Plant Molecular Biology **15**: 65–79, 1990.

Analysis of leaky viral translation termination codons *in vivo* by transient expression of improved β-glucuronidase vectors

James M. Skuzeski, Lindy M. Nichols and Raymond F. Gesteland
The Howard Hughes Medical Institute and Department of Human Genetics, 743 Wintrobe Bldg., The University of Utah School of Medicine, Salt Lake City, UT 84132, USA

Received 17 January 1990; accepted in revised from 19 March 1990

Key words: β-glucuronidase vectors, leaky stop codons, RNA viruses, tobacco, transient expression, translation

Abstract

Plant RNA viruses commonly exploit leaky translation termination signals in order to express internal protein coding regions. As a first step to elucidate the mechanism(s) by which ribosomes bypass leaky stop codons *in vivo*, we have devised a system in which readthrough is coupled to the transient expression of β-glucuronidase (GUS) in tobacco protoplasts. GUS vectors that contain the stop codons and surrounding nucleotides from the readthrough regions of several different RNA viruses were constructed and the plasmids were tested for the ability to direct transient GUS expression. These studies indicated that ribosomes bypass the leaky termination sites at efficiencies ranging from essentially 0 to ca. 5% depending upon the viral sequence. The results suggest that the efficiency of readthrough is determined by the sequence surrounding the stop codon. We describe improved GUS expression vectors and optimized transfection conditions which made it possible to assay low-level translational events.

Introduction

With few exceptions, translation starts on eukaryotic mRNAs at 5′ proximal AUG codons in appropriate nucleotide contexts [34] and the evolutionary constraints imposed by this canonical initiation mechanism are apparent in plant RNA virus genomes and their expression strategies. Genome segmentation, production of subgenomic mRNAs, synthesis of polyprotein precursors and other strategies employed by these viruses are compatible with expression of the viral genome in the form of one or more functionally monocistronic mRNAs [15].

The expression of internal open reading frames is achieved in a variety of plant RNA viruses, however, with the use of leaky stop codons which separate two in-frame protein coding cistrons [8, 37, 39, 41, 47, 48, 55]. Recent studies [37, 52, 54] indicate that some plant viruses may also utilize ribosomal frameshifting to couple the translation of protein coding regions which have short overlaps in different reading frames as is found in many animal retroviruses [22, 23].

The most thoroughly studied leaky termination site is at the end of the 5′ proximal cistron in the genomic RNA of tobacco mosaic virus (TMV). Termination at the leaky UAG (amber) stop

codon results in the synthesis of a 126 kDa protein whereas occasional readthrough leads to the synthesis of a C-terminal extension, resulting in a 183 kDa product [47; see below, Fig. 3B]. Studies on site-directed mutants of TMV [20] show that the leaky amber codon is essential for productive infection and strongly suggest that readthrough regulates the expression of the putative viral RNA-dependent RNA replicase encoded by the downstream cistron [29].

It is generally assumed that the leaky amber of TMV is read by a suppressor tRNA primarily as a result of codon-anticodon interactions and evidence for this type of mechanism has been presented [2, 3]. However, the recent findings of 16S rRNA mutations which suppress certain stop codons in *E. coli* [40] and yeast mitochondria [51] raise the possibility that the leakiness of the TMV stop could be caused by a unique interaction between rRNA or polypeptide release factor and a sequence element in the readthrough region. Such a model would be consistent with the fact that the sequence around the leaky TMV amber is also found at the readthrough sites of some other viral RNAs (see below, Tables 2 and 3). Alternatively, the ability of ribosomes to bypass the stop could be due to an extended shift of the viral RNA during translation of the readthrough region sequence. Precedence for this type of mechanism, which would involve the dissociation of the peptidyl tRNA from the RNA template and reengagement on a downstream codon, comes from studies on *E. coli* in which 'tRNA hopping' has been demonstrated by protein sequence analysis [42, 53]. The RNA sequence around the leaky TMV termination site [14], 5′...-CAA-UAG-CAA-...3′, could support such a process since the stop is flanked by identical glutamine codons which would provide compatible 'take-off' and 'landing' sites for a hopping peptidyl-glutamine tRNA [42, 53].

Here, we describe a system for investigating translational events, such as stop codon readthrough, which can be coupled to the transient expression of reporter mRNAs in plant cells. We have utilized the reporter gene, *E. coli uidA*, whose protein product, β-glucuronidase (GUS) [25], is very stable in plant cells and is easily detected by sensitive, fluorometric methods [24, 26]. Studies indicate that changes in the NH_2-terminal coding region on the GUS gene may not interfere with enzyme activity [21, 24, 30] and we have constructed GUS expression vectors which can be specifically altered to insert sequences of interest into the 5′ end of the gene. We show that with improved expression vectors and the use of appropriate gene transfer conditions, readthrough over stop codons can be monitored *in vivo* in tobacco protoplasts by transient expression of GUS. Furthermore, we present an analysis of the ability of ribosomes to bypass 'leaky' translation termination signals derived from several different RNA viruses, and the results suggest that the different stops may be bypassed by distinct mechanisms.

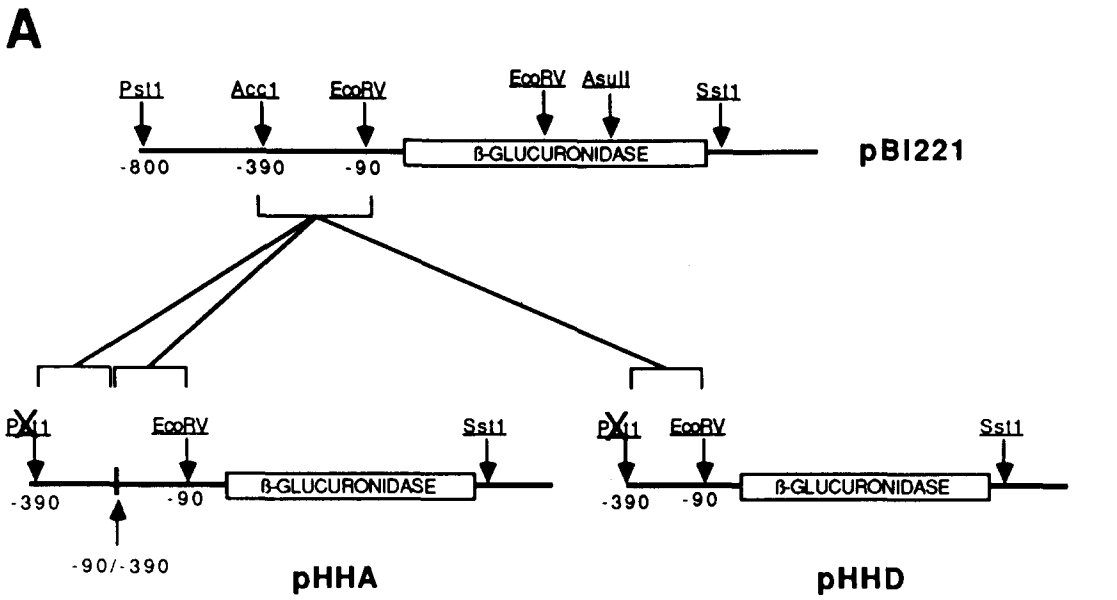
Materials and methods

Restriction enzymes, alkaline phosphatase, bovine serum albumin, S1 nuclease and T4 DNA ligase were obtained from either Boehringer Mannheim, New England Biolabs, or Promega Biotec. Sequenase was purchased from United States Biochemical Corporation. Oligodeoxyribonucleotides were synthesized on either an Applied Biosystems 380A or 380B DNA synthesizer. Before use in ligations or for DNA sequencing, the crude oligonucleotides were ethanol-precipitated, dissolved in 10 mM Tris HCl-pH 8, 0.1 mM EDTA and quantitated by absorbance at 260 nm. The plasmids, pBI221 and pRAJ275, and substrate (4-methylumbelliferyl-β-D-glucuronide [MUG]) for GUS assays were obtained from Clontech.

