Identification of Novel 5′ and 3′ Translation Enhancers in Umbravirus-like Coat-Protein-Deficient RNA Replicons

Jingyuan Liu and Anne E. Simon*

Department of Cell Biology and Molecular Genetics

University of Maryland College Park

College Park, MD 20742

*Corresponding Author

Phone: 301-405-8975
Email: simona@umd.edu

Running Title: New type of 3′CITE

Keywords: RNA structure; 3′CITEs; Non-canonical translation; translation enhancer; eIF4G-binding structure

Abstract

Translation of plant plus-strand RNA viral genomes that lack a 5′ cap frequently requires the use of cap-independent translation enhancers (CITEs) located in or near the 3′ UTR. 3′CITEs are grouped based on secondary structure and ability to interact with different translation initiation factors or ribosomal subunits, which assemble a complex at the 3′ end that is nearly always transferred to the 5′ end via a long distance kissing-loop interaction between sequences in the 3′CITE and 5′ hairpins. We report here the identification of a novel 3′CITE in coat-protein-deficient RNA replicons that are related to umbraviruses. Umbra-like associated RNAs (ulaRNAs), such as citrus yellow vein-associated virus (CYVaV), are a new type of subviral RNA that do not encode movement proteins, coat-proteins, or silencing suppressors, but can independently replicate using their encoded RNA-dependent RNA polymerase. An extended hairpin structure containing multiple internal loops in the 3′ UTR of CYVaV is strongly conserved in the most closely related ulaRNAs and structurally resembles an I-shaped structure (ISS) 3′CITE. However, unlike ISS, the CYVaV structure binds to eIF4G and no long-distance interaction is discernible between the CYVaV ISS-like structure and sequences at or near the 5′ end. We also report that the ~30 nt 5′ terminal hairpin of CYVaV and related ulaRNAs can enhance translation of reporter constructs when associated with either the CYVaV 3′CITE, the 3′CITEs of umbravirus PEMV2, and even independent of a 3′CITE. These findings introduce a new type of 3′CITE and provide the first information on translation of ulaRNAs.

Importance

Umbra-like associated RNAs (ulaRNAs) are a recently discovered type of subviral RNA that use their encoded RNA-dependent RNA polymerase for replication but do not encode any coat proteins, movement proteins or silencing suppressors yet can be found in plants in the absence of any discernable helper virus. We report the first analysis of their translation using Class 2 ulaRNA citrus yellow vein associated virus (CYVaV). CYVaV uses a novel eIF4G-binding I-shaped structure as its 3′ cap-independent translation enhancer (3′CITE), which does not connect with the 5′ end by a long-distance RNA:RNA interaction that is typical of 3′CITEs. ulaRNA 5′ terminal hairpins can also enhance translation in association with cognate 3′CITEs or those of related ulaRNAs, and to a lesser extent with 3′CITEs of umbraviruses, or even independent of a 3′CITE. These findings introduce a new type of 3′CITE and provide the first information on translation of ulaRNAs.

Introduction

Translation initiation in eukaryotes is a precisely controlled, conserved process that requires the coordinated action of over 10 eukaryotic translation initiation factors (eIFs) (1). Canonical cap-dependent translation initiation begins with eIF4 recognition of the $m^7G(5')ppp(5')N$ cap structure at the 5′ end of a mRNA, followed by scaffold protein eIF4G binding to eIF4E to form the eIF4F complex (2). eIF4G bridges the mRNA 5′ and 3′ ends by also interacting with poly(A)-binding protein that is bound to the 3′ poly(A) tail, which is followed by recruitment of the 43S preinitiation complex (PIC) (40S ribosomal subunit loaded with the eIF2-$tRNA_i^{met}$-GTP complex and other eIFs) to the mRNA 5′ end (3). Once loaded, the PIC scans along the mRNA until recognition of a downstream start codon in a favorable initiation context. At this point, the selected start codon base pairs with the $tRNA_i^{met}$ anticodon followed by dissociation of eIFs and joining of the 60S ribosomal subunit to form the 80S initiation complex (4). The elongation phase then begins with entry of the correct aminoacyl-tRNA into the ribosome A-site.

Many RNA plus-strand viral genomic (g)RNAs have no 5′ cap or 3′ poly(A) tail and thus must attract the ribosomal machinery through non-canonical mechanisms that use cis-acting elements such as 5′ bound VPg or internal structured RNA elements like internal ribosome entry sites. For plant viruses in this category, ribosome attraction commonly depends on cap-independent translation enhancers (3′CITEs) located in the 3′ region of the genome that facilitate highly efficient translation initiation at the 5′ end (5-7). This occurs when 3′CITEs bind to eIFs that subsequently recruit ribosomal subunits (8) or bind directly to ribosomal subunits (9, 10), which is usually followed by a long-distance interaction (LDI) between a CITE-associated hairpin and 5′ end sequences to transfer bound components to the 5′ end for translation initiation

(5, 10-15).

3′CITEs have been identified in several plant virus families and are generally grouped according to their secondary structure, presence of conserved nucleotides, and interaction with specific translation components (5, 7, 16). For example, the CITE known as the "I-shaped structure" (ISS) has been identified in viruses from four genera in the *Tombusviridae* (zeavirus, aureusvirus, tombusvirus, and carmovirus) (5, 7, 16, 17). ISS are composed of a single stem-loop with one or two internal loops and binds efficiently to eIF4F through a direct interaction with the eIF4E subunit (18). Nearly all ISS share 16 conserved residues around the internal loop region, and some of these bases are required for binding to eIF4F (18). Residues in the ISS 5 to 8 nt apical loop engage in a long-distance kissing-loop interaction with 5′ proximal hairpins in the gRNA and subgenomic (sg)RNA, assisted by eIF4F (19). Subsequent binding of the 40S ribosomal subunit to the ISS leads to a transfer of the complex to the 5′ proximal initiation site where translation begins (19).

Coat protein-deficient RNA replicons (CdRr) are viruses or virus-like RNAs that can replicate autonomously using their encoded RdRp but are dependent on a helper virus for at least encapsidation and vector transmission. CdRr include members of the umbravirus genus within the *Tombusviridae*, which encode two replication-required proteins (ORF1 and ORF2) and two movement proteins (ORF3 and ORF 4). Umbraviruses lack encoded capsid proteins or RNA silencing suppressors, which must be supplied by a helper virus that is usually a virus from the polerovirus or enamovirus genera (20, 21). Umbraviruses are unusual in containing multiple 3′CITEs, most commonly a barley yellow dwarf-like translation enhancer (BTE) in the central region of their long (~700 nt) 3′ UTR that binds to eIF4G (12, 22-24) and a 3′TSS located near the 3′ end that binds to 60S ribosomal subunits and 80S ribosomes (10, 25). BTE participate in

long-distance kissing-loop interactions with 5′ proximal hairpins in the gRNA and sgRNA and similar connections between 3′TSS and viral RNA 5′ ends are absent. The umbravirus PEMV2 contains three 3′CITEs, two of which are required for efficient translation of the gRNA, and all three are needed for efficient translation of the sgRNA (26). The two used by both gRNA and sgRNA are: (*i*) the kissing-loop T-shaped structure (kl-TSS), which binds to 80S ribosomes and ribosomal subunits and participates in LDIs with 5′ proximal hairpins in the gRNA and sgRNA (10, 27); and (*ii*) the panicum mosaic virus-like translation enhancer (PTE), whose internal pseudoknot is postulated to form an eIF4E-binding pocket (10, 28, 29). The third 3′CITE, the 3′TSS, is exclusively used for sgRNA translation (26).

Members of a recently recognized type of CdRr known as umbra-like associated (ula)RNAs have been suggested to be progenitors of umbraviruses based on: (i) related RdRps that contain an umbravirus-specific motif (A.E.Simon, unpublished); (ii) RdRp that are synthesized by-1 programmed ribosomal frameshifting (-1PRF) just upstream of the stop codon for the 5′ proximal replication-required protein (30, 31); and (iii) similar 5′ and 3′ terminal structures. ulaRNAs range in size from 2.7 to 4.6 kb, and, like umbraviruses, do not encode any capsid proteins or silencing suppressors [(30); Simon, Liu, and Gao, unpublished] (Fig. 1A). However, in contrast with umbraviruses, most if not all ulaRNAs do not encode any known movement proteins, and thus additionally do not require the umbravirus ORF3 long-distance movement protein that also suppresses nonsense mediated decay of viral and cellular RNAs (32). Furthermore, many ulaRNAs are found in symptomatic or asymptomatic plants in the absence of any definable helper virus (33-37), unlike umbraviruses, which are always associated with a helper virus in nature.

ulaRNAs have been divided into three classes based on phylogenetic relationships and

other features (30): Class 1 ulaRNAs only encode the two umbravirus-like replication-required proteins and have an expansive 1.9 kb 3′ UTR (35, 36, 38, 39); Class 2 ulaRNAs are the shortest (2.7-3.1 kb), and all except citrus yellow vein associated virus (CYVaV) have an additional ORF (ORF5) that overlaps with the 3′ end of the RdRp ORF (30, 33, 34, 37, 40); the single member of the Class 3 ulaRNAs also has an additional ORF in a similar but not identical location, and its origin differs from the one found in Class 2 ulaRNAs.

We report here the identification of a 3′CITE in CYVaV that is strongly conserved in other Class 2 ulaRNAs and structurally resembles an ISS. However, unlike ISS, the 3′CITE binds to eIF4G and no LDI is discernible between the ISS-like structure and sequences at or near the 5′ end. We also report that the ~30 nt 5′ terminal hairpin of CYVaV and related Class 2 ulaRNAs opuntia umbra-like virus (OULV) and fig umbra-like virus (FULV) enhance translation of reporter constructs when associated with either the CYVaV 3′CITE, the 3′CITEs of PEMV2, and even when independent of a 3′CITE. These findings thus introduce a new type of 3′CITE and provide the first information on translation regulation of ulaRNAs.

Results

Conserved sequences and structures in CYVaV Domain 3

The full-length structure of CYVaV gRNA was previously determined using SHAPE (selective 2′-hydroxyl acylation analyzed by primer extension) structure probing, phylogenetic comparisons, and assistance of computational algorithms and the RNA structure drawing program RNA2Drawer (30, 41, 42). As shown in Figure 1B, the genome-length structure subdivides into three domains, with Domain 1 (D1) containing the region from the 5′ end to just past the ribosome recoding site; D2 encompassing an extended central portion that includes a

142 portion of the 3′ UTR and is flanked by long-distance base-paired bridging sequences (i.e., “the
143 bridge”); and D3, which includes only 3′ UTR sequences. In Class 2 ulaRNAs that contain
144 ORF5, D3 constitutes the vast majority of 3′ UTR sequence. Sequence and structure alignments
145 of ulaRNAs revealed that the absence of ORF5 in CYVaV was due to deletion of two extended
146 hairpins as well as the loss of the initiation codon and presence of multiple stop codons within
147 the analogous reading frame (30). Recent GenBank submissions of two Class 2 ulaRNA
148 sequences from cannabis share ~90% sequence identity with CYVaV and still contain ORF5
149 (MT893740 and MT893741), suggesting that CYVaV was derived from a ulaRNA that also
150 contained the additional ORF, with D3 then encompassing the majority of the 3′UTR.

151 Based on the location of 3′CITEs in umbravirus 3′UTRs, it was likely that translation
152 elements in CYVaV were also located in D3. Examination of the RNA structures in CYVaV D3,
153 and comparing them with structures previously identified for umbravirus PEMV2, revealed
154 similar hairpins only near the 3′ ends (Fig. 1C). PEMV2 and six additional umbraviruses contain
155 hairpins H4a, H4b, and H5, along with two pseudoknots (ψ_2 and ψ_3) that altogether fold into the
156 3′TSS 3′CITE (12, 25). In addition, all umbraviruses, except for groundnut rosette virus and
157 ixeridium yellow mottle virus 2, contain 3′ terminal hairpin Pr and pseudoknot ψ_1, which forms
158 between the penultimate H5 hairpin apical loop and the 3′ terminal four bases downstream of Pr
159 (Fig. 1C). CYVaV D3 contains hairpins in identical locations as H4a, H4b, H5, and Pr in
160 addition to pseudoknot ψ_1. ψ_2 and ψ_3 are not discernable, suggesting that a TSS 3′CITE does not
161 form in this location.

