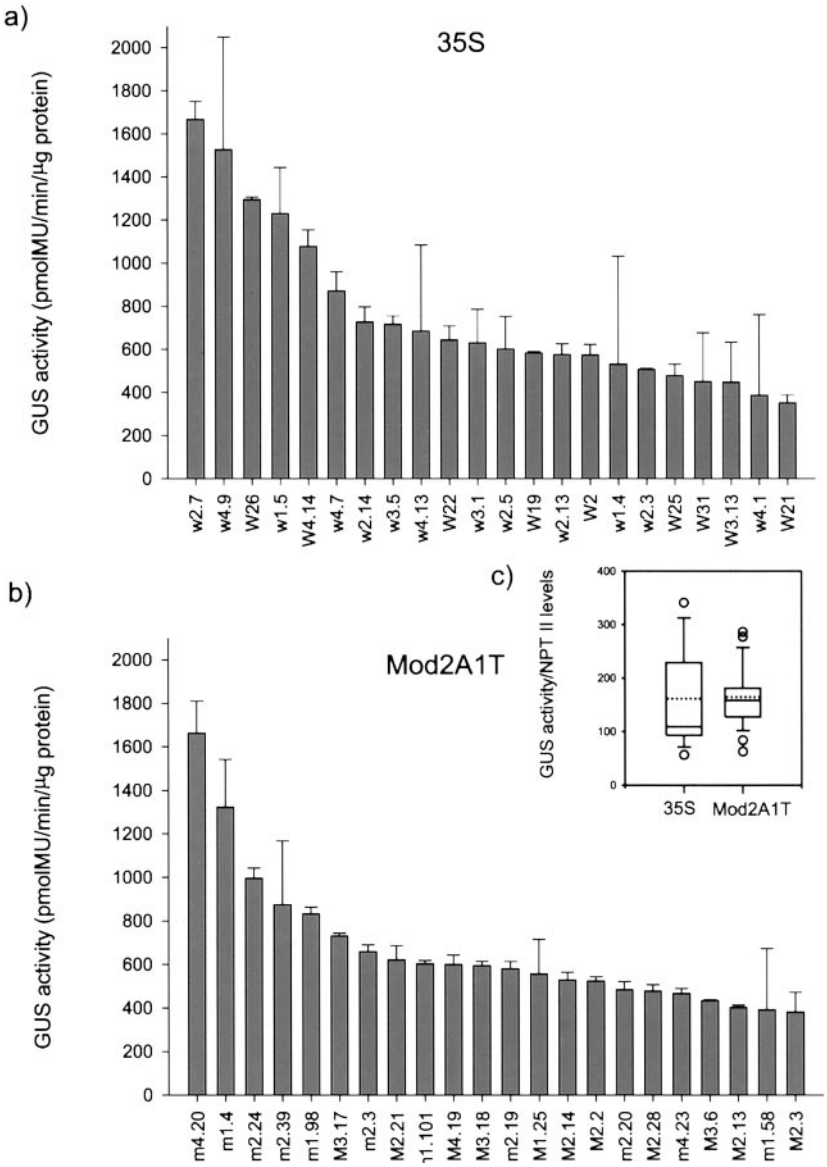

Figure 5. GUS activity of 35S (a) and Mod2A1T (b) promoters in the leaves of transgenic tobacco plants. The independent transgenic lines were grown under controlled conditions in a growth chamber. Transgenic lines with the 35S promoter-driven *gus* construct used as control were grown simultaneously under similar growth conditions (16-h day and 8-h night, 28°C ± 2°C, relative humidity 70%) with the transgenic plants containing the Mod2A1T-*gus* construct. Promoter activity was measured at two time points i.e. from leaves of 45- and 60-d-old (after transfer to the soil) plants. Each bar represents the promoter activity of a single transgenic plant. The value in the bars is the mean ± SE of the two independent experiments (with internal replicates). c, Comparison of normalized promoter activity (GUS activity/NPT II levels) of 35S and Mod2A1T has been depicted in a Box and Whisker plot. The horizontal lines in the Box and Whisker plot represent 10, 25, 50, 75, and 90 percentiles. The dotted line represents the mean of the data. Extreme values are depicted as small circles at the top and bottom of the plot.

as reference. Normalized GUS activity of these sets of plants has been represented in a Box and Whisker plot in Figure 5c to compare the normalized distribution and mean of the 35S and Mod2A1T activity. We observed that the strength of Mod2A1T was similar to that of the 35S. On the other hand, the expression of the Mod3A1T promoter in leaves was higher than 35S if mean values were compared (Fig. 6b). It was also observed that transgenic lines expressing *gus* under the Mod3A1T promoter showed lower levels of variations (3.3-fold) between independent lines as seen by both the GUS activity (Fig. 6a) as well as normalized GUS activity (Fig. 6b).