DNA constructions

DNA manipulations were carried out by standard methods [36]. All constructions described in this paper were subjected to DNA sequence analysis [49] in accordance with the instructions supplied in the USB Sequenase kit in order to verify the presence of the desired modifications.

The GUS vector, pBI221 [26, see Fig. 1], was digested with *Pst* I and *Acc* I in order to remove all cauliflower mosaic virus (CaMV) sequences that are located more than 390 nucleotides upstream of the 35S RNA promoter transcription initiation site [17]. The ends of the digested DNA were made blunt by treatment with S1 nuclease and the plasmid was then either recircularized to



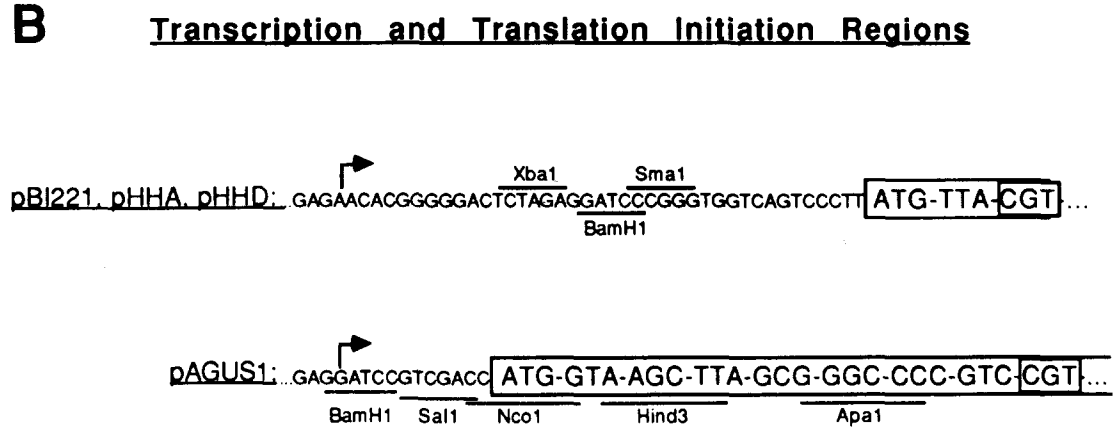


Fig. 1. Constructions derived from pBI221. The features of pBI221 relevant to the construction work done in this study are indicated at the top in A. (Note that the sites for *Xba* I, *Bam* HI and *Sma* I between the transcription and translation initiation sites are shown in B.) The plasmids, pHHA and pHHD, derived directly from pBI221 by modification of the promoter region are shown below. Thin horizontal lines represent non-protein coding sequences (the 35S RNA promoter at the left and the nopaline synthase 3′ region at the right) while the GUS coding region is depicted as a box. Sites in the 35S RNA promoter region are numbered according to their position relative to the RNA start site. The *Pst* I site of pBI221 was destroyed in all derived clones as indicated by the X. Sequences around the 5′ end of the GUS gene in different DNAs are shown in B. Nucleotides in the GUS coding region are indicated with large letters starting with the ATG translation initiation codon. The arginine codon (CGT) that is boxed in all of the constructs corresponds to codon three of the wild-type GUS coding region. Arrows pointing to the right show the positions of the RNA start site in the different DNAs. Figures were not drawn to scale. For more details, see the text or Materials and methods.

construct pHHD (Fig. 1) or else ligated to the purified, blunt-ended *Acc* I-*Eco* RV fragment derived from positions − 390 to − 90 of the same promoter. A plasmid (pHHA, Fig. 1A) which contained one extra, directly repeated copy of the − 390 to − 90 segment was identified by analysis of plasmids from *E. coli* transformants.

In order to construct a derivative of pHHA which contained an accessible transcription initiation region, it was necessary to remove a relatively large *Eco* RV fragment from the DNA as there were no convenient sites which flank the RNA start site. Complete digestion removed the region between position − 90 of the promoter and position 790 of the GUS coding region in the plasmid. The deleted region was replaced with a 1.5 kilobase pair (kb) *Eco* RV fragment from pDO432 [45], a plasmid which contains a luciferase gene attached to a modified 35S promoter with a *Bam* HI site that ovelaps the transcription initiation site. The 1.5 kb fragment extended from position − 90 of the promoter (as in pHHA above) to an *Eco* RV site about 1.3 kb into the luciferase coding region. The GUS coding region in the resulting DNA was then reconstructed by replacing the region from the *Bam* HI site in the promoter to the *Sst* I site at the 3′ end of the GUS coding region with the *Bam* HI-*Sst* I GUS coding region fragment from pBI221 (Fig. 1A and 1B). These manipulations produced the plasmid pHHAB (not shown).

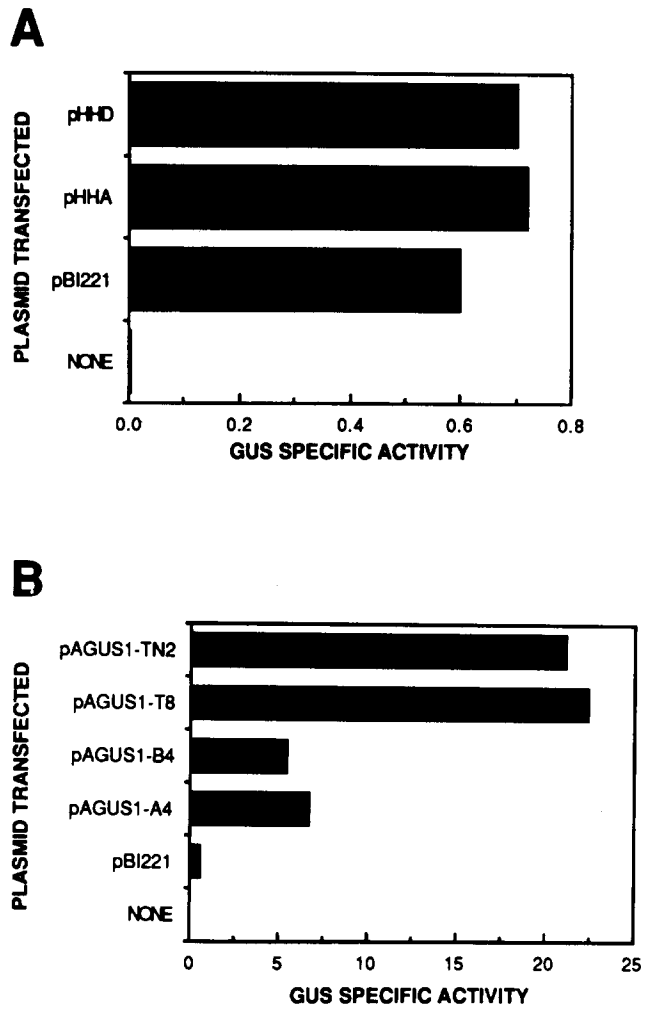


Fig. 2. Transient activity of GUS expression vectors in BY2 protoplasts. The data in 2A and 2B were obtained in different experiments. Each of the indicated plasmids (100 μg) was introduced into 3 × 10^6 protoplasts by electroporation using a 240 μF capacitor charged to 400 V. GUS specific activity shown for each construct is the mean of three electroporation experiments carried out with the same batch of protoplasts.

