The EMBO Journal vol.9 no.7 pp.2029 – 2035, 1990

En/Spm encoded tnpA protein requires a specific target sequence for suppression

Sarah R.Grant, Alfons Gierl and Heinz Saedler

Max-Planck-Institut für Züchtungsforschung, Abteilung Molekulare Pflanzengenetik, Carl-von-Linne Weg 10, 5000 Köln 30, FRG

Communicated by H.Saedler

The En/Spm encoded suppressor function has been reconstituted in transgenic tobacco protoplasts. The suppressor affects genes which contain an En/Spm responsive transposable element in the transcribed sequences. The En/Spm encoded protein tnpA binds a defined *cis* element in the inserted transposon, repressing expression of the adjacent gene. This was shown by monitoring transient expression of a bacterial marker gene (GUS) expressed from a strong plant viral promoter. Suppressible variants of the marker gene were produced by inserting I element sequences into the untranslated sequences of the GUS transcript. Comparison of transient expression of these variants in wildtype tobacco protoplasts with their expression in protoplasts transgenic for tnpA protein demonstrates that tnpA is the suppressor. In addition, the minimal *cis* element required for suppression has been defined as a dimer consisting of two 12 bp tnpA binding sequences in a particular inverted orientation. One of these dimers occurs in each En/Spm end close to the characteristic 13 bp terminal inverted repeat. TnpA binding sites in different arrangements do not respond as well to tnpA. The implications of this observation are discussed. This system can be used to analyse tnpA – DNA interactions involved in gene regulation further.

Key words: En/repression/Spm/suppression/tnpA/transient expression/transposable element

Introduction

Maize transposable elements of the En/Spm family can be either autonomous, controlling their own transposition, or they can be non-autonomous internal deletion derivatives (I elements) which cannot encode transposition functions but respond to En/Spm transposase (Peterson, 1953; Pereira *et al.*, 1986). In addition to the factors required for transposition, autonomous En/Spm elements encode a *trans*-acting regulatory function which inhibits expression of genes with inserted I elements. McClintock named this function 'Suppressor' (McClintock, 1954). In this discussion, we use the word 'suppression' as McClintock did, to mean reduction of gene expression similar to repression of bacterial genes. We are not referring to classical suppression in which the phenotype of a mutant allele is restored to wildtype by the effect of another mutation.

En/Spm suppression probably affects all genes with I element insertions. However, suppression can only be observed phenotypically when it affects genes which remain detectably expressed even though they contain I element inserts. Such a supressible allele is shown in Figure 1. The *a1-m1* (5719A1) allele of the *A1* gene encoding an enzyme involved in synthesis of the purple pigment, anthocyanin, contains a 789 bp I element (Schwarz-Sommer *et al.*, 1985). Unlike some other I element inserts, its presence reduces, but does not abolish, *A1* expression. Kernels homozygous for *a1-m1* (5719A1) have a pale purple colour. The effect of an inserted I element on gene expression depends on its internal sequence and on its position within the affected gene (Schwarz-Sommer *et al.*, 1987). Several other suppressible alleles which contain I elements and yet are expressed in the absence of autonomous En/Spm elements, have been described (McClintock, 1965; Reddy and Peterson, 1976; Nelson and Klein, 1984). Analysis of transcripts of suppressible alleles of the *Bz* gene shows that almost all of the I element sequences are removed by splicing, restoring the *Bz* reading frame (Kim *et al.*, 1987). Similar effects may allow expression of other suppressible maize alleles (Tacke *et al.*, 1986).

In the presence of an autonomous En/Spm element, expression of alleles with I element inserts is fully repressed regardless of the structure of the I element or its position in the gene, as long as it is located in the transcribed sequence. In the case of the *a1-m1* (5719A1) allele shown in Figure 1, the suppressor from an autonomous En/Spm, element introduced by crossing inhibits residual *A1* expression so that the majority of aleurone cells are unpigmented. Pigment is produced only in cells where the transposase function has removed the I element, thereby restoring the structure of the *A1* gene and its En/Spm independent expression.

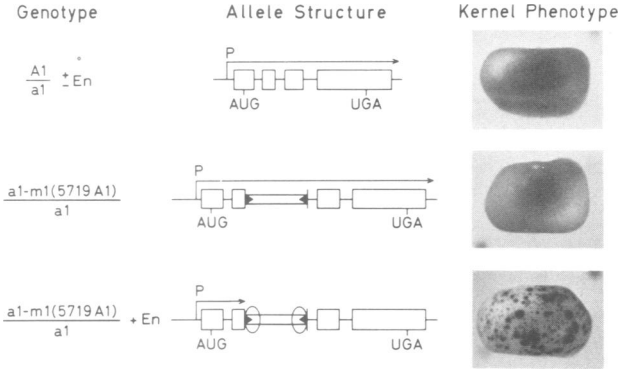


Fig. 1. Example of suppression. Genotypes are presented at the left, followed by the exon – intron structures of two alleles of the maize *A1* gene, the wildtype allele above and the suppressible allele *a1-m1*(5719A1) below. Large open boxes represent exons, P with an arrow below represents transcription from the *A1* promoter. The inserted I element is bound by filled triangles representing the 13 bp En/Spm terminal inverted repeats. Open ovals represent the En/Spm tnpA protein binding to the ends of the I element. Maize kernels demonstrating the phenotypes of these alleles with or without an autonomous En element *in-trans* are shown to the right.