162 At the 5′ end of D3 are three small hairpins (H_a, H_b, and H_c) followed by an extended
163 unbranched hairpin originally designated as Structure (S)14, which contains a long stem-loop
164 with multiple internal loops (Fig. 1C) (30). Most D3 structures are well-conserved in Class 2

ulaRNAs despite poor sequence conservation. Noted exceptions are the absence of H4a in OULV and an insert in FULV2 that folds into two hairpins just downstream of S14 (33). As with CYVaV, the other Class 2 ulaRNAs have ψ_1, but lack discernible ψ_2 and ψ_3 (Fig. 1C).

Mapping regions important for translation in the CYVaV 3′ UTR

To aid in determining locations of translation element(s) in CYVaV, stepwise deletions were generated from both ends of the complete 3′ UTR in a construct containing full-length CYVaV cDNA (Fig. 2), and *in vitro* synthesized transcripts were subjected to translation in wheat germ extracts (WGE). As shown for translation of full-length CYVaV, the -1PRF event is unusually efficient, with a recoding rate of nearly 30% (Fig. 2B lane 1) (30). This implies that any deletions or other alterations that primarily affect -1PRF could result in additional ribosomes terminating at the p21 stop codon, thereby increasing levels of p21. If levels of p21 remain constant or decrease when -1PRF levels are substantially reduced, this suggests that the alteration may also be affecting translation from the 5′ end.

When the entire 3′ UTR was deleted (O1, Δ2162-2692), translation of p21 decreased to 33% of full-length CYVaV and the -1PRF product p81was not detectable (Fig. 2). This suggests that the CYVaV 3′ UTR contains elements that facilitate both translation initiation and -1PRF. Deletion of just Structures 12 (S12) and 13 (S13) at the 5′ end of the 3′ UTR (O2, Δ2162-2294 [see Fig. 1B for location of the structures]) decreased p21 levels to 45% of full length and p81 levels to 14% of full-length. While suggestive that a translation element is impacted by this deletion, multiple small hairpins and unstructured sequences subsequently inserted into this region significantly affected RNA structure throughout CYVaV (but especially proximal to the 3′

187 end), and caused comparable reductions in translation of p21 and p81 (E.Carino and A.E.Simon,
188 unpublished). This suggests that structural disruptions distal to the deletion and not necessarily
189 structures directly impacted by the deletion could be negatively affecting translation. Deletion
190 O3 extended into the bridge sequence (Δ2162-2374) and still allowed for 5% of full-length p81
191 levels, while deletions further downstream into the beginning of D3 and beyond (O4-O6)
192 eliminated detectable p81. This suggests the presence of a possible -1PRF element at the
193 beginning of D3, which is proximal to the upstream frameshifting site due to the long-distance
194 bridge. This element will be the subject of another report (E. Carino, J. Liu, F. Gao and A.E.
195 Simon, manuscript in preparation).

196 Deletions were also generated from the CYVaV 3′ end towards the 5′ end of the 3′ UTR.
197 Deleting the 3′ terminal region (CSacI, Δ2600-2692 and CD8, Δ2578-2692) eliminated p81
198 synthesis while increasing p21 levels by 2.5 to 3.5-fold (Fig. 2B). Loss of -1PRF was likely
199 caused by deletion of the sequence between H5 and Pr that is complementary to the apical loop
200 of the recoding stimulatory element (RSE), as a LDI between the RSE and 3′ end has been
201 identified for many members of the *Tombusviridae* (31,43-45). When the deletion was extended
202 into S14 (CD7, Δ2537-2692), p81 translation remained undetectable and p21 levels were reduced
203 by 8.6-fold. Extending the deletions further upstream (CD2, CNheI, CD4, CD5, CD6) yielded
204 similar results as CD7. These results suggested that the 3′ border of a translation element might
205 be between positions 2537 and 2577, which correlates with the 3′ end of S14.

206

207 **CYVaV S14 is important for translation of p21 and p81 in vitro**

208 Since S14 was a possible 3′CITE, a sequence alignment for eight Class 2 ulaRNAs
209 (CYVaV, CYVaV-RioBlanco, CYVaV-Delta, FULV1, FULV2, OULV, Ethiopian maize

associated virus 1 [EMaV1] and 2 [EMaV2]) was generated using ClustalW, and bases that were invariant or found in all but one of the ulaRNAs are shown in Figure 3A (dark green and light green residues, respectively). All Class 2 S14 consist of a long stem-loop containing (from top to bottom) (i) an apical hairpin region with an asymmetrical internal loop consisting of mainly variable bases; (ii) a C-rich internal loop; (iii) a six base-pair central stem with covariations in the upper portion; (iv) a highly conserved region just below the central stem containing two asymmetrical internal loops (B1 and B2); and (v) a lower stem with mainly variable base-pairs.

To investigate the importance of specific sequences and features within CYVaV S14, site-directed mutations were engineered throughout the element (Fig. 3B) and WT and mutant CYVaV transcripts were subjected to *in vitro* translation using WGE. Alteration of residues on both sides of the central stem (m1 and m2, positions 2486-2490 and 2526-2530) reduced p21 translation by 76% and 89%, respectively, and virtually eliminated p81. Combining the two sets of mutations (m1m2), which were designed to be compensatory, restored p21 and p81 levels to 100% and 88% of WT, respectively. These results support S14 as being critical for translation with the central stem as a necessary feature. Interestingly, the reduction in translation caused by m1 and m2 was comparable or greater than translation reductions caused by deletions that eliminated S14 (see O5, O6 and CNheI, Fig. 2B). This suggests that destabilizing the S14 central stem while maintaining the rest of the S14 sequence has an enhanced repressive effect on translation. This is similar to results obtained for PEMV2, where translation was repressed more when the PTE was maintained but the LDI was disabled, compared with deletions that removed the PTE and the LDI (26).

Mutations located in the C-rich internal loop adjacent to the central stem that extended the central stem by two C:G base-pairs (**CC**2524-2525**GG**) reduced p21 and p81 levels by 59%

and 55%, respectively. However, mutations on both sides of the C-rich loop that did not lead to additional base pairing enhanced translation of both p21 and p81 by 45% and 34%, respectively, for **C**2492**G**/**C**2496**G**, and 54% and 84%, respectively, for **C**2524**U**. This suggests that this region of S14 may harbor a translation repressor. **C**2498**A**, located in a C-rich asymmetric loop sequence just above the C-rich internal loop, also increased p81 translation (by 36%). However **G**2518**C**, located on the opposite side of this alteration, decreased p21 and p81 levels by 34% and 35%, respectively, suggesting that this mutation had a negative effect on translation of p21 that led to reduced synthesis of p81. These results suggest that the putative repressor region in the S14 translation element may extend beyond the C-rich loop, but only on the 5′ side of the hairpin that is also C-rich. We also altered resides in the S14 apical loop since most 3′CITEs participate in a LDI with 5′ sequences, and the pairing residues are frequently in the apical loop of a CITE-associated hairpin,. Surprisingly, **GGA**2509-2511**CCU** in the apical loop (and upper stem) and **C**2505**G**/**U**2508**A** in the apical loop did not substantially affect p21 and p81 levels in WGE.

One of the defining features of S14 are the two asymmetric internal loops on opposite sides of the structure (B1 and B2), located just below the central stem. Both asymmetric loops have invariant residues that extend an additional 5 or 6 bases downstream (B1: GAUAGCACUGU; B2:AGAUUUGUGAA). Despite strict sequence conservation, **U**2479**C** within B1 did not notably affect p21 or p81 levels *in vitro*. Mutating the two unpaired adenylates on the opposite side of B1 (**AA**2540-2541**UU**) also did not negatively impact p21 levels while reducing p81 synthesis by 45%. Mutations in B2 just upstream of the adenylates (**AU**2533-2534**UA and U**2536**C)** also reduced p81 synthesis by 31% and 45% respectively without negatively impacting levels of p21. **GG**2472-2473**CC**, located in the lower stem, led to a similar

41% decrease in p81 synthesis with p21 levels remaining essentially unchanged. These results suggest that B2 and the adjacent lower stem may play a role in suppressing -1PRF. Despite the significant conservation of the B1 and B2 segments in Class 2 ulaRNA S14, neither S14 nor these conserved elements are discernable in Class 1 or Class 3 ulaRNAs.

S14 is important for accumulation of CYVaV in vivo

Since CYVaV accumulates efficiently in Arabidopsis thaliana protoplasts in the absence of any helper virus (30), selected S14 mutants were also assayed for accumulation in protoplasts. Mutations m1 and m2 that disrupted either side of the central stem and were detrimental for translation *in vitro* reduced CYVaV accumulation in protoplasts to 18% of WT levels. Mutation m1m2, which re-established the stem, restored both *in vitro* translation and accumulation in protoplasts to near WT levels. All mutations below the central stem, regardless of their effect on *in vitro* translation (if any), reduced CYVaV accumulation to between 16% and 30% of WT levels *in vivo*. Notable was **U**2479**C** within B1, which did not discernibly affect translation *in vitro* yet reduced accumulation *in vivo* to 28% of WT, a level identical to disruption of the central stem. Mutations in the C-rich loop, which either reduced or enhanced *in vitro* translation, were detrimental *in vivo*, reducing accumulation to 10 to 20% of WT. Apical loop mutations **GGA**2509-2511**CCU**, which had no notable effect on p21 or p81 levels *in vitro*, never-the-less reduced CYVaV accumulation in protoplasts to 24% of WT. These results indicate that mutations throughout S14, which had varying effects on translation *in vitro*, were important for efficient accumulation of the ulaRNA *in vivo*. All together, these results strongly suggest that S14 is a translation element essential for efficient accumulation *in vivo* and translation *in vitro*.

278

279 **S14 is a 3′CITE that interacts with eIF4G**

280 When fragments containing 3′CITEs are added *in trans* to WGE, most inhibit translation
281 of viral templates by sequestering limited translation factors (24, 45, 46). To determine if
282 CYVaV S14 can inhibit translation *in trans*, *in vitro* synthesized transcripts containing either the
283 entire 3′ UTR, the complete S14 (fragment ISSLS; 2452-2559) or just the apical portion of S14
284 ($ISSLS_{\Delta B}$; 2484-2532; see Fig. 3B) were added to *in vitro* translation reactions containing full-
285 length WT CYVaV, and levels of p21 synthesized were determined. Addition of the CYVaV 3′
286 UTR fragment at a 10- or 25-fold excess decreased translation by the same 92%, indicating that
287 one or more translation factor(s) present in a limiting amount was likely being sequestered by
288 one or more elements in the 3′ UTR. Addition of 10- or 25-fold excess ISSLS fragment similarly
289 decreased translation by 84% and 96%, respectively. In contrast, addition of 10- or 25-fold
290 ISSLS containing mutation m1 decreased translation by only 34% and 40%, respectively.
291 Similarly, addition of 10- or 25- fold $ISSLS_{\Delta B}$ decreased translation by only 17% and 32%,
292 respectively. These results suggest that the complete S14 inhibits translation *in trans* in WGE
293 with similar efficiency as the complete 3′ UTR, and thus likely comprises a 3′CITE.