The activity of the two modified promoters vis-à-vis that of the 35S promoter was also analyzed in the stem and roots of primary transformants grown in tissue culture. The data from a minimum of 20 independent transgenic lines in each case are summarized in a Box and Whisker plot in Figure 7, a through d. The modified promoters were found to be active in both of these tissues. Although the expression profile of Mod2A1T across the transgenic lines was similar to that of the 35S in the stem, in case of root expression, Mod2A1T showed expression levels comparable with the second quartile of that of 35S. Mod3A1T, on the other hand, functioned at par with that of the 35S promoter in both stem and root tissues. Modified promoters were also found to be as active as 35S in callus tissue raised from both stem and leaves (Fig. 7, e and h) as well as in the floral tissues (data not shown).

Expression Analysis of the Mod2A1T Promoter in the Seedlings of the T_1 Generation

To test whether the expression levels driven by modified promoter Mod2A1T in primary transfor-

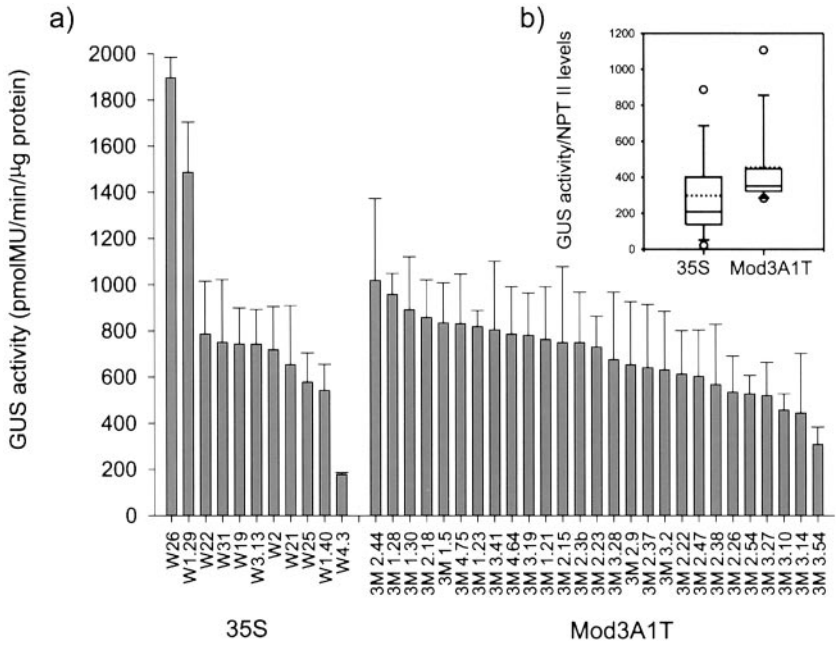

Figure 6. GUS activity of the 35S and Mod3A1T promoter in the leaves of primary transgenic tobacco plants. The experimental design was as described in Figure 5. Each bar represents the promoter activity of a single transgenic plant. The values in the bars represent the mean ± SE of two independent experiments. b, Comparison of normalized promoter activity (GUS activity/NPT II levels) of 35S and Mod3A1T has been depicted in a Box and Whisker plot. The horizontal lines in the Box and Whisker plot represent 10, 25, 50, 75, and 90 percentiles. The dotted line represents the mean of the data. Extreme values are depicted as small circles at the top and bottom of the plot.

mants are also reflected in the next generation, GUS expression in the seedlings of representative transgenics containing the Mod2A1T and 35S promoter constructs was quantified. Segregation analysis based on the GUS expression (histochemical staining) of T_1 seedlings from Mod2A1T transgenics revealed that all of the plants tested showed insertions at a single locus. In the case of transgenics with the 35S promoter, seven of nine plants tested segregated for a single-locus insertion, whereas two plants segregated for two loci (data not shown). Extracts made from approximately 150 7-d-old seedlings from each line were used to analyze promoter activity. Normalized GUS activities represented in Figure 8 show an expression pattern of Mod2A1T promoter similar to that of 35S.

DISCUSSION

The present work originated from the need to develop diverse promoters to drive transgene expression to circumvent the problem of HBGS. Diverged promoters could be developed either by domain swapping or by placing cis-elements in synthetic stretch of DNA. These possibilities were tested by modifications of domain A of the 35S promoter.