The *Bam* HI-*Asu* II fragment which extends from the RNA start site to position 1095 of the GUS coding region in pHHAB was replaced with a *Bam* HI-*Asu* II fragment from pHHBA. This latter plasmid (not shown) is a derivative of pRAJ275 [24] in which the region around the 5′ end of the GUS gene was modified to include several restriction sites. The restriction sites and sequence of this modified GUS coding region can be seen in the resulting plasmid, pAGUS1 (Fig. 1B).

The 5′ untranslated leader region in pAGUS1 was replaced with viral leader sequences by digesting the plasmid with *Bam* HI, removing the single-stranded ends of the DNA with S1 nuclease, cutting with *Hin*d III and then inserting the viral sequences encoded by oligonucleotides with flush 5′ ends and 3′ ends that were compatible with *Hin*d III. Synthetic DNAs encoding the different viral readthrough region sequences were inserted into either pAGUS1-T8 or -TN2 (Fig. 3B) using the *Nco* I and *Apa* I sites of the vector DNA.

The restriction sites located upstream of the 35S promoter regions in all plasmids except pBI221, pHHA, and pHHAB were changed by

insertion of an oligonucleotide to provide sites for *Eco* RI, *Bgl* II, and *Sph* I (see pAGUS1 at the top of Fig. 3B).

Plant material and electroporation

The *Nicotiana tabacum* L. cell line, BY2 [32], was maintained in Murashige-Skoog medium containing 0.2 mg/l 2,4-dichlorophenoxyacetic acid (MS.2). Cells were subcultured by inoculating fresh medium with 0.02 volumes of a week-old culture. Cultures to be used for the preparation of protoplasts were started by inoculating fresh medium with 0.05 volumes of a week-old culture and cells were harvested three days later.

Protoplasts were prepared as described by Bradshaw *et al.* [7] and counted with a hemacytometer. Samples of 3×10^6 protoplasts in one ml of 10 mM HEPES-KOH pH 7, 150 mM KCl, 4 mM $CaCl_2$, and 200 mM sorbitol were placed in disposable polystyrene cuvettes (Fisher Scientific Co., # 14-385-991F) lined with aluminum foil strips (4 mm wide, 1 cm apart) and subjected to electroporation with an 'old sparky' electroporator [7], generally, by discharge of a 240 μF capacitor charged to 400 V. After incubation on ice for 15 min and then at room temperature for 15 min, protoplast samples were diluted with 10 ml of MS.2 containing 200 mM sorbitol and incubated in a dark growth chamber held at 25 °C for 24 h which we found to be optimal for the accumulation of transient GUS activity (data not shown).

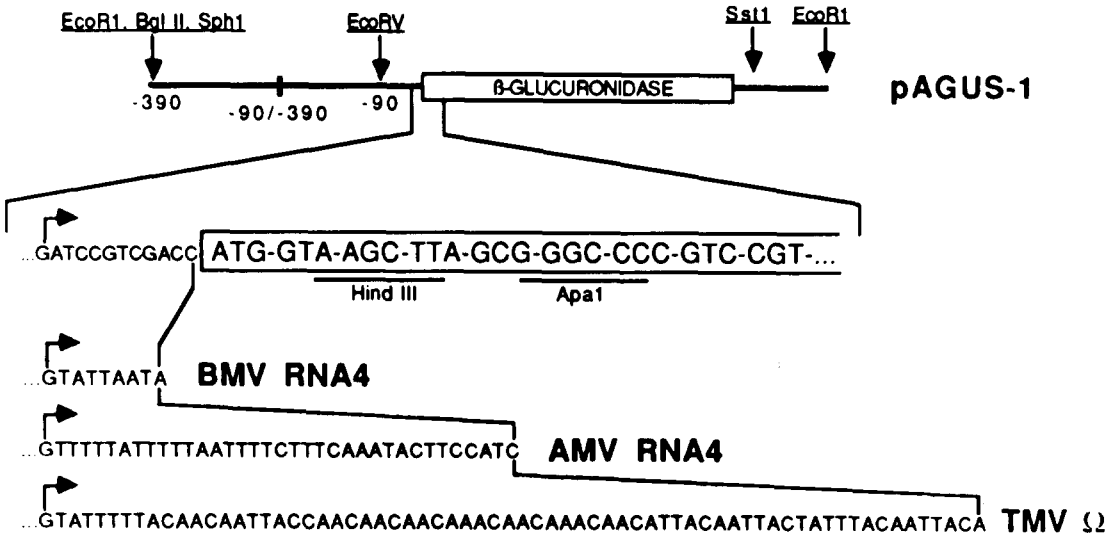
Protoplasts were harvested by centrifugation and lysed by one freeze/thaw cycle and sonication in 500 μl of GUS lysis buffer [24]. Extracts were clarified by centrifugation at $12000 \times g$ for 4–5 min and assayed for protein content [6]. GUS activity was measured using a kinetic, fluorogenic procedure [24, 26]. Typically, 50 μl of lysates from cells transfected with constructs containing stop codons (or pBI221, pHHA and pHHD) were incubated at 37 °C in a total volume of 300 μl containing 1 mM MUG in GUS lysis buffer. After equilibration for 5 minutes, a zero time aliquot (50 μl) was removed to 0.5 ml of 0.2 M Na_2CO_3 for standardization and then additional aliquots were taken after 15, 30, 45 and 60 minutes. Lysates from cells transfected with constructs lacking stop codons were generally diluted 30-fold and then 50 μl was assayed. Fluorometric analyses were performed with the use of a Perkin-Elmer LS-5 fluorometer, generally with the slits = 3 nm, excitation at 360 nm and emission measured at 450 nm. The GUS specific activity (pmol 4-methylumberlliferone/min/μl/μg protein) shown for each construct represents the mean of three electroporation experiments carried out with the same batch of protoplasts. With any one batch of protoplasts, the standard deviations of the average GUS specific activities were between 5–15%, but generally less than 10%, of the mean values.

Results

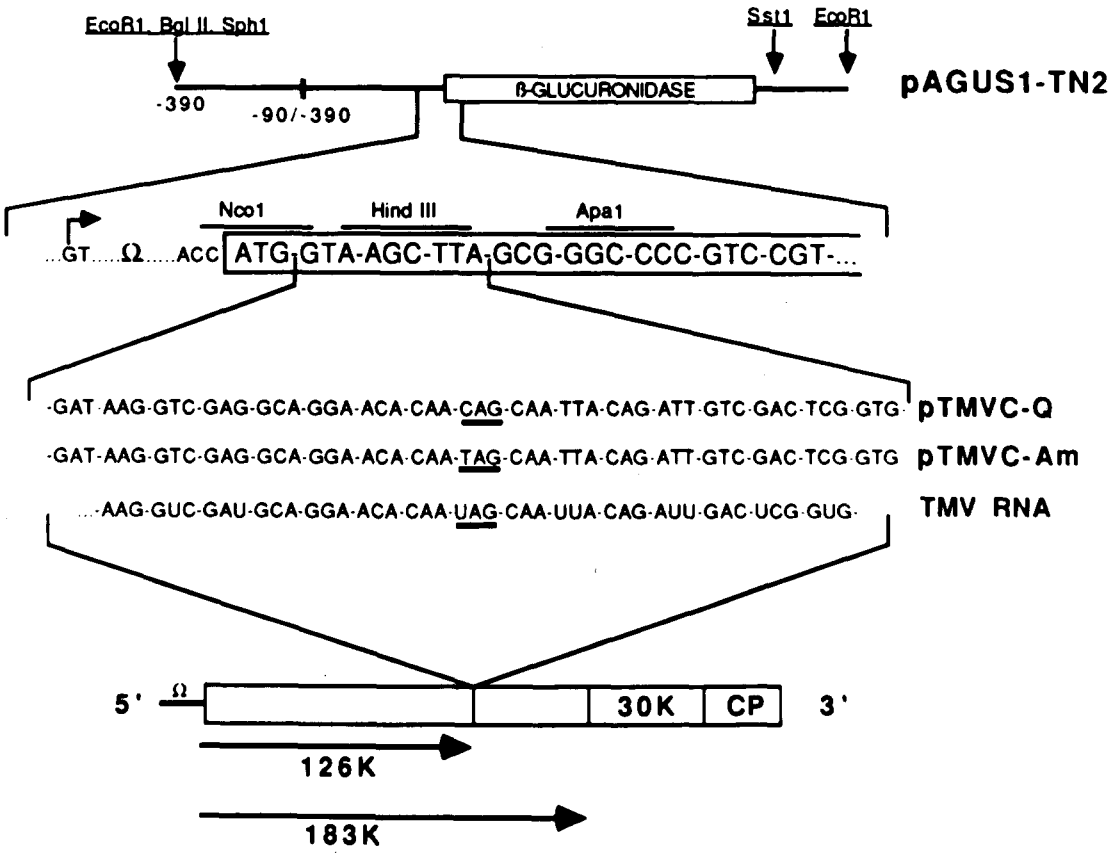
In vitro translation experiments [47, 48] suggested that ribosomes bypass the leaky stop codons of TMV and tobacco rattle virus (TRV) with efficiencies of 10–20%. We felt that it should be possible to monitor readthrough *in vivo* when coupled to the translation of GUS reporter mRNAs. However, it was necessary to optimize expression of our reporter constructs in transfected protoplasts in order to detect leaky termination events without background interference.

