

## Effects of rho-independent terminators on the expression of the prokaryotic $\beta$ -glucuronidase gene in tobacco

Shih-Tong Jeng\* and Chia-Yu Yen

Department of Botany, National Taiwan University, Taipei 106, Taiwan, Republic of China

\*Corresponding author, e-mail: stjeng@ccms.ntu.edu.tw

Received 10 August 1999; revised 1 April 1999

The threonine (*thr*) attenuator with a dyad symmetrical structure is from the regulatory region of the *thr* operon of *Escherichia coli*, and encodes RNA with a stem-and-loop structure followed by a stretch of uridine residues. The *thr* attenuator and its variants were subcloned into the region between the 35S promoter and  $\beta$ -glucuronidase (GUS) coding region, and the transient expression of GUS gene in tobacco protoplast was treated as a reporter for gene regulation in plants. Results from the 14 variants in the stem region of the *thr* attenuator indicated that both base pairing and sequence specificity in the G + C-rich region of the stem were important for the GUS expression, but 1 base mismatch in the A + U-rich region of stem did not affect the GUS expression in plants. Seven variants with nested deletion in the stretch of

uridine residues were also analyzed, and the results suggested that the variants with the shorter uridine stretch produced more GUS protein than those with the longer stretch. Transgenic tobacco plants with the *thr* attenuator and its variants located between the 35S promoter and GUS coding region were also generated, and their steady state RNAs were hybridized with 2 radioactive antisense RNA probes which bound 5' and 3' of the *thr* hairpin, respectively. After the digestion of S1 nuclease, the amount of the nuclease-resistant transcript from the protection of the 5' antisense RNA probe was much more than that from the protection of the 3' probe in all tested variants. This result suggests that these dyad symmetries may affect transcription of plant RNA polymerase II.

### Introduction

The primary sequence of DNA can determine the secondary structures of nucleic acids, and the dyad symmetrical or palindromic DNA encoding RNA hairpin structures are prevalent within the prokaryotic and eukaryotic genomes. The palindromic DNA has been well studied as a termination signal, and is involved in the attenuation control of amino acid biosynthetic operons in *Escherichia coli* (Yanofsky 1981, Gardner 1982). In animals, DNA with the inverted repeat has been implicated in the transcription termination of the chicken  $\beta^H$ -globin gene (Pribyl and Martinsen 1988), and DNA encoding an RNA with a hairpin followed by a uridine-rich region has been noted in the transcription termination of RNA polymerase II in the SV40 late gene (Hay et al. 1982). Incorrect transcription also caused the accumulation of nuclear RNA in cells (Whitelaw and Proudfoot 1986, Wiest and Hawley 1990). Premature termination of transcription has been reported, and is an important method of gene regulation in the case of *c-fos* (Bonnieu et al. 1989), *c-myc* (Bentley and Groudine 1988), and *c-myb* (Bender et al. 1987). Also, genes of Hsp70 and

$\beta$ -globin have been reported to produce attenuated mRNAs due to the termination within genes (Rougvie and Lis 1988, Whitelaw and Proudfoot 1986). Therefore, DNA with dyad symmetry is involved in the regulation of gene expression.

DNA with the dyad symmetry has been found in plant genomes. It is present in the 5' untranslated region of gene *rbcS* (Poulsen et al. 1986), the 3' untranslated region of gene *ids-1* (Okumura et al. 1991), among poly(A) signals of genes *ara* and *his* (Legerton and Yanofsky 1985, Matsui et al. 1989), and after the poly(A) signals of gene *ara* (Matsui et al. 1989). The functions of these dyad symmetries within genomes are unknown. However, several synthetic palindromic DNAs were transcribed by purified wheat germ RNA polymerase II, and results indicate that these dyad symmetrical DNAs can terminate the transcription in vitro (Job et al. 1991).

The threonine (*thr*) attenuator contains a G + C-rich dyad symmetry followed by a run of deoxythymidine, encodes an RNA hairpin with a stretch of uridine, and is from the regulatory region of the *thr* operon of *E. coli*. The effects

of the *thr* attenuator on transcription termination have been well studied with *E. coli* (Yang and Gardner 1989), bacteriophage (Jeng et al. 1990, 1992, 1997), and chloroplast (Chen et al. 1995) RNA polymerases. Results from these studies indicate that the G:C base pairing in the stem region of the hairpin and a certain length of the polyuridine tract of the *thr* RNA hairpin are required for these RNA polymerases to terminate well. Moreover, not only the base pairing, but also the sequence by itself in the G + C-rich region of the *thr* RNA hairpin are critical in the transcription termination of bacteriophage RNA polymerases.

The functions of dyad symmetrical DNA in bacteria and animals have been well documented, but little is known about its effects in plant. This study is designed to use the *thr* attenuator and its variants to study their effects on gene expression in plants. These variants include point substitutions of dyad symmetry in the G + C-rich region (Lynn et al. 1987, 1988, Burton 1989) and the A + T-rich region (Yang and Gardner 1989), as well as a set of nested deletions in the polythymidine tract (Lynn et al. 1988). Results from this study can be compared with previous studies from bacteria and bacteriophages with the same DNA dyad symmetries.

## Materials and methods

### Plasmid construction

DNA oligonucleotides with sequences 5'-CTAGAGCAAA-GATGTTT-3' and 5'-GATCCAAACATCTTTGCT-3' were chemically synthesized, and inserted into the *Xba*I and *Bam*HI sites of pBI221 (Clontech, Palo Alto, CA, USA, Fig. 1B) to produce plasmid pBI221SI (Fig. 1C).

The 109-bp fragments isolated from the *Bam*HI digestion of plasmids pTZ19thr (Fig. 1A) and its variants containing the mutations in the *thr* attenuator (Jeng et al. 1990) were cloned into the *Bam*HI site of pBI221 (Fig. 1B) to produce pBI221thr (structure not shown) and its variants. The 109-bp fragments from the *Xba*I digestion of pBI221thr and its variants were then put into the *Xba*I site of pBI221SI to form pBI221SIthr (Fig. 1D) and its variants.

The 104- to 97-bp fragments from plasmids pTZ19T8 to pTZ19T1 (Jeng et al. 1990) digested by *Bam*HI were cloned into the *Bam*HI site of pBI221 to produce pBI221T8 to pBI221T1. The 104- to 97-bp fragments from the pBI221T8 to pBI221T1 treated by *Xba*I were subcloned into the corresponding site of pBI221SI, and the final pBI221SI T8 to pBI221SI T1 were used in this study.