The modified 35S promoters developed in this study by domain swapping were found to be not as efficient as the wild-type 35S promoter. This was observed in modifications of both mp regions and subdomain A1 of 35S. In the case of the mp region, the fall in activity of the three domain-swapped promoters could be due to either (a) differences in the TATA element core sequence or (b) divergence in the sequences flanking the TATA box. Two of these mps (1AT and 3AT) were selected for domain swapping because they carried a plant consensus TATA element (TATATATA; Joshi, 1987; Sawant et al., 1999), which could possibly increase the strength of the 35S promoter. The third mp (22RT) had a non-consensus TATA box (TAAATAAA) but was closer to the 35S TATA element (TATATAAG). Our results with the above promoters suggest that the viral TATA element functions better in the 35S context as compared with the plant consensus TATA element. Because ModT promoter functions at par with the 35S promoter, it shows that sequence alterations in TATA-flanking regions seem to be tolerated, and an entirely novel synthetic stretch of DNA can provide the necessary framework for the TATA binding protein to bind and regulate the promoter activity.

On the basis of our observations, we propose that plant viral promoters have a different preferred TATA element (TATATAAG) that is more efficient than the plant consensus TATA element (TATATATA) in its own promoter context. This is further supported by a survey of several plant viral promoters where we observed that all viral promoters have TATA elements similar to or identical to that of 35S but different from those of plant promoters (data not shown). Conversely, out of a survey of 79 plant genes (Joshi, 1987), none of the promoters showed viral-like TATA boxes. This opens up an interesting question as to the efficiency of viral versus plant TATA element and also as to whether expression levels of plant promoters will improve if these carried a viral TATA element.

The results with the transient expression of Mod1A1 and Mod2A1 promoters show that the intervening 7 bp between TGACG repeats, as present in the *as-1* element of the 35S promoter, is necessary for optimal activity. In the case of subdomain A1,

