

Unusual dicistronic expression from closely spaced initiation codons in an umbravirus subgenomic RNA

Feng Gao¹, Olga M. Alekhina^{2,3}, Konstantin S. Vassilenko² and Anne E. Simon^{1,*}

¹Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD 20742, USA,

²Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region 142290, Russia and

³Federal Research and Clinical Centre of Physical-Chemical Medicine, Moscow 119435, Russia

Received March 16, 2018; Revised August 24, 2018; Editorial Decision September 13, 2018; Accepted September 19, 2018

ABSTRACT

Translation commencing at closely spaced initiation codons is common in RNA viruses with limited genome space. In the subgenomic RNA (sgRNA) of *Pea enation mosaic virus 2*, two closely spaced, out-of-frame start codons direct synthesis of movement/stability proteins p26 and p27. Efficient translation from AUG²⁶/AUG²⁷ is dependent on three 3'-proximal cap-independent translation enhancers (3'CITEs), whereas translation of the genomic (gRNA) requires only two. Contrary to strictly scanning-dependent initiation at the gRNA, sequence context of AUG²⁶/AUG²⁷ does not conform with Kozak requirements and insertion of efficient upstream AUGs had pronounced effects for AUG²⁶ but only moderate effects for AUG²⁷. Insertion of a hairpin within an extended 5' UTR did not significantly impact translation from AUG²⁶/AUG²⁷. Furthermore, AUG²⁷ repressed translation from upstream AUG²⁶ and this effect was mitigated when inter-codon spacing was reduced. Addition of a stable hairpin to the very 5' end of the sgRNA severely restricted translation, testifying that this 3'CITE-driven initiation is 5' end-dependent. Similar to gRNA, sgRNA reporter transcripts were nearly exclusively associated with light polysomes and 3'CITE-promoted long-distance interaction connecting the sgRNA ends affected the number of templates translated and not the initiation rate. We propose a non-canonical, 3'CITE-driven mechanism for efficient dicistronic expression from umbravirus sgRNAs.

INTRODUCTION

Protein biosynthesis, a fundamental process in all living organisms, is often strictly regulated at the stage of translation initiation (1). Prokaryotic and eukaryotic messenger RNAs (mRNAs) have distinct mechanisms for how ribo-

somes access templates and select start codons for initiation. Prokaryotic polycistronic mRNAs recruit ribosomes directly to start codons via 16S ribosomal RNA base-pairing to Shine–Dalgarno sequences positioned just upstream of initiation codons (2,3). In contrast, monocistronic eukaryotic mRNAs possess a 5' m⁷GpppN cap at their 5' termini, which, when bound to eIF4F, attracts the 43S preinitiation complex (PIC) that contains a 40S ribosomal subunit along with a ternary eIF2–GTP–transfer RNA (tRNA)^{Met} complex and eIF3, eIF1, eIF1A and eIF5 (1,4). The 43S complex then scans in the 5' to 3' direction as a 48S complex tethered to eIF4F until reaching an initiation codon in an appropriate context, which frequently is the 5' proximal AUG, thus rendering translation of most eukaryotic mRNAs monocistronic (5). This general model proposed by Kozak (6) is widely accepted and genome-wide evidence for 40S scanning of 5' UTRs was recently obtained in living yeast cells (7). In addition, the main elements of the scanning model have been employed to explain alternative modes of 5' end-promoted initiation, since transcriptome-wide studies revealed high incidence of cap-independent initiation of translation of eukaryotic cellular mRNAs (8).

Initiation of translation of some RNAs, however, does not involve a scanning stage. Internal ribosome entry sites, found mainly in the 5' UTR of animal viruses and some cellular mRNAs, directly recruit ribosomes to the vicinity of the initiation codon in the absence of some or all initiation factors, promoting internal cap- and even 5' end-independent initiation (9–11). Ribosomes are also proposed to use non-scanning mechanisms when mRNAs have extremely short or no 5' UTR (leaderless mRNAs) as is commonly found in the Archaea and some bacteria and mammalian mitochondria (9,12,13). In mammalian cells, leaderless RNAs directly bind eukaryotic 80S ribosomes and initiate translation in the presence of Met-tRNA_i without a requirement for initiation factors (14). Ribosome shunting, first described for *Cauliflower mosaic virus* (15) and subsequently found in other plant and animal viruses as well as some cellular mRNAs (16–19), allows ribosomes to bypass part of the 5' UTR to access the AUG start codon. Moreover, evidence also supports some percentage of cap-

*To whom correspondence should be addressed. Tel: +1 301 405 8975; Fax: +1 301 314 9489; Email: simona@umd.edu

dependent translation of mammalian mRNAs occurring in the absence of base-by-base 5' to 3' directional scanning (20). Using a capped leader RNA and an uncapped reporter RNA annealed by a stable dsRNA bridge, ribosomes could initiate translation of the reporter RNA without melting the bridge. This 'RNA looping', scanning-free translation mechanism was proposed to involve random collisions between ribosomes and the initiation site governed by the length, structure and flexibility of the 5' UTR as well as by accessibility of the initiation codon (21).

Even in the framework of the scanning model, the 5' proximal AUG can be bypassed or poorly used for initiation, usually because the codon is in a weak context for ribosome recognition. While the optimal context for an initiation codon can vary among mammalian and plant mRNAs, efficient initiation usually requires a purine at position -3 and/or a guanylate at position +4 (relative to the AUG initiation codon that is numbered from +1 to +3). Those nucleotides are thought to stabilize the conformational changes that occur upon codon-anticodon base-pairing during 48S initiation complex formation by interacting with the components of the complex (2,22-25). Even slight deviations from an optimal context can cause some scanning ribosomes to bypass an initial AUG codon and initiate at a downstream AUG in a better context, a process known as leaky scanning (2,26). The 'leakiness' of ribosomes at the initial AUG can also depend on additional factors including the length of the 5' leader, structural accessibility of the AUG and proximity of the two AUGs. For example, extremely short 5' leaders (<12 nt) potentiate efficient initiation at a downstream AUG in a strong context (27,28).

Since the leaky scanning model was first proposed, it has become widely accepted and usually referenced as the most likely explanation for dicistronic translation of overlapping Open reading frame (ORFs), which are commonly found in RNA viruses with limited genome space. Leaky scanning allows for the expression of co-C-terminal isoforms of a protein if the alternative start codons are in-frame, or different proteins if they are out-of-frame. Since ribosomes translating from the first AUG could potentially mask accessibility of downstream AUGs, close proximity of the downstream initiation codon is thought to minimize the masking effect leading to more efficient initiation at the second AUG (29). This has been suggested to occur for translation of closely spaced initiation codons in *Turnip yellow mosaic virus* (TYMV) and influenza virus B (separated by 7 and 4 nt, respectively) (30,31). In addition, since translation from the upstream TYMV AUG is influenced by the presence of the downstream AUG, the suggestion was made that the movement of scanning ribosomes is not unidirectional; i.e. ribosomes were proposed to transit with forward and backward oscillations known as 'fluttering', such that a second AUG codon within a particular upstream range (<16 nt in this study) captures a portion of the oscillating ribosomes that might otherwise have scanned back to the upstream initiation codon (30).

Umbraviruses (family *Tombusviridae*) are small, plus-sense RNA viruses with a single genomic (gRNA) that has no 5' cap or 3' poly(A) tail (Figure 1A). For umbravirus *Pea enation mosaic virus 2* (PEMV2), co-N terminal p33 and p94

(the RNA-dependent RNA polymerase [RdRp]) are translated from the 5' end of the gRNA, with the latter protein produced by a -1 ribosome frameshift event that extends the C-terminus of p33 (32). Efficient translation of p33 and p94 requires the cooperation of two 3' cap-independent translation enhancers (3'CITEs) that are centrally located within the 703-nt 3' UTR: the kl-TSS that binds to ribosomes and participates in a long-distance RNA-RNA interaction with a 5' proximal coding region hairpin and the Panicum mosaic virus-like translation enhancer (PTE), which binds to eIF4G by first binding directly to eIF4E (33-35) (Figure 1B). All umbraviruses also produce a subgenomic (sgRNA) in infected cells for translation of the two overlapping out-of-frame ORFs that are preceded by a short 5' UTR. The 5' proximal ORF of PEMV2 encodes p26, which is a multifunctional long-distance movement protein required to stabilize the gRNA by forming ribonucleoprotein particles (36) that are critical for virus accumulation in protoplasts (37). The ORF that nearly completely overlaps the p26 ORF encodes the cell-to-cell movement protein p27 (38,39). AUG²⁶, the initiation codon for p26, is 16-nt upstream of AUG²⁷ (Figure 1B), a distance that is longer than that required for coupled initiation in TYMV (30).

We recently reported that the Kozak contexts for both AUG²⁶ and AUG²⁷ are very weak and that mutating AUG²⁷ enhances translation of upstream AUG²⁶ (37). This implied that AUG²⁷ might be suppressing translation from AUG²⁶ and that PEMV2 sgRNA translation may use a mechanism other than simple unidirectional ribosome scanning. We now report that translation from AUG²⁶ and AUG²⁷ was differentially affected by highly efficient upstream AUGs incorporated into an extended 5' UTR sequence, but were virtually unaffected when the upstream AUG produced a comparable levels of transcripts. Furthermore, insertion of a stable hairpin in the center of an extended 5' UTR did not significantly impact translation from AUG²⁶ or AUG²⁷. However, translation initiation remained 5' end-dependent since blocking the sgRNA 5' end severely restricted translation. We also provide additional evidence that AUG²⁷ represses translation of upstream AUG²⁶, which is mitigated when the spacing between the initiation codons is reduced or when the AUG²⁶ context is strengthened. In addition, improvements in the AUG²⁶ context led to enhanced accumulation of stalled ribosomal complexes at the start codon in the absence of increased translation, which also implies the existence of a specific initiation mechanism. Similar to previous results obtained for PEMV2 gRNA, enhancement of sgRNA translation by 3'CITEs boosted the number of templates translated but not the number of ribosomes per template. Overall, our data are not consistent with mechanisms proposed to explain translation from closely spaced initiation codons in viral genomes. We propose a new non-canonical, 5' end-dependent 3'CITE-driven mechanism that is used by umbraviruses for efficient dicistronic expression from their sgRNA.

MATERIALS AND METHODS

Vector construction

Full-length PEMV2 sgRNA was placed downstream of a T7 promoter and cloned into pUC19 with EcoR I and Sma