Plasmids for RNA protection assay were constructed as follows. A *Hind*III-*Rsa*I fragment (71 bp) from pTZ19thr (Fig. 1A, Jeng et al. 1990) was cloned into the *Hind*III and *Sma*I digested pTZ18u (Mead et al. 1986) to form pTZ18HM, which contains a reversed orientation of the DNA sequence between the 35S promoter and *thr* attenuator in pBI221SIthr and its variants. The *Bam*HI-*Bcl*II fragment (113 bp) from pBI221SI (Fig. 1C) was inserted into the *Bam*HI sites of pTZ18u in order to form pTZ18BP, which contains part of the GUS gene in a reversed orientation.

The 2.2-kb fragment isolated from the *Xba*I and *Eco*RI digestion of pBI221SIthr was cloned into the *Xba*I and

*Eco*RI sites of pTZ19thr containing T7 promoter (Mead et al. 1986) to form p19thrGUS, by which the RNA produced by T7 RNA polymerase was used as a control during the S1 protection assay. All plasmids were purified by CsCl-ethidium bromide isopycnic centrifugation before electroporation (Maniatis et al. 1982).

### Electroporation

The method to isolate protoplast from the leaves of tobacco (*Nicotiana tabacum*) has been described (Watanabe et al. 1992, Hung et al. 1993). In order to compare the  $\beta$ -glucuronidase (GUS) expression from pBI221SIthr and its variants quantitatively, plasmid MTC301, which contains the 35S promoter and firefly luciferase (kindly provided by Dr. Su-May Yu, Institute of Molecular Biology, Academia Sinica, Taiwan), was treated as an internal control during electroporation. Isolated protoplast was adjusted by an electroporation solution (0.4 M D-mannitol, 8 mM MES-KOH, and 5 mM CaCl<sub>2</sub>) to obtain  $2 \times 10^6$  cell ml<sup>-1</sup>, and treated with heat shock at 45°C for 5 min. The mixture containing 8  $\mu$ g of pBI221SIthr or its variants, 8  $\mu$ g of plasmid

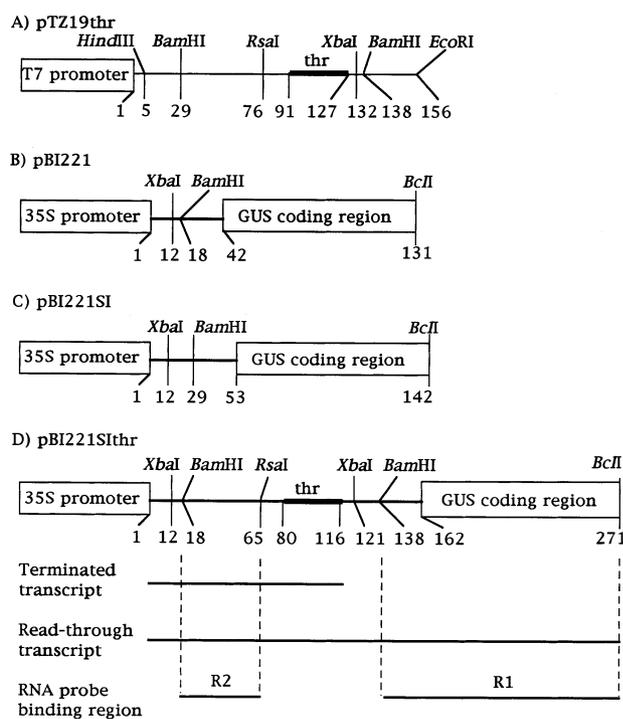


Fig. 1. Schematic representation of DNA templates, RNA transcripts, and RNA probes. Only the T7 promoter and part of 5' untranslated region of pTZ19thr and the 35S promoter, 5' untranslated region, and part of  $\beta$ -glucuronidase (GUS) coding region of pBI series plasmids are shown. Transcription initiation at the promoter in each template is assigned as position +1. The position of the *thr* attenuator (*thr*) is indicated. A, plasmid pTZ19thr; B, plasmid pBI221; C, plasmid pBI221SI; D, plasmid pBI221SIthr. Transcription terminating at the *thr* attenuator is indicated as terminated transcript, and reading through *thr* attenuator is shown as read-through transcript. After the digestion of S1 nuclease the nuclease-resistant regions of transcripts from the hybridization of RNA probes BP and HM are indicated as R1 and R2, respectively.

MTC301, 10 µg of calf thymus DNA, and 50 µl of PEG solution (24% [w/v] polyethylene glycol 4000, 0.4 M D-mannitol, 30 mM MgCl<sub>2</sub>, pH 5.6) was added to 100 µl of protoplast after the heat shock. Electroporation was performed by Baekon 2000 electroporator (Baekon, Inc, Saratoga, CA) with conditions recommended by the manufacturer (2 kV, 22 pulse, 0.4 s burst time, 5 cycles, and 62.5 µs). After centrifugation at 800 g for 5 min, cell pellets were resuspended in an incubation solution (MS salt, 1% sucrose, and 0.4 M D-mannitol) at 25°C for 16 h.

### GUS assay

The GUS assay procedure was described by Jefferson et al. (1987) and modified as follows. Protoplast after 16-h incubation was collected by centrifugation at 800 g for 5 min, and were resuspended in 75 µl of extraction buffer [100 mM potassium phosphate buffer (pH 7.7), 1 mM EDTA, 8 mM 2-mercaptoethanol, 0.5% Triton X-100, and 10% glycerol]. Cells were broken down by alternate freezing and thawing 3 times, and the supernatant of the cells was collected by centrifugation at 10000 g for 5 min. Supernatant (50 µl) was incubated with 150 µl of 4-MUG assay buffer at 37°C (Jefferson et al. 1987), and 45 µl of this reaction mixture was taken out at 0-, 10-, 20-, and 40-min time points by the addition of 1 ml of 0.2 M sodium carbonate. The fluorescence of the samples was measured with the excitation wavelength at 365 nm and emission wavelength at 455 nm using a Hitachi Spectrofluorometer F2000. A solution of 100 mM 4-methyl umbelliferone (MU) in 0.2 M sodium carbonate was used to calibrate the fluorescence intensity.

### Luciferase assay

The assay method for firefly luciferase was described by Ow et al. (1987), and the luciferase activity was measured using the Optocomp 1 Bacterial System (GEM Biomedical Inc., Hamden, CT, USA). The average value with standard deviation (SD) from all experiments in this transient expression study is  $15 \times 10^6 \pm 6 \times 10^6$  using the above system.

### Protein assay

The concentration of protein in each cell extract was measured by the Coomassie blue binding method (Sedmak and Grossberg 1977, Spector 1978).