294 The ISS 3′CITE, which is structurally similar to S14, binds proficiently to eIF4F through
295 its eIF4E subunit, but binds inefficiently to eIF4E and eIF4G individually (18, 19). To determine
296 which factors bind the CYVaV S14, electrophoretic mobility shift assays (EMSA) combined
297 with UV cross-linking were conducted using radiolabeled ISSLS and $ISSLS_{\Delta B}$ fragments and
298 wheat eIF4F, eIF4G, and eIF4E that were expressed in *E.coli* and purified by affinity
299 chromatography (47). Since the BTE 3′CITE is known to bind to eIF4G, and the PTE 3′CITE
300 binds to eIF4E (28, 29), opium poppy mosaic virus (umbravirus) BTE and PEMV2 PTE

fragments were used as positive controls.

As expected, retarded migration in gels was found for the OPMV BTE fragment bound to eIF4G and eIF4F, and the PEMV2 PTE bound to eIF4E and eIF4F (Fig. 3D). Migration of the ISSLS fragment was retarded only in the presence of eIF4G and eIF4F, with no detectable binding to eIF4E. Truncated ISSLS$_{\Delta B}$ did not bind detectably to any of these proteins, suggesting that the upper portion of S14 by itself does not interact with eIF4G or eIF4F. Purified eIF4G and eIF4F were also added back to the *trans*-inhibition assays to determine if sequestering eIF4G and/or eIF4F was responsible for the inhibitory activity of fragment ISSLS. As shown in Figure 4, supplementing extracts with 200/400 nM eIF4G or eIF4F substantially restored BTE-inhibited translation, consistent with previous findings using the BYDV BTE (24). In contrast, translation that was inhibited 84% with addition of the ISSLS improved only slightly when supplemented with 400 nM eIF4G or eIF4F. The ISSLS$_{\Delta B}$ fragment did not negatively impact translation of WT CYVaV in the presence and absence of translation factors. These results suggest that in addition to binding to eIF4G and eIF4F, the CYVaV ISSLS may sequester additional factors or interact in a negative manner with the CYVaV gRNA. Altogether, these results support the designation of S14 as a unique 3′CITE, which we have named an ISS-like structure or ISSLS.

The CYVaV 5′ terminal S1 hairpin is required for efficient translation of reporter constructs *in vivo*

Most 3′CITEs, including all ISS, require a LDI with a 5′ terminal sequence to enhance translation (5, 7, 16, 17). The 5′ ends of all Class 2 ulaRNAs contain two small hairpins followed by a large extended stem-loop [Structures (S)1, 2 and 3; Fig. 1B and Fig. 5A) (30). To

investigate the importance of these structures in translation enhancement by the ISSLS, firefly luciferase (F-Luc) reporter constructs were generated containing different lengths of CYVaV 5′ sequence upstream of the F-Luc ORF, followed by the full length CYVaV 3′ UTR. Construct C5′33+C3′U contained positions 1 to 33, encompassing S1; construct C5′60+C3′U contained positions 1 to 60, encompassing S1 and S2; and construct C5′201+C3′U contained positions 1 to 201, encompassing S1, S2 and a portion of S3 that included its apical hairpin (Fig. 5B). The PEMV2 5′89 nt with the CYVaV 3′ UTR (P5′89+C3′U) was used as a negative control (Fig. 5A and B).

In vitro synthesized transcripts of the reporter constructs were transfected into *A. thaliana* protoplasts and luciferase activity was measured 18 h after transfection. Translation of C5′33+C3′U was 5.3-fold higher than P5′89+C3′U, suggesting that S1 enhanced translation of the reporter *in vivo*. Extending the length of the 5′ end to include Structure 2 (construct C5′60+C3′U) did not further improve luciferase activity levels. Construct C5′201+C3′U produced 3.8-fold less luciferase activity than P5′89+C3′U , indicating that this extension of 5′ sequence was inhibitory.

To determine if translation enhancement mediated by S1 requires the CYVaV 3′ UTR, the 3′ UTR in construct C5′33+C3′U was replaced with 250 nt of vector-derived sequence (C5′33+V). C5′33+V luciferase activity was 3-fold lower than that of C5′33+C3′U, suggesting that the CYVaV 3′ UTR was contributing to elevated luciferase activity of C5′33+C3′U. Taken together, these results suggest that CYVaV S1 facilitates *in vivo* translation of the reporter gene by interacting directly or indirectly with the CYVaV 3′ UTR. However, it should be noted that luciferase activity of C5′33+C3′U was 23-fold lower than that of P5′89+P3′U, a previously generated construct containing the 5′ 89 nt and 3′ UTR of PEMV2 (10). Since CYVaV and

PEMV2 full-length gRNAs generated similar levels of ORF1 protein in WGE (30), C5′33+C3′U is likely missing sequences that are necessary for full translation activity.

Specific residues within CYVaV S1 are important for translation of reporter constructs *in vivo*

Since translation did not improve significantly with inclusion of 5′ downstream sequences, the parental construct used for mutagenesis was C5′33+C3′U. To determine if the sequence of CYVaV S1 contributes to translation enhancement, most of the unpaired residues in S1 were altered generating construct C5′33mp+C3′U (Fig. 5C and D). Luciferase levels of C5′33mp+C3′U were 52% lower than C5′33+C3′U but still 2.5-fold higher than P5′89+C3′U , suggesting that unpaired residues are not the only feature of S1 contributing to translation. Single bases in the apical loop were important for translation as **A**13**U** and **A**13**G** reduced luciferase activity of the parental construct by 2.3-fold and 3.6-fold, respectively (Fig. 5C and D). **C**15**U** also reduced translation by a similar 3.2-fold, suggesting that the apical loop of S1 is important for translation of reporter constructs *in vivo*.

CYVaV ISSLS apical bases are not likely involved in a LDI with S1

In many plant viruses, at least four residues in apical loops of 3′CITE hairpins and apical loops of 5′ proximal hairpins are involved in a LDI (5, 7). In CYVaV, only three consecutive residues in the apical loops of S1 and the ISSLS are capable of Watson-Crick pairing: the UCC at positions 14-16 in S1 and GGA at positions 2509-2511 in the ISSLS (Fig. 6A). As shown above, mutation of this ISSLS GGA (**GGA**2509**CCU**) did not negatively impact translation of the gRNA *in vitro* but reduced accumulation of full-length CYVaV to basal levels in protoplasts

(Fig. 3B). Addition of these same ISSLS alterations to the parental reporter construct [C5′33+C3′U(**GGA**2509**CCU**] reduced translation to below the level of a random 3′ sequence (5C5′33+V), suggesting that these residues play a key role in ISSLS translation enhancement (Fig. 6B). Mutating the complementary residues in the S1 apical loop [construct C5′33(**UCC**14**AGG**)+C3U] reduced translation by 53%. Combining both sets of mutations, which were designed to be compensatory, did not improve translation efficiency, which remained at background (C5′33+V) levels. These results suggest either that the ISSLS apical bases are not base pairing with S1, or that residues involved in the LDI are sequence specific. To help distinguish between these possibilities, two single mutations were individually generated in the S1 (positions 14 and 15) and ISSLS apical loops (positions 2510 and 2511). The individual S1 mutations reduced translation by 76% and 68% and ISSLS mutations reduced translation by 60% and 75% (Fig. 6B). Combining mutations to re-establish base-pairing did not improve translation. While the possibility remains that a very sequence-specific LDI is forming, these results did not provide evidence for a LDI connecting S1 and the ISSLS 3′CITE.

Other Class 2 ulaRNA S1 enhance translation more effectively than CYVaV S1 in the presence of the CYVaV 3′ UTR

Sequences in the S1 stems of OULV and FULV2 are mainly well conserved with the CYVaV S1, but apical loops are not conserved (Fig. 7A). As with CYVaV, FULV2 and OULV lack Watson Crick pairing of four or more nucleotides between their S1 and ISSLS apical loops. To investigate whether S1 of OULV and FULV2 are capable of facilitating translation in the context of the CYVaV 3′ UTR (in the absence of discernable base pairing), CYVaV S1 in construct C5′33+C3′U was replaced with S1 of OULV and FULV2, generating constructs

O5′39+C3′U and F5′34+C3′U, respectively. Surprisingly, O5′39+C3′U and F5′34+C3′U generated 2.2-fold and 5.5-fold more luciferase activity, respectively, than C5′33+C3′U (Fig. 7B). These results support a lack of LDI between ISSLS and S1 in CYVaV.

To determine whether increased translation associated with O5′39+C3′U and F5′34+C3′U (compared with C5′33+C3′U) was inherent in the S1 hairpins or whether these hairpins interacted more efficiently with the CYVaV 3′ UTR, the CYVaV 3′ UTR in constructs O5′39+C3′U and F5′34+C3′U was replaced with the 250 nt vector-derived sequence, generating constructs O5′39+V and F5′34+V. As shown in Figure 7B, luciferase activities of O5′39+V and F5′34+V were 2-fold and 2.5-fold higher than that of C5′33+V, suggesting that increased translation in constructs containing the CYVaV 3′ UTR was at least partially due to stronger S1 inherent enhancer activity.

To examine whether S1 can enhance translation when associated with different 3′CITEs, the 3′ portions of C5′33+C3′U, O5′39+C3′U and F5′34+C3′U were replaced with the PEMV2 3′ UTR, generating constructs C5′33+P3′U, O5′34+P3′U and F5′39+P3′U (Fig. 7B). C5′33+P3′U, O5′34+P3′U and F5′39+P3′U generated 1.5-, 1.9- and 2.9-fold more luciferase activity compared with C5′33+V, O5′34+V and F5′39+V, respectively. These results suggest that all S1 can enhance translation when associated with heterologous 3′CITEs, but to a lesser extent than with the CYVaV ISSLS.

To examine the functionality of the FULV2 S1 hairpin in a more natural setting, the 5′ 33 bases in the full-length CYVaV gRNA construct was replaced with the FULV2 S1 sequence (construct $CYVaV_{FS\text{-}1}$). As shown in Figure 7C, $CYVaV_{FS\text{-}1}$ was functional in this context, although accumulation levels were unusually variable and reduced by an average of about 50% compared with WT CYVaV. This result supports the findings from the reporter assays that

heterologous pairings of S1 and the ISSLS are still functional despite a lack of discernable pairing sequences.

Strength of S1 in promoting translation of reporter constructs correlates with ability to inhibit translation *in trans*

CYVaV, OULV, and FULV2 S1 transcripts were added to WGE to determine their capacity to reduce WT CYVaV translation. As shown in Figure 8A, 10-fold and 25-fold molar excess of CYVaV S1 reduced translation by 35% and 39%, respectively. In contrast, 10-fold excess FULV2 S1 reduced translation by 42%, which increased to 82% using a 25-fold molar excess. Similarly, 10-fold excess OULV S1 reduced translation by 15% whereas a 25-fold excess reduced translation by 81%. To determine if OULV and FULV S1 repressed translation by sequestering eIF4F and/or eIF4G, eIF4F and eIF4G (200 nM or 400 nM each) were added to *in vitro* translation reactions containing full-length WT CYVaV along with CYVaV, OULV, or FULV2 S1 transcripts at a 25-fold molar excess. As shown in Figure 8B, there was little or no enhancement of translation with the addition of either concentration of eIF4G and eIF4F. This suggests that translation inhibition by OULV and FULV2 S1 is through sequestering other translation factors or interacting directly with the CYVaV template RNA.

Discussion

Identification and characterization of a 3′CITE in Class 2 ulaRNAs

Many plant RNA viruses lacking 5′ caps have evolved 3′CITEs that attract translation

437 components such as eIFs or ribosomal subunits to compete effectively with host cap-dependent
438 translation. 3′CITEs have been divided into seven major classes based on their secondary
439 structure and association with specific translation factors (5, 7, 16, 17). However, there are
440 likely far more classes of these elements as none of the currently identified 3′CITEs are
441 discernable in a number of carmoviruses and umbraviruses, while other members of these genera
442 contain BTE, PTE, TSS, ISS, and TED 3′CITEs (Simon, A.E., unpublished). ulaRNAs, which
443 are newly discovered subviral RNA replicons that are clearly related to umbraviruses (30), are
444 expected to also contain one or more 3′CITEs in their 3′ UTRs. However, despite exhaustive
445 searching, none of the ulaRNAs contain the conserved sequences or signature structures found in
446 known 3′CITEs.