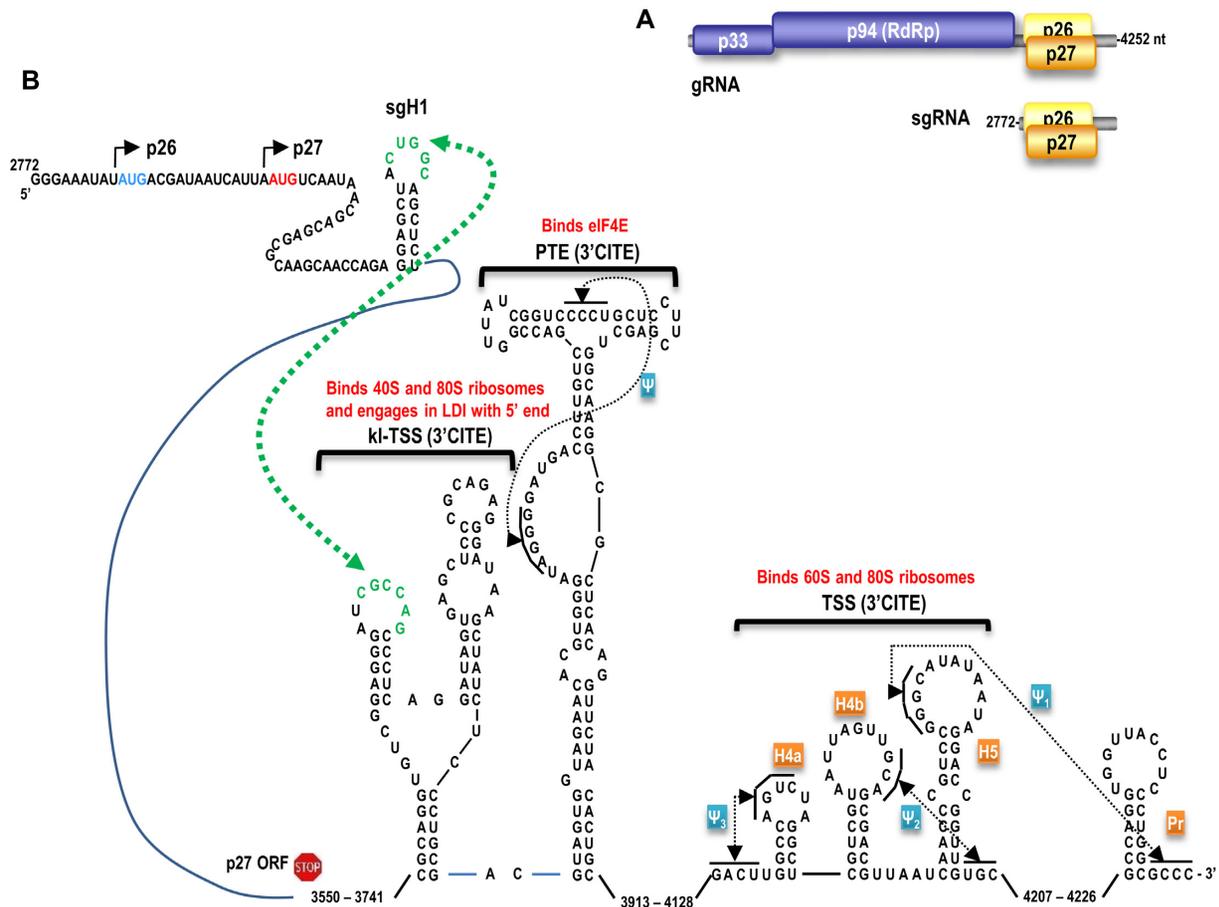


Figure 1. 5' and 3' sequences involved in translation of the PEMV2 sgRNA. (A) Genome organization of PEMV2 gRNA and sgRNA. (B) Sequences at the 5' and 3' ends of the sgRNA. Initiation codons for AUG²⁶ and AUG²⁷ are in blue and red, respectively. The long-distance interaction (LDI) is shown in green. The kl-TSS 3'CITE binds to 40S, 60S and 80S ribosomes and can simultaneously engage in the LDI (34,41). The PTE 3'CITE binds to eIF4E, which is followed by binding to eIF4G (35). The TSS 3'CITE binds to 60S and 80S ribosomes (42).

I restriction sites to generate pUC19-PEMV2 sgRNA as previously described (37). pUC19-PEMV2 sgRNA substitution, deletion or insertion mutants were generated using quick-change one-step site-directed mutagenesis (40). Luciferase reporter constructs p26-LUC and p27-LUC contain the initial 49 codons of the p26 ORF or 44 codons of the p27 ORF fused to the firefly luciferase ORF flanked by the sgRNA 5' UTR and the full-length 3' UTR as previously described (37). The gRNA luciferase reporter construct contains the 5' 89 nt and the full-length 3' UTR, as previously described (34). All mutations were confirmed by sequencing (Eurofins Genomics).

RNA transcription and *in vitro* translation

Linearized plasmids served as templates for RNA transcription using bacteriophage T7 RNA polymerase. RNA transcripts were purified with lithium chloride and quantified using NanoDrop 1000 (Thermo Scientific). For sedimentation analysis, RNA transcripts were fluorescently labeled as described in (33). For *in vitro* translation, RNA transcripts (0.5 pmol) were combined in 10 μ l translation mixtures with 5 μ l of wheat germ extract (WGE, Promega), 0.8 μ l of 1 mM amino acids mix (Met-), 100 mM potassium acetate and

0.5 μ l [5 μ Ci] ³⁵S-methionine. The translation mixture was incubated at 25°C for 45 min and then resolved on a 10% sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) gel. Gels were dried, exposed to a phosphorimager screen and analyzed using a FLA-5100 fluorescent image analyzer (Fujifilm) with data quantified using Multi Gauge Ver. 2.0 (Fujifilm). Relative efficiency of p26 and p27 synthesis was calculated taking into account the number of methionines in their sequences, 3 and 6, respectively.

Ribosome toeprinting

Complete amino acids mixtures were used in WGE translation reactions supplemented with 1 mM cycloheximide (CHX), and the reactions were incubated at 25°C for 45 min. The WGE reaction (3.5 μ l) was then mixed with 5 μ l of primer annealing buffer (1 \times Superscript III reverse transcriptase buffer, 10 mM dithiothreitol (DTT), 0.5 mM deoxynucleotides (dNTPs), 1 mM CHX and 1 U/ μ l RNase-out ribonuclease inhibitor) and incubated at 55°C for 2 min. One picomole of [γ -³²P] ATP-labeled primer complementary to positions 2875–2898 was annealed to the template, and the reactions incubated at 37°C for 5 min. Reverse tran-

scription was performed using Superscript III reverse transcriptase (25 U, Invitrogen) at 37°C for 30 min. Complementary DNA products were resolved on a 8% denaturing acrylamide gel, which was dried and exposed to a phosphorimager screen and quantified as described above.

Protoplast transfections and *in vivo* luciferase assays

Protoplasts were prepared from callus cultures of *Arabidopsis thaliana* (ecotype Col-0) and transfected with *in vitro* transcribed RNA transcripts using a modified polyethylene glycol-mediated transformation protocol as previously described (37). Briefly, 4×10^5 protoplasts were transfected with 4 μ g of luciferase reporter transcripts along with 1.3 μ g Renilla luciferase transcripts as an internal control. Luciferase activity was assayed 18 h later using a Dual-Luciferase[®] Reporter Assay System (Promega) and Modulus microplate multimode reader (Turner BioSystems).

Statistical analysis

Data from three independent experiments were statistically analyzed using ANOVA or *t*-test as indicated (Graphpad prism 7.0).

Sedimentation analysis of polyribosomes

Labeled reporter RNAs (20 μ l) were translated in WGE for 25 min. The translation reaction mixtures were then chilled on ice, supplemented with CHX up to 0.01 mg/ml and layered atop a linear 15–45% sucrose gradient in 12 ml Ultra-Clear Beckman tubes, containing 25 mM Tris–HCl (pH 7.6), 5 mM MgCl₂, 100 mM KCl, 0.1 mM ethylenediaminetetraacetic acid and 0.01 mg/ml CHX. Samples were subjected to centrifugation for 2 h 45 min in a SW-41 rotor in an Optima L-90K (Beckman-Coulter) ultracentrifuge at 37 000 rpm at 4°C. Gradients were fractionated from the bottom of the tubes at 0.5 ml/min with continuous measurement of the optical density at 254 nm with a UV Cord 2238 (Pharmacia) and fluorescence ($\lambda_{ex} = 492$ nm, $\lambda_{em} = 518$ nm) with an RF-5031PC fluorometer (Shimadzu) equipped with a 120 μ l LC flow cell.

RESULTS

Dependence of translation on the 5' end of PEMV sgRNA

In contrast with PEMV2 gRNA, which requires two 3'CITEs, efficient translation of the sgRNA requires three 3'CITEs: the kl-TSS, the adjacent downstream PTE and additionally, the downstream 3' proximal, T-shaped structure (TSS), which binds 60S ribosomal subunits (41,42). Similar to the gRNA, a long-distance RNA–RNA interaction connects the terminal loop of one kl-TSS hairpin with the terminal loop of a coding region hairpin, which in the sgRNA is designated as sgH1 (Figure 1B). This long-distance interaction (LDI) is proposed to relocate 3'-bound translational machinery that includes the 40S subunit to the 5' end (37,43). These similarities and differences between the gRNA and sgRNA for 3'CITE requirements were evident using both full-length templates as well as luciferase reporter constructs that contained the full-length 3' UTR and

either the gRNA 5' 89 nt or the sgRNA 5' 141 or 142 nt (for translation of p26 or p27, respectively) (33,37). Therefore, the requirement for an additional 3'CITE by the sgRNA reflects differences contained within the 5' terminal regions of these RNAs.

To ascertain whether gRNA and sgRNA 5' sequences direct different modes of translation initiation, we examined requirements for unidirectional ribosome scanning from the 5' end. One commonly used method to ascertain whether 40S ribosomal subunits scan through the 5' UTR is by placement of new initiation codons upstream of the natural initiation codon. 40S scanning in the 5' to 3' direction encounters the new initiation codon and begin translation, substantially reducing translation from downstream initiation codons. To determine if scanning is used for translation of p33 from the gRNA, a single base alteration at position 7 (U7G) was generated in the luciferase reporter construct producing an out-of-frame AUG codon upstream of the natural AUG initiation codon located at position 21 (Figure 2A). As a control, the same residue was also altered to an adenylate (U7A). RNA transcribed from the reporter constructs was transformed into *A. thaliana* protoplasts and luciferase activity was assayed 18 h later. As shown in Figure 2B, the presence of an upstream AUG reduced luciferase levels by 16-fold, whereas control luciferase levels were reduced by only 19%. This result is consistent with the 40S ribosomal subunit accessing the p33 initiation codon in the gRNA by a canonical scanning mechanism.

Since the PEMV2 sgRNA leader is only 9-nt long, placement of a new AUG upstream of AUG²⁶ might alter the Kozak context of AUG²⁶. We therefore extended the length of the 5' UTR to allow for a greater separation between new initiation codons and downstream AUG²⁶. In addition, because of the inherent complexity of having two closely spaced initiation codons (AUG²⁶ and AUG²⁷), we generated upstream AUGs in-frame with the p26 ORF in full-length sgRNA constructs to allow all translation products to be examined. This required that translation assays be conducted *in vitro* using wheat germ extracts (WGE).

Addition of 37 nt at the 5' end of the sgRNA produced little change in p26 synthesis and a 70% decrease in p27 synthesis (Figure 2D). The 37-nt sequence was also added to previously described luciferase reporter constructs p26-LUC and p27-LUC for *in vivo* translation assays in protoplasts. p26-LUC contains the PEMV2 sgRNA 5' UTR and N-terminal 49 codons of p26, fused in-frame with the firefly luciferase ORF followed by the full-length 3' UTR of PEMV2. p27-LUC was similar, but contained the N-terminal 44 codons of p27 (37). The presence of the +37 sequence in the reporter constructs had no effect on translation in protoplasts from AUG²⁶ and led to a 42% reduction in translation from AUG²⁷ (Figure 2E).

Initiation codons in different contexts and positions were next individually generated within the +37 sequence to determine their effects on *in vitro* translation from downstream AUG²⁶ and AUG²⁷ (Figure 2C). If translation of AUG²⁶ and AUG²⁷ is by leaky scanning, an upstream AUG was expected to significantly and proportionally reduce translation of p26 and p27. When the inserted AUG was located 15-nt upstream of AUG²⁶ in a weak Kozak context