### Data calculation

The GUS assays of each sample contain 4 time points. The slope, with the units of pmol MU min<sup>-1</sup>, can be obtained from the linear regression of these 4 points (Jefferson et al. 1987). This GUS result was adjusted by the protein concentration for the amount of cells in each sample and by the luciferase activity for the efficiency of electroporation, and the final adjusted data with the unit of pmol MU mg<sup>-1</sup> protein min<sup>-1</sup> were used in this study. All results were repeated at least 6 times, and represented as mean  $\pm$  SD.

### Production of transgenic plants

The *Hind*III and *Eco*RI digestion of pBI221SI containing the *thr* attenuator and its variants were inserted into the pBI121 (Clontech) digested by *Hind*III and *Eco*RI. *Agrobacterium tumefaciens* strain LBA 4404 cells harboring the recombinant pBI121 series vectors were used to inoculate leaf discs of *N. tabacum* cv. W38, and transformed callus tissues were selected by their resistance to karamycin at 100 µg ml<sup>-1</sup> (Horsch et al. 1985). Shoots were regenerated from transformed calli, and, after root formation, plantlets were grown aseptically. The presence of gene constructs in transgenic plants was confirmed by Southern blot analysis and GUS assay.

### S1 nuclease protection assay

Transgenic plants with 1 copy of transferred gene were selected, and the total RNAs of their leaves were isolated as described (Chomczynski and Sacchi 1987). Using the *Hind*III-linearized plasmids pTZ18BP and pTZ18HM, radioactive RNA probes BP and HM were produced by in vitro transcription with the T7 RNA polymerase in the presence of [ $\alpha$ -<sup>35</sup>S]CTP, respectively. RNAs of transgenic plants, in vitro transcripts, and radioactive probes were treated with RQ-1 DNase (Promega, Madison, WI, USA) at the final concentration of 0.02 unit µl<sup>-1</sup> for 10 min before hybridization. Hybridization reactions (15 µl) contained 10 µg of RNA from transgenic plants or in vitro transcription, radioactive RNA probes 1 and 2 ( $\sim 3 \times 10^5$  cpm each, 20 pmol), 100 mM NaCl, and 50% formamide. The hybridization mixture was incubated at 94°C for 5 min, and then at 40°C for 16 h. To the 15 µl-hybridization mixture, 165 µl of water, 20 µl of 10  $\times$  S1 reaction buffer [30 mM sodium acetate (pH 4.5) and 1 mM zinc chloride], 5 µl of yeast carrier tRNA (10 mg µl<sup>-1</sup>), and 10 µl of S1 nuclease (20 units µl<sup>-1</sup>, Promega) were added. After incubation at 37°C for 60 min, reactions were stopped by adding 150 µl of phenol/chloroform (1:1). RNA was precipitated with ethanol and dried under vacuum, and latter analyzed by electrophoresis on a 6% polyacrylamide gel containing 8 M urea.

Radioactive gels were displayed on the Molecular Dynamics Phosphorimager, and their autoradiographs were printed out using a Kodak XLS 8600 PS Printer. The relative radioactivity of RNA bands was quantified using the Molecular Dynamics Imagequant software.

### Results

The prokaryotic GUS gene has been widely used as a reporter for transcription and translation regulations in eukaryotic cells. In terms of transient assay, the expression of a reporter is influenced by the efficiency of the electroporation process. In particular, for the quantitative analysis, a second reporter gene is required as an internal reference of electroporation for the comparison among samples. In this study we used the luciferase gene as the second reporter in all assays, and found that it was required to obtain repeatable results (data not shown). The concentration of protein

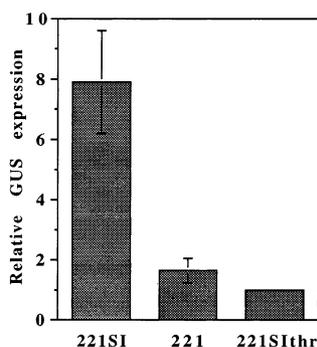


Fig. 2. GUS expressions of pBI221, pBI221SI, and pBI221SIthr. GUS gene expression of the plasmid pBI221SIthr is treated as the normalized reference, with a value of 1, for determining the relative amount of GUS activities. The error bars are indicated as the SD for at least 6 assays with each construct. The structures of these plasmids are shown in Fig. 1.

was used to normalize the amount of protoplast used among samples, and this control can eliminate the deviation in analyses due to the difference from the cell mass used in each sample.

### The construction of pBI221SI

Although the 35S promoter in plasmid pBI221 is a constitutive promoter, the DNA sequence surrounding the ATG, the start codon, the area of GUS coding region does not contain the conserved sequence for efficient translation in eukaryotic cells (Kozak 1986, Joshi et al. 1997). A conserved sequence, GCAAAGATGTTT, around the start codon of 4 soybean lipoxygenase genes has been reported (Shibata et al. 1987, Yenofsky et al. 1988), and this sequence was inserted into the *Xba*I and *Bam*HI sites of pBI221 to produce pBI221SI (Fig. 1). The GUS expression of pBI221SI was  $424 \pm 84$  pmol MU mg<sup>-1</sup> protein min<sup>-1</sup>, which is about 5-fold more than that expressed by pBI221 (Fig. 2). Similar results were also reported by Kato et al. (1991), and they showed a 7- to 11-fold more GUS expression than that of

rice and tobacco cell suspensions with the related constructions. Plasmid pBI221SI contains a new ATG start codon and an extra 10 amino acid codons in front of the original GUS coding region in pBI221. Although this soybean conserved sequence, GCAAAGATGT, is apparently different from the suggested conserved sequence, GCCGCCPuC-CATGG, surrounding the start codon in eukaryotes (Kozak 1986, 1987, Joshi et al. 1997) the construction pBI221SI rather than pBI221 was used in this study.

### Variants in the G + C-rich region of the hairpin

Previous studies have shown that some point variants in the G + C-rich region of the *thr* attenuator decrease the transcription termination of *E. coli* (Lynn et al. 1985, 1987, 1988, Burton 1989, Yang and Gardner 1989), bacteriophage (Jeng et al. 1990, 1992, 1997), and chloroplast (Chen et al. 1995) RNA polymerases. To analyze the effects of these altered *thr* attenuators on plant gene expression, variants containing the transcription-termination structure of the *thr* operon regulatory site were cloned into the region between the 35S promoter and GUS coding region of pBI221SI to become a pBI221SIthr series (Fig. 1D). These variant DNAs were then electroporated into the tobacco protoplast, and the transient expressions of GUS gene were treated as a reporter for gene regulation in plants. Due to the location of these dyad symmetries, the expression of GUS gene may be regulated by both transcription and translation.

