Plant Molecular Biology **23**: 97–109, 1993.

Plant viral leaders influence expression of a reporter gene in tobacco

Mandy J.Dowson Day[1], Jennifer L.Ashurst[1], Shelley F.Mathias[1], John W.Watts[2], T.Michael A.Wilson[3] and Ray A.Dixon[1,*]

[1] *AFRC-IPSR Nitrogen Fixation Laboratory, University of Sussex, Brighton, Sussex, BN1 9RQ, UK (* author for correspondence); [2] AFRC-IPSR John Innes Centre for Plant Science Research, Colney Lane, Norwich, NR4 7UH, UK; [3] Department of Virology, Scottish Crop Research Institute, Dundee, Tayside, DD2 5DA, UK*

Received 24 September 1992; accepted in revised form 22 June 1993

Key words: cauliflower mosaic virus, β-glucuronidase, 35S promoter, tobacco mosaic virus, translation efficiency

Abstract

In order to optimise expression of a foreign protein in transgenic plants we investigated the potential benefits of including a viral untranslated leader sequence within a plant transformation vector. A variety of 5 leaders, including the tobacco mosaic virus (TMV) leader sequence and 31 nucleotides of the cauliflower mosaic virus (CaMV) 35S RNA leader, were compared. Viral leader constructs employing the 35S promoter and the reporter β-glucuronidase (GUS) were tested by electroporation into tobacco mesophyll protoplasts and against a cointroduced chloramphenicol acetyl transferase (CAT) gene in transgenic tobacco leaves. In the transient assay system, GUS activities from the viral leaders were compared with those from either a short, random leader or a translational fusion of the CaMV 19S RNA ORF VI to GUS. A two- to-three-fold enhanced level of expression resulted when these leaders were substituted with either the 35S RNA or the TMV leader sequences. This enhancement was further increased, to four- to five-fold, by inclusion of four or seven of the bases from the 35S transcription initiation site adjacent to the TMV leader. In transgenic tobacco the improved GUS levels were maintained from constructs including either the TMV leader (eight-fold) or this sequence with the addition of the 35S transcription initiation site bases (ten-fold). A comparison of GUS enzyme amounts with GUS mRNA amounts, using the CAT gene as an internal standard, revealed that TMV leader-bearing mRNA was translated from four- to six-fold more efficiently than the random leader control.

Introduction

Optimal expression of foreign genes in transgenic eukaryotes requires not only that the coding sequence is under the control of a strong promoter [33, 21] but also that the resultant mRNA is efficiently translated and is stable. Although a good terminator signal can help to ensure message stability, other factors within the message can enhance or diminish this effect. In order to ensure adequate expression of a transgene, factors within both the coding sequence and the 5′-untranslated leader must also be addressed. The latter can have significance in terms of message stability,

export from the nucleus, ribosome recruitment and translation initiation [31].

For 40S ribosome attachment most cytoplasmic mRNAs require a 5′ cap (m^7 GpppG) [38] and have untranslated leaders of 20 to 100 bases [26] which tend to form secondary structures that are relatively unstable. The 40S ribosome complex is believed to scan this leader sequence until reaching the most 5′ AUG where 80S ribosome assembly and translation initiation occur [28], although this may depend on the base sequence around the AUG [34, 26] and on secondary structure [25, 29].

In vivo and *in vitro* studies have revealed that a number of viral RNAs are translated at relatively high levels. This is primarily a consequence of their 5′ leaders, as demonstrated by comparisons of the translation rates of various heterologous mRNAs with substituted leaders [24, 15, 5, 40]. To date one of the most effective plant viral leaders is the 68 base genomic leader sequence of the U_1 strain of TMV RNA [15, 41]. Assays *in vitro* of RNA bearing a TMV-derived leader, in wheat germ extracts and rabbit reticulocyte lysates, have resulted in considerably enhanced expression levels of CAT, GUS and neomycin phosphotransferase (*nptII*) [14, 39]. Experiments, where luciferase, GUS and CAT mRNAs have been electroporated into tobacco protoplasts, have shown similar or even greater levels of TMV leader-associated enhancement [14, 15, 16, 17].

With the intention of constructing a vector for plant transformation that should ensure efficient ribosome recruitment and translation initiation by a foreign mRNA, we investigated the effects of the TMV leader, as well as other 5′ leaders, on the expression of a complete reporter gene in plant cells. Since TMV RNA is not normally transcribed from DNA but is replicated by viral replicase it is not apparent whether this leader may permit adequate transcription initiation when inserted into a plant gene. Therefore constructs bearing additional bases from the initiation site of the promoter were also tested. Our results demonstrate that the TMV RNA leader sequence can enhance levels of a reporter protein when incorporated into the 5′ leader of complete plant genes. We show that this enhancement is primarily due to increased translation of the hybrid mRNA.

Materials and methods

Strains, vectors and oligonucleotides

Escherichia coli host strain 71-18 (*thi*$^-$ *pro*$^-$ (F' *pro*$^+$ *lacZ* ΔM15 *lacI*^{q})) was used for both plasmid and template preparation. The plasmid vector pTZ18R (Pharmacia) was used for all cloning steps and the helper phage M13KO7 for template preparation. Oligonucleotide sequences were verified after cloning by dideoxynucleotide sequencing.

Plasmid constructs

To construct the promoter-untranslated leader (UTL) fusions, a *Pst* I-*Eco* RV fragment derived from pBI120 (M. Bevan, unpublished) and an *Eco* RV-*Pst* I fragment from pCaP35J [43] were cloned into the *Pst* I site of pTZ18 to reconstruct ca. 940 bp of the original CaMV 35S promoter (CabbS strain) [10] with the addition of a *Stu* I site around the putative transcription initiation site [20]. The promoter was truncated to 526 bp by *Sph* I and *Nco* I digestion and then religated in the presence of a linker which destroyed the *Sph* I and *Nco* I sites whilst creating unique *Sna* BI, *Spe* I, *Asu* II and *Hpa* I sites. 74 bp and 40 bp double-stranded oligonucleotides, designated Ω* and 35S* respectively, were ligated into the *Stu* I site at +1 (Fig. 1A and B). To restore the 35S transcription initiation site, four and seven bases of the 5′ end of the 35S untranslated leader, as determined by Guilley *et al.* [20], were reintroduced upstream of Ω* by oligonucleotide directed mutagenesis (Fig. 1B, Ω* + 4 and Ω* + 7).