447 To gain information on how these replicons are translated, we studied CYVaV, an
448 unusual Class 2 ulaRNA that only encodes two replication-related proteins (30, 34). Using a
449 deletion analysis of the 3′UTR, the 3′ border of a translation element was mapped to the 3′ end of
450 an extended hairpin previously known as S14 and now referred to as an ISS-like structure or
451 ISSLS. The ISSLS, which is conserved in all Class 2 ulaRNAs, is currently proposed to form an
452 unbranched hairpin with four internal loops, three of which are composed of conserved or mostly
453 conserved residues. Most of the CYVaV ISSLS residues were important for translation *in vitro*
454 and/or accumulation *in vivo*, suggesting that the overall 3-D conformation is likely critical for
455 function. Based on the results of the mutagenesis study, several portions of the ISSLS are of
456 particular interest for further investigation. For example, mutations in the C-rich internal loop
457 above the central stem that did not extend the length of the stem enhanced translation of both p21
458 and p81 (Fig. 3B). This suggests that the C-rich region acts as a repressor, possibly through
459 base-pairing with G-rich sequences *in cis* or proteins *in trans* to suppress the functioning of the

460 ISSLS, which may be required to inhibit translation and allow replication to take place. Another
461 observation requiring further exploration is the number of alterations, including in the apical loop
462 and in invariant B1, that did not negatively impact translation of CYVaV in WGE yet reduced
463 accumulation of CYVaV or translation of reporter constructs to basal levels *in vivo*. This
464 difference could be attributed to the optimized translation environment provided by WGE, in
465 which negative effects of mutations on translation might be compensated, at least partially, by
466 increased abundance of translation factors compared with the cellular environment.
467 Alternatively, these alterations could have unexpected effects on replication or stability. A third
468 observation requiring future study is the relatively weak translation of the reporter constructs
469 containing CYVaV 5′ sequences and 3′ UTR. Extension of the 5′ sequences to include
470 downstream elements either had no effect or were strongly negative for translation, unlike
471 similar extensions for PEMV2, where translation improved ~20-fold by allowing the LDI to take
472 place with a coding region hairpin (10). Weak translation of the CYVaV reporter constructs
473 suggests either (i) additional translation elements may exist in the coding region that remain to
474 be identified; and/or (ii) that using the exact 3′ UTR, which disrupts S12 and eliminates
475 substantial downstream base-pairing between 3′ UTR and coding region sequences (see Fig. 1B)
476 likely affects RNA structure throughout the remainder of the 3′ UTR, which in turn may be
477 detrimental for translation.

478

479 **CYVaV ISSLS represents a novel class of 3′CITE**

480 The ISS 3′CITE also adopts an unbranched hairpin structure, but with fewer and smaller
481 internal loops than the ISSLS. ISSLS and ISS have different conserved nucleotides within their
482 most prominent internal loops (Fig. 9) and interact with different translation factors. The ISSLS

interacts with eIF4G/eIF4F (Fig. 3D) whereas the ISS interacts with eIF4F and inefficiently with eIF4E and eIF4G (18, 19, 48). ISSLS conserved sequences have not been identified in other eIF4G-binding 3′CITEs such as the BTE and thus the interaction between ISSLS and eIF4G or eIF4F is likely mediated by a specific tertiary fold, as postulated for the PEMV2 PTE interaction with eIF4E (28, 29), and the turnip crinkle virus (TCV) TSS with 60S ribosomal subunit (9, 49, 50). Since *trans*-inhibition assays with supplemental eIF4G and eIF4F did not restore translation that was repressed by addition of the ISSLS, the ISSLS may also be interacting directly or indirectly with other translation factors present in limited quantities or may directly interfere with the translated template. Altogether, these observations strongly suggest that the CYVaV ISSLS represents a new class of 3′CITE that differs from the previously reported ISS.

CYVaV S1 likely supports translation in a manner that does not involve direct RNA:RNA interaction with the ISSLS through a LDI

LDI involving four or more base-pairs are commonly associated with 3′CITEs (5, 7) to enhance the number of templates selected for translation (51), and many LDIs connect 3′CITEs with 5′ terminal hairpins. Only three consecutive residues can form Watson-Crick base pairs between the apical loops of any ulaRNA ISSLS and their 5′ terminal hairpins (S1), which were sufficient for optimal translation activity in the reporter constructs (Fig. 5). Mutating these three residues in either the ISSLS or S1 reduced reporter gene translation *in vivo*, however single and three base compensatory mutations did not restore efficient translation (Fig. 6). In addition, reporter constructs containing OULV and FULV S1 enhanced translation when associated with both the CYVaV 3′ UTR and the heterologous 3′ UTR from PEMV2 (Fig. 7), despite lacking discernable base-pairing with PEMV2 3′CITEs. Similar synergy in the absence of base-pairing

between a 5′ translation element and CITE-containing 3′ UTR was previously reported for TCV, whose pyrimidine-rich 5′ UTR was functionally replaceable with that of related carmovirus cardamine chlorotic fleck virus (52). FULV2 and OULV S1 also stimulated reporter gene translation in the absence of any 3′CITE (Fig. 7), suggesting that S1 may independently attract translation factors, similar to the TCV 5′ translation enhancing element that binds to 40S ribosomal subunits (52). For TCV, the proposal was made that 60S subunits binding to the 3′TSS connect with 40S subunits binding to the 5′ end, thus circularizing the template. Interestingly, the CYVaV S1 apical portion contains the same putative 18S rRNA-complementary sequence in the BYDV BTE (GGAUCCU) (8), and other Class 2 ulaRNA S1 also have potential pairing sites of six consecutive bases or more in the identical region. These postulated interactions remain to be verified, and whether some type of protein bridge connects the ends of Class 2 ulaRNAs for translation initiation remains to be investigated.

Conclusions

We have identified two translation enhancers in Class 2 ulaRNAs: the 5′ terminal hairpin and the ISSLS 3′CITE. Class 1, 2 and 3 ulaRNAs share similar structures at their 5′ ends extending hundreds of bases upstream into their recoding sites (30), and additionally at their 3′ ends beginning with the penultimate hairpin. Abrupt sequence and structural divergence among the classes of ulaRNAs correlates with the beginning of the additional ORF near the end of the RdRp ORF found in nearly all Class 2 ulaRNAs (and a different inserted ORF in the Class 3 ulaRNA). This is likely due to different recombination events leading to different classes of ulaRNAs, with the event giving rise to Class 2 ulaRNAs including the ISSLS 3′CITE (30). Whether Class 1 and 3 ulaRNAs also harbor one of more unique 3′CITE will be the subject of

529 future investigations.

530

531 **Materials and Methods**

532

533 **Plasmid construction**

534 Full-length CYVaV (Genbank accession number JX101610) was generated in vector pET17b
535 downstream of a T7 promoter by Biomatik (Wilmington, Delaware) and was kindly provided by
536 Georgios Vidalakis (UC Riverside) (30). CYVaV mutants were constructed using one-step site-
537 directed mutagenesis PCR (53) using the appropriate DNA oligonucleotides (IDT, Coralville,
538 Iowa). The F-Luc construct with PEMV2 1-89 and the PEMV2 3′ UTR was constructed
539 previously (10). F-Luc constructs were generated by addition of a *Bgl*II restriction site at the 5′
540 end of the Luc ORF and a *Swa*I restriction site at the 3′ end. 3′UTRs were added using ligation-
541 independent cloning (56) with a PmeI site added at the 3′ terminus for plasmid linearization. All
542 constructs and mutations were verified by sequencing (Eurofins Genomics, LLC). Plasmids were
543 linearized and used as templates for *in vitro* transcription with T7 RNA polymerase.

544

545 ***In vitro* translation in WGE**

546 Uncapped RNAs synthesized by *in vitro* transcription were purified using Monarch RNA
547 Cleanup Kit (New England Biolabs) and subjected to translation in a 10 µl WGE reaction mix
548 (Promega) in the presence of ^{35}S-methionine and reagents according to the manufacturer′s
549 instructions. Translation reactions were incubated at 25°C for 45 min. 2X SDS loading buffer
550 (90 mM Tris-HCl at pH 6.8, 2% SDS, 20% glycerol, 0.1% bromophenol blue, 200 mM DTT)

was added and the mixture separated on 10% SDS-polyacrylamide gels, which were then dried and exposed to a phosphorimaging screen followed by scanning using a Typhoon image analyzer (Amersham). Radioactive bands were quantified using GelQuantNET software. For *trans*-inhibition assays, excess competitor RNAs or eIF4G or eIF4F were added to WGE prior to the incubation. eIF4G and eIF4F were expressed in E. coli and purified (47). Translation initiation factors were a kind gift from Karen Browning (UT Austin).

Protoplasts preparation and transfection

Arabidopsis protoplasts were prepared and transfected as described previously (10). Briefly, protoplasts were prepared from seed-started callus cultures of *A. thaliana* (Col-0). Protoplasts ($4x10^5$) were transfected with 4 µg of purified *in vitro* transcribed RNA using a polyethylene glycol-mediated transformation method as previously described (10). Cells were collected at 18 h post-transfection and luciferase activity was assayed with a Dual-Luciferase Reporter Assay System (Promega) using a Modulus microplate multimode reader (Turner Biosystems).

EMSA cross-linking assay

RNAs that were internally labeled with α-^{32}P-CTP (2 nM) by *in vitro* transcription and purified were incubated with eIF4G or eIF4F (200 or 400 nM each) in 4F binding buffer (28 mM HEPES-KOH at pH 7.6, 57 mM KCl, 2.3 mM $MgAc_2$, 0.114 mg/ml BSA, 85 µg/ml yeast tRNA, 2.8% glycerol, 2.4 mM DTT) in a 15 µl mixture and incubated at 30°C for 15 min. Samples were then divided in two, with one subjected to 254 nm UV light at a distance of 8 cm for 15 min using a Spectrolline UV crosslinker. Samples were then mixed with 2X SDS loading buffer and subjected to 8% or 10% SDS-PAGE, the gels then dried and subsequently exposed to a

phosphorimaging screen.

RNA structure drawing

All RNA structures were drawn using the RNA2drawer online drawing tool at https://rna2drawer.app/ (42).

Acknowledgements

We thanks Georgios Vidalakis for the CYVaV full-length construct and Karen Browning for the wheat translation initiation factors. This work was supported by the National Science Foundation (MCB-1818229) and USDA NIFA Emergency Citrus Disease Research and Extension Program 2020-08455 to AES.

Figure legends

FIG 1 Structures in CYVaV Domain 3. (A) Genome organization of umbravirus PEMV2 and three Class 2 ulaRNAs. ORF1 and the -1PRF extension of OR1 (ORF2) encode replication-required proteins including the p81 RdRp. OULV and FULV2 contain an additional ORF (ORF5) that contains motifs found in some movement proteins (30). CYVaV has two deletions in the analogous ORF5 region that along with additional changes, eliminate translation of the ORF. (B) Structure of full-length CYVaV. The three domains (D1, D2, and D3) are indicated. Green asterisk denotes location of the start codon for the p21 ORF1 and the two red asterisks denote locations of stop codons for ORF1 and ORF2. The bridge stem at the base of D2 that juxtaposes D1 and D3 is indicated. Numbers are from (30) and refer to specific secondary structure elements. (C) Structure and sequence of CYVaV D3. Residues are colored according to their SHAPE reactivity (30). Pseudoknot 1 (ψ_1) is shown. Names of other structures are from (30). Line drawings at right are the putative structures for the same region in FULV2 and

OULV determined by comparative modeling based on the CYVaV D3 structure. An extra segment in FULV2 D3, which is not found in other Class 2 ulaRNAs, is bracketed. Inset at right is the 3′ end of PEMV2. The three hairpins and two pseudoknots that comprise the 3′TSS 3′CITE are shown.