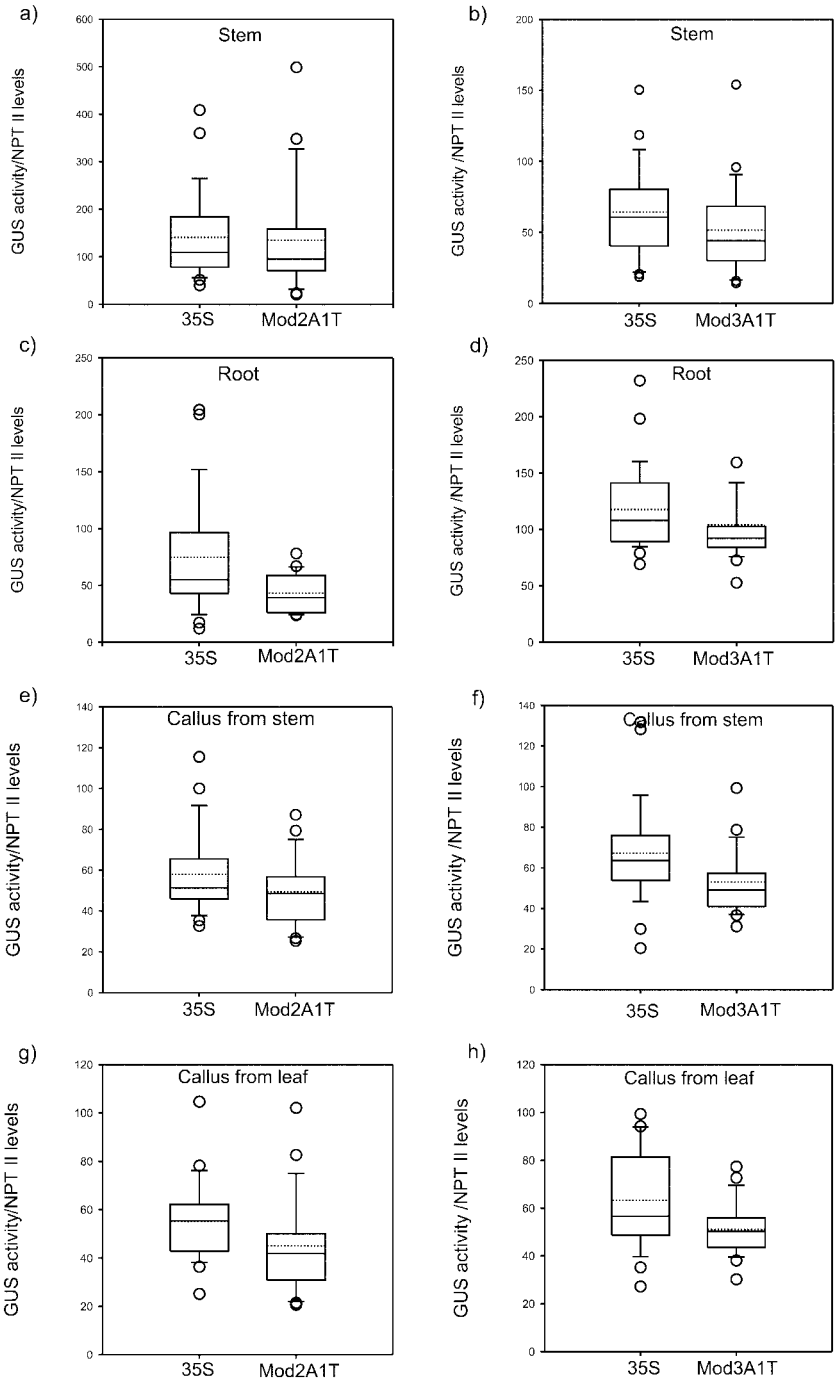

Figure 7. Comparison of normalized promoter activity (GUS activity/NPT II levels) between different tissues of transgenic plants with 35S and Mod2A1T promoter constructs (a, c, e, and g) as well as with 35S and Mod3A1T promoter constructs (b, d, f, and h). A minimum of 20 independent primary transformants grown in tissue culture has been used for expression analysis in stem (a and b) and root (c and d) tissues. Expression levels in callus raised from the stem (e and f) and callus raised from leaves (g and h) are represented. The comparisons have been depicted in a Box and Whisker plot. The horizontal lines in the Box and Whisker plot represent 10, 25, 50, 75, and 90 percentiles. The dotted line represents the mean of the data. Extreme values are depicted as small circles at the top and bottom of the plot.

Mod1A1 with a modified spacer and flanking sequence shows a 3-fold drop in activity. Mod2A1 with the *as-1* element and intervening sequence as present in the 35S promoter shows higher activity than 35S. The results obtained with Mod1A1 and Mod2A1 thus highlight the requirement for an accurate and functional definition of cis-elements in the design of synthetic promoters. Although the TGACG repeat element is important for ASF-1 binding (Lam et al., 1989), the intervening spacer sequence also plays a significant role in the functioning of the cis-element. On the basis of the present study and also an earlier observation (Bouchez et al., 1989), the functional moiety of subdomain A1 can be described as TGACG *TAAGGGA* TGAGC, rather than TGACG N_7 TGACG.

Overall, our data on transient expression with different constructs (Fig. 1) suggest that domain swapping may not be the most efficient method to develop promoters that are functionally equivalent. This is also reflected by the expression levels driven by the modified promoter ocsA1. Although ASF-1 can bind to the *ocs* element (Bouchez et al., 1989; Fromm et al.,

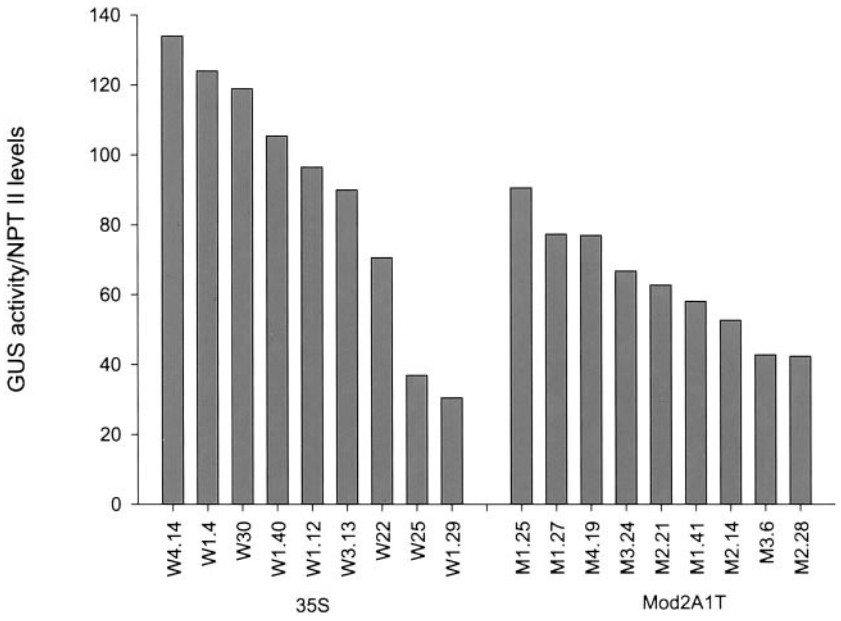

Figure 8. Comparison of normalized promoter activity (GUS activity/NPT II levels) of the 35S and Mod2A1T promoters in seedlings of the T_1 generation obtained from independent primary transformants. Each bar represents promoter activity observed in extracts made from approximately 150 seedlings (7 d old) from each line.