Derivatives of each of the promoter-UTL constructs were prepared containing the GUS coding and nopaline synthase (*nos*) terminator sequences (Fig. 1B). This was achieved by ligating *Nco* I-*Asu* II and *Asu* II-*Eco* RI fragments from pRAJ275 (GUS with a Kozak consensus) and

pBI201.3 (a pUC derivative of pBI101.3) [23], respectively, into *Nco* I-*Eco* RI-cut vectors.

The plasmid pJIT75 (P. Mullineaux, unpublished) is a translational fusion containing the 35S promoter (bases −394 to +3), the CaMV 19S leader (bases +7 to +21) with the ORF VI ATG (at +13) fused, in frame, to the GUS coding sequence (Fig. 1B) with the CaMV 35S polyadenylation signal (equivalent to positions 7435 – 126 in the CaMV genome). The *nos* terminator derivative, pMJD84, was constructed by ligating the 35S promoter-GUS *Kpn* I-*Sac* I fragment from pJIT75 into *Acc* I-*Sac* I cut pMJD60 (Fig. 1B) using T4 DNA polymerase to fill in and blunt the linear DNA which was then religated. The *nos* terminator in pMJD85 (Fig. 1B) was substituted by the CaMV polyadenylation signal by ligating a 730 bp *Sac* I-*Eco* RV fragment from pJIT75 into *Sac* I-*Eco* RI-cut pMJD85 in the presence of a *Nar* I-*Eco* RI linker to produce pMJD88.

A short, random leader-GUS fusion, in pBI221.2, was used as the standard in all experiments. pBI221.2 is a pUC-based plasmid containing the 35S promoter (−831 to +5) and an 18 base untranslated leader (Fig. 1B), followed by the GUS coding and *nos* terminator sequences [23].

All vectors for plant transformation were constructed in the binary vector pBIN19 [4]. The CAT gene, included in each of the plant vectors as an internal standard, was provided by pJIT54 (P. Mullineaux, unpublished). To remove unwanted restriction sites from a polylinker within the CAT gene, the *Kpn* I-*Sac* I and *Sac* I-*Hin*d III CAT gene fragments from pJIT54 were ligated together. The *Hin*d III-*Eco* R1 GUS gene from pMJD62 (Fig. 1B) was ligated into pBIN19 together with the *Kpn* I-*Hin*d III fragment carrying the CAT gene and a 900 bp *Kpn* I-*Hin*d III linker producing pSFM8, with GUS and CAT genes arranged in a back-to-back orientation (Fig. 1C). The other plant vectors pSFM9, 10, 11 and 12 (Fig. 1B), were prepared by ligating the *Hin*d III-*Eco* RI GUS genes from pMJD60, pMJD62, pBI221.2 and pMJD85, respectively, with the *Kpn* I-*Hin*d III CAT gene plus a 25 bp *Eco* RI-*Kpn* I oligonucleotide linker into *Hin*d III-cut pBIN19 (Fig. 1C).

DNA preparation for electroporation

DNA recovered from CsCl gradients was phenolchloroform-extracted and reprecipitated before being linearized by restriction enzyme digestion, at a site immediately 3 to the terminator. Digestion was in the presence of RNase at 0.1 mg/ml. Complete digestion and the absence of chromosomal DNA and RNA contamination were verified by agarose gel electrophoresis. Linearized plasmid DNA was extracted with phenolchloroform and ethanol-precipitated before being resuspended at 25 μg/ml in 0.7 M mannitol, as determined by the absorbance at 260 nm.

Transient expression in protoplasts

Leaf mesophyll protoplasts of tobacco (*Nicotiana tabacum* cv. Xanthi) were prepared and cultured by the method of Watts *et al.* [42]. Protoplasts ($2–5 \times 10^5$/ml) were electroporated in 0.7 M mannitol containing 25 μg DNA/ml at 6 °C with a single, square voltage pulse, at 500 V/cm, of 1–3 ms duration. After electroporation, protoplasts were allowed to stand for 10 min, washed in 0.7 M mannitol and cultured in a simple medium supplemented with 50 μg/ml augmentin in an illuminated incubator at 26 °C [42]. After 40 h the protoplasts were resuspended, washed and lysed into GUS extraction buffer [23] using a 25–26 gauge syringe and centrifuged at 13000 rpm for 5 min. GUS activities of the supernatants were determined using 4-methylumbelliferyl-β-*D*-glucuronide (MUG) as substrate, by the method of Jefferson [23].

All data were calculated as GUS activity (pmol methylumbelliferone (MU) produced per hour per μg protoplast protein) per μg DNA electroporated. Three independent electroporation experiments were carried out each with triplicate 25 μg samples for each plasmid tested. Each replicate DNA sample was independently linearized and