The GUS expressions on altered attenuators are presented in Table 1. The single-position variants (Fig. 3) creating mismatch in the G + C-rich region (BD10, 138U, 140A, 151A, 153A, and AD1) increased the GUS expression by at least 2-fold more than the GUS gene with a wild-type *thr* attenuator did. This suggests that helix integrity is required at these positions to diminish gene expression. Helix-stability values calculated by the method of Freier et al. (1986) indicate that the stability of these variants decreased by 5–28% with respect to the structure of the wild-type (Table 1). Therefore, there is no direct correlation with the decrease

Table 1. GUS expression of *thr* attenuator variants. <sup>a</sup>The positions correspond to those in Fig. 3. <sup>b</sup> $\Delta G$  values (kJ mol<sup>-1</sup> at 37°C) were calculated according to Freier et al. (1986). <sup>c</sup>The GUS gene expression of the wild-type *thr* attenuator was treated as the normalized reference, with a value of 1, for determination of the relative amounts of GUS expression in each experiment as described under 'Materials and methods'.

Variant	Position <sup>a</sup> altered	Base pair change	$\Delta G$ value (kJ mol <sup>-1</sup> ) <sup>b</sup>	Relative GUS expression <sup>c</sup>
Wild-type			-96	1.00 ± 0.18
BC10	98	CG → GG	-67	2.54 ± 0.47
138U	100	CG → UG	-88	2.55 ± 0.87
140A	102	CG → AG	-71	4.16 ± 0.95
151A	113	CG → CA	-71	2.26 ± 0.53
153A	115	CG → CA	-67	2.43 ± 0.62
AD1	117	CG → CC	-67	2.56 ± 0.43
BD16	98, 117	CG → GC	-96	1.57 ± 0.30
160A	122	AU → AA	-88	1.50 ± 0.53
160C	122	AU → AC	-88	1.41 ± 0.26
160G	122	AU → AG	-88	1.32 ± 0.23
131G	93	AU → GU	-92	1.64 ± 0.27
1316UA	93, 122	AU → UA	-96	0.99 ± 0.12
1316GC	93, 122	AU → GC	-105	0.81 ± 0.19
156AC	115, 122	CG → CA AU → AC	-59	2.65 ± 0.55

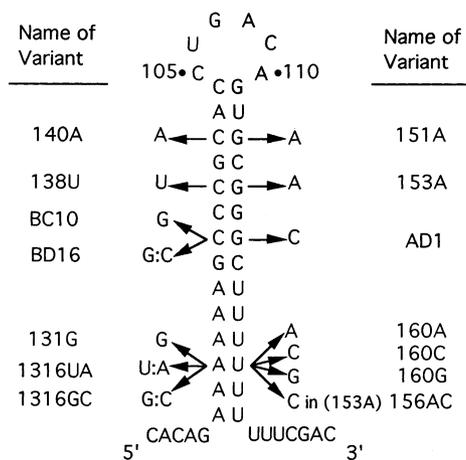


Fig. 3. The *thr* attenuator and its variants. The potential RNA structure encoded by the wild-type *thr* attenuator is presented. The nucleotide positions are relative to the transcription initiation site (+1) of 35S promoter in pBI221SI*thr* (Fig. 1D). The changed bases in each variant are indicated (see also Table 1).

in GUS expression and the decrease in helix stability, and similar effects were also noted previously with the bacterial and bacteriophage RNA polymerases on transcription termination with these same variants (Lynn et al. 1985, Jeng et al. 1990, 1997).

The compensating variant BD16 changes C:G to G:C (Fig. 3) and still keeps the hairpin intact. The calculated stability values for this variant and the wild-type are both  $-96 \text{ kJ mol}^{-1}$  (Table 1) but the GUS expressions are 1.57- and 1.00-fold, respectively. Furthermore, according to free energy suggested by Freier et al. (1986), variant 138U changing a C:G to an U:G pair (Fig. 3) did not create a mismatch in the stem region (Fig. 3), and shows 2.55-fold GUS expression. In other words, the GUS expression of 138U with a relative *thr* hairpin was dramatically increased compared to that of GUS gene with a wild-type *thr* attenuator. This further suggests that not only the stability but also the sequence specificity in the G + C-rich region of the *thr* attenuator are required to eliminate the expression of plant genes. Similar effects have been observed with bacteriophage RNA polymerases (Jeng et al. 1990, 1997).

#### Variants of the A + U-rich region of the hairpin

To determine the effects of changes in the sequence of the A + T-rich region of the attenuator, variants which encode both the runs of A and U residues in the transcript were tested. Some of these variants disrupted each stretch individually, whereas others altered both sequences but left them capable of forming a potential helix between the 2 regions (Fig. 3). Disrupting the U tract (160A, 160C, or 160G) by substituting A, C, or G for U slightly increases the GUS expression by 1.32- to 1.50-fold compared to that of the GUS gene with a wild-type *thr* attenuator (Table 1). If the SD was taken into consideration, their expressions were close to that of the GUS gene with the

wild-type *thr* attenuator. This may indicate that 1 base change in the U-rich region of the hairpin does not affect gene expression in plants.

To test the effect of modifying the run of A residues preceding the G + C-rich hairpin in the terminator transcript (Fig. 3), several variants were assayed. Variant 131G, changing an A:U to a U:G pair (Fig. 3), showed 1.64-fold of GUS expression (Table 1). As mentioned above, the U:G pair does not create a mismatch in the stem of the hairpin (Freier et al. 1986). However, both variants 131G and 138U creating a G:U pair showed more GUS expression than that of the GUS gene with the wild-type *thr* attenuator. It seems that a G:U or U:G pair in the stem of the hairpin is not sufficient for decreasing gene expression in plants. Alterations, 1316UA and 1316GC change both the A and U stretches to allow potential base pairing (Fig. 3). Variant 1316UA shows the same GUS expression as that of the GUS gene with the wild-type *thr* attenuator, while 1316GC changing the same A:U pair to G:C pair slightly decreases GUS expression.

#### Variants with nested deletion in the polyuridine tract

Lynn et al. (1988) showed that the length of the uridine tract behind the hairpin affected transcription termination of *E. coli* RNA polymerase. These variants with the shortened continuous deoxythymidine tract, or their corresponding uridine stretch in RNA, were cloned into plasmid pBI221SI to form pBI221SIT1 to pBI221SIT8 (Table 2). Variants pBI221SIT8 with 8 uridines after the G + C-rich region already show 2.54-fold GUS expression compared to that of the wild-type *thr* attenuator with 9 uridines. Furthermore, the *thr* attenuators with less than 8 uridines increase the GUS expression by at least 2.97-fold, and the shorter the uridine track is in the *thr* attenuator, the more GUS product the variants express (Table 2). This may mean that a long uridine track is required to reduce the gene expression in the plant.