1989), in the modified promoter ocsA1, the spatial arrangements with respect to other transcription factors may not be conducive for proper protein-protein interaction. This is a possible explanation for the observed drop in the promoter activity.

In comparison, the development of synthetic promoters by placing properly characterized cis-elements in a novel stretch of DNA (i.e. promoter ModT and Mod2A1) seems to be a viable strategy for creating promoters with comparable activity to the parent wild-type 35S promoter. To confirm the results obtained with transient assays, we extended the study by developing constructs Mod2A1T and Mod31AT by modifications of domain A of the 35S. We tested the activity of the two promoters in different tissues of the primary tobacco transformants. We observed that both the modified promoters were active in all of the different tissues that were tested. The stability of expression was confirmed with the study of the T_1 generation of transgenics containing the Mod2A1T promoter.

Several synthetic promoters have been created earlier by (a) developing hybrids between two existing promoters (Comai et al., 1990; Ni et al., 1995), (b) placing identified cis-elements in conjunction with heterologous promoters (Rushton et al., 2002), and (c) bringing together cis-elements from different promoters (Last et al., 1991; Sawant et al., 2001). Such promoters were developed either to enhance promoter activity or to add additional properties to the already existing promoters. However, in all of these studies, multiple copies of cis-elements have been used to achieve promoter activity. The repetitious use of the same cis-element in a promoter may impede the expression of endogenous genes because such promoters may titrate away transcription factors. Therefore, we feel the best alternative is to create synthetic promoters whose organization in terms of cis-elements is similar to the already existing promoters but shows no homology in the intervening regions.

An interesting feature of the modified promoters Mod2A1T and Mod3A1T is the removal of a recombination hotspot in the mp region of the 35S promoter. This hotspot has been reported to cause rearrangements in the transgene cassettes and thus impairs gene expression (Kohli et al., 1999; Iyer et al., 2000). This feature can also make these synthetic promoters more suitable for stable gene expression.

The work reported here also has an interesting implication on the evolution of promoters. Because intervening regions between the cis-elements in promoters show great flexibility toward nucleotide changes as observed in the present study, new cis-elements could evolve in these regions without impairing the existing function of the promoter. This may lead to changes in both temporal and spatial expression of the promoter leading to novel evolutionary possibilities.

Our observations of domain A can be extended to modifications of domain B once the cis-elements of this domain are properly identified. In conclusion, our study shows that it is possible to create a series of functional 35S-like promoters that can be used to drive different transgenes in a transgenic plant. This will help to express a large number of independent genes while reducing the possibility of HBGS as these promoters would have reduced sequence homology. Promoter inactivation has been shown to occur by de novo DNA methylation (Mette et al., 2000). However, it is still not clear how this sequence-specific methylation gets extended to homologous regions at independent loci in the genome. Several mechanisms like ectopic DNA-DNA pairing (Muskens et al., 2000) and

methylation via aberrant RNA (Wassenegger and Pelissier, 1998) have been postulated. The baseline of all of these mechanisms is having homology between sequences. Our approach of designing functionally equivalent promoters with reduced homology could be an important step toward circumventing the problem of HBGS. However, such synthetic promoters will have to be tested for their functionality and also for their susceptibility to silencing by crosses with genetic stocks containing silenced loci.

MATERIALS AND METHODS

Assembly of Modified Promoters

Modified promoters were assembled using recursive PCR (Dillon and Rosen, 1990) followed by subsequent cloning in pPCR Script SK(+). DNA sequence of the synthesized promoters was confirmed by sequencing. All of the clonings followed standard protocols (Sambrook et al., 1989). The modified promoters (*Ptest*; Fig. 2) were cloned as *Hin*dIII-*Nco*I fragments upstream to the *gus* reporter gene with a 35S poly(A) signal in pPCR Script SK(+) (Stratagene, La Jolla, CA). The *Ptest gus* expression cassette thus developed was cloned as a *Sac*I-*Sal*I fragment in the binary vector pPZP200 (Hajdukiewicz et al., 1994) containing a *Pnos*-driven *npt II* gene as a selection marker for plant transformation at an *Eco*RI site and *cat* gene under the control of 35S (*P35S*) promoter at the *Hin*dIII site (Fig. 2).

Transient Expression Analysis

Protoplasts were isolated from leaves of axenic cultures of 4- to 5-week-old tobacco (*Nicotiana tabacum* cv Xanthi) plants with a few modifications in protocols described earlier (Negrutiu et al., 1987, 1990; Prols et al., 1988; Morgan and Ow, 1995). All procedures pertaining to isolation and DNA uptake were performed under sterile conditions. DNA samples used for uptake were purified on CsCl density gradients.