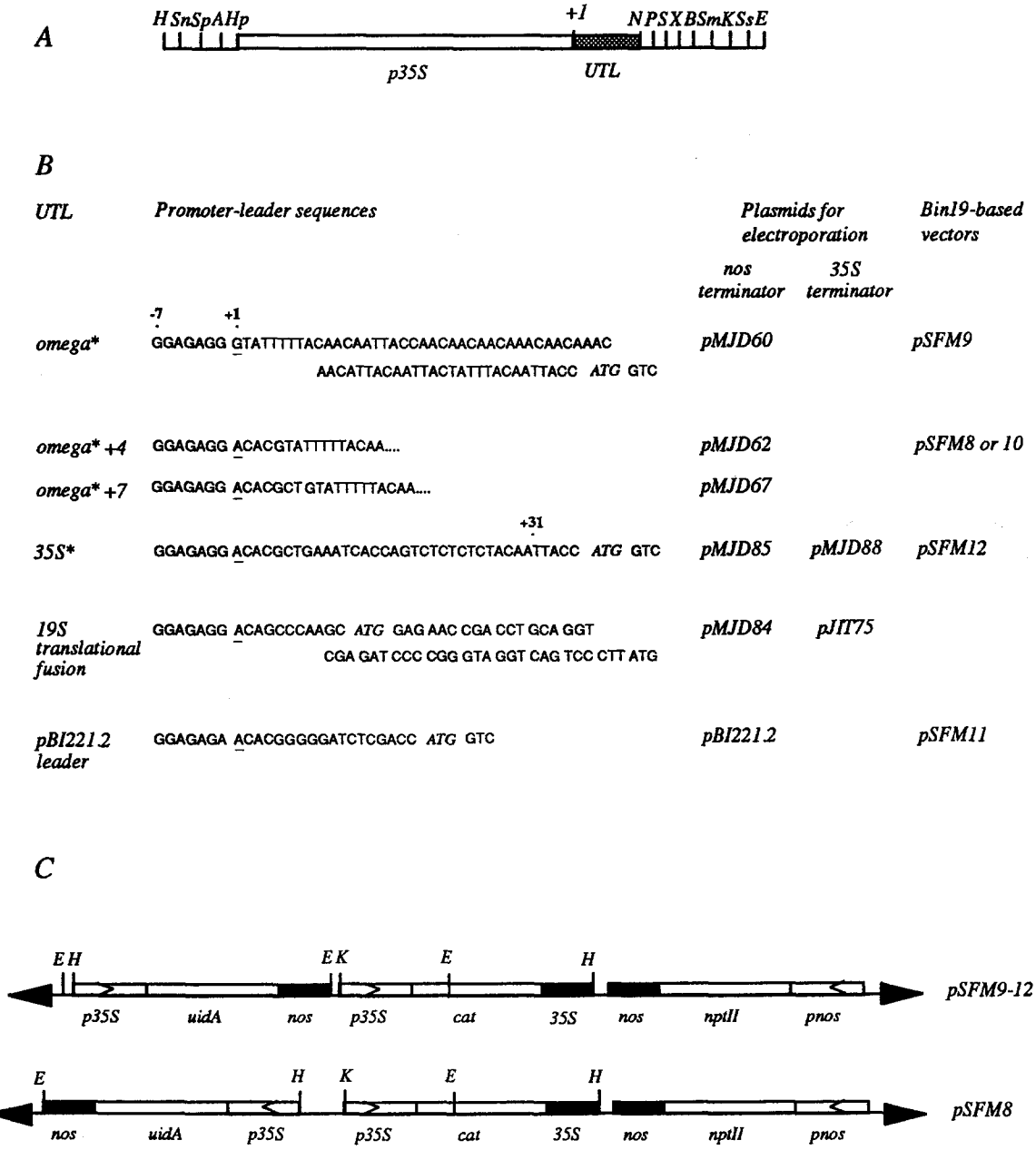

Fig. 1. Construction of viral leader vectors. **A.** The 35S promoter, truncated to 526 bp, with 5 and 3 multiple cloning sites and UTL ligated into previous *Stu* I site at + 1 [20]: H, *Hin*d III; Sn, *Sna* BI; Sp, *Spe* I; A, *Asu* II; Hp, *Hpa* I; N, *Nco* I; P, *Pst* I; S, *Sal* I; X, *Xba* I; B, *Bam* HI; Sm, *Sma* I; K, *Kpn* I; Ss, *Sst* I; E, *Eco* RI. **B.** Promoter-leader sequences with UTL and construct names. The equivalent map position, relative to the 35S TATA box, of the 5′ end of the major native 35S RNA is underlined [20]. Base positions are numbered relative to this. The Ω* leader is identical to the native U_1 TMV sequence [15] except, to create a *Nco* I site, it bears an A to C mutation one bp upstream of the ATG. The Ω* + 4 and Ω* + 7 leaders contain the first four and first seven bases of the 35S untranslated leader, respectively, 5′ to the TMV leader sequence Ω*. The 35S leader, designated 35S*, is identical to the native sequence as far as + 31, the sequence thereafter is the same as that for Ω* to maintain a consensus around the AUG initiation codon. The 19S-GUS translational fusion includes 3 codons from the 19S gene VI followed by 13 codons before the first ATG of GUS. pBI221.2 contains an 18 base random leader (as communicated by A. Goldsborough). **C.** pBin-19 based vectors for plant transformation: the upper diagram represents the tandemly-arranged GUS and CAT genes in pSFM9, 10, 11 and 12 and the lower diagram the back to back arrangement in pSFM8. p35S, 35S promoter; p*nos*, *nos* promoter; *uidA*, GUS-

the concentration determined. Since GUS expression levels were related to the amount of DNA added, GUS activity curves were produced to determine that the assay was not saturated.

Mean GUS activities for the Ω* + 4 leader ranged from 0.042 to 0.18 nmol MU per minute per mg soluble protein for the three experiments, therefore data were calculated as a percentage of the mean activity for pBI221.2 (100%) and significant differences were tested for by analysis of variance (anova) and a multiple range test.

Plant transformation

Plant transformation vectors pSFM8, 9, 10, 11 and 12 were mobilised to *Agrobacterium tumefaciens* LBA4404 by triparental mating [4]. Transformation of tobacco (*Nicotiana tabacum* cv. Samsum) leaf discs by *Agrobacterium* cultures was achieved according to the method of Horsch *et al.* [22]. First generation transformed plants were maintained in Magenta GA-7 pots on MS medium containing carbenicillin (200 μg/ml) and kanamycin (100 μg/ml), with an 18 h photoperiod, at 25 °C.

Protein extraction and enzyme assays

Young leaves were ground in liquid nitrogen with a pestle and mortar and thawed into GUS extraction buffer [23] supplemented with 0.5 mM leupeptin, in the ratio 1 g tissue/2 ml buffer. Insoluble material was removed by centrifugation at 13000 rpm for 5 min, at 4 °C, and the supernatant was mixed with an equal volume of 100% glycerol prior to storage at −2O°C.

GUS activities were determined using the Jefferson fluorogenic assay protocol [23] with several modifications. Assay buffer (2 mM MUG) was dispensed as 100 μl aliquots to which 10 μl of leaf extract were added. At 3 min intervals up to 24 min the reactions were stopped by adding 900 μl 0.2 M Na_2CO_3. The fluorescence of 100-fold dilutions in 0.2 M Na_2CO_3 were determined and points on the linear portion of the rate curve were used to calculate the relative GUS activities. Mean rates were determined from six independent reactions.

CAT activities were determined using a fluor diffusion assay [35]. Leaf samples were diluted 1:30 in GUS extraction buffer containing 0.5 mM leupeptin and incubated at 65 °C for 15 min, to inactivate interfering acylase activity, before adding chloramphenicol to 0.25 mM. The acetylation reaction was initiated by the addition of (acetyl-3H)acetyl coenzyme A (10 mCi/mmol) (NEN Dupont) to 0.2 mM final concentration. The reaction was gently overlaid with Econofluor (NEN Dupont) and the vials counted sequentially at timed intervals in a liquid scintillation counter. Mean activities were calculated from five linear rate curves.

RNA preparation and primer extension

Total RNA was extracted from young leaves as described by Prescott and Martin [37] except that LiCl-pelleted RNA was redissolved in TE and insoluble material was removed by centrifugation at 13000 rpm for 5 min at 4 °C before ethanol precipitation.