In preliminary experiments, we found that pBI221 [26], a construct that contains the wild-type GUS coding sequence and a highly active CaMV 35S RNA promoter, directed GUS activity that was only 10-fold above background in protoplasts from the tobacco cell line, BY2 (data not shown). These results were obtained using electroporation conditions (240 μF capacitor charged to 200 V and 25 μg of plasmid DNA per 3×10^6 protoplasts) that have been shown to be suitable for the detection of chloramphenicol acetyltransferase reporter gene activity in BY2 protoplasts [7]. Expression of the GUS reporter constructs needed to be improved at least tenfold, thus, we investigated the expression vector and a range of electroporation-mediated gene transfer conditions.

A


EcoR1, Bgl II, Sph1
EcoRV
Ssi1
EcoR1
-390
-90/-390
-90
ß-GLUCURONIDASE
pAGUS-1
...GATCCGTCGACC ATG-GTA-AGC-TTA-GCG-GGC-CCC-GTC-CGT-...
Hind III
Apa1
...GTATTAATA BMV RNA4
...GTTTTTATTTTTAATTTTCTTTCAAATACTTCCATC AMV RNA4
...GTATTTTTACAACAATTACCAACAACAACAAACAACAAACAACATTACAATTACTATTTACAATTACA TMV Ω

B


EcoR1, Bgl II, Sph1
Ssi1
EcoR1
-390
-90/-390
ß-GLUCURONIDASE
pAGUS1-TN2
Nco1
Hind III
Apa1
...GT.....Ω.....ACC ATG-GTA-AGC-TTA-GCG-GGC-CCC-GTC-CGT-...
-GAT-AAG-GTC-GAG-GCA-GGA-ACA-CAA-CAG-CAA-TTA-CAG-ATT-GTC-GAC-TCG-GTG- pTMVC-Q
-GAT-AAG-GTC-GAG-GCA-GGA-ACA-CAA-TAG-CAA-TTA-CAG-ATT-GTC-GAC-TCG-GTG pTMVC-Am
....-AAG-GUC-GAU-GCA-GGA-ACA-CAA-UAG-CAA-UUA-CAG-AUU-GAC-UCG-GUG- TMV RNA
5'
Ω
30K
CP
3'
126K
183K

Construction of improved GUS expression vectors

The work of Kay *et al.* [31] suggested that the activity of the 35S RNA promoter could be augmented 10- to 100-fold in transgenic tobacco by duplicating the region from positions −343 to −90. We therefore compared construction pHHA in which the segment from positions −390 to −90 of the promoter was duplicated and pHHD, which contained only a single copy of the upstream promoter sequences present in pHHA (Fig. 1A). The sequence duplication had little effect on GUS expression (Fig. 2A). However, the possibility remained that any improvements in the pHHA vector were simply obscured by saturation of the protoplast gene expression system. That hypothesis was ruled out by the finding that GUS expression does not plateau, but increases in proportion to the amount of pHHA transfected up to at least 200 μg/transfection (data not shown), twice the amount of DNA used for the experiment displayed in Fig. 2A. Similar results were obtained with other plasmids as well.

As a second approach to obtain increased expression, we tested whether the 5′ untranslated leader regions from several plant viruses could enhance the translation of GUS mRNAs. Previous studies [10, 11, 13, 27] have shown that the viral sequences significantly increase expression of mRNAs synthesized *in vitro*. To direct specific transcription *in vivo* it was necessary to incorporate the leader sequences into the vector DNA in proper juxtaposition to the transcription initiation site on the 35S RNA promoter. Thus, pHHA was modified through a series of steps (see Materials and methods) in order to construct pAGUS1, which can accommodate synthetic DNAs encoding a variety of start sites, 5′ leaders and NH_2-terminal coding region sequences (Fig. 1B).

Oligonucleotides encoding the 5′ untranslated leaders of alfalfa mosaic virus (AMV) RNA-4 (pAGUS1-A4), brome mosaic virus (BMV) RNA-4 (pAGUS1-B4), and TMV (pAGUS1-T8) were inserted into pAGUS1 (Fig. 3A). Transcription of these plasmids *in vivo* should generate mRNAs with the authentic 5′ ends, untranslated leader regions and translation start sites of the viruses although the GUS coding region is the same in all of these constructs.

The ability of the different DNAs to direct GUS synthesis was analyzed by transient expression in tobacco protoplasts. Figure 2B shows that all of the derivatives of pAGUS1 with viral leader regions directed greater GUS expression than pBI221. The most striking increase was obtained using pAGUS1-T8 with the TMV leader sequence, also called omega [35]. In this experiment the stimulation was *ca.* 35-fold relative to pBI221, whereas the increase measured using a different batch of protoplasts but similar transfection conditions was *ca.* 49-fold (data not shown). These results are in good agreement with Gallie *et al.* [12] who, using *in vitro*-synthesized mRNAs, also found the omega fragment to be more stimulatory than sequences from several other viruses. The concordance between our results implies that the expected regions of our plasmids are actually transcribed; however, we have not directly mapped the 5′ ends of the transcripts produced *in vivo*.

We have also prepared a slightly modified version of pAGUS1-T8 in which the sequence around the ATG translation initiation codon, 5′...ACATG-GTA...3′, was changed to

◀ *Fig. 3.* Structures of GUS expression vectors with different viral leader sequences and NH_2-terminal GUS coding regions. In 3A, the sequence surrounding the transcriptional and translational start sites in different constructs and the origin of each particular leader sequence are shown. The leader from BMV RNA-4 [1] was introduced to construct pAGUS1-B4, the AMV RNA-4 sequence [33] is in pAGUS1-A4, the TMV leader is in pAGUS1-T8 and a mutant TMV sequence is in pAGUS1-TN2 (see Fig. 2B). In 3B the structure of pAGUS1-TN2 is shown along with the NH_2-terminal portion of the GUS coding region which was modified as indicated to incorporate codons from the common strain of TMV. The lower portion of the figure shows a highly generalized genetic map of TMV. The location of the leaky amber stop codon in the virus is indicated in addition to the polypeptides which arise from termination at (126K) or readthrough past (183K) the stop. CP (coat protein) and 30K refer to the other protein coding regions of the virus. The sequence of the viral RNA around the leaky stop codon is shown above the map of the virus. Figures are not drawn to scale.