Table 2. GUS expression of the *thr* attenuator variants in the poly(U) track region. <sup>a</sup> Sequences only show the loop and right part of the *thr* hairpin (Fig. 3). The stem region of the hairpin is underlined. <sup>b</sup> The GUS expressions of the wild-type *thr* attenuator was treated as the normalized reference, with a value of 1, for determination of the relative amounts of transcript in each experiment as described under 'Materials and methods'.

Variant	Corresponding sequence in the <i>thr</i> attenuator <sup>a</sup>	Relative GUS expression <sup>b</sup>
<i>thr</i>	...CUGACAGUGCGGGCUUUUUUUUU...	0.99 ± 0.26
T8	...CUGACAGUGCGGGCUUUUUUUUU...	2.54 ± 0.81
T6	...CUGACAGUGCGGGCUUUUUUU...	2.97 ± 0.45
T5	...CUGACAGUGCGGGCUUUUU...	3.38 ± 0.42
T4	...CUGACAGUGCGGGCUUUUU...	3.60 ± 0.83
T3	...CUGACAGUGCGGGCUUU...	3.31 ± 0.85
T2	...CUGACAGUGCGGGCUU...	5.35 ± 1.11
T1	...CUGACAGUGCGGGCU...	12.20 ± 2.12

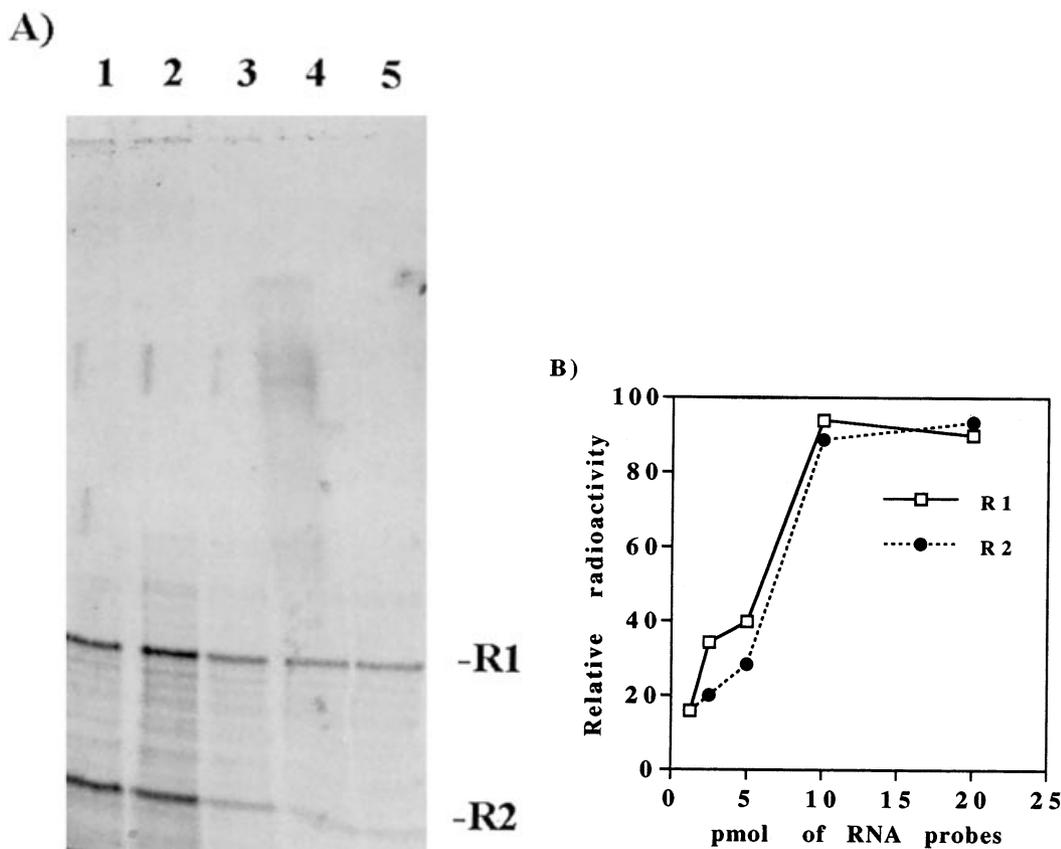


Fig. 4. Titration of transcript with RNA probes. Using the *Eco*RI-linearized plasmid p19thrGUS, RNA produced by T7 RNA polymerase in vitro was hybridized with radioactive antisense probes HM and BP, which hybridized to the 5' and 3' position of the RNA hairpin, respectively. After the treatment of S1 nuclease, parts of the 92-bp HM and 167-bp BP complementary to the RNA transcripts were nuclease-resistant, and became 47- and 113-bp probe products R2 and R1, respectively. An autoradiograph was produced from a Phosphorimager (panel A), and the relative radioactivity of RNA bands was quantified by the Molecular Dynamics Imagequant software (panel B). Lane 1, 20 pmol; lane 2, 10 pmol; lane 3, 5 pmol; lane 4, 2.5 pmol; lane 5, 1.25 pmol of each radioactive probes BP and HM added.

### S1 nuclease protection assay

So far, the effects of the *thr* attenuator on the GUS gene expression were analyzed by the transient expression of  $\beta$ -glucuronidase, whose activity results from the regulation of both transcription and translation. In order to analyze the roles of these dyad symmetries in transcription alone, steady state RNA hybridized with antisense probes were analyzed by S1 nuclease. These 2 antisense RNA probes, HM and BP, were produced by T7 RNA polymerase in vitro, and are able to bind upstream and downstream of the *thr* hairpin transcript (Fig. 1D), respectively. Although the sequence and composition of RNA probes HM and BP are different, both RNA probes were incubated within the same tube during the S1 nuclease assay. Therefore, it was important to show that both RNA probes binding with their complementary transcript have the similar resistance to S1 nuclease, and Fig. 4 did indicate this result.