Oligonucleotides for the simultaneous primer extension of both GUS and CAT mRNA species were selected on the basis that the products should not comigrate and the T_m values were similar: 5′-AGACTGAATGCCCACAGGCC-3′ for priming on GUS mRNA and 5′-GCCATTGGGATATATCAACGGT-3′ for priming on CAT mRNA. The oligonucleotide kinasing and primer extension reactions were carried out as described by Dunsmuir *et al.* [9], except that MuMLV reverse transcriptase was used for the latter reaction. Annealing of the primer to RNA

coding sequence; *cat*, CAT-coding sequence; *nptII*, neomycin phosphotransferase II; *nos*, *nos* terminator; 35S, CaMV 35S terminator; the filled arrows represent the left and right T-DNA borders.

was carried out, without vanadyl, at 52 °C for 3 h.

The relative amounts of GUS and CAT primer extended products were determined by densitometric measurement of specific band intensities on preflashed film using a Molecular Dynamics computing densitometer. Data for film density within a standard boxed area were compiled from tracks within the linear range for film response from exposures of replicate gels of duplicate primer extension reactions.

GUS and CAT protein and mRNA levels were converted into percentages relative to the mean for the pBI221.2 GUS gene transformants for duplicate RNA samples and protein extracts from each transformant. Statistical analysis was carried out as described above for the electroporation experiments.

Results

Efficiency of the TMV and CaMV leaders in tobacco protoplasts

Both the TMV (a single-stranded RNA virus) and CaMV (a double-stranded DNA virus) genomes are expressed at high levels in the infected plant cytoplasm. We wanted to test whether the 5′ leaders of the RNA of these viruses may enhance expression levels in plant cells when incorporated into a foreign reporter gene.

The first 66 and 31 bases, respectively, of the TMV and the CaMV 35S RNA sequences, designated Ω* and 35S*, were ligated adjacent to the 35S promoter (Fig. 1B). The leaders were designed to have an identical sequence around the initiating GUS ATG to eliminate differences within the sequence necessary for ribosome assembly. Similarly, the sequence around the GUS ATG in the control gene from pBI221.2 does not differ at any of the major consensus bases (see Fig. 1B) [26, 34]. The constructs were first compared by electroporation of the linear DNA into tobacco mesophyll protoplasts, followed by GUS activity assays. The mean results from three experiments on different batches of protoplasts, each experiment using triplicate DNA samples, are presented in Fig. 2. In these experiments the Ω* and 35S* leaders resulted in a 2.6- to 2.8-fold greater level of GUS expression than the short, random leader in pBI221.2 (Fig. 2). There was no significant difference in expression levels between the Ω* and 35S* leaders.

For most eucaryotic promoters, including the 35S, the contribution of the sequence surrounding the transcription initiation site to the levels of transcribed RNA is not known. The main consensus at the cap site is a purine at position + 1 which is understood to facilitate the addition of

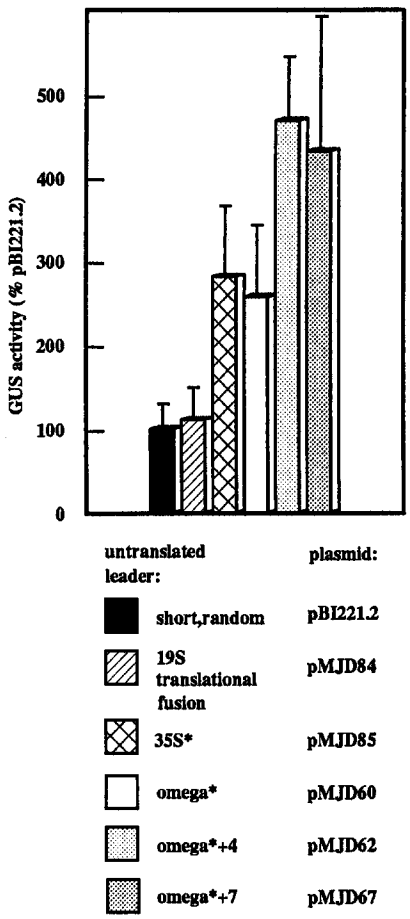

Fig. 2. Mean GUS activities in electroporated protoplasts. Data are the percentage means relative to pBI221.2 (100%) from three independent experiments with triplicate DNA samples. Significant differences were determined using the Student's *t* test to compute 95% confidence intervals (error bars).

a 5′ cap [38]. Since it is not apparent whether the TMV RNA leader sequence, when inserted directly into the putative CaMV 35S cap site [20], may influence transcription initiation, constructs were prepared containing the Ω* sequence 3′ to either four or seven bases of the original 35S leader (Fig. 1B), the fifth and eighth G of the 35S leader being substituted by the 5′ G of Ω*. In transient expression assays the inclusion of the extra four (Ω* + 4) or seven (Ω* + 7) transcription initiation site bases was found to result in statistically significant increases in GUS activity relative to Ω* alone, of 1.8- and 1.7-fold, respectively (Fig. 2).

Comparison with a translational fusion vector

As a second control, in addition to pBI221.1, we chose to compare a viral translational fusion vector, which employed the same promoter and terminator sequences, with the viral leader constructs. It has been considered that translational fusions employing plant initiation codons and 5′ leaders may facilitate translation initiation, and hence expression levels, in transgenic plants [21, 19]. The fusion comprised the CaMV 19S ORF VI leader and amino terminal coding sequence, preceding a polylinker which maintained GUS in frame (Fig. 1B). The activity of GUS was not expected to be affected by the amino terminal extension [23]. In transient expression assays GUS levels from the 19S translational fusion (Fig. 2) were not significantly greater than those for the pBI221.2 GUS gene, possibly endorsing the reported efficiency of expression from the latter construct [23].

It has been suggested that there may be a cooperative interaction between the 5′ and 3′ sequences of a mRNA. Consequently, we tested both the 35S* leader and the 19S translational fusion, with the CaMV 35S terminator sequence in place of the *nos* terminator. Results from two electroporation experiments did not indicate there was any specific interaction between either of the termination and leader sequences. Furthermore, the pooled data failed to show a statistically significant advantage attributable to either one of the terminators (data not shown).