FIG 2 Effect on translation of CYVaV 3′ UTR deletions. Names of deletion mutants and the deleted regions are shown. The positions of RNA structures in CYVaV D3 are indicted on top. Portions included in the constructs are denoted by dark grey lines. (B) *In vitro* translation of full-length and deletion mutants of CYVaV in WGE. Positions of p21 and p81 are shown. Average values and standard deviations of p21 and p81 translation levels were obtained from three independent experiments and are normalized to WT p21 and p81 levels.

FIG 3 S14 is a eIF4G-binding 3′CITE. (A). Base conservation in Class 2 ulaRNA S14. Conserved residues are indicated, with dark green and light green denoting conservation in eight or seven ulaRNAs, respectively. The ulaRNAs used for this alignment were: CYVaV [JX101610]; CYVaV-Delta [MT893741]; CYVaV-RioBlanco [MT893740]; OULV [MH579715]; FULV1 [MW480892]; FULV2 [MW480893]; Ethiopian maize associated virus 1 and 2 [EMaV1/2, MF415880 and MN715238]. Features referred to in the text are indicated. (B) Mutation analysis of CYVaV S14. Full-length CYVaV WT and mutant templates (mutations in red) were subjected to i*n vitro* translation in WGE. Numbers in black denote levels of p21 and p81 obtained in WGE. Selected constructs were also assayed for accumulation in protoplasts (in green; A, accumulation). For both assays, values are presented as a percentage of WT with

620 standard deviations obtained from three independent experiments. End points of the ISSLS$_{\Delta B}$
621 fragment used in C and D are indicated. (C) *Trans*-inhibition assay. Wild-type and mutant S14
622 fragments (ISSLS: positions 2452-2559, ISSLS$_{\Delta B}$: positions 2484-2532) were added in a 10- or
623 25-fold molar excess along with full-length CYVaV gRNA template to WGE. Values are a
624 percentage of the levels of p21 and p81 obtained in the reaction with no added fragments (lane -)
625 from three independent experiments with standard deviations. (D) EMSAs using 2 nM
626 radiolabeled RNA fragments. OPMV BTE (binds to eIF4G and eIF4F) and PEMV2 PTE (binds
627 to eIF4E and eIF4F) were included as controls. Fragments were incubated with 200 nM BSA or
628 200/400 nM wheat eIF4F, eIF4G, or eIF4E at 30°C for 15 min and then exposed to UV light for
629 15 min. 8% SDS-PAGE was used for ISSLS and OPMV BTE, and 10% SDS-PAGE was used
630 for ISSLS$_{\Delta B}$ and PEMV2 PTE.

631

632 **FIG 4** ISSLS inhibition of translation *in trans* is not restored by addition of eI4G or eIF4F.
633 CYVaV gRNA template and 10-fold excess OPMV BTE, CYVaV ISSLS, or ISSLS$_{\Delta B}$ fragments
634 were added to WGE with and without 200/400 nM eIF4G or eIF4F. p21 levels are averages with
635 standard deviations obtained from three independent experiments and are presented as a
636 percentage of that obtained with no added RNA fragments or proteins (lane -).

637

638 **FIG 5** CYVaV S1 promotes translation in the presence of the CYVaV 3′ UTR. (A) Secondary
639 structure of CYVaV positions 1-60 and the PEMV2 5′ 89 nt. Start codons are in green. (B)
640 Luciferase reporter construct ORFs are indicated by colored bars. Blue, CYVaV; green, PEMV2.
641 UTRs are open rectangles the same color as the corresponding ORFs. Vector-derived sequence is

642 denoted by a black bar. Relative luciferase activities were obtained from at least three
643 experiments. (C) Mutations incorporated into S1 in construct C5′33+C3′U. (D) Reporter
644 constructs containing the mutations shown in (C) were assayed for translation in protoplasts as
645 described in (B).

646

647 **FIG 6** No LDI is discernible between the terminal loops of CYVaV S1 and ISSLS. (A) Base
648 alterations in CYVaV S1 (left) and ISSLS (right) terminal loops. Mutated bases are in red and
649 base numbers are indicated. (B) Luciferase reporter constructs used to assay translation of
650 reporter constructs in protoplasts. See legend to Fig. 5 for details.

651

652 **FIG 7** Class 2 ulaRNA S1 are translation enhancers. (A) OULV and FULV2 S1. Bases
653 conserved with CYVaV S1 are in red. (B) Luciferase constructs containing S1 of CYVaV,
654 OULV or FULV2 upstream of F-luc reporter gene and either the CYVaV 3′ UTR (C3′), PEMV2
655 3′ UTR (P3′) or vector-derived sequence (V). Data represent mean ± standard deviation from at
656 least three independent experiments. (C) Northern blot analysis of WT CYVaV and CYVaV with
657 S1 from FULV2 ($CYVaV_{F\text{-}S1}$) at 18 h post inoculation. Each lane represents RNA extracted from
658 different plants.

659

660 **FIG 8** OULV and FULV SI can inhibit translation in WGE. (A) *Trans*-inhibition assay using
661 CYVaV gRNA template and 10- or 25-fold molar excess of CYVaV, FULV2 or OULV S1. (B)
662 *Trans*-inhibition assay using CYVaV gRNA template and 25-fold molar excess of CYVaV,
663 FULV2 or OULV S1 with and without 200/400 nM of eIF4G or eIF4F. p21 translation levels

are averages with standard deviations of values obtained from at least three independent experiments and are presented as a percentage of that obtained with no added RNA fragments or proteins (lane -).

FIG 9 Structural alignments between some Class 2 ulaRNAs ISSLS (left) and previously reported ISS (boxed, right). EMaV, Ethiopian maize associated virus; MNSV, maize necrotic streak virus; MWLMV, maize white line mosaic virus; MNSV-264, melon necrotic spot virus. Bases conserved with CYVaV are in red. Circled residues denote ISS-conserved sequences. Bases in ISS that engage in long-distance pairing with 5′ sequences are shaded in blue.

References

1. Jackson RJ, Hellen CU, Pestova TV. 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol 11:113-27.

2. Hinnebusch AG, Lorsch JR. 2012. The mechanism of eukaryotic translation initiation: new insights and challenges. Cold Spring Harbor Perspect Biol 4: a011544.

3. Sesma A, Castresana C, Castellano MM. 2017. Regulation of translation by TOR, eIF4E and eIF2 alpha in plants: current knowledge, challenges and future perspectives. Front Plant Sci 8: 644.

4. Pestova TV, Lorsch JR, Hellen CUT. 2007. The mechanism of translation initiation in eukaryotes, p 87-128. In Mathews MB, Sonenberg N, Hershey JWB (ed), Translational Control in Biology and Medicine. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

5. Simon AE, Miller WA. 2013. 3' Cap-independent translation enhancers of plant viruses. Annu Rev Microbiol 67:21-42.

6. Jaafar ZA, Kieft JS. 2019. Viral RNA structure-based strategies to manipulate translation. Nature Rev Microbiol 17:110-123.

7. Miras M, Miller WA, Truniger V, Aranda MA. 2017. Non-canonical translation in plant RNA viruses. Front Plant Sci 8:494.

8. Das Sharma S, Kraft JJ, Miller WA, Goss DJ. 2015. Recruitment of the 40S ribosomal 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.

9. Stupina VA, Meskauskas A, McCormack JC, Yingling YG, Shapiro BA, Dinman JD, Simon AE. 2008. The 3' proximal translational enhancer of Turnip crinkle virus binds to 60S ribosomal subunits. RNA 14:2379-2393.

10. Gao F, Kasprzak W, Stupina VA, Shapiro BA, Simon AE. 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.

11. Guo L, Allen EM, Miller WA. 2001. Base-pairing between untranslated regions facilitates translation of uncapped, nonpolyadenylated viral RNA. Mol Cell 7:1103-1109.

12. Ilyas M, Du Z, Simon A. 2021. Opium poppy mosaic virus has an Xrn-resistant, translated subgenomic RNA and a BTE 3' CITE. J Virol 9: e02109-20

13. Nicholson BL, White KA. 2014. Functional long-range RNA-RNA interactions in positive-strand RNA viruses. Nat Rev Microbiol 12:493-504.

14. Blanco-Perez M, Perez-Canamas M, Ruiz L, Hernandez C. 2016. Efficient translation of pelargonium line pattern virus RNAs relies on a TED-Like 3'-translational enhancer that communicates with the corresponding 5'-region through a long-distance RNA-RNA interaction. PloS One 11:e0152593.

15. Chattopadhyay M, Shi K, Yuan X, Simon AE. 2011. Long-distance kissing loop interactions between a 3' proximal Y-shaped structure and apical loops of 5' hairpins enhance translation of saguaro cactus virus. Virology 417:113-25.

16. Truniger V, Miras M, Aranda MA. 2017. Structural and functional diversity of plant virus 3 '-cap-independent translation enhancers (3'-CITEs). Front Plant Sci 8: 2047.

17. Nicholson BL, White KA. 2011. 3' Cap-independent translation enhancers of positive-strand RNA plant viruses. Curr Opin Virol 1:373-380.

18. Nicholson BL, Wu B, Chevtchenko I, White KA. 2010. Tombusvirus recruitment of host translational machinery via the 3' UTR. RNA 16:1402-19.

19. Liu Q, Goss DJ. 2018. The 3' mRNA I-shaped structure of maize necrotic streak virus binds to eukaryotic translation factors for eIF4F-mediated translation initiation. J Biol Chem 293:9486-9495.

20. Yoo RH, Lee S-W, Lim S, Zhao F, Igori D, Baek D, Hong J-S, Lee S-H, Moon JS. 2017. Complete genome analysis of a novel umbravirus-polerovirus combination isolated from Ixeridium dentatum. Arch Virol 162:3893-3897.

21. Taliansky ME, Robinson DJ. 2003. Molecular biology of umbraviruses: phantom warriors. J Gen Virol 84:1951-1960.

22. Wang DY, Yu CM, Liu SS, Wang GL, Shi KR, Li XD, Yuan XF. 2017. Structural alteration of a BYDV-like translation element (BTE) that attenuates p35 expression in three mild Tobacco bushy top virus isolates. Sci Rep 7: 4213.

23. Wang Z, Kraft JJ, Hui AY, Miller WA. 2010. Structural plasticity of Barley yellow dwarf virus-like cap-independent translation elements in four genera of plant viral RNAs. Virology 402:177-186.

24. Treder K, Kneller ELP, Allen EM, Wang ZH, Browning KS, Miller WA. 2008. The 3' cap-independent translation element of Barley yellow dwarf virus binds eIF4F via the eIF4G subunit to initiate translation. RNA 14:134-147.

25. Gao F, Kasprzak WK, Szarko C, Shapiro BA, Simon AE. 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.

26. Gao F, Simon AE. 2017. Differential use of 3 ' CITEs by the subgenomic RNA of Pea enation mosaic virus 2. Virology 510:194-204.

27. Gao F, Gulay SP, Kasprzak W, Dinman JD, Shapiro BA, Simon AE. 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.

28. Wang Z, Treder K, Miller WA. 2009. Structure of a viral cap-independent translation element that functions via high affinity binding to the eIF4E subunit of eIF4F. J Biol Chem 284:14189-202.

29. Wang Z, Parisien M, Scheets K, Miller WA. 2011. The cap-binding translation initiation factor, eIF4E, binds a pseudoknot in a viral cap-independent translation element. Structure 19:868-880.