A model of suppression has been proposed (Gierl *et al.*, 1985, 1988a; Schwarz-Sommer *et al.*, 1985). Suppression resembles a negative regulatory circuit in which an En/Spm encoded protein acts as a repressor. The repressor recognizes and binds a specific *cis* element located in the repetitive sequences characteristic of En/Spm element termini. The bound protein is thought to hinder sterically progression of RNA polymerase through the gene, resulting in prematurely terminated transcripts.

The autonomous En element has been cloned and sequenced (Pereira *et al.*, 1986) and four distinct RNA products identified, a 2.5 kb transcript named *tnpA* and and three transcripts 5–6 kb in size (Pereira *et al.*, 1986; Masson *et al.*, 1989). The *tnpA* encoded protein has been expressed in *Escherichia coli* and shown to bind DNA specifically *in vitro*. It recognizes a 12 bp sequence repeated several times in the subterminal repeated sequences found at the ends of all En/Spm elements (Gierl *et al.*, 1988a).

Two mutant En/Spm elements have been described which have sustained a deletion which affects the 5–6 kb transcripts but does not prevent tnpA expression. Because both of these are defective in transposition but still encode the suppressor function, tnpA protein is thought to perform the suppressor function (Masson *et al.*, 1987; Gierl *et al.*, 1988b). However, it is not clear from genetic studies that these mutants do not produce some aberrant proteins in addition to tnpA which may compensate for the activities of the 5–6 kb transcripts. In addition, these mutants can stimulate expression of cryptic En/Spm elements, found in almost every maize line (Banks *et al.*, 1988). It is possible that they accomplish suppression indirectly, by stimulating expression of genes from formerly inactive En/Spm elements. Therefore, in order to define *cis*- and *trans*-acting components of the suppressor function, we chose to reproduce suppression in tobacco, which has no obvious En/Spm homologous elements and is easily transformed and regenerated.

Results

To recreate suppression in tobacco, we combined the putative En/Spm suppressor factor, tnpA protein, with a responsive marker gene carrying a tnpA binding *cis* element from En/Spm. We compared transient expression of the marker gene in protoplasts in the presence or absence of tnpA protein.

A tnpA expressing line of tobacco

The cloned cDNA (Pereira *et al.*, 1986) for the tnpA transcript was inserted in either orientation between the TR2 promoter and the *nos* poly(A) addition sequences in pPCV720, a bifunctional plasmid carrying a plant specific hygromycin resistance gene (gift of Csaba Koncz and Jeff Schell). M-phase synchronized tobacco protoplasts (Meyer *et al.*, 1985) were transformed with these two constructs. Hygromycin resistant regenerants were selected and screened for the presence of tnpA homologous transcripts by Northern blot hybridization with single stranded probes (data not shown). A plant was chosen which produced a tnpA homologous transcript of the correct size. A second plant was chosen which produced tnpA antisense trancripts to initiate a negative control line. To analyse the regulatory effects of tnpA, we compared transient expression of the marker gene in protoplasts from the tnpA expressing line (tnpA protoplasts) with expression in protoplasts from the original untransformed tobacco line (wildtype protoplasts) and the antisense tnpA expressing line (anti-tnpA).

A suppressible marker gene

As in maize, suppression could only be monitored in tobacco if the responsive marker gene was detectably expressed in the absence of the suppressor. We chose a strongly expressed marker gene to measure the effects of suppression. Expression of the *E.coli* β-glucuronidase gene (GUS) (Jefferson *et al.*, 1986) from the promoter of the 35S transcript of the cauliflower mosaic virus can be measured in a quantitatively reliable assay of enzymatic activity (Jefferson *et al.*, 1987). Since this assay is very sensitive, it would be possible to detect GUS activity even if insertion of putative *cis* elements reduced gene expression in wildtype protoplasts to $<1\%$ of the activity of the original 35S–GUS construct. On average, 35S–GUS activity was in the range of 1000 pmol methyl umbelliferone released per mg plant extract protein per minute time. Activities as high as 20 000 pmol/mg/min were seen in the most active protoplasts. The GUS gene without a plant promoter produced activities of 5 pmol/mg/min or less. To avoid complete disruption of GUS expression, we inserted the elements in the untranslated RNA leader 5′ of the open reading frame, or 3′, between the termination codon and the 35S polyadenylation site. The structure of the 35S–GUS gene is shown on the pRT102 plasmid in Figure 2.

Comparison of expression and response to tnpA

Each construct was transformed into protoplasts from the wildtype and tnpA plants, and, in two experiments, into anti-tnpA protoplasts. In each experiment, tnpA and wildtype protoplasts were prepared from plants of the same age. Protoplasts of each genotype were pooled and divided into aliquots to transform with individual constructs. One aliquot of each preparation was transformed with pRT102 and one with pNP-GUS (see Figure 4) as standards for competence to express GUS and for background GUS activity, respectively. In two independent determinations of several different constructs transformed in triplicate into

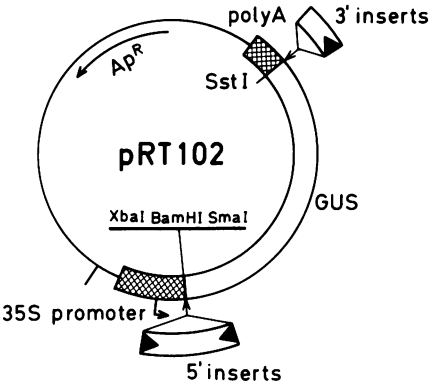


Fig. 2. The plasmid pRT102 used as the source of the wildtype GUS allele. Unique restriction sites used for insertion of I element sequences 5′ or 3′ of the GUS open reading frame (represented as an open box) are shown. The hatched boxes represent the 35S promoter and polyadenylation sequences. Ap^R refers to the bacterial β-lactamase gene. I element sequences are represented as in Figure 1.