Efficiency of the viral leaders in transgenic tobacco

The control GUS gene from pBI221.2 plus the viral leader-GUS constructs, 35S*, Ω* and Ω* + 4, were introduced into tobacco in tandem with a downstream CAT gene (Fig. 1C). The CAT gene, which is also transcribed from the 35S promoter, was included as an internal standard against which GUS message and enzyme levels were measured. The mean ratio of GUS to CAT enzyme activities for three or more independent tobacco transformants was related to the mean for four control transformants containing the pBI221.2 GUS gene (Table 1, column A). Of the 24 plants assayed, 6 gave results that were strongly at variance from the average (greater than five standard deviations from the mean) and therefore these data were not included (see Fig. 4 lanes 8 and 13 which lack significant amounts of CAT and GUS primer extended products respectively).

The results presented in Table 1 confirmed those from the electroporation experiments (Fig. 2) showing increased GUS activity levels from the viral leader vectors compared to the short, random leader derived from pBI221.2 in pSFM11. The 35S* leader gave the same 2.8-fold increase over the random leader as measured in the transient system, however, it was significantly less well expressed than both of the TMV leader constructs. For the latter constructs the magnitude of the increase was greater than in the transient assays, at eight-fold for Ω* and almost ten-fold for Ω* + 4 respectively. Under these conditions GUS was being expressed to as much as 0.2 to 0.6% of soluble leaf protein (the range for the four pSFM10 transformants).

In addition to the tandemly arranged GUS and CAT genes, a further construction was introduced in which the Ω* + 4 construct was arranged in a back-to-back orientation with the CAT gene (Fig. 1C). Since the promoter arrangement for this construct (pSFM8) differs from the tandemly

Table 1. Ratio of GUS/CAT enzyme activities, GUS/CAT mRNA amounts and enzyme/mRNA amounts in transgenic tobacco. Data are the percentage means calculated relative to the GUS gene from pBI221.2 (100%).

Vector	UTL	A GUS/CAT enzyme ratio		B GUS/CAT mRNA ratio		C enzyme/mRNA	
		mean	SEM	mean	SEM	mean	SEM
pSFM11	random	100.0	3.6	100.0	2.1	100.0	5.3
pSFM12	35S*	278.0	64.9	127.2	15.5	226.3	55.2
pSFM9	Ω*	811.3	69.9	204.1	27.0	420.6	36.0
pSFM10	Ω* + 4	976.4	118.1	231.3	40.6	452.4	59.2
pSFM8	Ω* + 4	(456.1)	(61.6)	(84.3)	(9.8)	(562.9)	(62.6)

SEM = standard error of the mean, where $n = 3$ to 7. Results for pSFM8 are in parenthesis since this construct contains a different promoter arrangement

arranged genes (pSFM10) a comparison of the relative GUS and protein message levels, which reflect transcription rates, is not appropriate (Table 1, pSFM8 data in parenthesis). However it is of interest that the back to back promoter arrangement was at least as stable as the tandemly arranged genes since none of the transgenic plants derived from pSFM8 gave results that were strongly at variance from the average.

Analysis of mRNA

Mapping the 5 termini of the viral RNA leaders

The viral leader constructs, and also the GUS gene derived from pBI221.2 and the CAT internal standard, are expressed from the CaMV 35S promoter. To determine the 5′ termini of the mRNA species from these constructs the products of the GUS- and CAT-specific primer extension reactions were electrophoresed alongside sequencing reactions in which the same primers were hybridised to the corresponding DNA templates (Fig. 3). The major 35S RNA transcript from the CaMV genome was reported by Guilley *et al.* [20] to initiate at the A in the following sequence: AGAGGACACGCTGA (defined as + 1 in Fig. 1). For the viral leader constructs and the CAT gene the major primer extension prod-

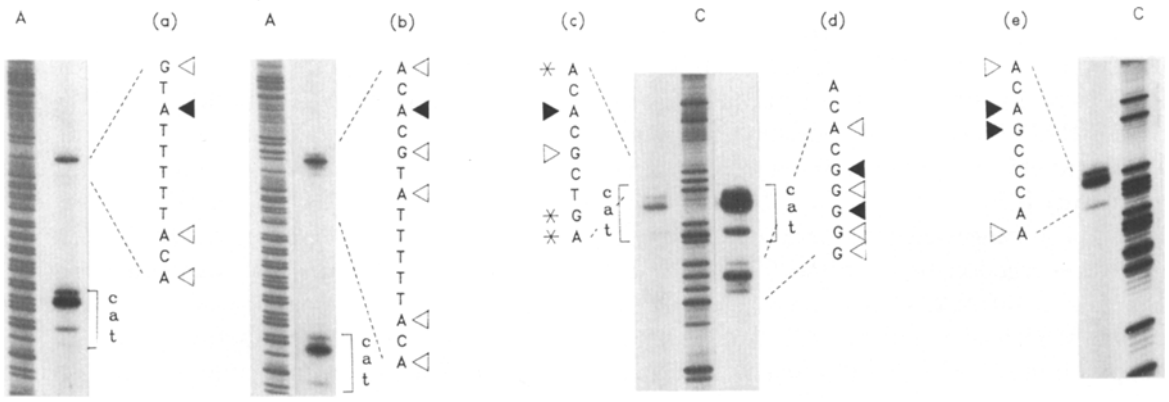

Fig. 3. GUS and CAT transcript 5′ termini in transgenic tobacco mapped by primer extension analysis. Primer extended products were electrophoresed alongside sequencing reactions on the corresponding DNA templates, marked A or C, using the same GUS or CAT specific primers. The sequences shown extend downstream from the same map position, relative to the 35S TATA box, as the major 35S RNA 5′ terminus [20]. Major (▸), minor (▹) and faintly visible (★) bands are marked. (a) Ω*-GUS; (b) Ω* + 4-GUS; (c) 35S*-GUS; (d) pBI221.2 leader-GUS products with 35S*-GUS sequencing reaction; (e) CAT.

ucts mapped to the adenine residue two bases downstream from this A, at the +3 position (Fig. 3a, b, c and e). However, in each case the most distal 5′ terminus, represented as a relatively minor primer extension product, mapped to the +1 position. For the pBI221.2 GUS gene the major band mapped to the next purine residue at +5 (Fig. 3d). It is perhaps significant that the 35S promoter in this construct has suffered a G to A mutation at position −1 (see Fig. 1B). The fact that there are several products for each construct, all of which extend to a purine residue, is in accordance with previously presented results from similar experiments on 35S-derived and other plant transcripts [20, 6].