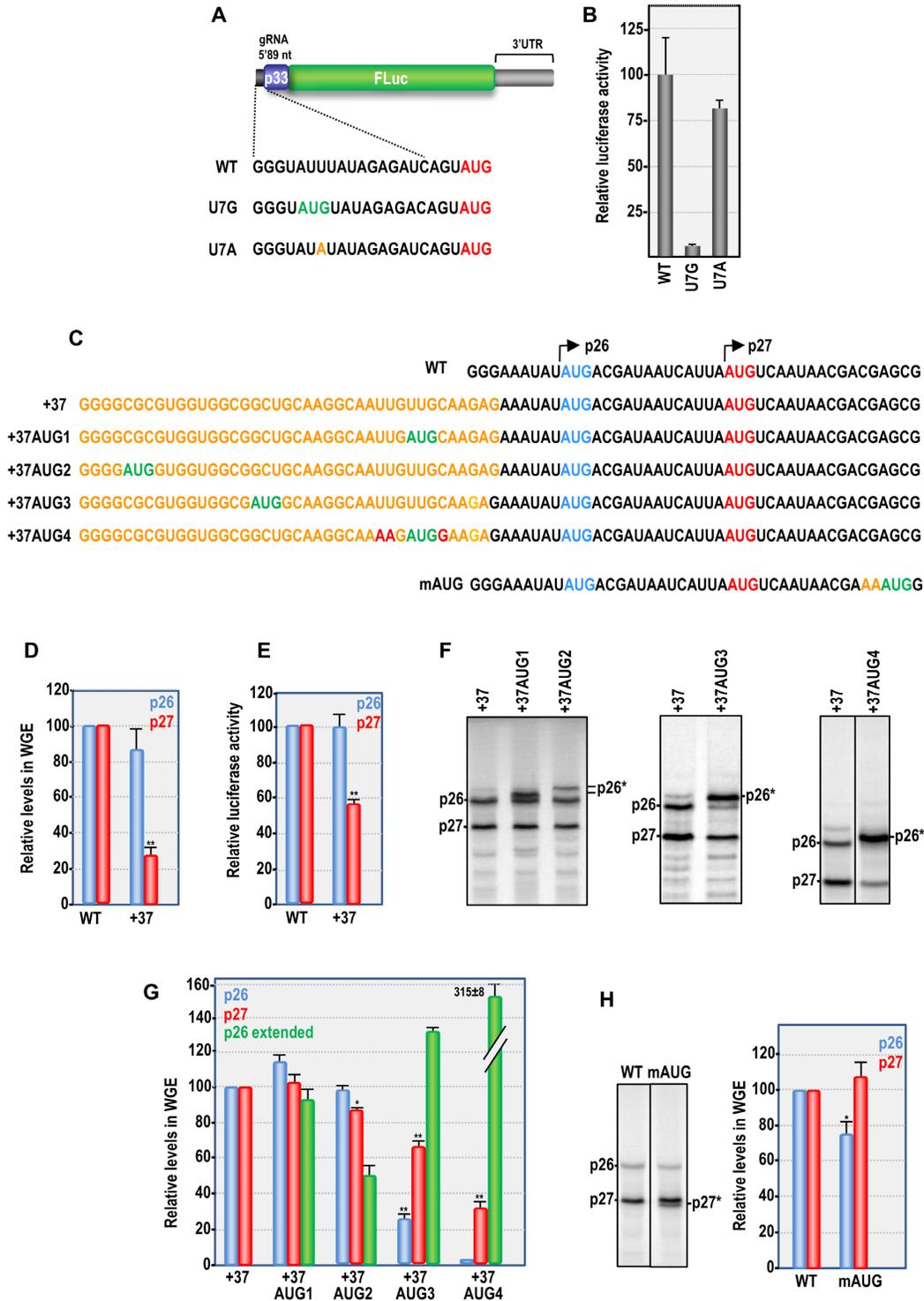


Figure 2. Effect of upstream AUG codons on translation initiation. (A) Alterations to the gRNA 5' UTR in a luciferase reporter construct. U7G generates an out-of-frame AUG upstream of the natural AUG at position 21. (B) *In vivo* translation of WT and mutant reporter constructs in *Arabidopsis thaliana* protoplasts. Throughout this report, mean values and standard error were calculated from at least three independent experiments. One-way ANOVA was used to analyze the statistical significance; * $P \leq 0.05$, ** $P \leq 0.01$. (C) Sequences of experimental and control full-length sgRNA constructs. Inserted and altered sequences are in orange. New initiation codons are in green. Additional altered residues in +37AUG4 to improve the context of the inserted AUG are in red. (D) Translation of WT and +37 sgRNA transcripts in WGE. (E) Translation of WT and +37 mutant reporter constructs in *Arabidopsis* protoplasts. (F) Typical SDS-PAGE gel showing levels of p26 and p27, and new p26 extended products (p26*) generated from the added in-frame initiation codons following translation in WGE. p26 and p27 migrate aberrantly, as previously described (37). (G) Quantification of the relative levels of p26*, p26 and p27 of +37 derived constructs in WGE. p26 and p27 protein levels were normalized to 100 for ease in comparisons. p26* levels were normalized to p26 for comparison. (H) (left) Typical SDS-PAGE gel showing levels of p26, p27 and the new product generated from the downstream, in-frame AUG (p26*); (right), quantification of the relative levels of p26 and p27 in WGE.

(+37AUG1), transcript levels generated from the new AUG were similar to those from WT AUG²⁶; however, there was no significant effect on translation from either AUG²⁶ or AUG²⁷ (Figure 2F and G). When the inserted AUG was located proximal to the 5' end (+37AUG2), translation levels from the new AUG were ~50% of WT AUG²⁶, and only a modest effect was found for AUG²⁷ translation. These data suggest that the ribosomal subunit efficiently missed the upstream AUGs in a weak context to provide the constant level of initiation at the target start codons. The additional translation from the upstream AUGs could be explained by spontaneous 3'CITE-independent initiation in the extended 5' UTR.

In contrast, when the inserted AUG was located in a good context 17 nt from the 5' end (+37AUG3), translation from the new AUG was 30% more efficient than WT AUG²⁶. Addition of this initiation codon unequally decreased translation of AUG²⁶ and AUG²⁷ by 73 and 34%, respectively. Improving the context of the inserted AUG in +37AUG1 from a weak to an optimal context (for Arabidopsis, AAAAUGG; +37AUG4) resulted in highly efficient translation from the new AUG, generating ~3-fold more product than +37AUG1 or WT AUG²⁶. The new AUG in +37AUG4 reduced translation from AUG²⁶ below the level of detection, but had a more limited effect on AUG²⁷, reducing its translation by only 66%. These data indicate that only highly efficient upstream initiation codons affect translation of p26 and p27 and initiation from AUG²⁷ is less sensitive than AUG²⁶ to the presence of these codons. Overall, these results suggest that translation of the sgRNA with an artificially extended 5' UTR is only partially consistent with expectations from canonical leaky scanning.

To determine the effect of insertion of a downstream initiation codon, one was inserted in an optimal context 15-nt downstream of AUG²⁷ (and in-frame with AUG²⁷) in WT full-length sgRNA (mAUG, Figure 2C). This initiation codon was efficiently utilized, generating a product of the expected molecular weight (asterisk, Figure 2H, left). Translation of p27 was not significantly affected by the downstream AUG, but translation of p26 decreased by 24%.

An assay frequently used to identify scanning ribosomes is insertion of a stable hairpin (HP) upstream of the initiation codon that blocks ribosome movement thus strongly affecting translation. To ascertain if upstream hairpins affect translation from AUG²⁶ and AUG²⁷, a 37-nt insert forming a hairpin ($\Delta G = -26$ kcal/mol; Figure 3A) was placed in three locations: (i) at the 5' end of the full-length sgRNA (5'HP); (ii) at the 5' end of the sgRNA with a 100-nt spacer to provide a separation distance between the hairpin and AUG²⁶ (+137-5'HP) and (iii) in the center of the 100-nt 5' extension (+137-cHP) (Figure 3B). The 'unstructured' +37 sequence (in orange in Figure 2C) was also placed at the 5' end of the 100-nt extension (+137) to provide a 137-nt control for the latter two constructs.

5'HP and +137-5'HP reduced translation from AUG²⁶ and AUG²⁷ to near background levels in WGE compared with the less structured sequences of equal length in the controls (+37 and +137) (Figure 3C). +137 reduced translation of p26 and p27 by 48 and 29%, respectively, compared with +37 (Figure 3C, left). When normalized to this control construct, +137-cHP was significantly less inhibitory

than +137-5'HP, reducing translation of p26 and p27 by 49 and 38% *in vitro* (Figure 3C, right). +137-cHP and control sequences were also added to p26-LUC and p27-LUC reporter constructs. As in WGE, 5'HP reduced translation from AUG²⁶ and AUG²⁷ to near background in protoplasts (Figure 3D, left). +137-cHP reduced translation from AUG²⁶ and AUG²⁷ by 71 and 73%, respectively, compared with the +137 control (Figure 3D, right). These results suggest that while 5' end accessibility is required for efficient translation, i.e. the initiation is a 5'-end-dependent process, a HP located downstream of the 5' end that may obstruct ribosome scanning still permits a moderate amount of translation to proceed, especially in WGE. All together, the above results suggest that translation of the PEMV2 gRNA and sgRNA may be occurring by different mechanisms directed by features in the 5' region of these RNAs.

Effect of sequence context and elimination of competitive AUG on translation from AUG²⁶ and AUG²⁷

We previously demonstrated that eliminating AUG²⁷ in WT sgRNA transcripts enhanced translation from AUG²⁶, suggesting that AUG²⁷ was suppressing translation from AUG²⁶ (32) (see m2, Figure 4B and C). Since increasing the length of the 5' UTR could alter the mechanism of translation from AUG²⁶ and AUG²⁷, we generated mutations that eliminated either AUG²⁶ or AUG²⁷ in the +37 background in full-length sgRNA and in the reporter constructs. Converting AUG²⁷ to CAG in the +37 construct (+37m2) enhanced translation from upstream AUG²⁶ by 36% *in vitro* and 139% *in vivo* (Figure 4D and E), suggesting that AUG²⁷ was still capable of suppressing translation from AUG²⁶ when the 5' UTR was expanded. Converting AUG²⁶ to CAG (+37m1) increased translation from AUG²⁷ by 6-fold *in vitro* and 4-fold *in vivo* (Figure 4D and E), which was higher than found for the comparable mutations in WT sgRNA (m1, Figure 4B and C). This difference can be explained by a significant decrease in the p27/p26 synthesis ratio (from 1.7 to 0.5) when increasing the length of 5' UTR, and, therefore by the stronger observable effect of switching the initiation from AUG²⁶ to AUG²⁷.

Unlike conventional leaky scanning where the downstream AUG is usually associated with an optimum Kozak context, the sequence contexts for both AUG²⁶ and AUG²⁷ are exceptionally weak [UAU(AUG²⁶)A and UUA(AUG²⁷)U; critical -3 and +4 positions are underlined]. To determine if improving the sequence context of AUG²⁶ or AUG²⁷ impacts translation of either p26 or p27, nucleotides surrounding individual initiation codons in WT sgRNA were converted to the optimal *A. thaliana* context (AAAAUGG; Figure 4A), and translation was assayed in WGE. Optimizing the AUG²⁶ context did not enhance translation of p26 (m3, Figure 4B); however, p27 levels were reduced by 11-fold. These results resembles the aforementioned data on 5' UTR extension (Figure 2D), suggesting that a very short 5' UTR may have an effect on initiation similar to that of a weak context. In contrast, optimizing AUG²⁷ had no discernable effect on translation of either p27 or p26 (m4, Figure 4B). m3 and m4 mutations were also introduced into p26-LUC and p27-LUC for *in vivo* translation assays in Arabidopsis protoplasts. Sim-