simultaneously prepared protoplasts of a unique or of different genotypes, GUS activity varied with standard deviations of <35% of the mean (data not shown). In contrast, the absolute values of GUS expression from each construct varied greatly between experiments performed on different days. However, the relative pattern of strongly or weakly expressing constructs was consistent in all experiments. Importantly, the pattern of response to tnpA conferred by the presence of the individual *cis* elements was also consistent in all experiments. Therefore, for each experiment, we represented the GUS activity of each construct as a percentage of the activity from the 35S−GUS construct in wildtype protoplasts in that experiment. The averages of these percentage values are presented in Figures 4 and 5. Because of the relatively high variation, only responses to tnpA of >50% reduction in GUS expression in tnpA protoplasts versus wildtype protoplasts in every experiment were considered to be significant.

In some experiments, a second plasmid, pUC9CAT, bearing a bacterial chloramphenicol acetyl transferase (CAT) gene expressed from the 35S promoter and termination signals was cotransformed with each GUS construct as a control for the viability and expression competence of each protoplast aliquot. After incubation, half of each transformed aliquot was tested for either GUS or CAT activity. The results were determined as a quotient of GUS activity divided by CAT activity in the same aliquot. Results of one representative experiment are shown in Figures 4 and 5. The value of the 35S−GUS/35S−CAT was set to 100 for easy comparison with the results expressed as a percentage of 35S−GUS activity. CAT activity in this experiment was quite variable between transformations, ranging from 700 to 3500 pmol chloramphenicol acetylated per mg extract protein per minute. However, it was not correlated to the tnpA response of the cotransformed GUS plasmids. Therefore, activity measured as GUS/CAT was similar to activity measured as a percentage of 35S−GUS activity (Figures 4 and 5).

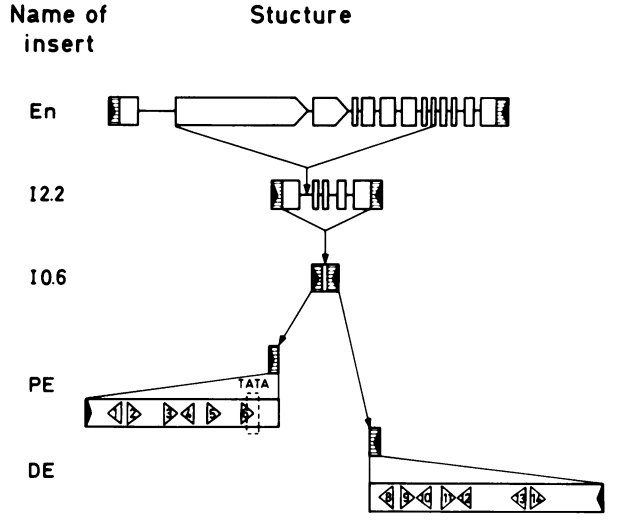


Fig. 3. I element sequences inserted into the 35S−GUS gene. The top line represents the exon−intron structure of an autonomous En element. Open boxes represent the exons of the tnpA transcript. Open arrows represent two open reading frames found with the tnpA exons in larger (5−6 kb) En/Spm transcripts. Striped boxes represent the En/Spm subterminal repeated sequences and filled triangles represent the 13 bp terminal perfect inverted repeat. Thin arrows indicate junctions between deleted sequences. PE and DE are reproduced at larger scale. The 12 bp long tnpA binding sites are represented as triangles containing the number assigned to each site by Gierl *et al.* (1988a). The TATA box of the En/Spm promoter is indicated to overlap the tnpA binding motif, no. 6.

The constructs which most strongly responded to tnpA were subcloned into the pUC9CAT plasmid, always in the same orientation. GUS and CAT expression of these plasmids were compared to ensure that the tnpA response was specific to the En/Spm sequence bearing GUS gene. Again, CAT expression was high (2000−4000 pmol chloramphenicol acetylated per mg protein per minute) and not correlated to the tnpA response of the linked GUS constructs (Figures 4 and 5 in parentheses).

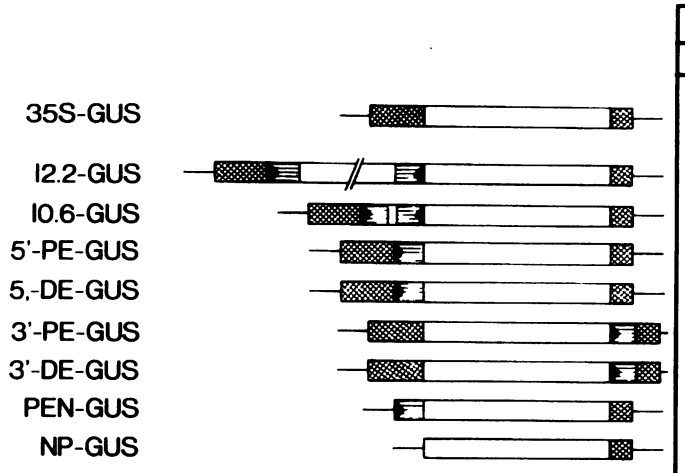
TnpA suppresses expression of a gene with a small I element insert

The largest insert tested was a naturally derived 2.2 kb I element, I2.2 (Figure 3). Insertion of this element 5′ of the GUS open reading frame reduced transient GUS expression in wildtype tobacco to 0.4% of 35S−GUS expression, or less. This expression was slightly higher than GUS activity from the NP−GUS construct which has no plant promoter (Figure 4). However, only a slight reduction of I2.2−GUS expression was observed in tnpA protoplasts. Thus, residual expression from this construct appears to be already too low to detect a measurable response to tnpA.