Enhanced mRNA levels from the viral leader constructs

The coextension of the GUS and CAT mRNA species by reverse transcriptase, from specific primers, was used to quantitate the relative levels of the GUS transcripts. An example of the primary data is shown in Fig. 4. All GUS or CAT specific bands (see Fig. 3) were included in the calculation of the total signal intensity which was determined by densitometry. The ratio of GUS to CAT signal intensity for each transformant (as determined by densitometry) was related to the mean for the pBI221.2 GUS gene (Table 1, column B). The relative mean steady-state levels of GUS message from the tandemly arranged Ω* and Ω* + 4 constructs were significantly greater than that for the pBI221.2 gene, at 2.0- and 2.3-fold respectively.

Translational advantage attributable to the TMV leaders

The ratio of enzyme activity to mRNA amount was determined for each transformant (Table 1, column C). The means, which reflect the relative translation rates of the different mRNA species compared to the pBI221.2 GUS gene in pSFM11, reveal that the Ω* and Ω* + 4 messages are translated more efficiently than the control random leader message, by 4.2- and 4.5-fold. The ratio of 5.6-fold for the back-to-back Ω* + 4 construct, pSFM8, which would be expected to be comparable to that for pSFM10, confirms the translational advantage attributable to the TMV leaders.

Fold predictions for the viral leader-GUS fusions

Secondary structure predictions for the expected mRNAs were generated using the Wisconsin FOLD program [8]. Although different lengths of

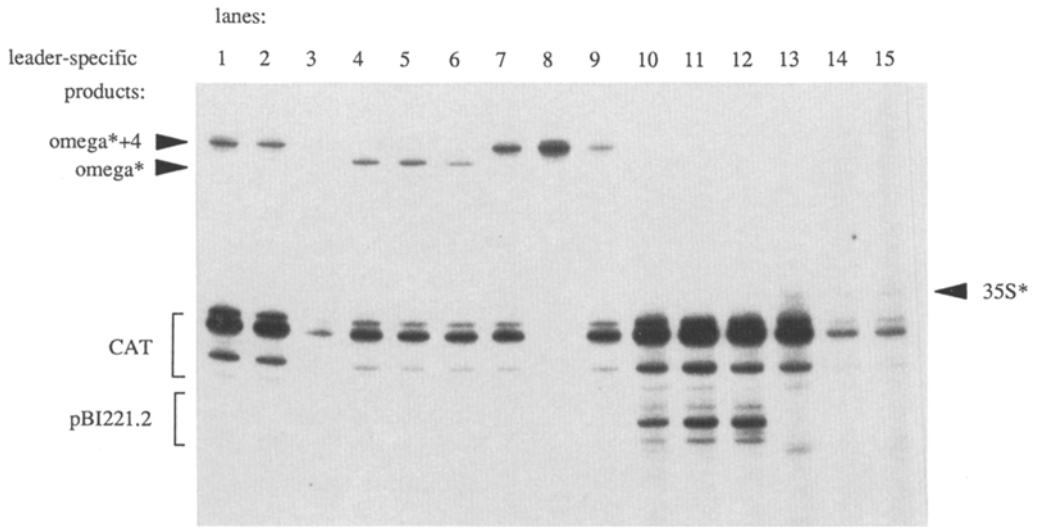

Fig. 4. Primer extension analysis of total RNA samples from independent transgenic tobacco plants using GUS and CAT specific primers in the same reaction. Lanes 1–3, pSFM8; lanes 4–6, pSFM9; lanes 7–9, pSFM10; lanes 10–12, pSFM11; lanes 13–15, pSFM12.

sequence affect the computer prediction, patterns of folding were often maintained, and the FOLDs presented here for 103 bases of sequence are representative of all three leaders. All fold predictions for the Ω* leader, when placed 5′ to various mRNAs, indicate that at least 44 of the 5′ bases remain unpaired (Fig. 5a). This is one of the features of the TMV leader sequence which is thought to promote high expression levels of the viral RNA [41, 39]. The fold for the 35S* leader showed more unpaired 5′ bases than that for the pBI221.2 random leader (Fig. 5b and c), however the ΔG value for this construct was slightly greater.

Discussion

In eukaryotes, apart from the sequence around the initiating AUG [26, 34] and the first transcribed base [38], there is no widespread consensus element within the untranslated leader of a mRNA which, like the prokaryotic Shine-Dalgarno sequence, might facilitate high expression levels. Rather it would appear that the control of the amount [31] and also pattern [27] of translation is primarily achieved through leader secondary structure [32], although evidence for the existence of regulatory sequences within UTLs is increasing [3, 7]. The advantage that relatively unstructured viral 5′ leaders can confer on heterologous mRNAs synthesised *in vitro* has already been described. Our results demonstrate that this advantage is maintained on transcription of the TMV RNA leader from a strong plant promoter.

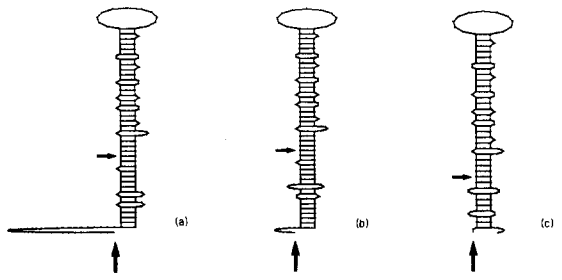

Fig. 5. FOLD predictions for RNA 5′-untranslated leaders and coding sequences extending 103 bases into the GUS-coding sequence. (a) Ω*-GUS, ΔG = − 117 kJ/mol; (b) 35S*-GUS, ΔG = − 139 kJ/mol and (c) pBI221.1 GUS gene, ΔG = − 113 kJ/mol. Large arrows indicate the most distal 5′ terminus of the mRNA relative to the translation initiation site (small arrows).

The Ω*-GUS construct was identical to the native U_1 strain TMV leader with the exception of an A to C change at −1 with respect to the ATG. This change, which affects neither of the main consensus bases, at −3 and +4, has been found to have no significant effect on expression levels in tobacco [19]. On electroporation into tobacco mesophyll protoplasts the Ω* construction was more strongly expressed than either the control random leader gene from pBI221.2 or a CaMV 19S leader translational fusion. Another viral UTL, derived from potato virus X and which shows some homology to the TMV leader, has been reported to similarly enhance expression levels in a transient assay system [36]. In transgenic tobacco, the enhancement attributable to Ω* was increased to eight-fold, the significance of which is not inconsiderable with respect to optimising foreign gene expression in plants.