5′...CCATG-GTA...3′, in order to introduce a site for *Nco* I (CCATGG). This constructuon, pAGUS1-TN2 (Fig. 3B), appeared to be slightly less active than pAGUS1-T8 in the experiment represented in Fig. 2B but the difference is probably not significant. In fact, we have isolated several plasmids with small (e.g. six nucleotide long) insertions or deletions in the TMV leader region and have found that they are very comparable to pAGUS1-T8 and -TN2 as vectors (data not shown). In general, the omega fragment seems very tolerant of minor mutations [13].

Establishment of optimal electroporation conditions

We focused on the readthrough process in TMV with a systematic analysis of electroporation-mediated transfection conditions using derivatives of pAGUS1-TN2 in which the NH_2-terminal coding region of GUS included several codons from the readthrough region of the common or U1 strain of TMV [14]. In the construct, pTMVC-Am, Am refers to the presence of an amber stop codon, whereas in pTMVC-Q, the amber stop is replaced by a glutamine codon as noted by the Q designation (Fig. 3B).

A wide range (*ca.* 35-fold) of GUS activity was observed when BY2 protoplasts were subjected to electroporation using a 240 μF capacitor charged to different voltages in the presence of different amounts of pTMVC-Q (Fig. 4). The results of this experiment indicated that the conditions which produced the highest GUS expression (100 μg plasmid/transfection, charge of 400 V) would allow monitoring of both high- and low-level translational events. These conditions appear to be optimal since the amount of pTMVC-Q or pTMVC-Am needed to saturate the system at the same electrical field strength is greater than (or equal to) 200 μg/transfection (data not shown).

Readthrough past the stop codon in pTMVC-Am

Figure 5 shows representative examples of GUS assays with extracts from cells transfected with

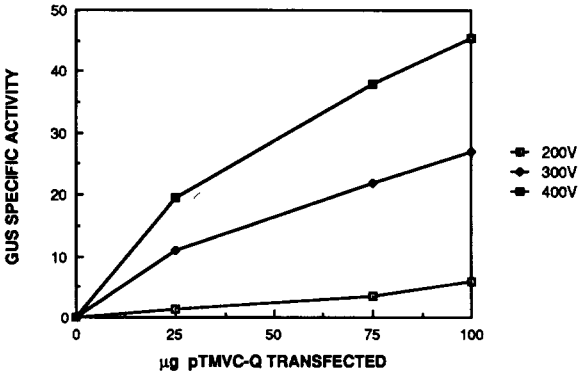


Fig. 4. Optimization of electroporation conditions. Samples of 3×10^6 protoplasts were subjected to electroporation at different voltages and in the presence of different amounts of pTMVC-Q. The indicated voltages correspond to the applied electrical field strengths since the electrodes were separated by a distance of 1 cm in the electroporation chambers. Each point represents the mean of three independent electroporation experiments.

either no DNA, 100 μg of pTMVC-Am, or 100 μg of pTMVC-Q. Clearly, the GUS activity directed by pTMVC-Am under optimum conditions is easily detected in undiluted protoplast lysates, despite the presence of the stop codon in the GUS reading frame. Similar GUS activity is exhibited by appropriate, i.e., 30-fold, dilutions of lysates from cells transfected with pTMVC-Q.

The efficiency of translational readthrough over the stop codon should be reflected in the ratio of GUS specific activities directed by pTMVC-Am and pTMVC-Q. These plasmids differ in only a single base pair (TAG → CAG) well removed from the promoter and translation initiation site. The GUS specific activities directed by these DNAs under optimum conditions in three independent experiments with different protoplast preparations and the ratios of activity are shown in Table 1. While the GUS activities produced from either construct varied over a range of about four-fold in the different experiments, the *relative* GUS specific activities were quite constant and indicate that ribosomes bypass the stop codon encoded by pTMVC-Am with an efficiency of 4–5%.

Table 1. Readthrough past the amber codon in pTMVC-Am.

Experiment	GUS specific activity directed by		Readthrough (Am/Q × 100)
	pTMVC-Q	pTMVC-Am	
1	29.9[a] (1.45)[b]	1.57 (0.13)	5.3
2	10.4 (0.08)	0.4 (0.08)	3.8
3	37.49 (3.7)	1.63 (0.04)	4.3

[a] Mean GUS specific activity.
[b] Standard deviation of the mean.

Stop codon readthrough in different viruses

Stop codon readthrough is utilized by a variety of unrelated viruses as shown in Table 2. In most of these cases, the presence of a leaky termination site has been demonstrated by *in vitro* translation studies and supported by nucleotide sequence data on the viral RNAs. The only exception is barley yellow dwarf virus (BYDV) for which only sequence data exist. Readthrough appears to provide a mechanism for the expression and/or regulation of the RNA-dependent RNA replicase domain in half of the viruses; many of the other viruses use this mechanism to produce coat proteins with C-terminal extensions of *ca.* 50 kDa.

Table 2 also shows that the readthrough regions can be divided into two major sequence classes. One is found in carnation mottle virus (CarMV), maize chlorotic mottle virus (MCMV), BYDV and beet western yellows virus (BWYV) while the other is found in TMV, turnip yellow mosaic virus (TYMV) and beet necrotic yellow vein virus (BNYVV). These two sequence classes appear to be functionally interchangeable as readthrough over either sequence type can lead to either replicase expression or extension of the coat protein, depending upon the virus. Also surprising is the fact that replicase synthesis in two distinct but related viruses, TRV RNA-1 and TMV [4, 18, 19], involves readthrough over different stop codons in very dissimilar sequence contexts.

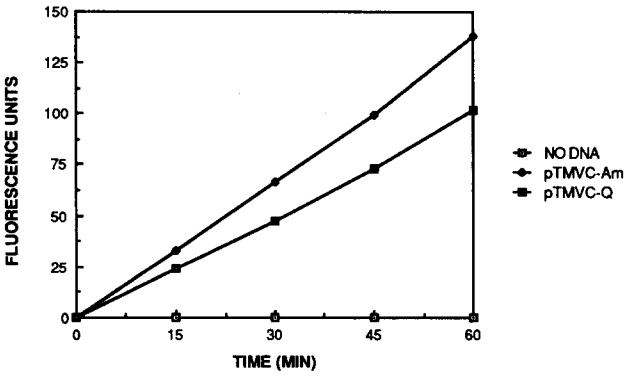


Fig. 5. Analysis of readthrough past the leaky stop codon in pTMVC-Am. Lysates from cells transfected with no DNA, 100 μg of pTMVC-Am or 100 μg of pTMVC-Q were assayed as described in Materials and methods. The lysate from pTMVC-Q transfected cells was diluted 30-fold prior to the assay.

The 'leakiness' of several of the viral termination signals was analyzed by incorporating the stop codons and surrounding nucleotides into the GUS expression vectors (Table 3). In each case, readthrough was determined by comparing the GUS activities generated from a pair of constructs which were identical except for one base pair difference that creates either an amber stop or a glutamine codon. GUS activities are shown in Fig. 6 and ratios of activity, i.e., readthrough efficiencies, are presented in Table 4.

The practice of using pairs of nearly identical constructs appears to be essential since GUS activity can be appreciably affected by the addition of certain NH_2-terminal sequences. In particular, a short sequence from BNYVV, and longer sequences from TRV and the tomato L strain of TMV (Table 3) appeared to decrease the activity of the enzyme (Fig. 6). Thus, the GUS specific activities directed by constructions encoding these sequences were adjusted by appro-

Table 2. Readthrough sites in different plant RNA viruses.