In order to create an I element which would allow higher expression of the surrounding transcript than I2.2, we removed an internal fragment from I2.2 leaving the subterminal repetitive sequences intact to form a 641 bp I element (I0.6, Figure 3). Variants containing this insert in either orientation in the 5′ untranslated region of the GUS gene expressed up to 10 times as much GUS activity as I2.2−GUS in wildtype protoplasts (Figure 4). Significantly, in tnpA protoplasts, expression of the I0.6−GUS allele was reduced to 20% (differences ranged from 10% to 30%) of its expression in wildtype (or anti-tnpA) protoplasts. The level of expression in the presence of tnpA is essentially equivalent to expression of I2.2−GUS.

One end of an I element is sufficient for tnpA suppression

To characterize the minimal sequences required *in-cis* for suppression, we began by inserting each of the repetitive ends of I0.6 into the 35S−GUS gene. Either the most 5′ 261 bp including the En promoter (PE for promoter end), or the most 3′ 242 bp (DE for distal end, Figure 3), were inserted into pRT102 in either of two positions, 5′ or 3′ of the GUS open reading frame (Figure 4, 5′-PE−GUS, 3′-PE−GUS, 5′-DE−GUS and 3′-DE−GUS). The ends were inserted with the 13 bp terminal inverted repeat closest to the 35S promoter (Figure 4) in order to mimic the natural situation in which RNA polymerase always encounters one of the 13 bp En/Spm termini first as it transcribes through a gene into an inserted element. Variants containing either end expressed more GUS protein in wildtype or anti-tnpA protoplasts than the I0.6−GUS variant. In tnpA protoplasts, GUS genes with either insert were consistently expressed at lower levels (Figure 4). The greatest response to tnpA was seen with the 5′-DE−GUS variant (Figure 4). Expression of this variant in tnpA protoplasts, was an average of 5% (ranging from 2% to 14%) of its expression



GUS EXPRESSION

	WILDTYPE		ANTI-TNPA	TNPA	
	%35S-GUS	GUS/CAT	%35S-GUS	%35S-GUS	GUS/CAT
35S-GUS	100 (100)	100 (100)	130	140 (115)	71 (75)
I2.2-GUS	0.19 (0.39)	0.15 (0.25)	ND	0.14 (0.14)	0.13 (0.11)
I0.6-GUS	2.0	0.78	2.4	0.44	0.16
5'-PE-GUS	6.7	ND	12	1.4	ND
5,-DE-GUS	5.8 (4.9)	18 (8.2)	7.8	0.28 (0.21)	0.27 (0.19)
3'-PE-GUS	12	19	ND	1.3	3.2
3'-DE-GUS	13	26	ND	1.5	1.8
PEN-GUS	1.9	ND	ND	2.6	ND
NP-GUS	0.042	0.0055	0.02	0.063	0.16

Fig. 4. Relative levels of GUS expression. The name and structures of GUS variants are shown on the left. Symbols used are as for Figures 2 and 3. The relative GUS expression in tobacco protoplasts derived from wildtype tobacco, the anti-tnpA line and the tnpA line is compared in the table to the right. Relative GUS expression is determined as a percentage of 35S−GUS expression in wildtype protoplasts (% 35S−GUS) or divided by the CAT expression from a cotransformed 35S−CAT gene (GUS/CAT). The % 35S−GUS values are an average of the results of four to five independent determinations, with the exception of those in anti-tnpA protoplasts which are the average of two determinations. GUS/CAT results are from one represenative experiment unless in parentheses. GUS/CAT of 35S−GUS are arbitrarily set to 100 and all other values were adjusted accordingly. Numbers in parentheses represent GUS activities from plasmids with the indicated GUS construct and a linked 35S−CAT gene, not cotransformed with pUC9CAT. Results are the average of five determinations in two experiments. ND refers to tests not done.

NAME OF INSERT SEQUENCE GUS EXPRESSION

		WILDTYPE		TNPA	
		%35S-GUS	GUS/CAT	%35S-GUS	GUS/CAT
mo-1	G CCGACACTCTTA A	55	44	73	18
mo-2	T TAAGAGTGTCGG C	50	45	62	18
ht-1	CCGACACTCTTA AAG CCGACACTCTTA A	134	70	134	76
ht-2	AT TAAGAGTGTCGG AAT TAAGAGTGTCGG A	105	ND	50	ND
t 1-2	ATT AAAGAGTGTCGG GG CCGACACTCTTA A	13 (30)	11 (22)	0.54 (2.5)	0.33 (1.0)
t 13-14	AT TAAGAGTGTCGG TTGA CCGACACTCTTA A	13 (25)	ND (10)	0.43 (0.81)	ND (1.1)
h 3-4	A CCGACACTCTTA ATT TAAGAGTGTCGG A	7.1 (14)	7.0 (16)	1.1 (1.9)	1.5 (1.1)
h 9-10	A CCGACACTCTTA TGT TAAAAGTGTCGG A	250	ND	121	ND
h 11-12	A CCGACGCTCTTA TGT TAAAAGTGTCGG A	158	ND	69	ND

Fig. 5. Relative GUS expression of alleles with oligonucleotide inserts. The mo-1 and mo-2 inserts contain a single consensus tnpA binding site inserted in opposite orientations with respect to the 35S−GUS gene. Ht-1 and ht-2 oligonucleotides contain two consensus tnpA binding sites in head-to-tail orientiation as shown. Oligonucleotides named 't' contain tail-to-tail dimers while 'h' refers to head-to-head dimers. The numbers in the names of the 't' and 'h' oligonucleotides refer to the number assigned to the tnpA binding sites in Figure 3. The 35S promoter would be to the left. Arrows above the sequence indicate the tnpA binding sites. The arrowheads are in the same direction as the points of triangles in Figure 3. The GUS activity in wildtype and tnpA protoplasts is presented as in Figure 4.

in wildtype and anti-tnpA protoplasts. Since the expression in tnpA protoplasts was as low as expression of the I2.2−GUS construct, the presence of the distal I element end, including the subterminal and terminal repetitive sequences, is sufficient for tnpA to suppress GUS expression completely. The interaction of both ends of an I element through DNA folding is not required for suppression.