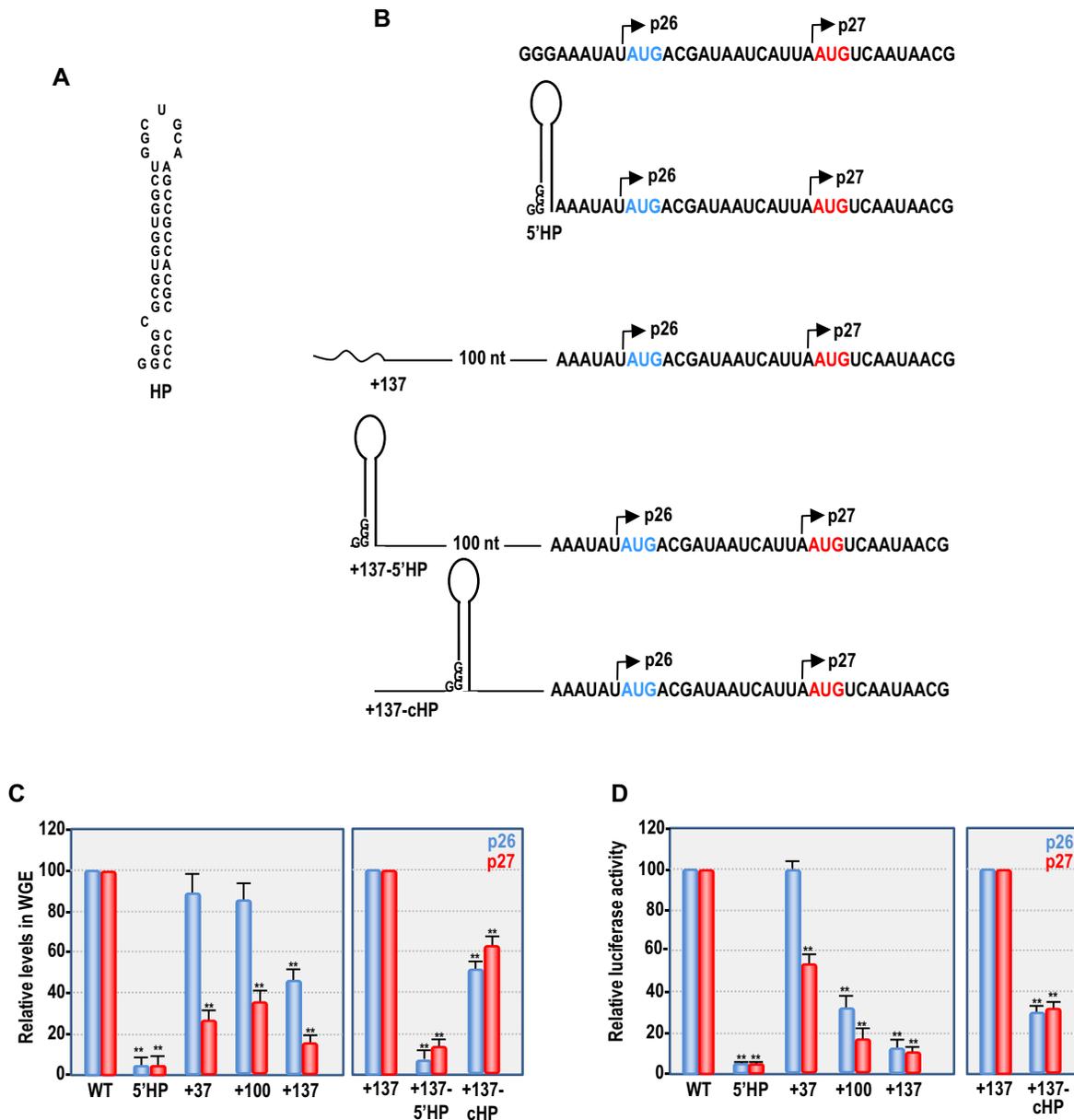


Figure 3. Effect of upstream stable hairpin on translation from AUG²⁶ and AUG²⁷. (A) Forty nucleotide hairpin (HP) generated from the 37 nt added sequence. (B) Constructs showing insertion of HP at the 5' end of WT sgRNA (5'HP), at the 5' end prior to a 100-nt insert (+137-5'HP); and in the center of the 100-nt insert (+137-cHP). Insertion of unstructured sequence at the 5' end prior to a 100-nt insert serves as a control (+137). (C) *In vitro* translation of wt and mutant sgRNAs. Data were either normalized to p26 and p27 levels in WT (left) or +137 (right). (D) Translation of p26-LUC and p27-LUC in protoplasts containing the 5' terminal HP or 5' terminal extensions (left) or +137cHP (right). One-way ANOVA was used to analyze the statistical significance; *P ≤ 0.05, **P ≤ 0.01. See legend to Figure 2 for more details.

ilar to results using WGE and full-length sgRNA, optimizing the context of AUG²⁶ did not significantly impact p26 translation *in vivo*, whereas p27 translation was reduced to near background levels (m3, Figure 4C). Optimizing AUG²⁷ slightly enhanced translation from both AUG²⁷ (1.2-fold) and AUG²⁶ (1.3-fold) (m4, Figure 4C). Improving the Kozak context of AUG²⁶ when the template contained the extended 5' UTR did not significantly affect translation of p26 while reducing p27 synthesis to near background levels *in vivo* and *in vitro* (+37m3; Figure 4E and F). Likewise, improving the context of AUG²⁷ in the extended 5'

UTR template had no significant effect on translation of p26 or p27 *in vitro* or *in vivo* (+37m4; Figure 4D and E). These results provide additional evidence that the behavior of AUG²⁶ and AUG²⁷ in translation remains similar when the 5' UTR is increased from 9 to 46 nt.

Reduced translation of p27, when AUG²⁶ was optimized, could have been due to either a negative impact of the altered residues on initiation at AUG²⁷ or enhanced ribosome association with AUG²⁶ (although without a corresponding increase in translation initiation) leading to reduced ribosome association with AUG²⁷. To distinguish be-

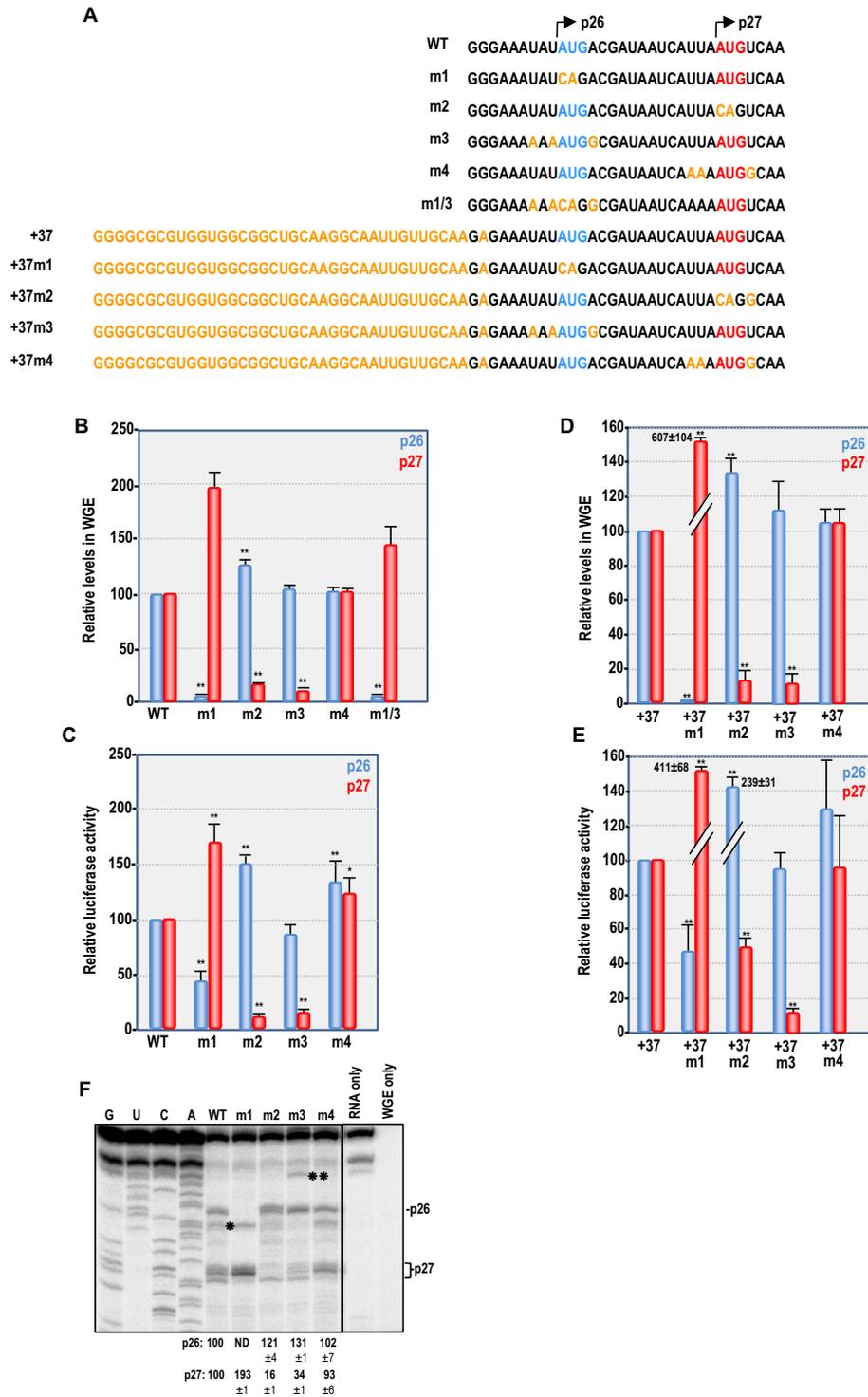


Figure 4. Effect of context on translation initiation from AUG²⁶ and AUG²⁷. (A) Mutations in and near AUG²⁶ and AUG²⁷ that either alter the AUGs (m1 and m2) or improve their Kozak context (m3 and m4) are shown for WT constructs and constructs with added 5' +37 nt. Data for m1 and m2 are from reference (37) and are repeated here for clarity. (B) Translation of WT and mutant sgRNAs in WGE. (C) Relative luciferase activity in protoplasts transfected with WT and mutant p26-LUC and p27-LUC. (D) Translation of +37 and +37 mutant sgRNAs in WGE. (E) Translation of WT, +37 and +37 mutant reporter constructs in Arabidopsis protoplasts. B, C, D, and E, one-way ANOVA was used to analyze the statistical significance; *P ≤ 0.05, **P ≤ 0.01. See legend to Figure 2 for more details. Luciferase activity that persists *in vivo* despite the lack of initiation codons in +37m1 and +37m2 is likely due to translation initiating from the natural luciferase AUG. (F) Ribosome toeprinting of WT and mutant using WGE. Toeprints corresponding to stalled ribosomes on AUG²⁶ and AUG²⁷ are labeled. Single and double asterisks denote additional AUG²⁷ and AUG²⁶ toeprints, respectively, described previously (37), which may represent improper fixing of the sgRNA in the mRNA-binding cleft of the 40S subunit, allowing for further extension by the reverse transcriptase (33). G, U, C, A are ladder lanes. Data are from three independent experiments. Line between lanes denotes removal of an intervening lane.

tween these two possibilities, AUG²⁶ was mutated to CAG while retaining the context alterations in the surrounding sequences (generating m1/3). As shown in Figure 4B, p27 levels increased by 1.5-fold over WT levels when translating m1/3 transcripts, indicating that the altered bases surrounding AUG²⁶ were not directly responsible for reduced translation initiation at AUG²⁷.

To determine if optimizing the AUG²⁶ context enhanced ribosome occupancy in the absence of increased translation initiation, ribosome toeprinting assays were performed using full-length sgRNA and WGE treated with CHX to stall ribosomes at the initiation sites. Toeprints corresponding to ribosomes stalled at AUG²⁶ or AUG²⁷ were either eliminated or reduced when the respective initiation codons were mutated (m1 and m2) (Figure 4F). Improving the sequence context of AUG²⁶ (m3) was associated with a 31% increase in the AUG²⁶ toeprint and a 3-fold reduction in the AUG²⁷ toeprint. Altering the AUG²⁷ context did not significantly impact ribosome occupancy at either AUG²⁶ or AUG²⁷ (Figure 4F). These results along with the *in vitro* translation experiments suggest that optimizing the AUG²⁶ context enhanced ribosome association at AUG²⁶ but not translation initiation, and reduced ribosome association at AUG²⁷. In contrast, optimizing the context of AUG²⁷ had no significant effect on ribosome association or translation at AUG²⁷ or AUG²⁶.

Residue in position –2 of AUG²⁶ influences translation start site selection

To further explore how 5' UTR residues impact start site selection, the PEMV2 sgRNA 5' UTR was replaced with 5' UTRs from sgRNAs of two other umbraviruses (Figure 5A). The 5' UTRs of most umbravirus sgRNAs are of similar size and overall sequence composition (38), due to the presence of a 5' terminal Carmovirus Consensus Sequence (G₂₋₃ A/U₃₋₉) that is required for synthesis of sgRNAs by RdRp of the related Betacarmovirus, *Turnip crinkle virus* (44,45). Replacement of the PEMV2 sgRNA 5' UTR with that of *Opium poppy mosaic virus* (OPMV) introduced 4- of 9-nt alterations yet had no significant impact on translation from AUG²⁶ and AUG²⁷ in WGE or in protoplasts (Figure 5B and C). In contrast, replacement of the PEMV2 sgRNA 5' UTR with that of *Tobacco bushy top virus* (TBTv), which is 1 nt shorter, resulted in a 29% increase in p26 levels and a 72% decrease in p27 levels in WGE. Similar results were also obtained using the reporter constructs in protoplasts, with a 60% increase in luciferase translated from AUG²⁶ and a reduction in translation from AUG²⁷ to near background levels (Figure 5C). To determine if this increased translation from AUG²⁶ was correlated with enhanced ribosome binding at AUG²⁶, ribosome toeprinting was conducted for WT sgRNA and sgRNA with the 5' UTR of TBTv sgRNA (Figure 5D). The 5' UTR of TBTv caused ribosome binding to increase 2-fold at AUG²⁶ and decrease by 67% at AUG²⁷. Since the increased levels of translation from AUG²⁶ were similar to the levels obtained when AUG²⁷ was absent (m2; Figure 4B and C), this suggests that alterations in the 5' UTR can alleviate the repressive effect of AUG²⁷ on AUG²⁶.