Virus	Group	Readthrough region sequence			Consequence	Reference
TMV:	Tobamovirus					
Common strain		-ACA -	CAA-UAG-CAA-UUA	-CAG-AUU-	Replicase expression	[14]
Tomato strain		-ACU -	CAA-UAG-CAA-UUA	-CAG-GUC-	Replicase expression	[44]
BNYVV	Furovirus	-GGA -	CAA-UAG-CAA-UUA	-GCU-GCU-	54 kDa extension of coat protein	[5]
TYMV	Tymovirus	-GUC -	CAA-UAG-CAA-UCA	-GCC-CCA-	14.8 kDa extension of polyprotein	[38]
TRV	Tobravirus	-GUC -UUA-UGA-CGG-UUU-CGG-UCU-			Replicase expression	[18]
CarMV	Carmovirus	-UUU -CCC -	AAA-UAG-GGG	-GGC-CUG-	Replicase expression	[16]
MCMV	Carmovirus	-GAG -UUG-	AAA-UAG-GGG	-UGU-CUU-	Replicase expression	[41]
BYDV	Luteovirus	-ACG -GCC-	AAA-UAG-GUA	-GAC-UCC-	50 kDa extension of coat protein	[37]
BWYV	Luteovirus	-AAC -CCC -	AAA-UAG-GUA	-GAC-GAG-	51.5 kDa extension of coat protein	[52]

Table 3. N-Terminal coding sequences of GUS constructs containing viral readthrough regions.[a]

Virus/construct	Sequence
TMV common/pTMVC-Am	ATG-GAT-AAG-GTC-GAG-GCA-GGA-ACA-CAA-TAG-CAA-TTA-CAG-ATT-GTC-GAC-TCT-GTG-GGC-CCC-GTC-CGT
TMV tomato/ pTMVL-201Am	ATG-GTA-AGC-TTA-GAC-ATG-TAC-AAA-GTA-GAT-GCA-GGT-ACT-CAA-TAG-CAA-TTA-CAG-GTC-GAC-TCT-GTG-GGC-CCC-GTC-CGT-
TMV tomato/ pTMVL-202Am	ATG-GTA-AGC-TTG-GAT-CCT-TTA-GTT-AGT-ATC-ATT-AGA-GAT-TTA-GAA-CGG-GTT-AGT-AGT-TAC-TTA-TTA-GAC-ATG-TAC-AAA-GTA-GAT-GCA-GGT-ACT-CAA-TAG-CAA-TTA-CAG-GTC-GAC-TCT-GTG-TTT-AAA-AAT-TTC-ATT-CTT-TTT-GTA-GCA-GCT-CCA-AAG-ACT-GCA-GTA-TTG-GGC-CCC-GTC-CGT-
BNYVV/pBNYVV-Am	ATG-CCA-CCC-GGA-CAA-TAG-CAA-TTA-GCT-GCT-GCT-GTC-GAC-TCT-GTG-GGC-CCC-GTC-CGT-
CarMV/pCARMV-Am	ATG-GTA-AGC-TTT-CCC-AAA-TAG-GGG-GGC-CTG-GGC-CCC-GTC-CGT-
MCMV/pMCMV-Am	ATG-GTA-AGC-TTT-GTT-AGG-CAC-CTG-GGT-GGG-CTG-GAT-GAT-CCA-GCC-CTA-TTC-CAG-GAG-TTG-AAA-TAG-GGG-TGT-CTG-CAG-TTG-GGC-CCC-GTC-CGT-
TRV/pTRV-Op	ATG-GTA-AGC-TTA-TTT-TTT-GTT-ACC-GAG-ACC-GTC-TTA-TGA-CGG-TTT-CGG-TCT-AGG-TTT-GAT-GTT-TTT-GCG-GGC-CCC-GTC-CGT-

[a] The column on the left indicates the origin of each readthrough region tested and the names of the construct in which they were incorporated. The sequences shown begin at the ATG translation start of the GUS gene and extend to the CGT arginine codon (double underlined) which corresponds to codon three of the wild-type GUS gene. Underlined regions in the sequences correspond exactly to viral readthrough regions. The stop codon in each construct is shaded. Constructs lacking stop codons which were used in conjunction with the amber-containing plasmids to estimate readthrough efficiencies are not shown.

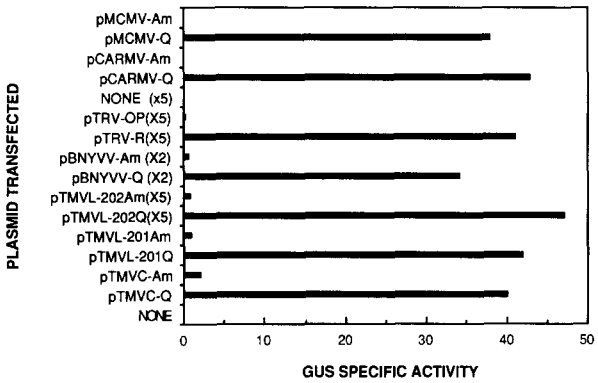


Fig. 6. GUS expression directed by constructs with leaky termination regions from different viruses and corresponding constructs lacking stop codons. BY2 protoplasts were subjected to electroporation under optimal conditions in the presence of 100 μg of each of the indicated plasmids. GUS specific activities shown are the means of three electroporation experiments done with the same batch of protoplasts. Note that GUS activities directed by some of the constructs were lower than others (see text) and were multiplied by the indicated factors so that all results could be visualized on the same chart.

priate factors for presentation, as indicated in Fig. 6.

In agreement with the values in Table 1, readthrough over the termination signal from the common strain of TMV was *ca.* 5% in this experiment (pTMVC-Am, Fig. 6 and Table 4). The TMV common strain sequence consistently supports a higher level of readthrough than the corresponding region from the tomato L strain which is represented in a short sequence in pTMVL-201Am and in a larger insert in pTMVL-202Am (Fig. 6 and Table 4). A short insert representing the readthrough site in BNYVV is bypassed with an efficiency comparable to that of the TMV tomato L strain region. Presumably, the differences in readthrough efficiency observed between the different constructs is a reflection of the codon contexts outside of the conserved 5′...-CAA-*UAG*-CAA-UYA-...3′ sequence. We have not directly tested the readthrough region from TYMV; however in a detailed analysis of mutations in the TMV common strain sequence we found that ribosomes are able to bypass a TYMV-like sequence (manuscript in preparation).

Table 4. Efficiency of readthrough past different viral stop signals.

Virus	Constructs used	Readthrough efficiency[a]
TMV common	pTMVC-Am & Q	5.3%
TMV tomato	pTMVL-201Am & Q	2.3%
TMV tomato	pTMVL-202Am & Q	1.7%
BNYVV	pBNYVV-Am & Q	1.7%
CarMV	pCarMV-Am & Q	0.02%
MCMV	pMCMV-Am & Q	0.01%
TRV	pTRV-Op & R[b]	0.2%

[a] Readthrough efficiencies were determined from the ratios of GUS specific activities directed by the appropriate pairs of constructs indicated in the center column.

[b] R refers to the CGA arginine codon in pTRV-R which replaces the UGA opal stop codon in pTRV-Op.

We observed essentially no readthrough over the sequences from MCMV and CarMV (Fig. 6 and Table 4). Furthermore, a sequence spanning the UGA opal stop from TRV in pTRV-Op (Op refers to opal) directed barely detectable readthrough. These data suggest that an essential readthrough determinant is absent, either from these constructions, or from BY2 protoplasts.

Discussion

The GUS gene has proven to be an ideal reporter for our studies. The ability to measure readthrough derives from the stability of GUS, the lack of critical assay condition requirements and the sensitivity with which the enzyme can be assayed [24]. In addition we found the placement of methionine codons in the GUS gene to be very convenient. The second methionine codon in the GUS reading frame is located 102 codons 3′ to the initiating ATG [25] and reinitiation at the downstream methionine codon, if it occurred, might be expected to be 'invisible' to us since the truncated protein is likely not to be enzymatically active. The absence of complications resulting from reinitiation in these studies is readily apparent in the lack of activity exhibited by constructs with the readthrough regions from CarMV, MCMV and TRV.