The importance of the 35S transcription initiation site was investigated by the construction of TMV Ω* derivatives bearing either four or seven of the 35S RNA leader bases 5′ to the Ω* sequence. In the transient assay system these constructs were significantly more highly expressed than Ω*-GUS, and this amounted to a ten-fold increase in GUS levels relative to the control random leader gene in whole plants.

A second viral leader construction, termed 35S*-GUS, was designed that would maintain the 35S RNA sequence adjacent to the promoter. This construct included only the first 31 bases of the native 600 nucleotide long leader which precedes ORF VII in the viral genome. It therefore comprised one of two CT-rich regions from the native leader followed by an identical sequence, around the initiating ATG, to the Ω* constructs. Rather than promoting expression levels of ORF VII, the complete leader, which may form a large stem loop structure [11], has been found to reduce expression from 8- to 100-fold in protoplasts [2, 12, 13]. In electroporated protoplasts the 35S*

leader was found to be equally as efficient as the Ω* leader. However, in whole plants the construct did not show the same level of enhancement attributable to the TMV leaders.

In order to determine whether the increased GUS protein levels may be attributed solely to increased translation from the TMV leaders, the steady-state levels of the GUS messages were measured in the transgenic plants. Results revealed that the mRNA from both of the TMV leader construct was twice as abundant as that for the pBI221.2-derived gene. Since the TMV RNA leader has not been found to enhance protein levels by stabilising a message or reducing 5′ to 3′ exonuclease degradation of a mRNA [39], the two-fold increased message level may be a more likely consequence of differences in transcription rate. However the effect of a different UTL on the stability of a message must in part depend on the intrinsic stability of the original or control UTL.

The most significant effect of the TMV leaders, however, was exerted at translation, with the Ω* and Ω* + 4 constructs showing a four- to six-fold increase in the protein/mRNA ratio relative to the pBI221.2-derived gene. The ratio for Ω* + 4-GUS is not significantly different to that for Ω*-GUS indicating that the 'translatability' of the two messages is similar. This may be expected since RNA fold predictions for the two constructs indicate that the secondary structure of the modified leader would be the same as that for Ω* (Fig. 5). From these predictions it would appear that the TMV leader, like the 5′ cap and some of the associated binding factors, reduces RNA secondary structure, thereby rendering the 5′ terminus more accessible to scanning by the 40S ribosomal subunits. This is supported by the facts that the 5′ cap has less effect on the TMV leader than other leaders [14, 39] and that Ω-CAT translation is eIF-4E-independent [1]. There is a similarly reduced requirement for certain initiation factors for an RNA bearing the alfalfa mosaic virus leader [5]. It has also been found that by increasing the length of an unstructured leader more 40S ribosomal subunits can accumulate resulting in an increased protein yield [30]. In addition, some of this enhancement may be sequence-dependent, a fact which, with the identification of a potential sequence-specific protein binding factor, has led Gallie and Walbot to conclude that there may be a core regulatory element within the TMV leader [18]. It therefore appears that the translational advantage conferred by the TMV leader may be attributed not only to the reduction of secondary structure at the 5′ end of the RNA and its length, of 68 nucleotides, but also to the existence of at least one initiation factor binding motif within it.

In conclusion, our results confirm that viral leader sequences influence the level of expression of an introduced gene in plant cells. Viral leaders may be particularly advantageous when introducing prokaryotic coding sequences into transgenic plants, especially where the RNAs may be less stable than GUS, since the beneficial effects of the TMV leader were found to be greater for unstable mRNAs such as the *nptII* message [14]. Furthermore, under certain stress conditions where transcription and/or translation of many genes may be repressed, the advantages of an unstructured RNA leader may be even more marked [27].

Acknowledgements

We wish to thank M. Bevan for providing pBI120, W. Liang and A. Goldsborough for pBI221.2 and sequence details, J. Yamaya for pCaP35J, W. Liang for pRAJ275 and pBI201.3 and P. Mullineaux for pJIT75. We are also grateful to M. Buck, M. Merrick, S. Austin and B. Smith for their comments on the manuscript and to Rosemary Foote and Carol Sterenberg for typing it.

References

1. Altmann M, Blum S, Wilson TMA, Trachsel H: The 5′-leader sequence of tobacco mosaic virus RNA mediates initiation factor-4E-independent, but still initiation factor-4A-dependent translation in yeast extracts. Gene 91: 127–129 (1990).
2. Baughman G, Howell SH: Cauliflower mosaic virus 35S