To ascertain if these results were due to the slightly shorter length of the TBTv 5' UTR, an additional adenine residue was inserted into the sequence that increased the length of the 5' UTR to that of the PEMV2 sgRNA 5' UTR (TBTv+1). This particular nucleotide addition was selected as the 5' UTR sequence now differed at only a single residue from the OPMV sgRNA 5' UTR sequence, which had a negligible impact on translation start site selection (OPMV: GGGUAUUGAAUG; TBTv+1: GGGUAUUC^AAAUG). The 5' UTR TBTv+1 produced a similar ratio of start site selection in WGE as the TBTv 5' UTR (Figure 5B) and maintained the uneven ratio of translation from AUG²⁶ and AUG²⁷ in protoplasts, but with WT levels of translation from AUG²⁶ and a lower reduction in translation from AUG²⁷ (Figure 5C). These results strongly suggest that enhanced levels of translation from AUG²⁶ coupled with reduced translation from AUG²⁷ was unrelated to the slight difference in 5' UTR length but correlated with a cytidine residue at position –2.

Reducing the distance between AUG²⁶ and AUG²⁷ blocks repression by AUG²⁷

In PEMV2 and most umbraviruses, AUG²⁶ and AUG²⁷ are separated by a 13-nt spacer (the exceptions are *Carrot mottle virus* and *Carrot mottle mimic virus* with 19- and 25-nt spacers, respectively) (37). Shortening the spacer by 3 nt in full-length sgRNA (m5, Figure 6A) increased synthesis of p26 by 1.7-fold and reduced p27 levels by 53% (Figure 6B). The same deletion in p26-LUC and p27-LUC increased p26 levels by 1.7-fold and decreased p27 levels by 72% when assayed *in vivo* (Figure 6C). To determine if the elevated level of p26 and reduced level of p27 were dependent on the presence of the other's initiation codon, AUG²⁶ or AUG²⁷ were altered to 'CAG' in the m5 background generating m5.1 and m5.2, respectively. Eliminating AUG²⁷ in the m5 background (m5.2) caused a similar increase in translation of AUG²⁶ *in vitro* or *in vivo* as did the 3-nt spacer deletion. This result suggests that elevated levels of translation from AUG²⁶ when only 13-nt separate the two initiation codons is due to ribosome access to AUG²⁶ no longer being affected by the presence of AUG²⁷. In contrast, elimination of AUG²⁶ in m5.1 increased translation from AUG²⁷ by ~60% over WT levels *in vitro*, similar to results obtained when AUG²⁶ was eliminated in the WT background (compare with m2, Figure 4B and C). This suggests that the 3-nt deletion did not directly affect AUG²⁷ translation levels, but rather that the reduction in AUG²⁷ translation was due to enhanced translation from AUG²⁶.

While the 3-nt deletion shortened the distance between AUG²⁶ and AUG²⁷, it also shortened the distance between AUG²⁶ and stem-loop sgH1, whose terminal loop engages in the LDI with the k1-TSS 3'CITE. Since the LDI is necessary for transferring initiation factors from the 3' UTR to the vicinity of translation initiation, we wanted to determine if the reduced distance between AUG²⁶ and sgH1 was involved in the loss of suppression by AUG²⁷. Insertion of 3-nt downstream of AUG²⁷ in the m5 background restored the distance between AUG²⁶ and sgH1 (m5.3; Figure 6A). As a control, the 3-nt insertion was also constructed in the WT background (m6). No significant difference in

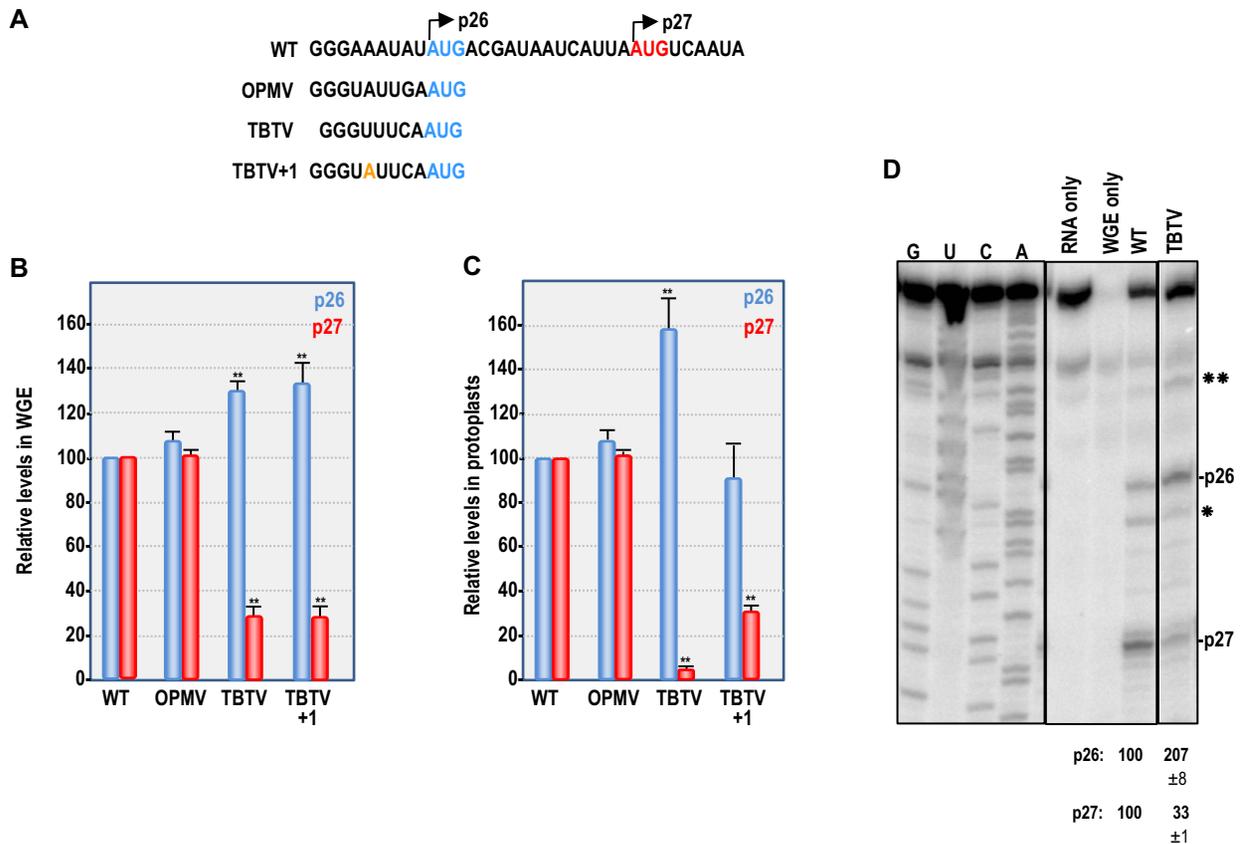


Figure 5. Replacing the 5' UTR of PEMV2 sgRNA with the 5' UTRs of other umbraviruses. (A) Sequences used to replace the 5' UTR of PEMV2 sgRNA. TBTV+1 has an insert of a single residue (in orange) such that the sequence is the same length as the 5' UTRs of PEMV2 and OPMV, and now differs at a single position from the OPMV 5' UTR sequence. (B) *In vitro* translation of WT and mutant sgRNAs in WGE. (C) *In vivo* translation of WT and mutant p26-LUC and p27-LUC in protoplasts. B and C, one-way ANOVA was used to analyze the statistical significance; * $P \leq 0.05$, ** $P \leq 0.01$. (D) Ribosome toeprinting of WT and mutant sgRNAs in WGE. Single and double asterisks denote additional AUG27 and AUG26 toeprints, respectively.

translation was obtained between WT and m6, or between m5 and m5.3 (Figure 6B and C), indicating that shortening the distance between AUG²⁶ and sgH1 in m5 did not directly contribute to loss of suppression by AUG²⁷. To determine if the critical factor was the removal of important residues and not the distance between AUG²⁶ and AUG²⁷, the deleted 'AAU' and the adjacent cytidylate were replaced with 'UUCG' (m7). No significant difference was found in translation between WT and m7 *in vitro* or *in vivo* (Figure 6B and C).

We also assayed the effect of increasing the distance between AUG²⁶ and AUG²⁷ by inserting 3 or 9 nt into the spacer region (m8 and m9, respectively, Figure 6A). As shown in Figure 6B and C, m8 had a positive effect on AUG²⁷ levels *in vivo* only, and m9 had a positive effect on AUG²⁶ *in vivo* only. Neither insertion caused a loss of repression by AUG²⁷, which is accompanied by reduced levels of translation from AUG²⁷ (e.g. m5). In addition, deletion of 3 or 9 nt between AUG²⁷ and hairpin sgH1 (m10 and m11, respectively) also had no effect on AUG²⁷ repression (Figure 6B and C). Altogether, these findings suggest that moving AUG²⁷ closer to AUG²⁶ (which also makes them more similar in terms of their positions relative to 5' end and sgH1), but not other insertions or deletions in the vicinity, blocks repression of AUG²⁶ by AUG²⁷.

The long-distance interaction enhances the number of templates translated but not the number of ribosomes per template

Translation of PEMV2 gRNA is enhanced by the kl-TSS and PTE 3'CITEs if the LDI is functional. We recently reported that this 3'CITE-mediated enhancement of translation in WGE was not due to an increase in the gRNA translation initiation rate, but rather correlated with an increase in the number of gRNA templates undergoing translation (33). In addition, gRNA reporter templates with and without the LDI were mainly occupied by monosomes and disomes when translated in WGE, unlike similar-sized control constructs that also associated with heavy polysomes (more than four ribosomes occupying the template).

Since the gRNA and sgRNA have different requirements for 3'CITEs and appear to have distinct modes of translation initiation, we wanted to determine if the LDI between the kl-TSS and sgH1 in the sgRNA also affected the number of templates utilized and not the number of ribosomes per template. p26-LUC and p27-LUC, with and without their LDIs (Figure 7A), were fluorescein-labeled and subjected to polysome sedimentation after 30-min incubation in WGE. Distribution of ribosomes in sucrose gradients was also determined by measuring ultraviolet absorbance (A_{254}). Con-