In agreement with other studies [21, 24, 30] we found that a variety of NH_2-terminal extensions do not abolish GUS activity. Nevertheless, our results indicate that certain N-terminal amino acid sequences alter the enzyme in an unpredictable manner. To minimize this effect, readthrough efficiencies were determined using pairs of constructs which would encode enzymes with similar NH_2-terminal ends. This approach appears to be valid, as indicated by the fact that readthrough over the leaky stop of the tomato L strain of TMV encoded by pTMVL-201Am and pTMVL-202Am takes place with similar efficiencies (Table 4). In these plasmids, the stop is embedded in N-terminal coding sequences that differ significantly in length (Table 3) and in the apparent extent to which they influence GUS activity (Fig. 6). (The effect of different N-terminal sequences can be seen best by examining the relative activities of constructs lacking stop codons in Fig. 6.)

In order to detect readthrough we found that it was helpful to construct more efficient GUS expression vectors and to establish appropriate gene transfer conditions. We estimate that the end result was to elevate transient expression levels by 1200 to 1500 fold. The transfection conditions we adopted to compare the different viral termination signals (3×10^6 protoplasts, 100 μg plasmid, 240 μF capacitor charged to 400 V) were 'optimum' in the sense that the protoplast gene expression system was not saturated and we could confidently measure GUS expression that was equivalent to *ca.* 2% of the GUS activity directed by our most active constructs, pAGUS1-TN2 or pTMVC-Q, for example.

The activity of the expression vectors was increased to the greatest extent by the presence of the TMV omega fragment. Based on the extensive studies by Gallie and associates [10, 11, 12], it seems likely that the 5′ leader sequence increases the utilization of mRNAs. The failure to obtain higher expression with a duplicated 35S RNA promoter 'enhancer' element in these studies is perhaps not surprising since the ability of multimers of the upstream promoter region to elevate gene expression appears to vary widely in both transgenic plants [9, 31] and protoplast systems [43, 45]. Nevertheless, we have chosen to use vectors with the modified promoter since the duplicated 'enhancer' could prove to be useful in stably transformed cells or plants [9, 31].

Our results strongly indicate that GUS activity directed by constructs containing the leaky amber from TMV truly reflects specific, translational readthrough. First, we demonstrate that of the three 'classes' of leaky termination sites tested, expression of appreciable reporter activity is observed with only the TMV-like readthrough region. This result indicates that UAG stops, in general, are not leaky, and moreover, suggests that leakiness is determined by the sequences which flank a stop codon. In addition, we show that the efficiency of readthrough over the TMV amber is quite reproducible.

Although the TMV-encoded 126 kDa and 183 kDa polypeptides have been detected in virus-infected tissues [2, 46, 50], we are unaware of any quantitative data which would provide an accurate estimate of *in vivo* readthrough efficiency. The apparent efficiency of readthrough over the TMV stop measured in this study (*ca.* 5%) is well in line with some stated values [14] but it is about 2–4-fold lower than efficiencies determined from *in vitro* translation experiments with full-length viral RNA [47]. It is not clear whether this difference is significant and it is interesting that at least one qualitative parallel can be drawn between the *in vivo* and *in vitro* experiments. Pelham [47] observed that readthrough past the TMV amber was more readily detected than readthrough over an amber from a phage Qβ mutant and this is similar to our finding of preferential readthrough over the TMV sequence.

In view of the apparent functional equivalence of the different leaky termination sites and the efficiency with which readthrough takes place *in vitro* [8, 48], our failure to detect appreciable readthrough over the stop codons from CarMV and TRV is somewhat puzzling. It is possible that these stop codons are bypassed at very low frequencies *in vivo* and that readthrough is for some reason exaggerated *in vitro*. Alternatively, readthrough over these stops could involve

mechanisms that require more extensive sequence elements than we have tested here. Recently, it has been shown that viral reproduction is inhibited by mutations which disrupt a potential stem and loop structure around the suppressible amber stop codon of Moloney murine leukaemia virus [28]. Although regions with substantial secondary structure can be identified around the leaky termination regions of CarMV, MCMV and TRV (not shown), the inclusion of these sequences in our constructs (pMCMV-Am and pTRV-Op) was not sufficient to direct detectable readthrough.

The immediate goals of our studies are to elucidate the mechanism(s) which permits ribosomes to bypass leaky stop codons and to determine how sequence elements around these termination sites facilitate readthrough. The fidelity displayed by our *in vivo* system has enabled us to identify the readthrough determinants around the TMV stop (manuscript in preparation). Eventually it should be possible to generate stably transformed cell lines from which readthrough products can be purified. N-terminal sequence analysis of these polypeptides will ultimately clarify the nature of readthrough events *in vivo* and provide insight into the workings of the translational apparatus.

Acknowledgements

We want to thank Toby Bradshaw for assistance in the construction of 'old sparky', for sending us BY2 cells and for advice on protoplast preparation and transient expression; David Ow for the plasmid, pDO432; John Atkins, Barry Condron and Norma Wills for many helpful discussions and for reviewing the manuscript; Bob Weiss for suggestions on oligonucleotide synthesis, vector construction and translation; Wayne Gerlach and Steve Lommel for communicating results prior to publication; Jeff Ives for help on computers and Diane Dunn and Ed Meenen for supplying some of the oligonucleotides used in these studies.

References

1. Ahlquist P, Luckow V, Kaesberg P: Complete nucleotide sequence of brome mosaic virus RNA 3. J Mol Biol 153: 23–38 (1981).
2. Beier H, Barciszewska M, Krupp G, Mitnacht R, Gross HJ: UAG readthrough during TMV RNA translation: Isolation and sequence of two $tRNAs^{Tyr}$ with suppressor activity from tobacco plants. EMBO J 3: 351–356 (1984).
3. Beier H, Barciszewska M, Sickinger HD: The molecular basis for the differential translation of TMV RNA in tobacco protoplasts and wheat germ extracts. EMBO J 3: 1091–1096 (1984).
4. Boccara M, Hamilton WDO, Baulcombe DC: The organisation and interviral homologies of genes at the 3′ end of tobacco rattle virus RNA 1. EMBO J 5: 223–229 (1986).
5. Bouzoubaa S, Ziegler V, Beck D, Guilley H, Richards K, Jonard G: Nucleotide sequence of beet necrotic yellow vein virus RNA-2. J Gen Virol 67: 1689–1700 (1986).
6. Bradford M: A rapid and sensitive method for the quantification of microgram quantities of protein using the principles of protein-dye binding. Anal Biochem 72: 248–254 (1976).
7. Bradshaw HD Jr, Parson WW, Sheffer M, Lioubin PJ, Mulvihill ER, Gordon MP: Design, construction, and use of an electroporator for plant protoplasts and animal cells. Anal Biochem 166: 342–348 (1987).
8. Carrington JC, Morris TJ: Characterization of the cell-free translation products of carnation mottle virus genomic and subgenomic RNAs. Virology 144: 1–10 (1985).
9. Fang RX, Nagy F, Sivasubramaniam S, Chua NH: Multiple cis regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. Plant Cell 1: 141–150 (1989).
10. 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).
11. Gallie DR, Sleat DE, Watts JW, Turner PC, Wilson TMA: The 5′-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts *in vitro* and *in vivo*. Nucl Acids Res 15: 3257–3273 (1987).
12. Gallie DR, Sleat DE, Watts JW, Turner PC, Wilson TMA: A comparison of eukaryotic viral 5′-leader sequences as enhancers of mRNA expression *in vivo*. Nucl Acids Res 15: 8693–8711 (1987).
13. 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).
14. Goelet P, Lomonosoff GP, Butler PJG, Akam ME, Gait MJ, Karn J: Nucleotide sequence of tobacco mosaic virus RNA. Proc Natl Acad Sci USA 79: 5818–5822 (1982).