30. Liu JY, Carino E, Bera S, Gao F, May JP, Simon AE. 2021. Structural analysis and whole genome mapping of a new type of plant virus subviral RNA: umbravirus-like associated RNAs. Viruses 13: 646.

31. Gao F, Simon AE. 2016. Multiple cis-acting elements modulate programmed -1 ribosomal frameshifting in Pea enation mosaic virus. Nucleic Acids Res 44:878-895.

32. May, JP, Johnson, PZ, Ilyas, M, Gao, F, and Simon, AE 2020. Disruption of nonsense-mediated decay by the multifunctional long-distance movement protein of Pea enation mosaic virus 2. Mbio 11:e00204-20. https://doi.org/10.1128/mBio

33. Wang, X., Olmedo-Velarde, A., Larrea-Sarmiento, A., Simon AE, Kong A, Borth, W., Suzuki, J.Y., Wall, M.M., Hu J, Mellzer, M. 2021. Genome characterization of fig umbra-like virus. Virus Genes https://doi.org/10.1007/s11262-021-01867-4

34. Kwon SJ, Bodaghi S, Dang T, Gadhave KR, Ho T, Osman F, Al Rwahnih M, Tzanetakis IE, Simon AE, Vidalakis G. 2021. Complete nucleotide sequence, genome organization, and

764 comparative genomic analyses of citrus yellow-vein associated virus (CYVaV). Front Microbiol
765 12: 1371.

766 35. Cornejo-Franco JF, Flores F, Mollov D, Quito-Avila DF. 2021. An umbra-related virus
767 found in babaco (Vasconcellea x heilbornii). Arch Virol 166: 2321–2324.

768 36. Cornejo-Franco JF, Alvarez-Quinto RA, Quito-Avila DF. 2018. Transmission of the
769 umbra-like Papaya virus Q in Ecuador and its association with meleira-related viruses from
770 Brazil. Crop Protec 110:99-102.

771 37. Felker P, Bunch R, Russo G, Preston K, Tine JA, Suter B, Mo XH, Cushman JC, Yim
772 WC. 2019. Biology and chemistry of an umbravirus like 2989 bp single stranded RNA as a
773 possible causal agent for Opuntia stunting disease (engrosamiento de cladodios) - A Review. J
774 Profess Assoc Cactus Dev 21:1-31.

775 38. Sa Antunes TF, Vionette Amaral RJ, Ventura JA, Godinho MT, Amaral JG, Souza FO,
776 Zerbini PA, Zerbini FM, Bueno Fernandes PM. 2016. The dsRNA virus papaya meleira virus
777 and an ssRNA virus are associated with papaya sticky disease. PloS One 11: e0155240.

778 39. Quito-Avila DF, Alvarez RA, Ibarra MA, Martin RR. 2015. Detection and partial
779 genome sequence of a new umbra-like virus of papaya discovered in Ecuador. Eur J Plant Path
780 143:199-204.

781 40. Tahir MN, Bolus S, Grinstead SC, McFarlane SA, Mollov D. 2021. A new virus of the
782 family Tombusviridae infecting sugarcane. Arch Virol 166:961-965.

783 41. Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction.
784 Nucleic Acids Res 31:3406-3415.

785 42. Johnson PZ, Kasprzak WK, Shapiro BA, Simon AE. 2019. RNA2Drawer: geometrically
786 strict drawing of nucleic acid structures with graphical structure editing and highlighting of
787 complementary subsequences. RNA Biol 16:1667-1671.

788 43. Cimino PA, Nicholson BL, Wu B, Xu W, White KA. 2011. Multifaceted regulation of
789 translational readthrough by RNA replication elements in a tombusvirus. PLoS Path 7:e1002423.

790 44. Kuhlmann MM, Chattopadhyay M, Stupina VA, Gao F, Simon AE. 2016. An RNA
791 element that facilitates programmed ribosomal readthrough in Turnip crinkle virus adopts
792 multiple conformations. J Virol 90:8575-8591.

793 45. Newburn LR, Wu BD, and White KA. 2020. Investigation of novel RNA elements in the 3'
794 UTR of tobacco necrosis virus-D. Viruses 8: 856 https://doi.org/10.3390/v12080856

795 46. Nicholson BL, Zaslaver O, Mayberry LK, Browning KS, White KA. 2013. Tombusvirus
796 Y-shaped translational enhancer forms a complex with eIF4F and can be functionally replaced
797 by heterologous translational enhancers. J Virol 87:1872-1883.

798 47. Mayberry LK, Dennis MD, Allen ML, Nitka KR, Murphy PA, Campbell L, Browning
799 KS. 2007. Expression and purification of recombinant wheat translation initiation factors eIF1,

800 eIF1A, eIF4A, eIF4B, eIF4F, eIF(iso)4F, and eIF5. Translation Initiation: Reconstituted Systems
801 and Biophysical Methods 430:397-408.

802 48. Miras M, Truniger V, Querol-Audi J, Aranda MA. 2017. Analysis of the interacting
803 partners eIF4F and 3 '-CITE required for Melon necrotic spot virus cap-independent translation.
804 Mol Plant Pathol 18:635-648.

805 49. Zuo XB, Wang JB, Yu P, Eyler D, Xu H, Starich MR, Tiede DM, Simon AE, Kasprzak
806 W, Schwieters CD, Shapiro BA, Wang YX. 2010. Solution structure of the cap-independent
807 translational enhancer and ribosome-binding element in the 3' UTR of turnip crinkle virus. Proc
808 Natl Acad Sci USA 107:1385-1390.

809 50. McCormack, J. C., Yuan, X., Yingling, Y. G., Zamora, R. E., Shapiro, B. A., and Simon,
810 A. E. 2008. Structural domains within the 3' UTR of Turnip crinkle virus. J Virol 82:8706-
811 8720.

812 51. Du Z, Alekhina OM, Vassilenko KS, Simon AE. 2017. Concerted action of two 3' cap-
813 independent translation enhancers increases the competitive strength of translated viral genomes.
814 Nucleic Acids Res doi: 10.1093/nar/gkx643.

815 52. Stupina VA, Yuan X, Meskauskas A, Dinman JD, Simon AE. 2011. Ribosome binding to
816 a 5' translational enhancer is altered in the presence of the 3' untranslated region in cap-
817 independent translation of turnip crinkle virus. J Virol 85:4638-53.

818 53. Liu H, Naismith JH. 2008. An efficient one-step site-directed deletion, insertion, single
819 and multiple-site plasmid mutagenesis protocol. BMC Biotechnol 8:91.

820 54. Jeong J-Y, Yim H-S, Ryu J-Y, Lee HS, Lee J-H, Seen D-S, Kang SG. 2012. One-step
821 sequence- and ligation-independent cloning as a rapid and versatile cloning method for
822 functional genomics studies. Appl Environ Microbiol 78:5440-5443.

823

824

825

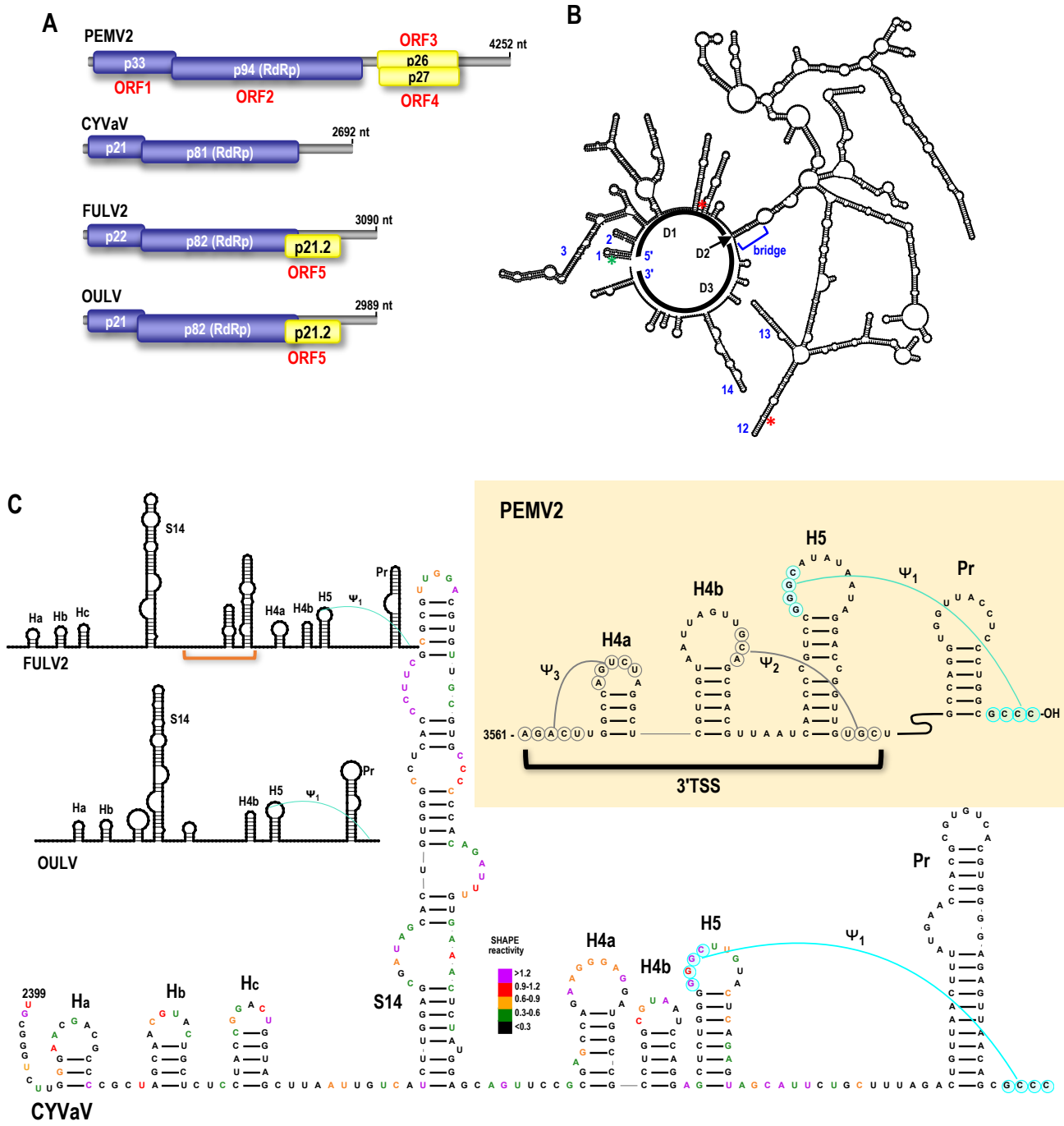

A
PEMV2
p33
p94 (RdRp)
p26
p27
4252 nt
ORF1
ORF2
ORF3
ORF4
CYVaV
p21
p81 (RdRp)
2692 nt
FULV2
p22
p82 (RdRp)
p21.2
3090 nt
ORF5
OULV
p21
p82 (RdRp)
p21.2
2989 nt
ORF5
B
D1
D2
D3
bridge
5'
3'
1
2
3
12
13
14
C
FULV2
S14
Ha
Hb
Hc
H4a
H4b
H5
Ψ1
Pr
OULV
PEMV2
H5
H4b
H4a
Ψ1
Pr
Ψ3
Ψ2
3561 -
-OH
3'TSS
SHAPE reactivity
>1.2
0.9-1.2
0.6-0.9
0.3-0.6
<0.3
2399
CYVaV