First, the optimum condition for the amount of probes used in the reaction was determined in order to quantify the RNA transcript by S1 nuclease. Plasmid p19thrGUS (Materials and methods), like pBI221SI*thr*, contains the *thr* attenuator, GUS coding region, and T7 promoter (Fig. 1D). Using the *Eco*RI-linearized plasmid p19thrGUS, RNA produced by T7 RNA

polymerase in vitro included 127 bp terminated and 2.2 kb read-through transcripts, and was hybridized with radioactive antisense probes HM and BP, which are able to hybridize the 5' and 3' positions of the RNA hairpin, respectively (Fig. 4A). After the treatment of S1 nuclease, parts of the 92-bp HM and 167-bp BP probes complementary to RNA transcripts were nuclease-resistant, and became 47- and 113-bp probe products R2 and R1 (Fig. 1D), respectively. Results showed that 0.2 pmol (0.16  $\mu$ g) of the RNA transcripts was saturated with 20 pmol each of probes BP and HM (Fig. 4B). Also, the radiation value of each band in the gel can be normalized according to the length and CMP composition of the probe products (Jeng et al. 1990). The relative molar ratio of probe products R1 and R2 in lane 1 of Fig. 4A is 0.9:1, which is close to the estimated ratio of the read-through to the total RNA produced by T7 RNA polymerase with p19thrGUS as a template (data not shown). Therefore, probes BP and HM (20 pmol each) were used and digested by S1 nuclease in the same tube to determine the amount of RNA transcripts from transgenic plants. Also, the amount of mRNA within tobacco cells is less than 10% of the steady state RNA and GUS mRNA is only part of cellular mRNA; therefore 10  $\mu$ g of the total RNA from a transgenic plant should be saturated with 20 pmol of probes BP and HM.

The *thr* attenuator and some of its variants were cloned into the region between the 35S promoter and GUS coding region of pBI121, and transgenic plants were generated by agrobacteria with pBI121 constructs (Horsch et al. 1985). The steady state RNA isolated from the transgenic plant was hybridized by the radioactive probes, HM and BP. After digestion with S1 nuclease, nuclease-resistant R2 from probe HM indicates the total transcript, and R1 from probe BP represents the read-through-hairpin transcript produced by the plant RNA polymerase II from the 35S promoter in vivo (Fig. 1D). There was no detectable product R1, when the dyad symmetry is wild-type or variant BD16 (lanes 2 and 4, Fig. 5). However, product R1 was present, when the dyad symmetry is either variant AD1 or T5 (lanes 1 and 3, Fig. 5). These may indicate that plant RNA polymerase II terminated more frequently at the dyad symmetry of the wild-type and variant BD16 than of variants AD1 and T5. Indeed, the GUS expressions of the wild-type and BD16 were lower than those of variants AD1 and T5 (Tables 1 and 2). However, since the steady state RNA was measured within the S1 protection assay, the influence of RNA stability cannot be excluded.

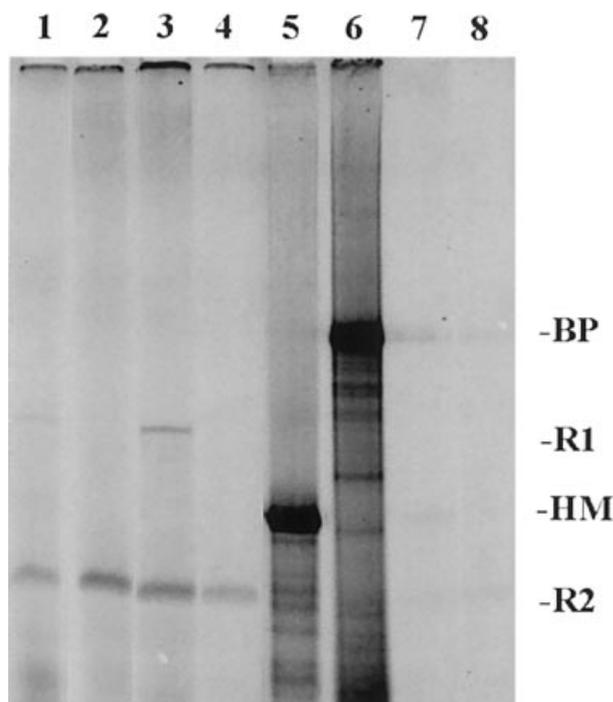


Fig. 5. The S1 nuclease assays of the transgenic plants with the *thr* attenuator and its variants. The steady state RNA from the leaves of transgenic plants grown aseptically was hybridized with RNA probes HM and BP, which were produced by the in vitro transcription of the T7 RNA polymerase using the DNA templates pTZ18HM and pTZ18BP, respectively. After the digestion of S1 nuclease, the nuclease-resistant products were assayed by polyacrylamide gel. Without the digestion of S1 nuclease, the positions of probes HM and BP from pTZ18HM and pTZ18BP were indicated as HM and BP, respectively. With the digestion of S1 nuclease, the nuclease-resistant products from the hybridization of transcripts with probes HM and BP were indicated as R2 and R1, respectively. Lane 1, AD1; lane 2, wild-type; lane 3, T5; lane 4, BD16; lane 5, probe 2 without S1 digestion; lane 6, probe 1 without S1 digestion; lane 7, probe 2 with S1 digestion; lane 8, probe 1 with S1 digestion.

## Discussion

In this study, dyad symmetries of the *thr* attenuator and its variants were inserted between the 35S promoter and GUS coding region, and the expression of the GUS gene through DNA electroporation was taken as an indicator of their effects on gene expression. Since the level of protein expression is the consequence of both translation and transcription regulation, the steady state RNA from transgenic plants with the *thr* attenuator and its variants were quantified by S1 nuclease in order to analyze transcription regulation alone. Results suggested that these dyad symmetries may affect transcription of plant RNA polymerase II, and may also influence the stability of RNA. However, since the *thr* attenuator is located at the 5' untranslated region (UTR) of the GUS gene, the involvement of translation control cannot be excluded.

The function of hairpin structures in the 5' UTR of the plant nuclear mRNA has been studied by the translational activity (Kozak 1989, 1994). A hairpin with free energy  $-125 \text{ kJ mol}^{-1}$  inhibited translation when it was 12 nucleotides (nt) from the 5' end of mRNA, but did not affect translation when it was 52 nt from the 5' end (Kozak 1989). This result suggested that ribosome, once bound to mRNA, had the ability to penetrate the secondary structure (Kozak 1989) and ribosome binding requires mRNA with a relatively long stretch of 5' end without a hairpin structure (Kozak 1994). According to this conclusion, the *thr* attenuator within the pBI221SI may not interfere with mRNA from translation, since the free energy of the wild-type *thr* attenuator is only  $-96 \text{ kJ mol}^{-1}$  and located at 80 nt from the 5' end of GUS mRNA (Fig. 1D).

Also, with the 5' UTR of plant nuclear mRNA the presence of a hairpin structure helps the recognition of the start codon, AUG, in a cell-free system using the chloramphenicol transferase mRNA (Kozak 1989). However, a hairpin structure from the PO149 mRNA of alfalfa contains the start codon as its stem region, and was proposed to repress translation (Wu et al. 1996). The choice of start codons in the mRNA of barley yellow dwarf virus encoding overlapping genes is also regulated by the hairpin structure (Dinesh-Kumar and Miller 1993). It seems that if the location of a hairpin structure is nearby the start codon, then the hairpin may be involved in the regulation of translation.