15. Goldbach RW: Molecular evolution of plant RNA viruses. Ann Rev Phytopath 24: 289–310 (1986).
16. Guilley H, Carrington JC, Balazs E, Jonard G, Richards K, Morris TJ: Nucleotide sequence and genome organization of carnation mottle virus RNA. Nucl Acids Res 13: 6663–6677 (1985).
17. Guilley H, Dudley RK, Jonard G, Balazs E, Richards KE: Transcription of cauliflower mosaic virus DNA: Detection of promoter sequences, and characterization of transcripts. Cell 30: 763–773 (1982).
18. Hamilton WDO, Boccara M, Robinson DJ, Baulcombe DC: The complete nucleotide sequence of tobacco rattle virus RNA-1. J Gen Virol 68: 2563–2575 (1987).
19. Harrison BD, Robinson DJ: Tobraviruses. In: Van Regenmortel MHV, Fraenkel-Conrat H (eds) The Plant Viruses, Vol. 2, The rod-shaped plant viruses. Plenum Press, New York (1986).
20. Ishikawa M, Meshi T, Motoyoshi F, Takamatsu N, Okada Y: *In vitro* mutagenesis of the putative replicase genes of tobacco mosaic virus. Nucl Acids Res 14: 8291–8305 (1986).
21. Iturriaga G, Jefferson RA, Bevan MW: Endoplasmic reticulum targeting and glycosylation of hybrid proteins in transgenic tobacco. Plant Cell 1: 381–390 (1989).
22. Jacks T, Power MD, Masiarz FR, Luciw PA, Barr PJ, Varmus HE: Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature 331: 280–283 (1988).
23. Jacks T, Varmus HE: Expression of the rous sarcoma virus pol gene by ribosomal frameshifting. Science 230: 1237–1242 (1985).
24. Jefferson RA: Assaying chimeric genes in plants: The GUS gene fusion system. Plant Mol Biol Rep 5: 387–405 (1987).
25. Jefferson RA, Burgess SM, Hirsch D: β-Glucuronidase from *Escherichia coli* as a gene-gusion marker. Proc Natl Acad Sci USA 83: 8447–8451 (1986).
26. Jefferson RA, Kavanagh TA, Bevan MW: GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J 6: 3901–3907 (1987).
27. Jobling SA, Gehrke L: Enhanced translation of chimaeric messenger RNAs containing a plant viral untranslated leader sequence. Nature 325: 622–625 (1987).
28. Jones DS, Nemoto F, Kuchino Y, Masuda M, Yoshikura H, Nishimura S: The effect of specific mutations at and around the gag-pol junction of Moloney murine leukaemia virus. Nucl Acids Res 17: 5933–5945 (1989).
29. Kamer G, Argos P: Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. Nucl Acids Res 12: 7269–7282 (1984).
30. Kavanagh TA, Jefferson RA, Bevan MW: Targeting a foreign protein to chloroplasts using fusions to the transit peptide of a chlorophyll *a/b* protein. Mol Gen Genet 215: 38–45 (1988).
31. Kay R, Chan A, Daly M, McPherson J: Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. Science 236: 1299–1302 (1987).
32. Koiwai A, Noguchi M, Tamaki E: Changes in the amino acid composition of tobacco cells in suspension culture. Phytochem 10: 561–566 (1971).
33. Koper-Zwarthoff EC, Brederode FT, Veeneman G, van Boom JH, Bol JF: Nucleotide sequences at the 5′ termini of the alfalfa mosaic virus RNAs and the intercistronic junction in RNA 3. Nucl Acids Res 8: 5635–5647 (1980).
34. Kozak M: The scanning model for translation: An update. J Cell Biol 108: 229–241 (1989).
35. Mandeles S: Location of unique sequences in tobacco mosaic virus ribonucleic acid. J Biol Chem 243: 3671–3674 (1968).
36. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1982).
37. Miller WA, Waterhouse PM, Gerlach WL: Sequence and organization of barley yellow dwarf genomic RNA. Nucl Acids Res 16: 6097–6111 (1988).
38. Morch MD, Boyer JC, Haenni AL: Overlapping open reading frames revealed by complete nucleotide sequencing of turnip yellow mosaic virus genomic RNA. Nucl Acids Res 16: 6157–6173 (1988).
39. Morch MD, Drugeon G, Benicourt C: Analysis of the *in vitro* coding properties of the 3′ region of turnip yellow mosaic virus genomic RNA. Virology 119: 193–198 (1982).
40. Murgola EJ, Hijazi KA, Goringer HU, Dahlberg AE: Mutant 16S ribosomal RNA: A codon-specific translational suppressor. Proc Natl Acad Sci USA 85: 4162–4165 (1988).
41. Nutter RC, Scheets K, Panganiban LC, Lommel SA: The complete nucleotide sequence of the maize chlorotic mottle virus genome. Nucl Acids Res 17: 3163–3177 (1989).
42. O'Connor M, Gesteland RF, Atkins JF: tRNA hopping: Enhancement by an expanded anticodon. EMBO J 8: 4315–4323 (1989).
43. Odell JT, Knowlton S, Lin W, Mauvais CJ: Properties of an isolated transcription stimulating sequence derived from the cauliflower mosaic virus 35S promoter. Plant Mol Biol 10: 263–273 (1987).
44. Ohno T, Aoyagi M, Yamanashi Y, Saito H, Ikawa S, Meshi T, Okada Y: Nucleotide sequence of the tobacco mosaic virus (tomato strain) genome and comparison with the common strain genome. J Biochem 96: 1915–1923 (1984).
45. Ow DW, Jacobs JD, Howell SH: 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 (1987).
46. Paterson R, Knight CA: Protein synthesis in tobacco protoplasts infected with tobacco mosaic virus. Virology 64: 10–22 (1975).
47. Pelham HRB: Leaky UAG termination codon in tobacco mosaic virus RNA. Nature 272: 469–471 (1978).
48. Pelham HRB: Translation of tobacco rattle virus RNAs

in vitro: Four proteins from three RNAs. Virology 97: 256–265 (1979).
49. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).
50. Siegel A, Hari V, Kolacz K: The effect of tobacco mosaic virus infection on host and virus-specific protein synthesis in protoplasts. Virology 85: 494–503 (1978).
51. Shen Z, Fox TD: Substitution of an invariant nucleotide at the base of the highly conserved '530-loop' of 15S rRNA causes suppression of yeast mitochondrial ochre mutations. Nucl Acids Res 17: 4535–4539 (1989).
52. Veidt I, Lot H, Leiser H, Scheidecker D, Guilley H, Richards K, Jonard G: Nucleotide sequence of beet western yellows virus RNA. Nucl Acids Res 16: 9917–9932 (1988).
53. Weiss RB, Dunn DM, Atkins JF, Gesteland RF: Slippery runs, shifty stops, backward steps, and forward hops: −2, −1, +1, +5, and +6 ribosomal frameshifting. CSH Symp Quant Biol 52: 687–693 (1987).
54. Xiong Z, Lommel SA: The complete nucleotide sequence and genome organization of red clover necrotic mosaic virus RNA-1. Virology 171: 543–554 (1989).
55. Ziegler V, Richards K, Guilley H, Jonard G, Putz C: Cell-free translation of beet necrotic yellow vein virus: Readthrough of the coat protein cistron. J Gen Virol 66: 2079–2087 (1985).