The leaves were incubated in K_3AS medium for 1 h before incubation in the digestion solution. MaCa solution (0.5 M mannitol, 15 mM $CaCl_2 \cdot 2H_2O$, and 0.1% [w/v] MES; Negrutiu et al., 1990) was used to resuspend and count the protoplasts. Aliquots of 1×10^6 protoplasts in 0.5 mL of MaCa were heat shocked at 45°C for 5 min and incubated at room temperature for 5 min. Plasmid DNA (10 μg) and sheared salmon sperm DNA (50 μg) were added to the suspension, followed by the addition of 0.5 mL of PEG solution (40% [w/v] PEG 1450 in 0.4 M mannitol and 0.1 M $Ca(NO_3)_2.4H_2O$, pH 7–9). Incubation was carried out at room temperature for 5 min. The sample was then diluted to 10 mL with W_5 medium and spun at 200*g* for 5 min. The pellet was resuspended in 10 mL of liquid Murashige and Skoog medium with 2 $\mu g\ mL^{-1}$ α-naphthalene-acetic acid and 0.5 $\mu g\ mL^{-1}$ 6-benzylamino-purine, transferred to petri plates, and incubated for 24 h (28°C) in dark, after which transient expression assays were carried out.

Development of Transgenics

Binary vectors were mobilized into the disarmed *Agrobacterium* sp. strain GV2260 by electroporation. *Agrobacterium* sp.-mediated transformation of leaf-disc explants was carried out following the protocol of Svab et al. (1995). With each construct, 22 to 27 independent transgenic lines were developed. Transgenics were grown in a growth chamber (16-h day and 8-h night, 28°C ± 2°C, relative humidity 70%).

For the analysis of expression of different tissues, mainly stem and roots, the plants were grown in tissue culture room (16-h day and 8-h night, 28°C ± 2°C) in glass bottles, and stem and roots were harvested 35 to 40 d after subculturing. Callus was raised from stem and leaf tissue by placing the explants on Murashige and Skoog agar supplemented with 2 $mg\ L^{-1}$ α-naphthalene-acetic acid and 0.5 $mg\ L^{-1}$ 6-benzylamino-purine. Extracts were prepared after 14 to 16 d of callusing.

Selfed (T_1) seeds were collected from growth chamber-grown primary transformants (T_0) and germinated on germination paper. Approximately 150 7-d-old seedlings were taken for making total protein extracts.

Enzyme Assays for Estimation of Promoter Strength

Total protein from protoplasts or different tissues of transgenic plants or seedlings was extracted in GUS extraction buffer (Jefferson, 1987). Protein concentration was estimated following Bradford (1976). Fluorometric GUS assays using 4-methyl umbelliferyl β-D-glucuronide substrate were done according to Bilang et al. (1994) in the case of the transient system and according to Jefferson (1987) for transgenics. The product released (methylumbelliferone) was estimated with DyNA Quant 200 Fluorometer (Hoeffer Pharmacia Biotech, San Francisco). GUS activity was expressed as picomoles of methylumbelliferone per minute per microgram of protein.

The amount of CAT and NPT II protein in total protein extracts was measured by ELISA using kits from Roche Diagnostics (catalog no. 1363727, Mannheim, Germany) and Agdia Incorporated (catalog no. PSP 73000, Indiana), respectively.

Distribution of Materials

Upon request, all novel materials described in this publication will be made available in a timely manner for noncommercial research purposes, subject to the requisite permission from any third-party owners of all or parts of the material. Obtaining any permission will be the responsibility of the requestor.

ACKNOWLEDGMENTS

We thank Dr. Ramneek Gupta for developing the program DNASEQ. We also thank Taru Gautam and Chandra Mohan Khantwal for their technical assistance.

Received January 18, 2003; returned for revision February 27, 2003; accepted March 4, 2003.

LITERATURE CITED

An G, Ebert PR, Yi B-Y, Choi C-H (1986) Both TATA box and upstream regions are required for the nopaline synthase promoter activity in transformed tobacco cells. Mol Gen Genet **203:** 245–250

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

Benfey PN, Ren L, Chua N-H (1989) The CaMV 35S enhancer contains at least two domains which can confer different developmental and tissue-specific expression patterns. EMBO J **8:** 2195–2202

Bilang R, Kloti A, Schrott M, Potrykus I (1994) PEG-mediated direct gene transfer and electroporation. *In* SB Gelvin, RA Schilperoort, eds, Plant Molecular Biology Manual, Ed 2, Supplement 1, Section A1. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1–16

Bouchez D, Tokuhisa JG, Llewellyn DJ, Dennis ES, Ellis LG (1989) The ocs-element is a component of the promoters of several T-DNA and plant viral genes. EMBO J **8:** 4197–4204