Furthermore, the 5' UTR in the late anther tomato gene was identified as repressing the expression of the luciferase reporter gene in both transient experiments and stable transformation, and a hairpin structure participated in this repression. This 5' UTR sequence dramatically decreased mRNA accumulation but without affecting translation rate and mRNA stability (Curie and McCormick 1997). This result agrees with our conclusion that if the position of a hairpin structure is located between the 5' end and the start codon of mRNA it may inhibit transcription.

The variability in expression of a transgene among the transgenic plant population was reported (Hauffe et al. 1993), and also observed in this study (data not shown). In this study, a transgenic tobacco population was selected by kanamycin, and further confirmed by Southern blot and GUS activity analyzes. The antisense RNA probes, HM and

BP, separately binding to the upstream and downstream of the *thr* attenuator, were digested by S1 nuclease, and the nuclease-resistant products R1 and R2 were used to estimate the relative amount of steady state RNA in each individual transgenic tobacco. Therefore, the variability of the amount of mRNA among the transgenic plant population should not affect the ratio of R1 to R2 of a specific transgenic construct.

The steady state RNA of transgenic tobacco was estimated by antisense RNA probes, and these 2 probes, which are different in the length and composition, were used in the same tube during S1 nuclease assay. Therefore, it is important to confirm that 2 RNA probes, once binding to their complementary RNA, show the same or similar ability to survive in the digestion of S1 nuclease. Probes HM and BP were chosen for the S1 assay (Fig. 4) after several probes were tested (data not shown).

The *thr* attenuator has been studied with *E. coli* and bacteriophage RNA polymerases, and reported as a transcription terminator (Lynn et al. 1985, 1987, Jeng et al. 1990, 1997). Compared to the wild-type *thr* attenuator, variants, like BC10, 140A, 151A, 153A, and AD1, with the mismatch in the G + C-rich region of the hairpin increased the RNA polymerase reading through the *thr* attenuator hairpin in *E. coli* and bacteriophages, and enhanced the GUS protein production in plants (Table 1). It seems that the base pairing in the G + C-rich region is critical for maintaining the function of the hairpin in both prokaryotic and eukaryotic cells. However, the predicted hairpin stability (Table 1) did not correlate with the termination efficiency of *E. coli* and bacteriophage RNA polymerases (Lynn et al. 1985, Jeng et al. 1990, 1997), and neither did GUS gene expression in plants (Table 1).

The studies of the *thr* variants with nested deletions in the poly(U) region with *E. coli* and bacteriophages (Lynn et al. 1988, Jeng et al. 1990, 1997) indicated that the longer the polyuridine tract, the more the RNA polymerase was terminated by the *thr* hairpin. Results from the transcription of T7 RNA polymerase showed that T7 RNA polymerase needed the *thr* hairpin with at least 5 uridine residues to begin to cease RNA synthesis, and that 8 or 9 are required for maximal termination (Jeng et al. 1990). Similar results were also found in plant gene expression, and hairpins with a longer uridine tract prevented more GUS from expression than that with a shorter uridine run (Table 2). These results may indicate that the poly(U) tract was involved in the gene regulation of both prokaryotes and plants.

Dyad symmetries are present in many plant genes. In this paper, the expression of GUS gene by tobacco can be controlled by the dyad symmetrical DNA structure encoding a hairpin structure of RNA. The sequence specificity in the G + C-rich region of the *thr* hairpin was identified to be an important factor in terminating bacteriophage RNA polymerase (Jeng et al. 1990, 1997) and decreasing the GUS protein from tobacco, but was not important for the *E. coli* RNA polymerase (Lynn et al. 1985). Also, the S1 nuclease assays of the steady state RNA from transgenic plants showed that these dyad symmetries affect transcription in plants. The results from this study enhance the understanding of the roles and functions of dyad symmetries within the 5' UTR region of plant genes.

**Acknowledgements** – We thank Drs M. T. Chan and S. M. Yu (Institute of Molecular Biology, Academia Sinica) for their generous gift of the plasmid MTC 301. We specially thank Drs C. C. Chen and Y.-Y. Kao (Department of Botany, National Taiwan University) for providing us space and equipment for handling the plant tissue culture. This work was supported by the National Science Council under grant 87-2311-B-002-058 (to S. T. J.).

## References

- Bender TP, Thompson CB, Kuehl WM (1987) Differential expression of *c-myc* messenger RNA in murine B lymphomas by a block to transcription elongation. *Science* 237: 1473–1476
- Bentley DL, Groudine M (1988) Sequence requirements for premature termination of transcription in the human *c-myc* gene. *Cell* 53: 245–256
- Bonnieu A, Rech J, Jeanteur P, Fort P (1989) Requirements for *c-fos* messenger RNA down regulation in growth stimulated murine cell. *Oncogene* 4: 881–888
- Burton, WS (1989) A Study of the Molecular Mechanism of the Attenuation Response in the Threonine Operon of *Escherichia coli* by Synthesis of Mutant Attenuators. Doctoral Thesis, University of Illinois at Champaign-Urbana.
- Chen L-J, Liang Y-J, Jeng S-T, Orozoco EM, Gumport RI, Lin C-H, Yang M-T (1995) Transcription termination at the *Escherichia coli thr* terminator by spinach chloroplast RNA polymerase in vitro is influenced by downstream DNA sequences. *Nucleic Acids Res* 23: 4690–4697
- Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162: 156–159
- Curie C, McCormick S (1997) A strong inhibitor of gene expression in the 5' untranslated region of the pollen-specific LAT59 gene of tomato. *Plant Cell* 9: 2025–2036
- Dinesh-Kumar SP, Miller WA (1993) Control of start codon choice on a plant viral RNA encoding overlapping genes. *Plant Cell* 5: 679–692
- Freier SM, Kierzek R, Jaeger JA, Sugimoto N, Caruthers MN, Neilson T, Turner DH (1986) Improved free-energy parameters for predictions of RNA duplex stability. *Proc Natl Acad Sci USA* 83: 9373–9377
- Gardner JF (1982) Initiation, pausing, and termination of transcription in the threonine operon regulatory region of *Escherichia coli*. *J Biol Chem* 257: 3896–3904
- Hauffe KD, Lee SP, Subramaniam R, Douglas J (1993) Combinatorial interactions between positive and negative cis-acting elements control spatial patterns of 4CL-1 expression in transgenic tobacco. *Plant J* 4: 235–253
- Hay N, Skolnik-David H, Aloni Y (1982) Attenuation in the control of SV40 gene expression. *Cell* 29: 183–193
- Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT (1985) A simple and general method for transferring genes into plants. *Science* 227: 1229–1231
- Hung CY, Lai YK, Feng TY, Chen CC (1993) Chloroplast segregation in somatic hybrids of *Nicotiana plumbaginifolia* and *N. sylvestris* having different ratios of parental nuclear genomes. *Plant Cell Rep* 13: 83–86
- Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusions:  $\beta$ -glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J* 6: 3901–3908
- Jeng S-T, Gardner JF, Gumport RI (1990) Transcription termination by bacteriophage T7 RNA polymerase at rho-independent terminators. *J Biol Chem* 265: 3823–3830
- Jeng S-T, Gardner JF, Gumport RI (1992) Transcription termination in vitro by bacteriophage T7 RNA polymerase: The role of sequence elements within and surrounding a rho-independent transcription terminator. *J Biol Chem* 267: 19306–19312
- Jeng S-T, Lay S-H, Lai H-M (1997) Transcription termination by bacteriophage T3 and SP6 RNA polymerases at rho-independent terminators. *Can J Microbiol* 43: 1147–1156
- Job D, Job C, Mercoyrol L, Shire D (1991) Transcription of synthetic DNA containing sequences with dyad symmetry by wheat-germ RNA polymerase II: Increased rates of product release in single-step addition reaction. *Eur J Biochem* 195: 831–840