Expression of the 5'-PE−GUS variant was also suppressed in tnpA protoplasts to 20% (5%−40%) of its expression in wildtype or anti-tnpA protoplasts. However, the level of expression from the 5'-PE−GUS variant in tnpA protoplasts was higher than from the 5'-DE−GUS construct. This expression, unresponsive to tnpA, could be initiated from the En/Spm promoter located in PE in the correct

orientation to initiate GUS transcription. In the presence of tnpA, 5′-PE−GUS expression was similar to expression of GUS from the En/Spm promoter in a related construct (PEN−GUS) in which PE is upstream of the GUS open reading frame but the 35S promoter is removed. Expression from the En promoter is not affected by tnpA in these assays (Figure 4).

Position independence

GUS expression was also suppressed in tnpA protoplasts if PE or DE was inserted 3′ of the GUS open reading frame. In tnpA protoplasts, both 3′-PE−GUS and 3′-DE−GUS were expressed at an average of 10% (4%−25%) of levels observed in wildtype protoplasts (Figure 4). Thus, as in maize, the response to suppression is independent of the position of the *cis* element within a gene. In contrast to the different magnitude of the tnpA response of GUS variants with either PE or DE 5′ of the GUS open reading frame, variants with either insert at the 3′ position were expressed at the same level in tnpA protoplasts.

One in vitro defined tnpA binding site is not a sufficient cis element for suppression

TnpA binds oligonucleotides specifically if they contain one copy of a 12 bp consensus binding sequence (Gierl *et al.*, 1988a). Therefore, we predicted that the 12 bp binding site could function as a minimal *cis* element for suppression. We inserted an oligonucleotide containing this sequence (mo-1, mo-2, Figure 5) in the *Sma*I site 5′ of the GUS open reading frame in pRT102 (Figure 2). However, the presence of the oligonucleotide in either orientation had no effect on GUS expression, either in tnpA or in wildtype protoplasts (as compared with expression of 35S−GUS with no inserted sequences) (Figure 5).

At least two tnpA binding sites, in inverted orientation are necessary to suppress gene expression

Oligonucleotides (Figure 5), containing dimers of the 12 bp tnpA binding sequence were inserted into pRT102 between the *Xba*1 site and the *Sma*1 sites at the GUS 5′ end (Figure 2). We designed two types of oligonucleotide dimers. One contains two direct repeats of the consensus 12 bp tnpA binding sequence. These repeats are separated by 3 bp (ht-1, ht-2, Figure 5). No similar directly repeated sequences exist in the ends of En/Spm elements. The other type consist of exact replicas of the five inverted-repeat-dimer tnpA binding sites found in the En/Spm subterminal repeats (t1−2, t13−14, h3−4, h9−10, h11−12, Figures 3 and 5). These were inserted into pRT102 in both possible directions. GUS expression is reported only for variants with sequences in the direction shown in Figure 5. The direction of the inserted oligonucleotides did not significantly affect the response to tnpA they conferred on the GUS variant (data not shown).

The relative orientation of the 12 bp tnpA binding sequences with respect to each other was the critical determinant for the response to tnpA. The presence of directly repeated tnpA binding sites (ht-1 and ht-2) had no significant effect on GUS expression in wildtype or tnpA protoplasts.

Variants containing the oligonucleotides illustrated in Figure 5 with tail-to-tail arrows (t1−2 and t13−14) were the most responsive to tnpA. In tnpA protoplasts, these produced ~5% (1.5%−15%) of the GUS activity produced in wildtype protoplasts. This response to tnpA was almost as strong as the response of the 5′-DE−GUS variant containing a complete I element subterminal repeat (Figure 4). In wildtype protoplasts, the presence of the tail-to-tail motifs also caused a significant inhibition of GUS expression compared with expression of 35S−GUS (which was further reduced in the presence of tnpA).

All other naturally derived dimer sequences have the inverted 12 bp tnpA binding sequences in the head-to-head orientation (h3−4, h9−10 and h11−12, Figure 5). These inserts had very different effects on GUS expression both in the presence and absence of tnpA. One head-to-head dimer (h3−4) responds to tnpA but more weakly than the tail-to-tail dimers. In tnpA protoplasts, a GUS gene containing h3−4 is expressed at ~20% (10%−30%) of its level of wildtype protoplasts. Similarly to the tail-to-tail dimers, h3−4 also reduces GUS expression in wildtype protoplasts. In contrast, the other two head-to-head dimers (H9−10 and h11−12) are not significantly responsive to tnpA mediated suppression. In addition, they have no negative effect on GUS expression in wildtype protoplasts.

Discussion

TnpA protein is the suppressor factor

By reproducing suppression in a transgenic, transient expression system, we have confirmed the proposal, based on analysis of mutant En/Spm elements in maize (Gierl *et al.*, 1988b), that tnpA protein performs the suppressor function. We have not yet addressed the functions of the products of the 5−6 kb En/spm transcripts in suppression. Although they are obviously not essential, one or all of them may play an accessory role.