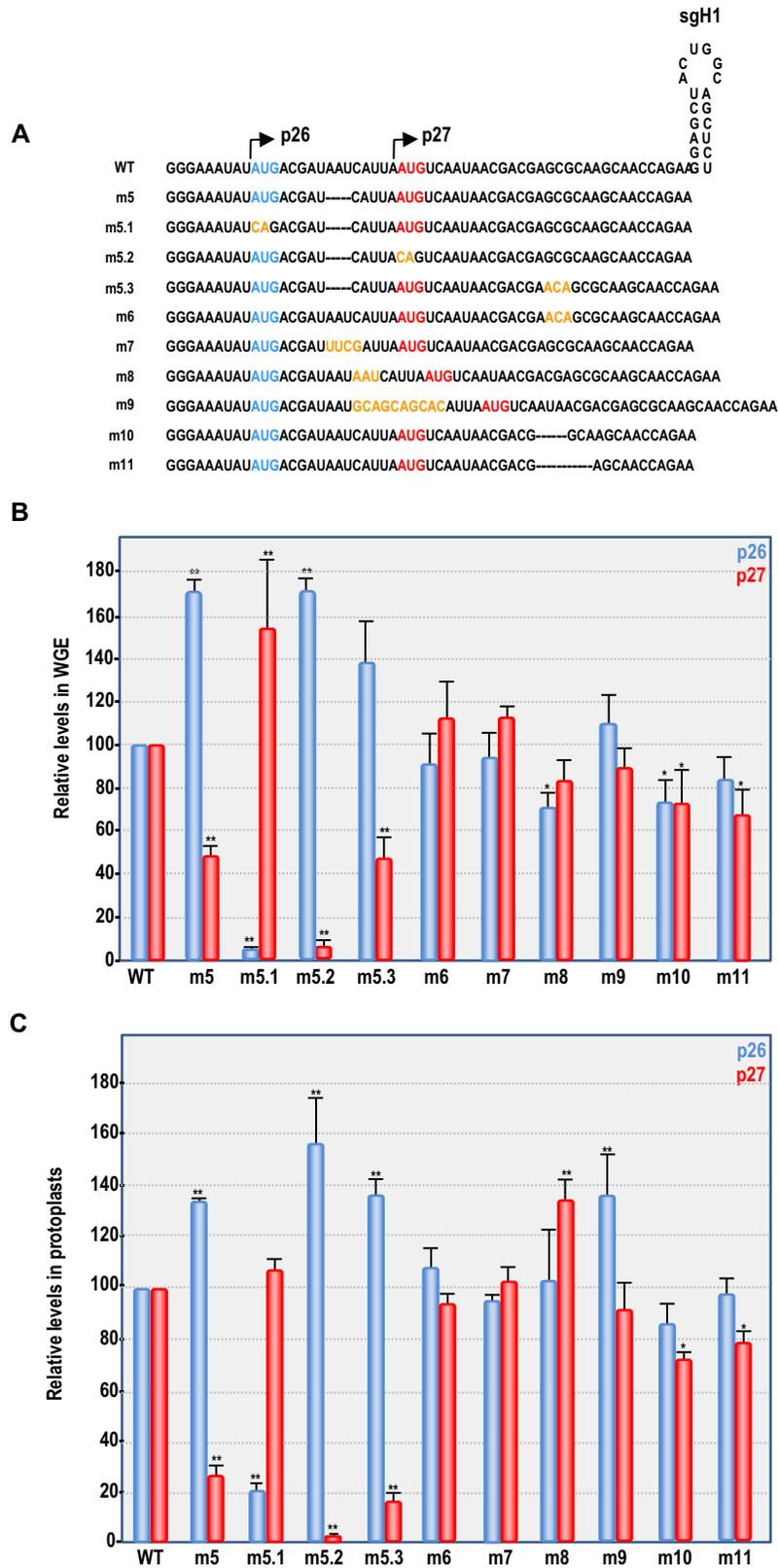


Figure 6. Effect of shortening and lengthening the distance between AUG²⁶ and AUG²⁷ on start codon selection. (A) Constructs showing deletions and insertions between AUG²⁶ and AUG²⁷, and deletions between AUG²⁷ and sgH1. (B) *In vitro* translation of WT and mutant sgRNAs in WGE. (C) *In vivo* translation of WT and mutant p26-LUC and p27-LUC in protoplasts. One-way ANOVA was used to analyze the statistical significance; *P ≤ 0.05, **P ≤ 0.01. See legend to Figure 2 for more details.

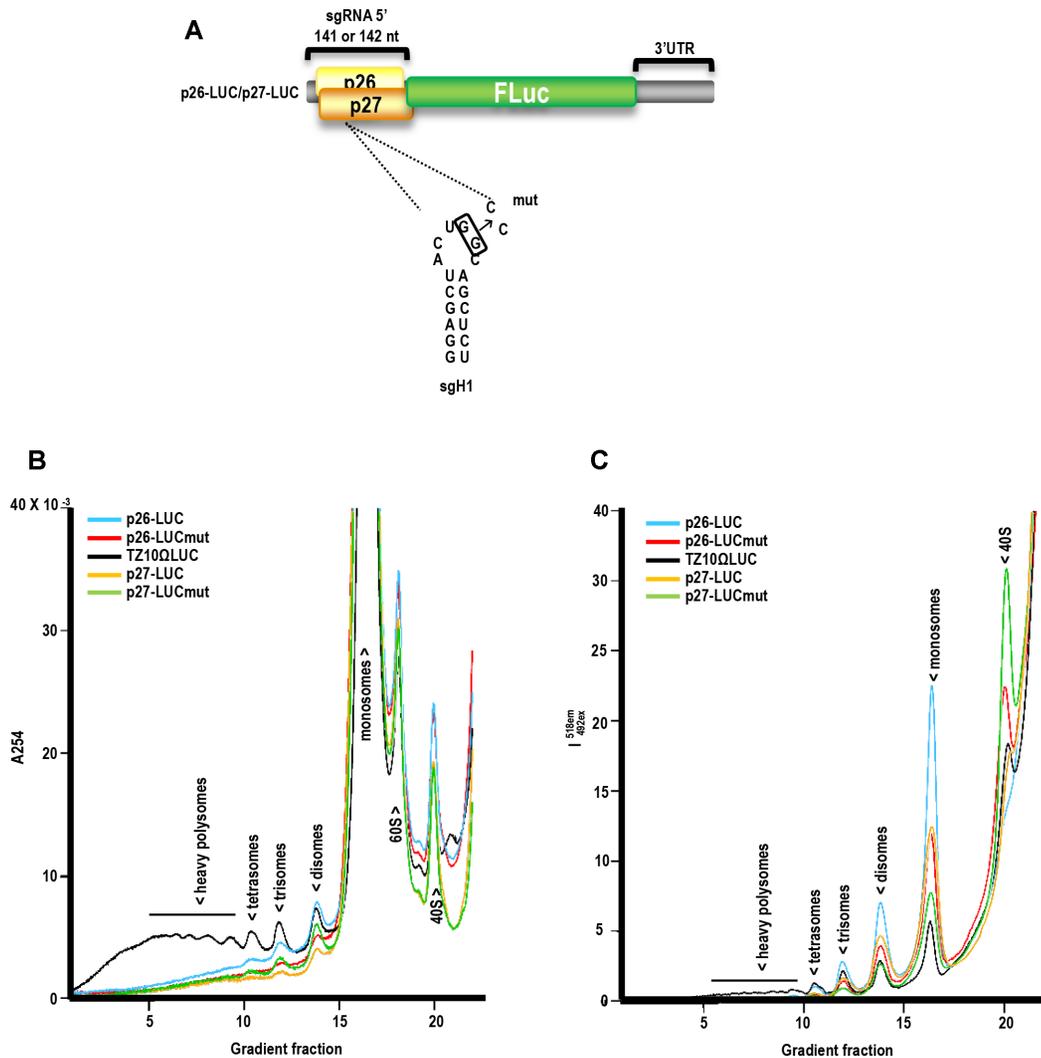


Figure 7. LDI enhances the number of templates translated but not the number of ribosomes per template. (A) Luciferase constructs p26-LUC and p27-LUC showing mutation in sgH1 that eliminates the LDI between the *kl*-TSS and sgH1, generating p26-LUCmut and p27-LUCmut. (B) Polysome distribution on uncapped p26-LUC, p26-LUCmut, p27-LUC, p27-LUCmut and TZ10ΩLuc templates. TZ10ΩLuc is a control template containing the TMV Ω translation enhancer in the 5' UTR and the TMV 3' UTR. RNAs (50 nM) were incubated in WGE for 30 min, then subjected to sedimentation analysis. Ultraviolet absorbance profiles reflect mainly the distribution of ribosomes. (C) Fluorescence profiles represent allocation of the labeled reporter RNAs along gradients.

control construct TZ10ΩLuc contains the *Tobacco mosaic virus* (TMV) Ω 5' translation enhancer and the 3' TMV UTR. As previously reported (33), TZ10ΩLuc associated with both light and heavy polysomes (Figure 7B and C). In contrast, both WT and mutant (p26-LUCmut and p27-LUCmut) sgRNA reporter constructs associated mainly with monosomes and disomes (Figure 7B). As with the gRNA, sgRNA templates containing the LDI had higher peaks than their mutant counterparts, indicating that more templates were being translated (Figure 7C). Also, more templates without a functional LDI were associated with 40S subunits, i.e. not participating in translation (Figure 7C), suggesting that 40S subunits sequestered by the 3' CITEs are more efficiently involved in translation if the LDI is intact, similar to what was found for the gRNA (33). These data suggest that, as with the gRNA, sequences at the 5' and 3' ends of the sgRNA moderate the initiation level thus restricting the number of

ribosomes per template while the LDI enhances the number of templates translated.

DISCUSSION

Translation of overlapping p26 and p27 ORFs in PEMV2 sgRNA requires precise regulation to balance the progress of critical late-stage events. p26 binds the nuclear protein fibrillarlin and is required for long-distance movement and for genome RNA stabilization in cells (46,47). p27 is likely the cell-to-cell movement protein based on examination of the Groundnut rosette umbravirus ortholog (38,39). Recently, we demonstrated that dicistronic expression of p26 and p27 from the sgRNA is a controlled process enhanced by three 3' CITEs in the 3' UTR when a long-distance RNA–RNA interaction connects the 5' and 3' ends (37). This differs from translation of the gRNA, which,

along with the LDI, requires only two 3'CITEs (33). These results strongly suggested that initiating translation on a virus's gRNA and sgRNA may involve different mechanisms, which may reflect the likelihood that only the gRNA is subjected to amplification.

Initiation codon selection at AUG²⁷ can be from a nonlinear process

If ribosomes simply scan linearly in the 5' to 3' direction along the PEMV2 sgRNA, then initiation codons inserted upstream of AUG²⁶ should capture a portion of the ribosomes that otherwise would reach AUG²⁶ or AUG²⁷, thus proportionately decreasing translation of p26 and p27. AUGs inserted upstream of AUG²⁶ in templates with an extended 5' UTR (+37 AUG1 and +37AUG2) where the inserted AUGs were used for translation at an efficiency that was at or below that of AUG²⁶, had no significant effect on translation from either AUG²⁶ or AUG²⁷ (Figure 2F left and G). This may reflect that initiation at weak upstream AUGs in the extended 5' UTR is a spontaneous 5' end-dependent process that does not compete with 3'CITE-driven initiation at AUG²⁶ and AUG²⁷. In contrast, the insertion of highly efficient upstream AUGs in good contexts inhibited translation from both AUG²⁶ and AUG²⁷; however, the decrease in translation was not proportional, with AUG²⁶ much more strongly affected than downstream AUG²⁷ (+37AUG3 and +37AUG4, Figure 2F and G). This suggests that, although 3'CITE-dependent initiation does involve reading of 5' UTR sequence, initiating ribosomes use a separate mechanism to select between AUG²⁶ and AUG²⁷, and this mechanism differs from canonical 5'→3' consecutive search for the start codon.

When a hairpin was placed in the center of the extended 5' UTR of PEMV2 sgRNA, translation from AUG²⁶ and AUG²⁷ only decreased between 30 and 50% in WGE. However, similar results have not been considered completely contradictory to ribosome scanning, as the inserted hairpin might be destabilized by the unwinding activity of ribosome-associated helicases (48,49). An alternative interpretation could be the 'RNA looping' model, where ribosomes can access initiation codons obstructed by a downstream hairpin as translation is independent of linear scanning (21,50). More importantly, accessibility of the RNA's 5' terminus was shown to be necessary for initiation from both AUG²⁶ and AUG²⁷ since placement of a stable hairpin at the very 5' end of the sgRNA with and without a 5' UTR extension virtually abolished detectable translation from AUG²⁶ and AUG²⁷ (Figure 3C).

Sequence context of AUG²⁶ and AUG²⁷ do not conform with the requirements for leaky scanning

Initiation at consecutive AUG codons is usually explained by a leaky scanning model, which posits that if one or more 5'-proximal initiation codons are surrounded by a weak Kozak context (absence of a purine at -3 and/or a guanylate at +4) some portion of the ribosomes bypass the site(s) and initiate at a downstream initiation codon(s). For dicistronic PEMV2 sgRNA, the AUG²⁶ and AUG²⁷ -3/+4 context combinations are U/A and U/U, which are very

weak and used at only 2.4 and 2.1%, respectively, in the Arabidopsis genome (51). In Arabidopsis, the -2 position is mainly occupied by an adenylate, followed by cytidylate, uridylate and guanylate (51). In WT PEMV2, the -2 positions for AUG²⁶ and AUG²⁷ are A and U, respectively, suggesting that overall, AUG²⁶ is in a stronger context than AUG²⁷. Despite the weaker context of AUG²⁷, synthesis of p27 is ~1.7-fold more efficient than that of p26 (e.g. see WT lane in Figure 2H), when taking into account the number of methionines in the corresponding proteins. This is a deviation from the canonical leaky scanning model, which implies that an upstream start codon in a stronger context should have a higher initiation potential. Improving the sequence context of AUG²⁶ in WT sgRNA enhanced ribosome association with AUG²⁶, but surprisingly did not increase translation of the encoded protein *in vitro* or *in vivo* (Figure 4). This suggests that while ribosomes were now associating more efficiently with AUG²⁶ in a strong context, a rate-limiting process at a later stage of initiation prevented the increased number of ribosomes from entering translation. Similarly, as previously shown, deletion of the TSS (which binds the 60S ribosomal subunit) suppressed sgRNA translation, but the amount of the corresponding initiation complex remained the same according to toeprinting analysis, which was not the case for deletion of the kl-TSS or PTE (37). It is possible that in both these cases, toeprinting was reporting on intermediate 48S complexes stalled by the inefficient delivery of 60S subunit as the rate-limiting step.