Figure 1

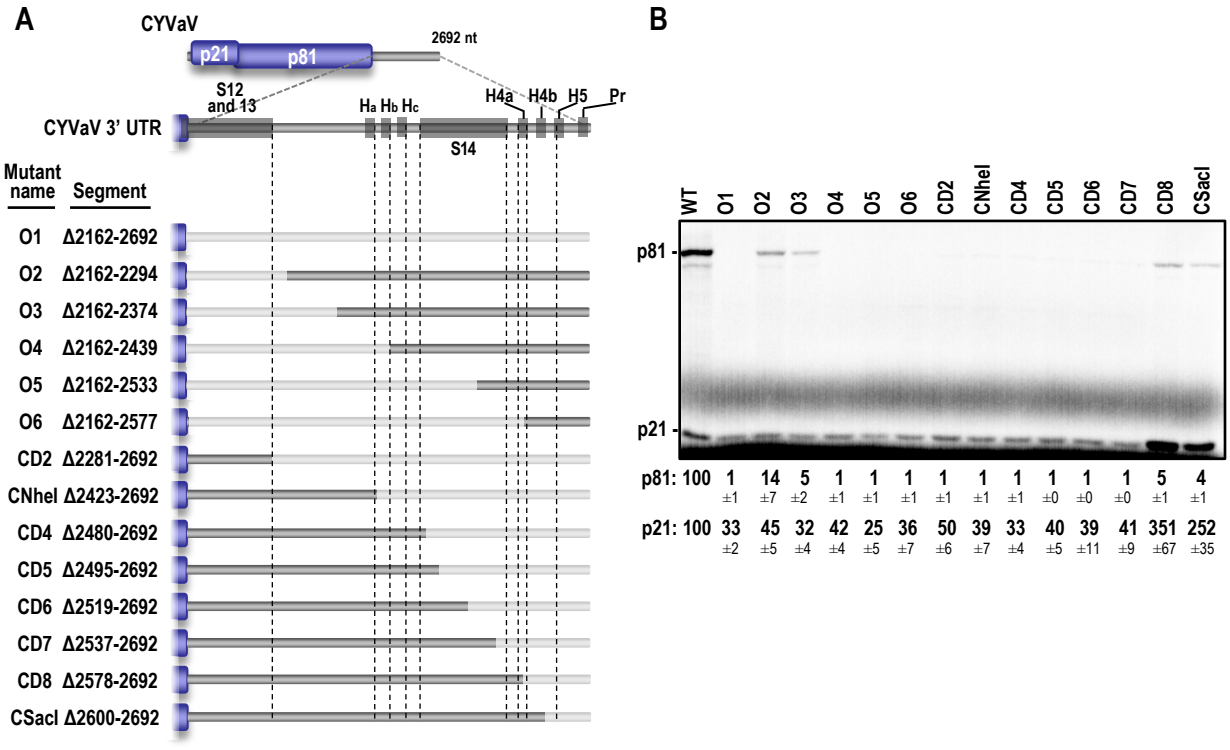

A
CYVaV
p21
p81
2692 nt
S12
and 13
Ha Hb Hc
H4a
H4b
H5
Pr
CYVaV 3' UTR
S14
Mutant name
Segment
O1 Δ2162-2692
O2 Δ2162-2294
O3 Δ2162-2374
O4 Δ2162-2439
O5 Δ2162-2533
O6 Δ2162-2577
CD2 Δ2281-2692
CNhel Δ2423-2692
CD4 Δ2480-2692
CD5 Δ2495-2692
CD6 Δ2519-2692
CD7 Δ2537-2692
CD8 Δ2578-2692
CSacI Δ2600-2692
B
WT
O1
O2
O3
O4
O5
O6
CD2
CNhel
CD4
CD5
CD6
CD7
CD8
CSacI
p81
p21
p81: 100 1 ±1 14 ±7 5 ±2 1 ±1 1 ±1 1 ±1 1 ±1 1 ±1 1 ±1 1 ±0 1 ±0 1 ±0 5 ±1 4 ±1
p21: 100 33 ±2 45 ±5 32 ±4 42 ±4 25 ±5 36 ±7 50 ±6 39 ±7 33 ±4 40 ±5 39 ±11 41 ±9 351 ±67 252 ±35

Figure 2

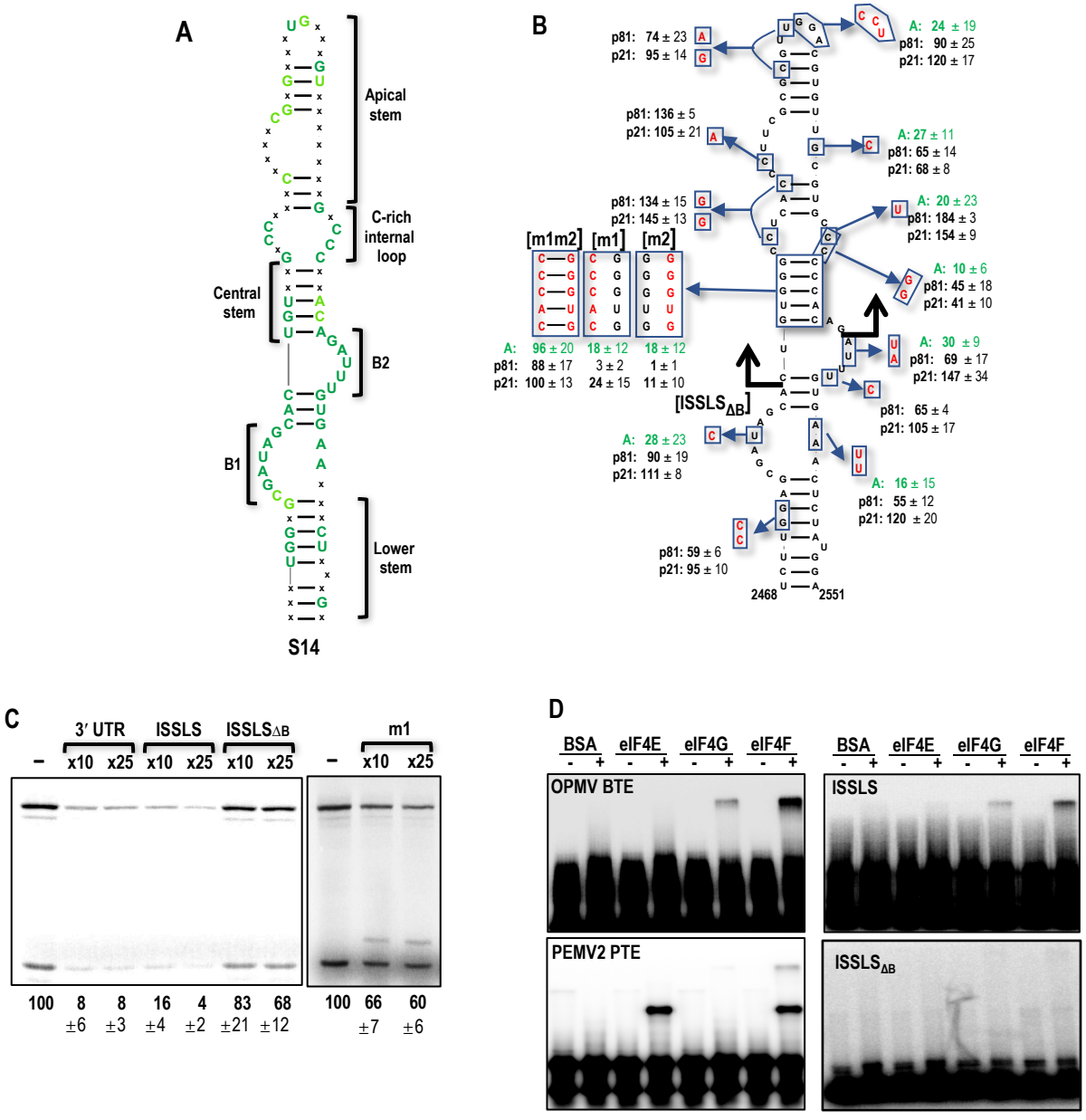

A
Apical stem
C-rich internal loop
Central stem
B2
B1
Lower stem
S14
B
p81: 74 ± 23
p21: 95 ± 14
A: 24 ± 19
p81: 90 ± 25
p21: 120 ± 17
p81: 136 ± 5
p21: 105 ± 21
A: 27 ± 11
p81: 65 ± 14
p21: 68 ± 8
p81: 134 ± 15
p21: 145 ± 13
A: 20 ± 23
p81: 184 ± 3
p21: 154 ± 9
[m1m2]
[m1]
[m2]
A: 10 ± 6
p81: 45 ± 18
p21: 41 ± 10
A: 96 ± 20
p81: 88 ± 17
p21: 100 ± 13
18 ± 12
3 ± 2
24 ± 15
18 ± 12
1 ± 1
11 ± 10
A: 30 ± 9
p81: 69 ± 17
p21: 147 ± 34
[ISSLSΔB]
p81: 65 ± 4
p21: 105 ± 17
A: 28 ± 23
p81: 90 ± 19
p21: 111 ± 8
A: 16 ± 15
p81: 55 ± 12
p21: 120 ± 20
p81: 59 ± 6
p21: 95 ± 10
2468
2551
C
3′ UTR
ISSLS
ISSLSΔB
m1
–
x10
x25
x10
x25
x10
x25
–
x10
x25
100
8 ±6
8 ±3
16 ±4
4 ±2
83 ±21
68 ±12
100
66 ±7
60 ±6
D
BSA
eIF4E
eIF4G
eIF4F
OPMV BTE
PEMV2 PTE
ISSLS
ISSLSΔB

Figure 3

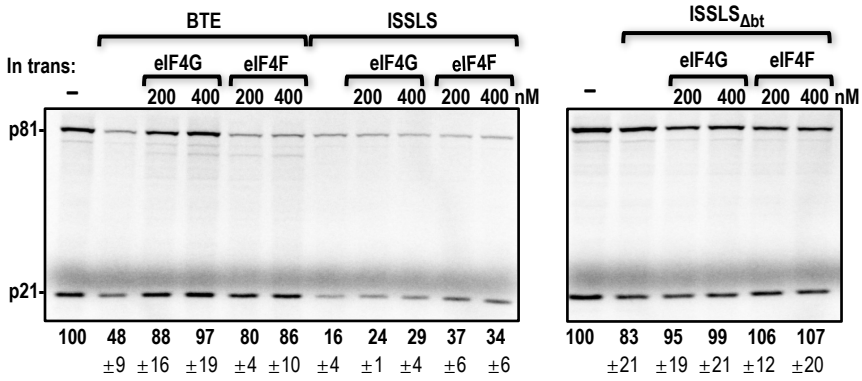

BTE
ISSLS
ISSLS$_{\Delta bt}$
In trans:
eIF4G
eIF4F
eIF4G
eIF4F
eIF4G
eIF4F
–
200 400 200 400 200 400 200 400 nM
–
200 400 200 400 nM
p81
p21
100 48 88 97 80 86 16 24 29 37 34
±9 ±16 ±19 ±4 ±10 ±4 ±1 ±4 ±6 ±6
100 83 95 99 106 107
±21 ±19 ±21 ±12 ±20