- Joshi CP, Zhou H, Huang X, Chiang VL (1997) Context sequences of translation initiation codon in plants. *Plant Mol Biol* 35: 993–1001
- Kato T, Shirano Y, Kawazu T, Tada Y, Itoh E, Shibata D (1991) A modified  $\beta$ -glucuronidase gene: Sensitive detection of plant promoter activities in suspension-cultured cells of tobacco and rice. *Plant Mol Biol Rep* 9: 333–339
- Kozak M (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44: 283–292
- Kozak M (1987) At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J Mol Biol* 196: 947–950
- Kozak M (1989) Circumstances and mechanisms of inhibition of translation by secondary structure in eukaryotic messenger RNA. *Mol Cell Biol* 9: 5134–5142
- Kozak M (1994) Features in the 5' non-coding sequences of rabbit alpha and beta-globin mRNAs that affect translational efficiency. *J Mol Biol* 235: 95–110
- Legerton TL, Yanofsky C (1985) Cloning and characterization of the multifunctional *his-3* gene of *Neurospora crassa*. *Gene* 39: 129–140
- Lynn SP, Bauer CE, Chapman K, Gardner JF (1985) Identification and characterization of mutants affecting transcription termination at the threonine operon attenuator. *J Mol Biol* 183: 529–541
- Lynn SP, Burton WS, Donohue TJ, Gould RM, Gumpert RI, Gardner JF (1987) Specificity of the attenuation response of the threonine operon of *Escherichia coli* is determined by the threonine and isoleucine codons in the leader transcript. *J Mol Biol* 194: 59–69
- Lynn SP, Kasper LM, Gardner JF (1988) Contributions of RNA secondary structure and length of the thymidine track to transcription termination at the *thr* operon attenuator. *J Biol Chem* 263: 472–479
- Maniatis T, Fritsch EF, Sambrook J (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Press, Cold Spring Harbor, NY
- Matsui M, Sasamoto S, Kuniade A, Nomura N, Ishizaki R (1989) Cloning of *ara*, a putative *Arabidopsis thaliana* gene homologous to the *ras*-related gene family. *Gene* 76: 313–320
- Mead DA, Skorupa ES, Kemper B (1986) Single-stranded DNA 'blue' T7 promoter plasmid: A versatile tandem promoter system for cloning and protein engineering. *Protein Eng* 1: 67–74
- Okumura N, Nishizawa NK, Umehara Y, Mori S (1991) An iron deficiency-specific cDNA from barley roots having two homologous cysteine-rich MT domains. *Plant Mol Biol* 17: 531–533
- Ow DW, Jacobs JD, Howell SH (1987) Functional regions of the cauliflower mosaic virus 35S RNA promoter determined by use of the firefly luciferase gene as a reporter of promoter activity. *Proc Natl Acad Sci USA* 84: 4870–4874
- Poulsen C, Fluhr R, Kauffman JM, Boutry M, Chua NH (1986) Characterization of an *rbcS* gene from *Nicotiana plumbaginifolia* and expression of an *rbcS*-CAT chimeric gene in homologous and heterologous nuclear background. *Mol Gen Genet* 205: 193–200
- Pribyl TM, Martinsen HG (1988) Transcription termination at the chicken beta-H-globin gene. *Mol Cell Biol* 8: 5369–5377
- Rougvie AE, Lis JT (1988) The RNA polymerase II molecule at the 5' end of the uninduced *hsp70* gene of *Drosophila melanogaster* is transcriptionally engaged. *Cell* 54: 795–804
- Sedmak JJ, Grossberg SE (1977) A rapid, sensitive, and versatile assay for protein using Coomassie Brilliant Blue G250. *Anal Biochem* 79: 544–552
- Shibata D, Steczko J, Dixon JE, Hermadoson M, Yazdanparast R, Axelrod B (1987) Primary structure of soybean lipoxygenase I. *J Biol Chem* 262: 10080–10085
- Spector T (1978) Refinement of the Coomassie Blue method of protein quantitation. *Anal Biochem* 86: 142–146
- Watanabe Y, Ogawa T, Okada Y (1992) In vivo phosphorylation of the 30-kDa protein of tobacco mosaic virus. *FEBS Lett* 313: 181–184
- Whitelaw E, Proudfoot NJ (1986) Alpha-Thalassemia caused by a polyadenylic acid site mutation reveals that transcriptional termination is linked to 3' end processing in the human alpha-2-globin gene. *EMBO J* 5: 2915–2922
- Wiest DK, Hawley DK (1990) In vitro analysis of a transcription termination site for RNA polymerase II. *Mol Cell Biol* 10: 5782–5795
- Wu Y, Qiu X, Du S, Erickson L (1996) PO149, a new member of pollen pectate lyase-like gene family from alfalfa. *Plant Mol Biol* 32: 1037–1042
- Yang M-T, Gardner JF (1989) Transcription termination directed by heteroduplex threonine attenuator templates. *J Biol Chem* 266: 2634–2639
- Yanofsky C (1981) Attenuation in the control of expression of bacterial operon. *Nature* 289: 751–758
- Yanofsky RL, Fine M, Liu C (1988) Isolation and characterization of a soybean (*Glycine max*) lipoxygenase 3 gene. *Mol Gen Genet* 211: 215–222

Edited by L. McIntosh