Effects of the position of the cis element are similar in maize and tobacco

TnpA mediated suppression is similar in tobacco and maize in that expression of genes containing the *cis* element is suppressed regardless of its position in the transcribed sequences. In addition, the position and structure of the inserts have a profound effect on expression in the absence of tnpA in both tobacco and maize (Figure 4; Schwarz-Sommer *et al.*, 1987).

TnpA interacts with two inverted in vitro defined tnpA binding sites to suppress gene expression

Although only one 12 bp tnpA binding site is required in a DNA fragment for it to be bound specifically by tnpA protein *in vitro* (Gierl *et al.*, 1988a), tnpA requires at least two closely juxtaposed binding sites in inverted orientation to suppress gene expression. The relative orientation of tnpA binding sites with respect to one another is more important to suppression (Figure 5) than the surrounding sequences or the homology of the binding sites to each other or to their consensus sequence. Both tail-to-tail dimers respond strongly to tnpA although they differ in sequence between the tnpA binding sites and they differ in the sequences of the binding sites themselves. One of the tnpA binding sites in t1−2 (the left site shown in Figure 5) deviates from the consensus tnpA binding site by a T to A transversion at its left end while the right binding site and both sites in t13−14 have the consensus sequence.

One interpretation of the observations in Figure 5 is that

tnpA binds tail-to-tail dimers with higher affinity than head-to-head dimers, which in turn are bound with higher affinity than tandem dimers or monomers. Studies are in progress to test this possibility *in vitro*.

Alternatively, both binding sites in a dimer may have to be occupied in order to create a significant steric block to transcription. Assuming that tnpA binds its target in a directionally dependent manner, it is possible that tnpA protein is sufficiently asymmetric that two molecules can best fit onto two closely juxtaposed sites if in the tail-to-tail orientation.

A third possibility is that two tnpA molecules can bind any two closely juxtaposed sites, but the tnpA bound to a tail-to-tail dimer forms the structure which most effectively impedes the progress of RNA polymerase. For example, such a structure may present a larger face towards incoming RNA polymerase and be more difficult to displace.

We cannot exclude the possiblity that tnpA may bind RNA to function. However, because tnpA binds DNA *in vitro* (Gierl *et al.*, 1988a), we favour that idea that tnpA interacts with DNA to suppress gene expression. Further *in vitro* analysis of tnpA binding to RNA and DNA will be necessary to distinguish between the possibilities listed above.

Only one of the head-to-head dimers (h3−4) responds strongly to tnpA mediated suppression. H3−4 contains two consensus 12 bp binding sites while both h9−10 and h11−12 have a deviation from the consensus in at least one binding site (Figure 5) which may account for the weaker responses of the latter two.

Responsive cis elements reduce GUS expression in the absence of tnpA

The ability of the dimer motifs to inhibit GUS expression in wildtype protoplasts is also correlated with the homology of their tnpA binding sequences to the consensus sequence. This is especially obvious when comparing the effects of h3−4 with those of h9−10 and h11−12 on GUS expression in wildtype protoplasts (Figure 5). Endogenous tobacco proteins many exist which also bind the tnpA consensus sequence reducing gene expression (although tnpA represses gene expression more strongly when bound at these sites).

Homology of the inverted binding sites also implies that they form perfect palindormic sequences. Palindromic sequences in untranslated RNA can interfere with translation in some cases (Kozak, 1986). However, comparison of the calculated free energy of formation of hairpin structures from the inverted dimers does not indicate that RNA secondary structure is the sole cause of reduced GUS expression. The most theoretically stable hairpin is t13−14. Its free energy of formation (Δg^0), calculated according to Salser (1977), is −23.5, which, theoretically, is not sufficiently stable to interfere with translation initiation (Kozak, 1986). In addition the free energy of formation of a hairpin from t1−2 (Δg^0 of −15.7) is almost the same as that of h9−10 (Δg^0 of −15.6) but there is a great difference in their effects on GUS expression (Figure 5).

Role of complex arrangement of tnpA binding sites in En/Spm subterminal repeats in suppression

Each En/Spm subterminal repeat contains several tnpA binding sites in a particular arrangement (Figure 3). Surprisingly, 35S−GUS genes with tail-to-tail dimer inserts are expressed with only slightly higher activities in tnpA protoplasts than the 5′-DE−GUS construct with a full subterminal repeat (0.5%−2% of 35S−GUS expression from t1−2 or t13−14 versus ~0.2% from 5′-DE−GUS). The binding of tnpA to additional motifs in the subterminal repeats probably reinforces the suppressor effect. However, most of the response appears to be mediated only by tnpA binding to the outermost dimer motifs (t1−2 or t13−14). Therefore, the complicated arangement of binding sites in the En/Spm ends appears to be less important for suppression than for the role tnpA plays in transposition (Masson *et al.*, 1987, 1989; Gierl *et al.*, 1988a, 1989) and in autoregulation (Banks *et al.*, 1988; Gierl, *et al.*, 1988a).

The position of the tail-to-tail dimers closest to the En/Spm termini is intriguing. It ensures that I elements remain substrates for suppression even though they are subject to internal deletions, (Scheifelbein *et al.*, 1985; Schwarz-Sommer *et al.*, 1985; Tacke *et al.*, 1986). Suppression isolates transcription of the element from that of flanking DNA. The En/Spm promoter is relatively weak (Pereira *et al.*, 1986). For example, PEN−GUS expresses ~1% of the GUS protein produced by 35S−GUS (Figure 4). Since very frequent transposition would cause mutations at high frequency, high En/Spm expression may be harmful to the host plant. Transcripts initiated at stronger promoters in sequences flanking an En/Spm element are prevented from reading into the element by suppression, avoiding overexpression of transposition functions.