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

Comai L, Moran P, Maslyar D (1990) Novel and useful properties of a chimeric plant promoter combining CaMV 35S and MAS elements. Plant Mol Biol **15:** 373–381

Dey N, Maiti I (1999) Structure and promoter/leader deletion analysis of mirabilis mosaic virus (MMV) full-length transcript promoter in transgenic plants. Plant Mol Biol **40:** 771–782

Dillon PJ, Rosen CA (1990). A rapid method for the construction of synthetic genes using the polymerase chain reaction. Biotechniques **9:** 298–299

Donald RGK, Cashmore AR (1990) Mutation of either G box or I box sequences profoundly affects expression from the *Arabidopsis rbcS-1A* promoter. EMBO J **9:** 1717–1726

Ellis JG, Llewellyn DJ, Walker JC, Dennis ES, Peacock WJ (1987) The *ocs* element: a 16 base pair palindrome essential for activity of the octopine synthase enhancer. EMBO J **6:** 3203–3208

Fire A (1999) RNA-triggered gene silencing. Trends Genet **15:** 358–363

Fluhr R, Moses P, Morelli G, Coruzzi G, Chua N-H (1986) Expression dynamics of the pea *rbcS* multigene family and organ distribution of the transcripts. EMBO J **5:** 2063–2071

Fromm H, Katagiri F, Chua N-H (1989) An octopine synthase enhancer element directs tissue-specific expression and binds ASF-1, a factor from tobacco nuclear extracts. Plant Cell **1:** 977–984

Gidoni D, Bond-Nutter D, Brosio P, Jones J, Bedbrook J, Dunsmuir P (1988) Coordinated expression between two photosynthetic petunia genes in transgenic plants. Mol Gen Genet **211:** 507–514

Gidoni D, Brosio P, Bond-Nutter D, Bedbrook J, Dunsmuir P (1989) Novel *cis*-acting elements in petunia *cab* gene promoters. Mol Gen Genet **215:** 337–344

Hajdukiewicz P, Svab Z, Maliga P (1994) The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. Plant Mol Biol **25:** 989–994

Hamilton AJ, Baulcombe DC (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. Science **286:** 950–952

Iglesias VA, Moscone EA, Papp I, Michalowski S, Phelan T, Spiker S, Neuhuber F, Matzke M, Matzke AJM (1997) Molecular and cytogenetic analysis of stably and unstably expressed transgene loci in tobacco. Plant Cell **9:** 1251–1264

Iyer LM, Kumpatala SP, Chandrasekharan MB, Hall TC (2000) Transgene silencing in monocots. Plant Mol Biol **43:** 323–346

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

Joshi CP (1987) An inspection of the domain between putative TATA box and translation start site in 79 plant genes. Nucleic Acids Res **15:** 6643–6653

Katagiri F, Lam E, Chua N-H (1989) Two tobacco DNA-binding proteins with homology to the nuclear factor CREB. Nature **340:** 727–730

Kim Y, Buckley K, Costa MA, An G (1994) A 20 nucleotide upstream element is essential for the nopaline synthase (*nos*) promoter activity. Plant Mol Biol **24:** 105–117

Kohli A, Griffiths S, Palacios N, Twyman RM, Vain P, Laurie DA, Christou P (1999) Molecular characterization of transforming plasmid rearrangements in transgenic rice reveals a recombination hotspot in the CaMV 35S promoter and confirms the predominance of microhomology mediated recombination. Plant J **17:** 591–601

Krebbers E, Seurinck J, Herdies L, Cashmore AR, Timko MP (1988) Four genes in two diverged subfamilies encode the ribulose-1,5-bisphosphate carboxylase small subunit polypeptides of *Arabidopsis thaliana*. Plant Mol Biol **11:** 745–759

Kuhlemeir C, Cuozzo M, Green PJ, Goyvaerts E, Ward K, 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

Lam E (1994) Analysis of tissue-specific elements in the CaMV 35S promoter. *In* W Hennig, L Nover, U Scheer, eds, Results and Problems in Cell Differentiation 20. Springer-Verlag, Berlin, pp 181–196

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

Last DI, Brettell RIS, Chamberlain DA, Chaudhary AM, Larkin PJ, Marsh EL, Peacock WJ, Dennis ES (1991) pEmu: an improved promoter for gene expression in cereal cells. Theor Appl Genet **81:** 581–588

Matzke MA, Matzke AJM (1995) How and why do plants inactivate homologous (trans) genes? Plant Physiol **107:** 679–685

Matzke MA, Matzke AJM (1998) Epigenetic silencing of plant transgenes as a consequence of diverse cellular defence responses. Cell Mol Life Sci **54:** 94–103