A 37-nt extension of the sgRNA 5' UTR (+37) did not change the results of AUG²⁶ context optimization (Figure 4). Interestingly, these results are in contrast with those obtained for an AUG inserted upstream of AUG²⁶ in the +37 construct (+37AUG1, Figure 2), where optimization of the context (+37AUG4) resulted in a 2.4-fold increase in translation (Figure 2G). When the 5' UTR was replaced with that of OPMV (Figure 5B and C), a C at position -2 significantly enhanced translation of p26 both in WGE and *in vivo*, also suggesting that the effect of AUG²⁶ context is more complicated than a mere improvement of 40S/AUG interaction.

When AUG²⁶ occupied the optimal Kozak context, both ribosome toe-prints at AUG²⁷ and translation of p27 were substantially reduced, a result consistent with a leaky scanning mechanism (Figure 4). Since insertion of an AUG downstream of AUG²⁷ was efficiently translated (Figure 2H), a canonical leaky scanning mechanism would also imply that a significant number of ribosomes are bypassing AUG²⁷. However, optimizing the sequence context of AUG²⁷ only minimally impacted translation of p27 or p26 (Figure 4), further suggesting that conventional Kozak context rules and a conventional leaky scanning mechanism do not apply.

The most detailed study addressing initiation from closely spaced AUG codons was confined to the analysis of TYMV gRNA translation (30). The initiation codons AUG⁶⁹ and AUG²⁰⁶ in the capped TYMV genome (which has no identified 3' translation enhancers) are separated by 7-nt. Since translation of the 5' AUG was affected by alterations at the downstream AUG, the authors proposed a 'backward scanning' model, whereby scanning ribosomes

engage in bidirectional oscillations that cover a distance of roughly 15 nt. The downstream initiation codon within the oscillation range is recognized by the fluctuating ribosomes that otherwise scan back to the upstream initiation codon, imposing a competition on ribosome selection. This leads to an inhibitory effect on the upstream AUG, which would be diminished when the downstream initiation codon is beyond the 15-nt oscillation range. As with PEMV2 sgRNA, the downstream initiation codon in TYMV gRNA has an inhibitory effect on the upstream AUG. In full agreement with the 'backward scanning' model, the inhibitory effect of AUG²⁰⁶ on AUG⁶⁹ decreased with an increase in their inter-codon spacing. On the contrary, suppression from the downstream initiation codon in PEMV2 sgRNA vanished when the distance between the two codons was reduced, while additional spacing did not substantially affect translation from either AUG (Figure 6). Also in contrast with PEMV2 sgRNA translation, initiation at either of the TYMV closely spaced AUGs could be boosted by bringing its sequence context in line with Kozak rules, simultaneously suppressing initiation at the another codon. Thus, even if this oscillation model is correct for TYMV, it fails to explain the peculiarities of PEMV2 sgRNA translation.

PEMV2 sgRNA and gRNA share similar strategies to fine tune translation efficiency

We recently showed that the kl-TSS/PTE do not enhance gRNA translation by increasing the translation initiation rate or the number of ribosomes translating the RNA, but rather they affect the number of templates being translated (33). Furthermore, reporter templates containing the 5' and 3' regions of the gRNA were mainly associated with monosomes and disomes. The PEMV2-derived templates were also able to efficiently outcompete a capped mRNA and a mRNA containing the TMV Ω 5' translation enhancer in translation, even though both control templates were associated with heavy ribosomes. Here we demonstrate that, as with the gRNA, sgRNA reporter templates were mainly associated with monosomes and disomes, with the LDI only enhancing the number of templates translated (Figure 7). The light density of ribosomes on the gRNA and sgRNA templates may be beneficial by allowing more time for refolding of critical RNA elements, such as the hairpins involved in ribosome recoding on the gRNA that are melted down each time by the elongating ribosomes (33). Therefore, despite different modes of translation initiation, gRNA and sgRNA share similar strategies that allow for more RNA templates to be translated without boosting the number of ribosomes on the template, which might also interfere with regeneration of the LDI that is melted by elongating ribosomes.

Model for translation of AUG²⁶ and AUG²⁷ in the PEMV2 sgRNA

A model for how the 43S (PIC recognizes AUG²⁶ and AUG²⁷ is presented in Figure 8. The model assumes that 40S ribosomal subunits are bound to sgRNA (most likely

as a part of PIC) through interaction with kl-TSS and PTE-bound initiation factor eIF4E (eIF4F). The long-range kl-TSS/sgH1 kissing loop interaction stimulates involvement of the RNA-bound 40S in the translation process (Figure 7B).

We propose that initiation starts with recognition of the sgRNA 5' end by the 3'CITE-bound PIC (Figure 8A). An accessible 5' terminus was shown to be a key factor since a stable hairpin placed at the 5' end of the sgRNA virtually blocked translation (Figure 3C). Next, the 40S subunit reads the 5' UTR sequence, most likely similarly to conventional scanning since initiating ribosomes can be intercepted by the insertion of upstream AUGs in strong Kozak contexts. At the same time, initiating 40S efficiently miss inserted upstream AUGs in a weak context in favor of AUG²⁶ and AUG²⁷ (Figure 2E and G) suggesting that in this case selection of the initiation codon can be more targeted.

We propose that directional scanning proceeds until the PIC reaches a certain position determined by sterical tensions and restrictions caused by 3'CITE-mediated interaction of the complex with hairpin sgH1 located only 29-nt downstream of AUG²⁷ (Figure 8B). Presumably, this position favors initiation at AUG²⁷, while local ribosome wandering leads to some level of initiation at AUG²⁶ since PIC initial association with a start codon is considered to be metastable and transient (52), allowing reversible sampling of different AUGs until binding is stabilized and GTP hydrolysis/P_i release occurs (53). Alterations that enhance the context of AUG²⁶ (Figure 4 and 5) or increase the distance to the 5' end (Figure 2 and 4) can stabilize binding of the PIC to AUG²⁶, thus reducing translation from AUG²⁷. Alternatively, improvement of AUG²⁶ context could lead to more efficient interception of scanning ribosomes before they even reach AUG²⁷. Minor effects of the alterations enhancing AUG²⁷ context can be explained assuming that its initially advantageous position masks further improvements. According to the model, reduction of the distance between the two codons diminishes differences in their initiation potentials by shifting AUG²⁶ closer to an optimal position. This improves initiation at AUG²⁶ at the expense of AUG²⁷ (Figure 6).

After completing the initiation process, newly formed elongating 80S ribosome will promptly melt hairpin sgH1, separating the sgRNA ends and delaying the next act of initiation (Figure 8C). Previously, we suggested that such delays in translation are necessary for the reduction in ribosome density along the viral gRNA to support sufficient levels of ribosomal frameshifting necessary for RdRp synthesis (37). Since the sgRNA shares a 3'CITE-dependent initiation mechanism like the gRNA, it could have inherited this ribosome dilution mechanism. Alternatively, there may be a need to recharge the 3'CITE's with ribosomal subunits and initiation factors that are hindered if the sgRNA ends are tightly bound. Once the 80S ribosome passes sgH1, the hairpin would be restored allowing formation of the inter-termini bridge and start of the next initiation cycle (Figure 8D).

The most important assumption made by the model is that the selection of a specific start codon by the scanning

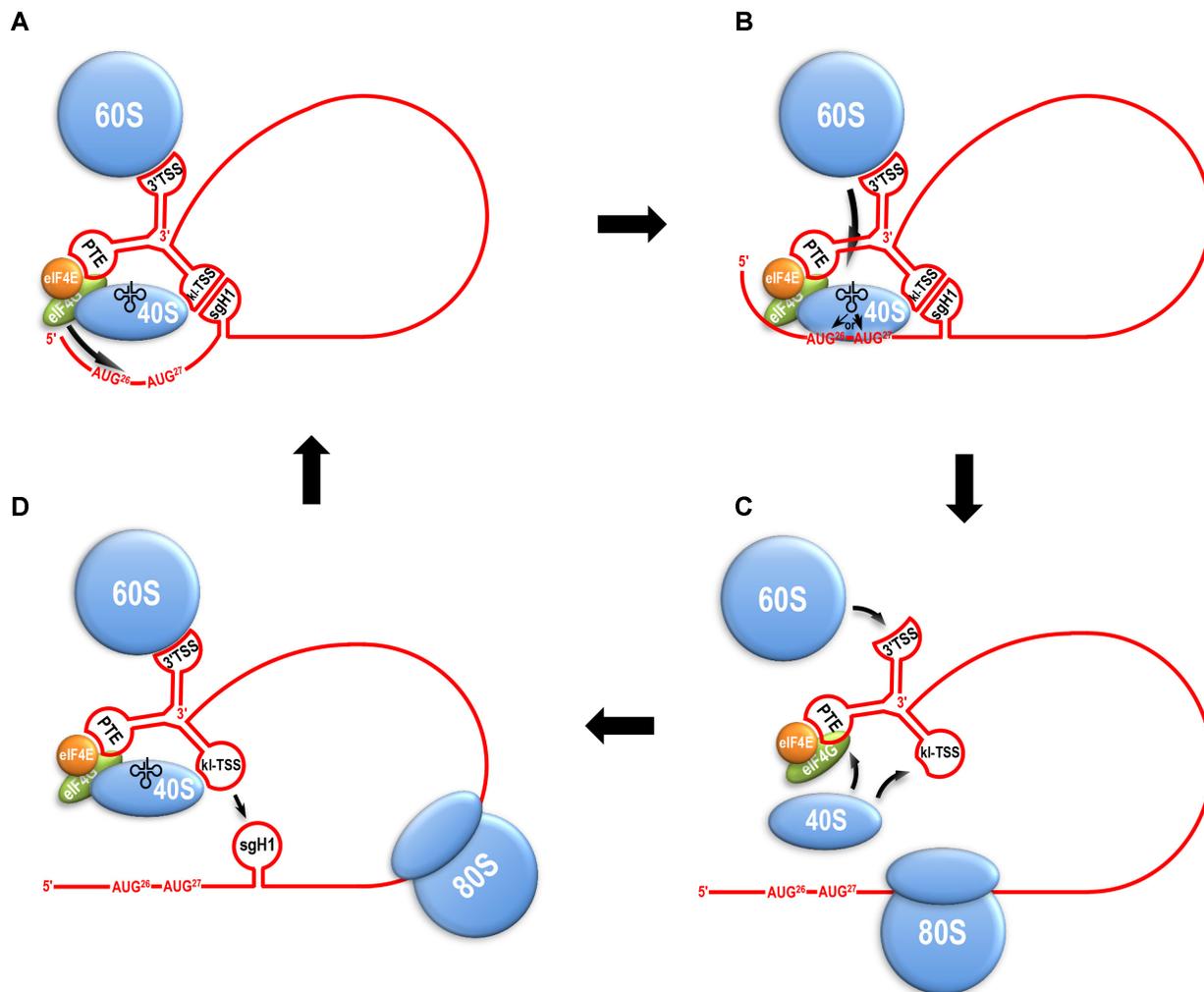


Figure 8. Model for translation of PEMV2 sgRNA. (A) Long-distance RNA–RNA interaction provides the delivery of 3′CITE-bound ribosomal subunits and initiation factors to 5′-termini of the sgRNA to engage the pre-formed initiation complex in recognition of the RNA 5′-end and further 5′ UTR scanning. (B) kl-TSS/sgH1 kissing-loop interaction allocates scanning 40S subunit at certain sterically favorable position along the RNA chain and engages it in transient sampling of nearby AUG₂₆ and AUG₂₇, with some preference for the latter. Following stable closed complex formation at one of the AUGs, the 40S subunit is joined by TSS-bound 60S. (C) Newly formed 80S ribosome starts translation of a reading frame and promptly melts the sgH1 stem-loop, tearing sgRNA ends apart and suspending the initiation process. Meanwhile, 3′CITEs are reloaded with new ribosomal subunits. (D) After the 80S has passed, sgH1 structure is restored again allowing inter-termini interaction to promote a new initiation cycle.

ribosome does not depend solely on the contexts and relative positions of the codons but can be achieved by direct coercive positioning of the initiation complex at certain location on the RNA chain. While this model helps to explain most of the experimental results obtained so far, it also leaves many unanswered questions that are currently being investigated.