Similarly, suppression prevents expression of En/Spm antisense transcripts from flanking promoters and stops En/Spm transcripts from proceeding outside of the element boundaries to inappropriately express flanking DNA. However, other transposable elements, such as the Ac-Ds family of maize, do not have an obvious suppressor function. For example, an allele of the waxy locus which contains a Ds element, *wx*-m9, is expressed at a similar, low level in the absence or presence of an autonomous Ac element (Wessler *et al.*, 1987, 1988). It remains unclear whether suppression is an essential En/Spm function or simply an artifact of the transposition mechanism.

TnpA−DNA interaction can be monitored in transgenic tobacco

We have recreated En/Spm suppression in a simplified model system. The effects of mutations in the *cis*-acting inverted repeat or in the tnpA protein sequence can be measured using this rapid and reproducible transient assay system. Comparison of the effects of mutations will illuminate the mechanisms of tnpA−DNA interaction. It will also be possible to combine mutant *cis* and *trans* factors to identify mutations in one which can compensate for mutations in the other. Development of similar assays to recreate other tnpA functions in transgenic plants will allow us to analyse tnpA activity even further.

Since the only maize specific protein required for suppression is tnpA, it should be possible to recreate suppression in many different plant species. The tnpA recognition site is a small, defined unit which appears to function at various positions within a gene. Using the tobacco suppressor assay, we hope to identify altered *cis* elements which interfere minimally with expression of a surrounding gene but still respond strongly to tnpA. Introduction of such a *cis* element into a target gene would allow conditional repression of gene expression when the *tnpA* gene was

introduced by crossing or expression from an inducible promoter.

Materials and methods

Plasmid constructions

The tnpA and anti-tnpA plants were transformed with the following plasmids. The cloned *tnpA* cDNA (Pereira *et al.*, 1986) was isolated as a *Sal*I fragment and inserted in both orientations into the *Sal*I site of pPCV720. The structural CAT gene from pBR325 (Bolivar, 1978) was inserted as a *Sau*3A fragment into the *Bam*HI site of the 35S expression cassette of pDH51 (the gift of Douglas Hannahan) which was then subcloned into the *Hinc*II site of pUC9 to make pUC9CAT. The I element, I2.2 has been cloned from the *a1-ml* (6078) allele of maize (Schwarz-Sommer *et al.*, 1987). Most of the *A1* flanking sequences were removed by *Bal*31 deletion and the resulting DNA fragment was ligated to *Bam*HI linkers and inserted into the *Bam*HI site of pUC9 (Vieira and Messing, 1982) (William Martin, personal communication). I0.6 was derived by cutting the I2.2 containing pUC9 plasmid with *Bss*H2 and religating to remove an internal fragment. *Bam*HI fragments containing either I0.6 or I2.2 were inserted in both orientations into the *Bam*HI site of the 35S−GUS gene from pRT99-GUS (Töpfer *et al.*, 1988). The 35S−GUS genes with their respective I elements were inserted as *Hin*dIII−*Sal*I fragments into pBR322 (Sutcliff, 1978) with the 35S−CAT gene from pUC9CAT in the *Cla*I site. PE, the 261 bp *Bam*HI−*Sal*I fragment or DE, the 242 bp *Bss*H2−*Bam*HI fragment from I2.2 was inserted into the 35S−GUS gene in pRT102 (the gift of Reinhard Töpfer and Jeff Schell) either at the *Bam*HI−*Sma*I sites or the blunted *Sac*I site (see Figure 2). The GUS containing *Hin*dIII−*Eco*RI fragment from pRT99.GUS.JD (Schultze-Lefert *et al.*, 1988) was inserted into pUC9 to make pNP-GUS. pPEN-GUS was made by insertion of the PE fragment into pNP-GUS between the *Bam*HI and the *Sal*I sites upstream of the GUS open reading frame. Some of the GUS constructs were inserted as *Hin*dIII fragments into pUC9CAT. Oligonucleotides were synthesized on an Applied Biosystems synthesizer (Model 380), annealed to complementary oligonucleotides and inserted into the *Sma*I site (for mo-1 and mo-2, Figure 5) or between the blunted *Xba*I site and the *Sma*I site of pRT102. The sequences of the inserts were verified by dideoxy sequencing (Sanger *et al.*, 1977).

Northern analysis

Poly(A)$^+$ RNA was prepared from 5 g of tobacco leaves from 8−10 week old, greenhouse grown plants according to the procedure of Pereira *et al.* (1986). Northern blotting was performed according to Cuypers *et al.* (1988).

Preparation of DNA

DNA for protoplast transformation was prepared using a modification of the Birnboim and Doly (1979) protocol and banded through CsCl.

Protoplasts transformation

Sterile shoot cultures of *Nicotiana tabaccum* cv. Havana SR1 (Maliga *et al.*, 1973) were grown on MS medium. Transgenic plants were selected by germination on MS with 50 mg/ml hygromycin sulphate. Non-synchronized protoplasts were prepared from 8−10 week old plants and transformed by a modified fusion technique (Hein *et al.*, 1983). In each independent experiment, protoplasts from each genotype to be tested were prepared simultaneously from plants of the same age, and transformed in the same time period.