Mette MF, Aufsatz W, van der Winden J, Matzke MA, Matzke AJM (2000) Transcriptional silencing and promoter methylation triggered by double-stranded RNA. EMBO J **19:** 5194–5201

Meyer P, Saedler H (1996) Homology-dependent gene silencing in plants. Annu Rev Plant Physiol Plant Mol Biol **47:** 23–48

Mikami K, Tabata T, Kawata T, Nakayama T, 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

Mol LJNM, Stuitje AR, van der Krol A (1989) Genetic manipulation of floral pigmentation genes. Plant Mol Biol **13:** 287–294

Morgan MK, Ow DW (1995) Polyethylene glycol-mediated transformation of tobacco leaf mesophyll protoplasts: an experiment in the study of Cre-*lox* recombination. *In* P Maliga, DF Klessig, AR Cashmore, W Gruissem, JE Varner, eds, Methods in Plant Molecular Biology, A Laboratory Course Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 1–16

Muskens MWM, Vissers APA, Mol JNM, Kooter JM (2000) Role of inverted DNA repeats in transcriptional and posttranscriptional gene silencing. Plant Mol Biol **43:** 243–260

Napoli C, Lemieux, Jorgensen R (1990) Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes *in trans*. Plant Cell **2:** 279–289

Negrutiu I, Dewulf J, Pietrzak M, Botterman J, Rietveld E, Wurzer-Figurelli EM, Ye D, Jacobs M (1990) Hybrid genes in the analysis of transformation conditions: II. Transient expression versus stable transformation: analysis of parameters influencing gene expression levels and transformation efficiency. Physiol Plant **79:** 197–205

Negrutiu I, Shillito R, Potrykus I, Biasini G, Sala F (1987) Hybrid genes in the analysis of transformation conditions: I. Setting up a simple method for direct gene transfer in plant protoplasts. Plant Mol Biol **8:** 363–373

Ni M, Cui D, Einstein J, Narasimhulu S, Vergara CE, Gelvin SB (1995) Strength and tissue specificity of chimeric promoters derived from the octopine and mannopine synthase genes. Plant J **7:** 661–676

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

Peach C, Velten J (1991) Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters. Plant Mol Biol **17:** 49–60

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

Rushton PJ, Reinstadler A, Lipka V, Lippok B, Somssich IE (2002) Synthetic plant promoters containing defined regulatory elements provide novel insights into pathogen- and wound-induced signaling. Plant Cell **14:** 749–762

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sanger M, Daubert S, Goodman RM (1990) Characteristics of a strong promoter from figwort mosaic virus: comparison with the analogous 35S promoter from cauliflower mosaic virus and the regulated mannopine synthase promoter. Plant Mol Biol **14:** 433–443

Sawant SV, Singh PK, Gupta SK, Madanala R, Tuli R (1999) Conserved nucleotide sequences in highly expressed genes in plants. J Genet **78:** 123–131

Sawant S, Singh PK, Madanala R, Tuli R (2001) Designing of an artificial expression cassette for the high-level expression of transgenes in plants. Theor Appl Genet **102:** 635–644

Singh K, Dennis ES, Ellis JG, Llewellyn DJ, Tokuhisa JG, Wahleithner JA, Peacock WJ (1990) OCSBF-1, a maize ocs enhancer binding factor: isolation and expression during development. Plant Cell **2:** 891–903

Steimer A, Amedeo P, Afsar K, Franz P, Mittelsten Scheid O, Paszkowski J (2000) Endogenous targets of transcriptional gene silencing in Arabidopsis. Plant Cell **12:** 1165–1178

Svab Z, Hajdukiewicz P, Maliga P (1995) Generation of transgenic tobacco plants by cocultivation of leaf discs with *Agrobacterium* pPZP binary vectors. *In* P Maliga, DF Klessig, AR Cashmore, W Gruissem, JE Varner, eds, Methods in Plant Molecular Biology. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 55–77

van der Krol AR, Mur LA, Beld M, Mol JNM, Stuitje AR (1990) Flavonoid genes in *Petunia*: Addition of a limited number of gene copies may lead to a suppression of gene expression. Plant Cell **2:** 291–299

Vaucheret H, Fagard M (2001) Transcriptional gene silencing in plants: targets, inducers and regulators. Trends Genet **17:** 29–35

Wassenegger M, Pelissier T (1998) A model for RNA-mediated gene silencing in higher plants. Plant Mol Biol **37:** 349–362

Ye X, Al-Babili S, Kloti A, Zhang J, Lucca P, Beyer P, Potrykus I (2000) Engineering the pro-vitamin A (β-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. Science **287:** 303–305