FUNDING

National Science Foundation (NSF) [MCB-1411836 to A.E.S.]; National Institutes of Health [R21AI117882–01 to A.E.S.]; Russian Science Foundation [18–14–00368 to K.S.V.]; Program ‘Molecular and Cell Biology’ RAS and RFBR [18–29–0804MK to O.M.A.]; NIH Institutional Training Grant [T32AI125186A to F.G.]. Funding for open access charge: NSF MCB-1411836.

Conflict of interest statement. None declared.

REFERENCES

1. Jackson, R.J., Hellen, C.U. and Pestova, T.V. (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.*, **11**, 113–127.
2. Kozak, M. (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene*, **234**, 187–208.
3. Gualerzi, C.O. and Pon, C.L. (1990) Initiation of mRNA translation in prokaryotes. *Biochemistry*, **29**, 5881–5889.
4. Sonenberg, N. and Hinnebusch, A.G. (2009) Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell*, **136**, 731–745.
5. Kozak, M. (1989) The scanning model for translation: an update. *J. Cell Biol.*, **108**, 229–241.
6. Kozak, M. (1978) How do eukaryotic ribosomes select initiation regions in Messenger-RNA. *Cell*, **15**, 1109–1123.
7. Archer, S.K., Shirokikh, N.E., Beilharz, T.H. and Preiss, T. (2016) Dynamics of ribosome scanning and recycling revealed by translation complex profiling. *Nature*, **535**, 570–574.
8. Shatsky, I.N., Dmitriev, S.E., Andreev, D.E. and Terenin, I.M. (2014) Transcriptome-wide studies uncover the diversity of modes of mRNA

- recruitment to eukaryotic ribosomes. *Crit. Rev. Biochem. Mol. Biol.*, **49**, 164–177.
9. Martinez-Salas, E., Pineiro, D. and Fernandez, N. (2012) Alternative mechanisms to initiate translation in eukaryotic mRNAs. *Comp. Funct. Genom.*, **2012**, 391546.
 10. Balvay, L., Soto Rifo, R., Ricci, E.P., Decimo, D. and Ohlmann, T. (2009) Structural and functional diversity of viral IRESes. *Biochim. Biophys. Acta*, **1789**, 542–557.
 11. Kieft, J.S. (2008) Viral IRES RNA structures and ribosome interactions. *Trends Biochem. Sci.*, **33**, 274–283.
 12. Akulich, K.A., Andreev, D.E., Terenin, I.M., Smirnova, V.V., Anisimova, A.S., Makeeva, D.S., Arkhipova, V.I., Stolboushkina, E.A., Garber, M.B., Prokofjeva, M.M. et al. (2016) Four translation initiation pathways employed by the leaderless mRNA in eukaryotes. *Sci. Rep.*, **6**, 37905.
 13. Benelli, D. and Londei, P. (2011) Translation initiation in Archaea: conserved and domain-specific features. *Biochem. Soc. Trans.*, **39**, 89–93.
 14. Andreev, D.E., Terenin, I.M., Dunaevsky, Y.E., Dmitriev, S.E. and Shatsky, I.N. (2006) A leaderless mRNA can bind to mammalian 80S ribosomes and direct polypeptide synthesis in the absence of translation initiation factors. *Mol. Cell. Biol.*, **26**, 3164–3169.
 15. Futterer, J., Kiss-Laszlo, Z. and Hohn, T. (1993) Nonlinear ribosome migration on cauliflower mosaic virus 35S RNA. *Cell*, **73**, 789–802.
 16. Cao, F. and Tavis, J.E. (2011) RNA elements directing translation of the duck hepatitis B virus polymerase via ribosomal shunting. *J. Virol.*, **85**, 6343–6352.
 17. Racine, T. and Duncan, R. (2010) Facilitated leaky scanning and atypical ribosome shunting direct downstream translation initiation on the tricistronic S1 mRNA of avian reovirus. *Nucleic Acids Res.*, **38**, 7260–7272.
 18. Schepetilnikov, M., Schott, G., Katsarou, K., Thiebauld, O., Keller, M. and Ryabova, L.A. (2009) Molecular dissection of the prototype foamy virus (PFV) RNA 5'-UTR identifies essential elements of a ribosomal shunt. *Nucleic Acids Res.*, **37**, 5838–5847.
 19. de Breyne, S., Monney, R.S. and Curran, J. (2004) Proteolytic processing and translation initiation: two independent mechanisms for the expression of the Sendai virus Y proteins. *J. Biol. Chem.*, **279**, 16571–16580.
 20. Paek, K.Y., Park, S.M., Hong, K.Y. and Jang, S.K. (2012) Cap-dependent translation without base-by-base scanning of a messenger ribonucleic acid. *Nucleic Acids Res.*, **40**, 7541–7551.
 21. Paek, K.Y., Hong, K.Y., Ryu, I., Park, S.M., Keum, S.J., Kwon, O.S. and Jang, S.K. (2015) Translation initiation mediated by RNA looping. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 1041–1046.
 22. Pisarev, A.V., Kolupaeva, V.G., Pisareva, V.P., Merrick, W.C., Hellen, C.U. and Pestova, T.V. (2006) Specific functional interactions of nucleotides at key -3 and +4 positions flanking the initiation codon with components of the mammalian 48S translation initiation complex. *Genes Dev.*, **20**, 624–636.
 23. Joshi, C.P., Zhou, H., Huang, X. and Chiang, V.L. (1997) Context sequences of translation initiation codon in plants. *Plant Mol. Biol.*, **35**, 993–1001.
 24. Lutcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F. and Scheele, G.A. (1987) Selection of AUG initiation codons differs in plants and animals. *EMBO J.*, **6**, 43–48.
 25. Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, **44**, 283–292.
 26. Kozak, M. (1989) Context effects and inefficient initiation at non-AUG codons in eukaryotic cell-free translation systems. *Mol. Cell. Biol.*, **9**, 5073–5080.
 27. Slusher, L.B., Gillman, E.C., Martin, N.C. and Hopper, A.K. (1991) mRNA leader length and initiation codon context determine alternative AUG selection for the yeast gene MOD5. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 9789–9793.
 28. Kozak, M. (1991) A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes. *Gene Expr.*, **1**, 111–115.
 29. Kozak, M. (1995) Adherence to the first-AUG rule when a second AUG codon follows closely upon the first. *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 2662–2666.
 30. Matsuda, D. and Dreher, T.W. (2006) Close spacing of AUG initiation codons confers dicistronic character on a eukaryotic mRNA. *RNA*, **12**, 1338–1349.
 31. Williams, M.A. and Lamb, R.A. (1989) Effect of mutations and deletions in a bicistronic mRNA on the synthesis of influenza B virus NB and NA glycoproteins. *J. Virol.*, **63**, 28–35.
 32. Gao, F. and Simon, A.E. (2016) Multiple cis-acting elements modulate programmed -1 ribosomal frameshifting in Pea enation mosaic virus. *Nucleic Acids Res.*, **44**, 878–895.
 33. Du, Z., Alekhina, O.M., Vassilenko, K.S. and Simon, A.E. (2017) Concerted action of two 3' cap-independent translation enhancers increases the competitive strength of translated viral genomes. *Nucleic Acids Res.*, **45**, 9558–9572.
 34. Gao, F., Kasprzak, W., Stupina, V.A., Shapiro, B.A. and Simon, A.E. (2012) A ribosome-binding, 3' translational enhancer has a T-shaped structure and engages in a long-distance RNA-RNA interaction. *J. Virol.*, **86**, 9828–9842.
 35. Wang, Z., Treder, K. and Miller, W.A. (2009) Structure of a viral cap-independent translation element that functions via high affinity binding to the eIF4E subunit of eIF4F. *J. Biol. Chem.*, **284**, 14189–14202.
 36. Taliansky, M., Roberts, I.M., Kalinina, N., Ryabov, E.V., Raj, S.K., Robinson, D.J. and Oparka, K.J. (2003) An umbraviral protein, involved in long-distance RNA movement, binds viral RNA and forms unique, protective ribonucleoprotein complexes. *J. Virol.*, **77**, 3031–3040.
 37. Gao, F. and Simon, A.E. (2017) Differential use of 3'CITEs by the subgenomic RNA of Pea enation mosaic virus 2. *Virology*, **510**, 194–204.
 38. Ryabov, E.V., Roberts, I.M., Palukaitis, P. and Taliansky, M. (1999) Host-specific cell-to-cell and long-distance movements of cucumber mosaic virus are facilitated by the movement protein of groundnut rosette virus. *Virology*, **260**, 98–108.
 39. Ryabov, E.V., Oparka, K.J., Santa Cruz, S., Robinson, D.J. and Taliansky, M.E. (1998) Intracellular location of two groundnut rosette umbravirus proteins delivered by PVX and TMV vectors. *Virology*, **242**, 303–313.
 40. Liu, H. and Naismith, J.H. (2008) An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.*, **8**, 91.
 41. Gao, F., Gulay, S.P., Kasprzak, W., Dinman, J.D., Shapiro, B.A. and Simon, A.E. (2013) The kissing-loop T-shaped structure translational enhancer of Pea enation mosaic virus can bind simultaneously to ribosomes and a 5' proximal hairpin. *J. Virol.*, **87**, 11987–12002.
 42. Gao, F., Kasprzak, W.K., Szarko, C., Shapiro, B.A. and Simon, A.E. (2014) The 3' untranslated region of Pea Enation Mosaic Virus contains two T-shaped, ribosome-binding, cap-independent translation enhancers. *J. Virol.*, **88**, 11696–11712.
 43. Sharma, S.D., Kraft, J.J., Miller, W.A. and Goss, D.J. (2015) Recruitment of the 40S ribosome subunit to the 3'-untranslated region (UTR) of a viral mRNA, via the eIF4 complex, facilitates cap-independent translation. *J. Biol. Chem.*, **290**, 11268–11281.
 44. Guan, H., Carpenter, C.D. and Simon, A.E. (2000) Requirement of a 5'-proximal linear sequence on minus strands for plus-strand synthesis of a satellite RNA associated with turnip crinkle virus. *Virology*, **268**, 355–363.
 45. Guan, H., Carpenter, C.D. and Simon, A.E. (2000) Analysis of cis-acting sequences involved in plus-strand synthesis of a turnip crinkle virus-associated satellite RNA identifies a new carmovirus replication element. *Virology*, **268**, 345–354.
 46. Kim, S.H., Macfarlane, S., Kalinina, N.O., Rakitina, D.V., Ryabov, E.V., Gillespie, T., Haupt, S., Brown, J.W. and Taliansky, M. (2007) Interaction of a plant virus-encoded protein with the major nucleolar protein fibrillarin is required for systemic virus infection. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 11115–11120.
 47. Ryabov, E.V., Robinson, D.J. and Taliansky, M. (2001) Umbravirus-encoded proteins both stabilize heterologous viral RNA and mediate its systemic movement in some plant species. *Virology*, **288**, 391–400.
 48. Kozak, M. (1989) Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. *Mol. Cell. Biol.*, **9**, 5134–5142.
 49. Sagliocco, F.A., Vega Laso, M.R., Zhu, D., Tuite, M.F., McCarthy, J.E. and Brown, A.J. (1993) The influence of 5'-secondary structures upon ribosome binding to mRNA during translation in yeast. *J. Biol. Chem.*, **268**, 26522–26530.

50. Chappell, S.A., Edelman, G.M. and Mauro, V.P. (2006) Ribosomal tethering and clustering as mechanisms for translation initiation. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 18077–18082.
51. Rangan, L., Vogel, C. and Srivastava, A. (2008) Analysis of context sequence surrounding translation initiation site from complete genome of model plants. *Mol. Biotechnol.*, **39**, 207–213.
52. Terenin, I.M., Akulich, K.A., Andreev, D.E., Polyanskaya, S.A., Shatsky, I.N. and Dmitriev, S.E. (2016) Sliding of a 43S ribosomal complex from the recognized AUG codon triggered by a delay in eIF2-bound GTP hydrolysis. *Nucleic Acids Res.*, **44**, 1882–1893.
53. Hinnebusch, A.G. (2014) The scanning mechanism of eukaryotic translation initiation. *Annu. Rev. Biochem.*, **83**, 779–812.