GUS transient expression

In each independent experiment, pooled prepared protoplasts (separated by genotype) were divided into aliquots of 10^6 protoplasts, each to be transformed with an individual plasmid construct. Ten micrograms of each construct was cotransformed with 30 μg of pUC9, both as supercoiled DNA. In some experiments, each aliquot was transformed with a mix of 10 μg of the GUS variant to be tested, 10 μg of pUC9 and 20 μg of pUC9CAT.

GUS assays

Protoplasts were harvested 18−22 h after transformation. Extracts were prepared and incubated with 4-methyl umbelliferyl glucuronide according to Jefferson *et al.* (1987). Fluorescence was measured as described on a Perkin-Elmer fluorimeter.

CAT transient expression

CAT expression was determined according to Pröls *et al.* (1988).

Acknowledgements

We thank Heinrich Cuypers for RNA analysis, Gerd Hombrecher for assistance with protoplast transformation, Andy Pereira for useful discussions and Jeffery Dangl, Susanne Kartzke and Peter Huijser for help with the manuscript.

References

Banks,J.A., Masson,P. and Fedoroff,N.V. (1988) *Genes Dev.*, **2**, 1364−1380.
Birnboim,H.C. and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513−1523.
Bolivar,F. (1978) *Gene*, **4**, 121−136.
Cuypers,H., Dash,S., Peterson,P.A., Saedler,H. and Gierl,A. (1988) *EMBO J.*, **7**, 2953−2960.
Gierl,A., Schwarz-Sommer,Zs. and Saedler,H. (1985) *EMBO J.*, **4**, 579−583.
Gierl,A., Lütticke,S. and Saedler,H. (1988a) *EMBO J.*, **7**, 4045−4053.
Gierl,A., Cuypers,H., Lütticke,S., Pereira,A., Schwarz-Sommer,Zs., Dash,S., Peterson,P.A. and Saedler,H. (1988b) In Nelson,O. (ed.), *Plant Transposable Elements*. Plenum Press, New York, pp. 115−119.
Gierl,A., Saedler,H. and Peterson,P.A. (1989) *Annu. Rev. Genet.*, **23**, 71−85.
Hein,T., Prezewozny,T. and Scheider,O. (1983) *Theor. Appl. Genet.*, **64**, 119−122.
Jefferson,R.A., Burgess,S.M. and Hirsh,D. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8447−8451.
Jefferson,R.A., Kavanagh,T.A. and Bevan,M.W. (1987) *EMBO J.*, **6**, 3901−3907.
Kim,H.-Y., Schiefelbein,J.W., Raboy,V., Furtek,D.B. and Nelson,O.E.,Jr (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 5863−5867.
Kozak,M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 2850−2854.
Maliga,P., Sz-Breznovitis,A. and Marton,L. (1973) *Nature New Biol.*, **244**, 29−30.
Masson,P., Surosky,R., Kingsbury,J.A. and Fedoroff,N.V. (1987) *Genetics*, **117**, 117−137.
Masson,P., Rutherford,G., Banks,J.A. and Fedoroff,N. (1989) *Cell*, **58**, 755−765.
McClintock,B. (1954) *Carnegie Institute of Washington Year Book*, **53**, 254−260.
McClintock,B. (1965) *Brookhaven Symp. Biol.*, **18**, 162−184.
Meyer,P., Walgenbach,E., Bussmanm,K., Hombrecher,G. and Saedler,H. (1985) *Mol. Gen. Genet.*, **201**, 513−518.
Nelson,O.E.,Jr and Klein,A.S. (1984) *Genetics*, **106**, 769−779.
Pereira,A., Cuypers,H., Gierl,A., Schwarz-Sommer,Zs. and Saedler,H. (1986) *EMBO J.*, **5**, 835−841.
Peterson,P.A. (1953) *Genetics*, **38**, 682−683.
Pröls,M., Töpfer,R., Schell,J. and Steinbiss,H.-H. (1988) *Plant Cell Reports*, **7**, 221−224.
Reddy,A. and Peterson,P.A. (1976) *Theor. Appl. Genet.*, **48**, 269−278.
Salser,W. (1977) *Cold Spring Harbor Symp. Quant. Biol*, **42**, 985−1002.
Sanger,F., Micklin,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463−5467.
Schiefelbein,J.W., Raboy,V., Fedoroff,N. and Nelson,O.E.,Jr (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 4783−4787.
Schulze-Lefert,P., Dangl,J.L., Becker-Andre,M., Halbrock,H. and Schulz,W. (1988) *EMBO J.*, **8**, 651−656.
Schwarz-Sommer,Zs., Gierl,A., Berndtgen,R. and Saedler,H. (1985) *EMBO J.*, **4**, 2439−2443.
Schwarz-Sommer,Zs., Shepherd,N., Tacke,E., Gierl,A., Rohde,W., Leclercq,L., Mattes,M., Berndtgen,R., Peterson,P.A. and Saedler,H. (1987) *EMBO J.*, **6**, 287−294.
Sutcliff,G. (1978) *Cold Spring Harbor Symp. Quaunt. Biol.*, **43**, 77−90.
Tacke,E., Schwarz-Sommer,Zs., Peterson,P.A. and Saedler,H. (1986) *Maydica*, **31**, 83−91.
Töpfer,R., Schell,J. and Steinbiss,H.H. (1988) *Nucleic Acids Res.*, **16**, 8725.
Vieira,J. and Messing,J. (1982) *Gene*, **19**, 259−268.
Wessler,S.R., Baran,G. and Varagona,M. (1987) *Science*, **237**, 916−918.
Wessler,S.R., Baran,G. and Vargona,M. (1988) In Nelson,O. (ed.), *Plant Transposable Elements*. Plenum Press, New York, pp. 293−303.

Received on December 21, 1989; revised on April 6, 1990