RNA leader region inhibits translation of downstream genes. Virology 167: 125–135 (1988).
3. Berry JO, Breiding DE, Klessig DF: Light-mediated control of translational initiation of ribulose-1.5-biphosphate carboxylase in amaranth cotyledons. Plant Cell 2: 795–803 (1990).
4. Bevan M: Binary *Agrobacterium* vectors for plant transformation. Nucl Acids Res 12: 8711–8721 (1984).
5. Browning KS, Lax SR, Humphreys J, Ravel JM, Jobling SA, Gehrke L: Evidence that the 5′-untranslated leader of mRNA affects the requirement for wheat germ initiation factors 4A, 4F, and 4G. J Biol Chem 263: 9630–9634 (1988).
6. Chandler PM, Huiet L: Primer extension studies on α-amylase mRNAs in barley aleurone 1. Characterization and quantification of the transcripts. Plant Mol Biol 16: 627–635 (1991).
7. Dawe RK, Lachmansingh AR, Freeling M: Transposon-mediated mutations in the untranslated leader of maize *adh1* that increase and decrease pollen-specific gene expression. Plant Cell 5: 311–319 (1993).
8. Devereaux J, Haeberli P, Smithies O: A comprehensive set of sequence analysis programs for the VAX. Nucl Acids Res 12: 387–395 (1984).
9. Dunsmuir P, Bond D, Lee K, Gidoni D, Townsend J: Stability of introduced genes and stability in expression. In: Gelvin SB, Schilperoort RA (eds) Plant Molecular Biology Manual, pp. C1: 1–17, Kluwer Academic Publishers, Dordrecht (1988).
10. Franck A, Guilley H, Jonard G, Richards K, Hirth L: Nucleotide sequence of cauliflower mosaic virus DNA. Cell 21: 285–294 (1980).
11. Fütterer J, Gordon K, Bonneville JM, Sanfaçon H, Pisan B, Penswick J, Hohn T: The leading sequence of caulimovirus large RNA can be folded into a large stem loop structure. Nucl Acids Res 16: 8377–8390 (1988).
12. Fütterer J, Gordon K, Pfeiffer P, Sanfaçon H, Pisan B, Bonneville JM, Hohn T: Differential inhibition of downstream gene expression by the cauliflower mosaic virus 35S RNA leader. Virus Genes 3: 45–55 (1989).
13. Fütterer J, Gordon K, Sanfaçon H, Bonneville JM, Hohn T: Positive and negative control of translation by the leader sequence of cauliflower mosaic virus pregenomic 35S RNA. EMBO J 9: 1697–1707 (1990).
14. Gallie DR, Sleat DE, Watts JW, Turner PC, Wilson TMA: The 5′ leader of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts *in vitro* and *in vivo*. Nucl Acids Res 15: 3257–3273 (1987).
15. Gallie DR, Sleat DE, Watts JW, Turner PC, Wilson TMA: A comparison of eucaryotic viral 5-leader sequences as enhancers of mRNA expression *in vivo*. Nucl Acids Res 15: 8693–8710 (1987).
16. Gallie DR, Sleat DE, Watts JW, Turner PC, Wilson TMA: Mutational analysis of the tobacco mosaic virus 5′-leader for altered ability to enhance translation. Nucl Acids Res 16: 883–893 (1988).
17. Gallie DR, Lucas WJ, Walbot V: Visualizing mRNA expression in plant protoplasts: factors influencing efficient mRNA uptake and translation. Plant Cell 1: 301–311 (1989).
18. Gallie DR, Walbot V: Identification of motifs within the tobacco mosaic virus 5′-leader responsible for enhancing translation. Nucl Acids Res 20: 4631–4638 (1992).
19. Guerineau F, Lucy A, Mullineaux P: Effect of two consensus sequences preceding the translation initiator codon on gene expression in plant protoplasts. Plant Mol Biol 18: 815–818 (1992).
20. Guilley H, Dudley RK, Jonard G, Balàzs E, Richards KE: Transcription of cauliflower mosaic virus DNA: detection of promoter sequences and characterization of transcripts. Cell 30: 763–773 (1982).
21. Harpster MH, Townsend JA, Jones JDG, Bedbrook J, Dunsmuir P: Relative strengths of the 35S cauliflower mosaic virus, 1, 2 and nopaline synthase promoters in transformed tobacco, sugarbeet and oilseed rape callus tissue. Mol Gen Genet 212: 182–190 (1988).
22. Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT: A simple and general method for transferring genes into plants. Science 227: 1229–1231 (1985).
23. Jefferson RA: Assaying chimeric genes in plants: the GUS gene fusion system. Plant Mol Biol Rep 5: 387–405 (1987).
24. Jobling SA, Gehrke L: Enhanced translation of chimaeric messenger RNAs containing a plant viral untranslated leader sequence. Nature 325: 622–625 (1987).
25. Kozak M: Influence of mRNA secondary structure on initiation by eukaryotic ribosomes. Proc Natl Acad Sci USA 83: 2850–2854 (1986).
26. Kozak M: An analysis of 5-noncoding sequences from 699 vertebrate messenger RNAs. Nucl Acids Res 15: 8125–8148 (1987).
27. Kozak M: Leader length and secondary structure moderate mRNA function under conditions of stress. Mol Cell Biol 8: 2737–2744 (1988).
28. Kozak M: The scanning model for translation: an update. J Cell Biol 108: 229–241 (1989).
29. Kozak M: Circumstances and mechanisms of inhibition of translation by secondary structure in eukaryotic mRNAs. Mol Cell Biol 9: 5134–5142 (1989).
30. Kozak M: Effects of long 5′ leader sequences on initiation by eukaryotic ribosomes *in vitro*. Gene Expression 1: 117–125 (1991).
31. Kozak M: Structural features in eukaryotic mRNAs that moderate the initiation of translation. J Biol Chem 266: 19867–19870 (1991).
32. Kuhlemeier C: Transcriptional and post-transcriptional regulation of gene expression in plants. Plant Mol Biol 19: 1–14 (1992).
33. Lawton MA, Tierney MA, Nakamura I, Anderson E, Komeda Y, Dubé P, Hoffman N, Fraley RT, Beachy RN: Expression of a soybean β-conglycinin gene under the control of the cauliflower mosaic virus 35S and 19S pro-

moters in transformed petunia tissues. Plant Mol Biol 9: 315–324 (1987).
34. Lütcke HA, Chow KC, Mickel FS, Moss KA, Kern HF, Scheele GA: Selection of AUG initiation codons differs in plants and animals. EMBO J 6: 43–48 (1987).
35. Nielsen DA, Chang T-C, Shapiro DJ: A highly sensitive, mixed-phase assay for chloramphenicol acetyltransferase activity in transfected cells. Anal Biochem 179: 19–23 (1989).
36. Pooggin MM, Skryabin KG: The 5′-untranslated leader sequence of potato virus X RNA enhances the expression of a heterologous gene *in vivo*. Mol Gen Genet 234: 329–331 (1992).
37. Prescott A, Martin C: A rapid method for the quantitative assessment of levels of specific mRNAs in plants. Plant Mol Biol Rep 4: 219–224 (1987).
38. Shatkin AJ: mRNA caps-old and newer hats. BioEssays 7: 275–277 (1987).
39. Sleat DE, Hull R, Turner PC, Wilson TMA: Studies on the mechanism of translational enhancement by the 5′-leader sequence of tobacco mosaic virus RNA. Eur J Biochem 175: 75–86 (1988).
40. Smirnyagina EV, Morozov SY, Radionova NP, Miroschnichenko NA, Solovyev AG, Fedorkin ON, Atabekov JG: Translational efficiency and competitive ability of mRNAs with 5′-untranslated $\alpha\beta$-leader of potato virus X RNA. Biochimie 73: 587–598 (1991).
41. Wilson TMA, Saunders K, Dowson Day MJ, Sleat DE, Trachsel H, Mundry KW: Effects of the 5′ leader sequence of tobacco mosaic virus RNA, or derivatives thereof, on foreign mRNA and native viral gene expression. In: McCarthy JEG, Tuite MF (eds) Post-transcriptional Control of Gene Expression, pp. 261–275, NATO-ASI Series, Vol H49 (1990).
42. Watts JW, Dawson JRO, King JM: Infection of plant protoplasts with two or more viruses. In: Ingram DS and Helgeson JP (eds) Tissue Culture Methods for Plant Pathologists, pp. 97–101, Blackwell Scientific Publications (1980).
43. Yamaya J, Yoshioka M, Meshi T, Okada Y, Ohno T: Expression of tobacco mosaic virus RNA in transgenic plants. Mol Gen Genet 211: 520–525 (1988).