Figure 4

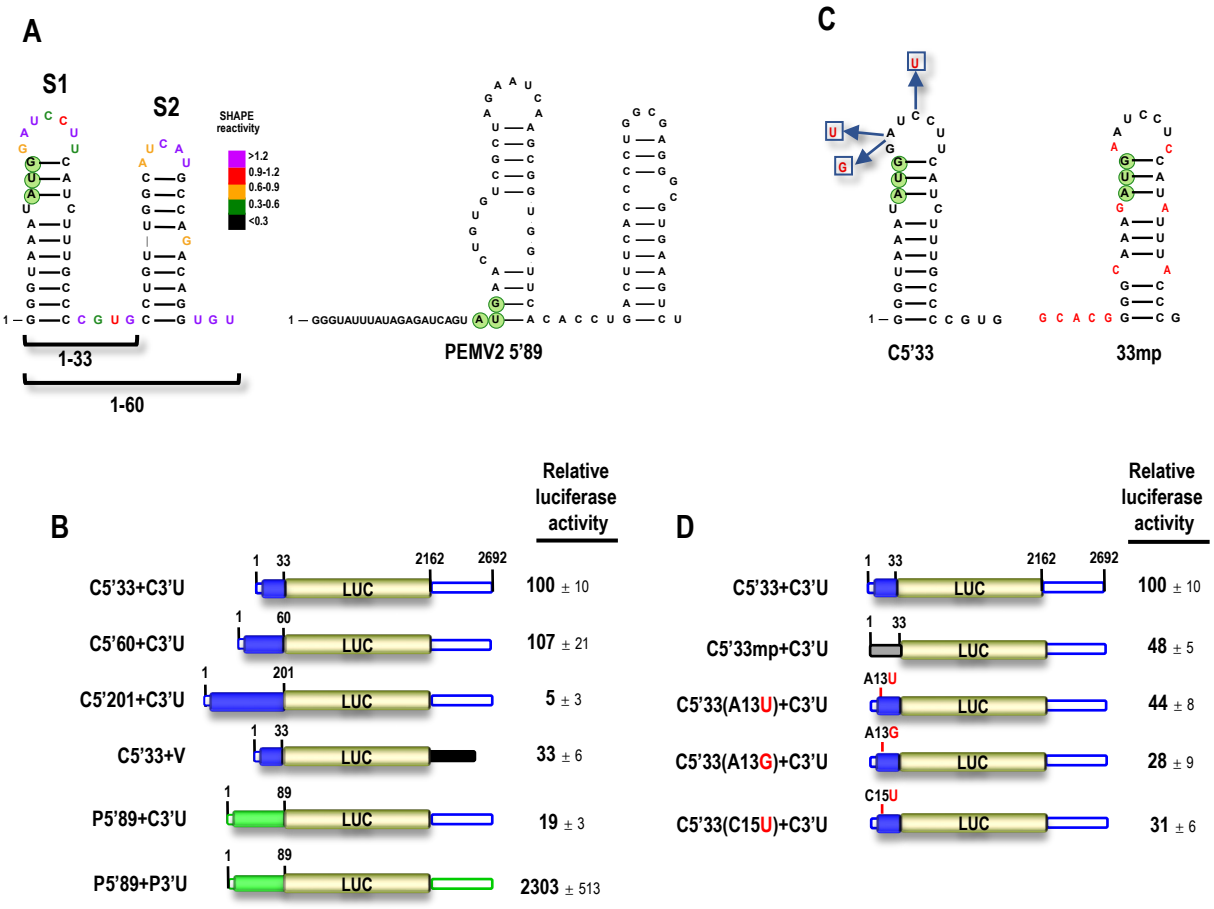

A
S1
S2
SHAPE reactivity
>1.2
0.9-1.2
0.6-0.9
0.3-0.6
<0.3
1-33
1-60
PEMV2 5'89
C
C5'33
33mp
B
Relative luciferase activity
C5'33+C3'U
100 ± 10
C5'60+C3'U
107 ± 21
C5'201+C3'U
5 ± 3
C5'33+V
33 ± 6
P5'89+C3'U
19 ± 3
P5'89+P3'U
2303 ± 513
LUC
D
Relative luciferase activity
C5'33+C3'U
100 ± 10
C5'33mp+C3'U
48 ± 5
C5'33(A13U)+C3'U
44 ± 8
C5'33(A13G)+C3'U
28 ± 9
C5'33(C15U)+C3'U
31 ± 6

Figure 5

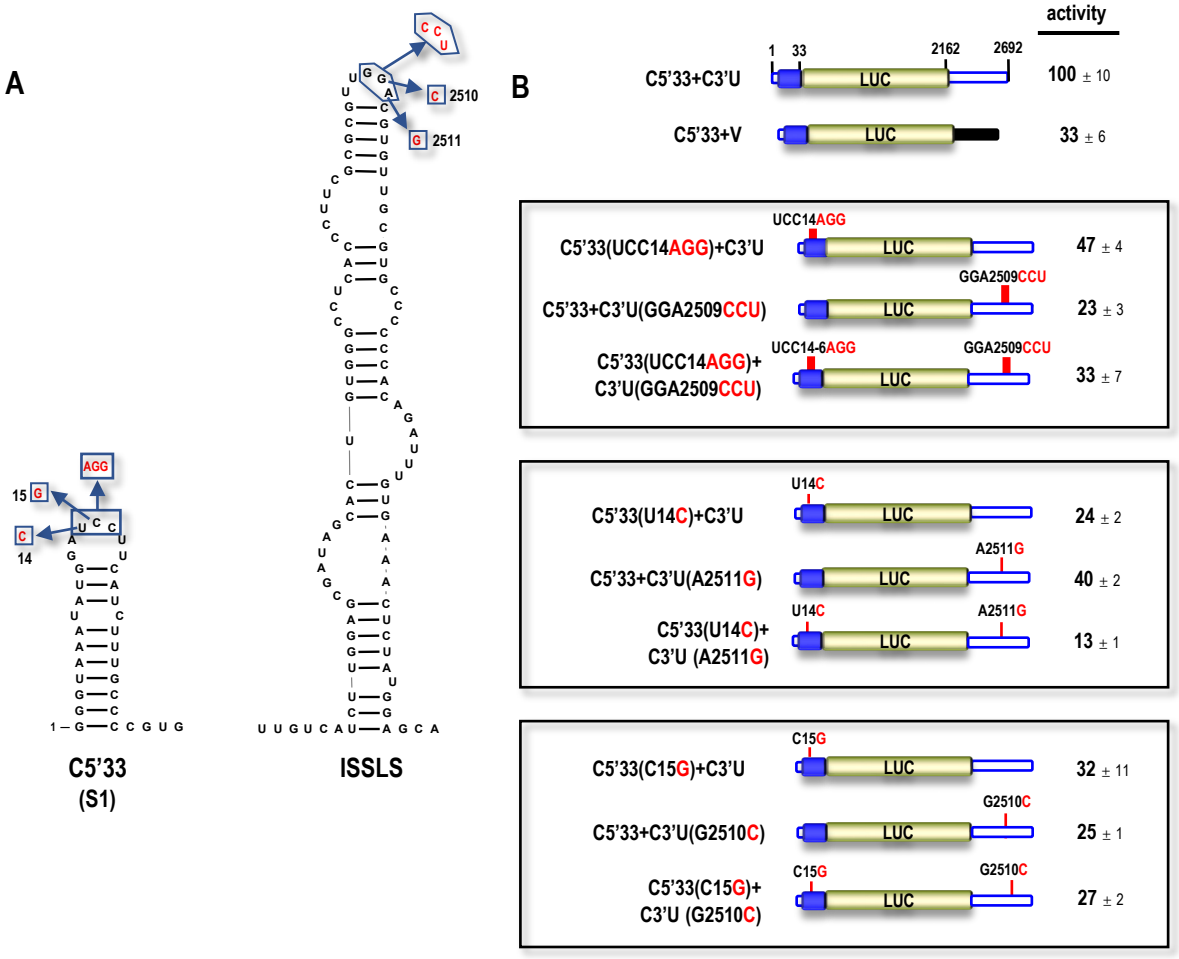

A
B
C5'33
(S1)
ISSLS
Relative luciferase activity
C5'33+C3'U
100 ± 10
C5'33+V
33 ± 6
C5'33(UCC14AGG)+C3'U
47 ± 4
C5'33+C3'U(GGA2509CCU)
23 ± 3
C5'33(UCC14AGG)+ C3'U(GGA2509CCU)
33 ± 7
C5'33(U14C)+C3'U
24 ± 2
C5'33+C3'U(A2511G)
40 ± 2
C5'33(U14C)+ C3'U (A2511G)
13 ± 1
C5'33(C15G)+C3'U
32 ± 11
C5'33+C3'U(G2510C)
25 ± 1
C5'33(C15G)+ C3'U (G2510C)
27 ± 2
LUC

Figure 6

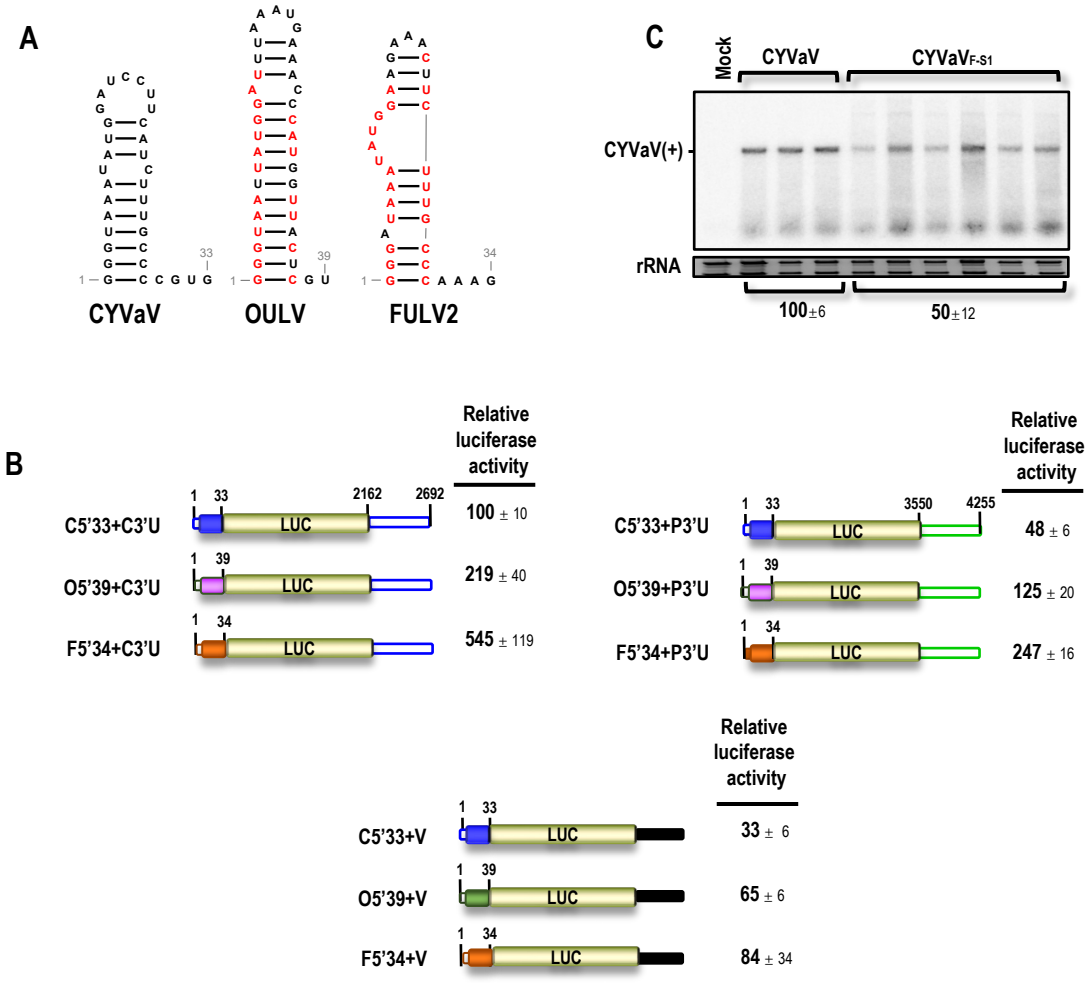

A
CYVaV
OULV
FULV2
C
Mock
CYVaV
CYVaV F-S1
CYVaV(+)
rRNA
100 ±6
50 ±12
B
Relative luciferase activity
C5'33+C3'U
O5'39+C3'U
F5'34+C3'U
LUC
100 ± 10
219 ± 40
545 ± 119
C5'33+P3'U
O5'39+P3'U
F5'34+P3'U
48 ± 6
125 ± 20
247 ± 16
C5'33+V
O5'39+V
F5'34+V
33 ± 6
65 ± 6
84 ± 34

Figure 7

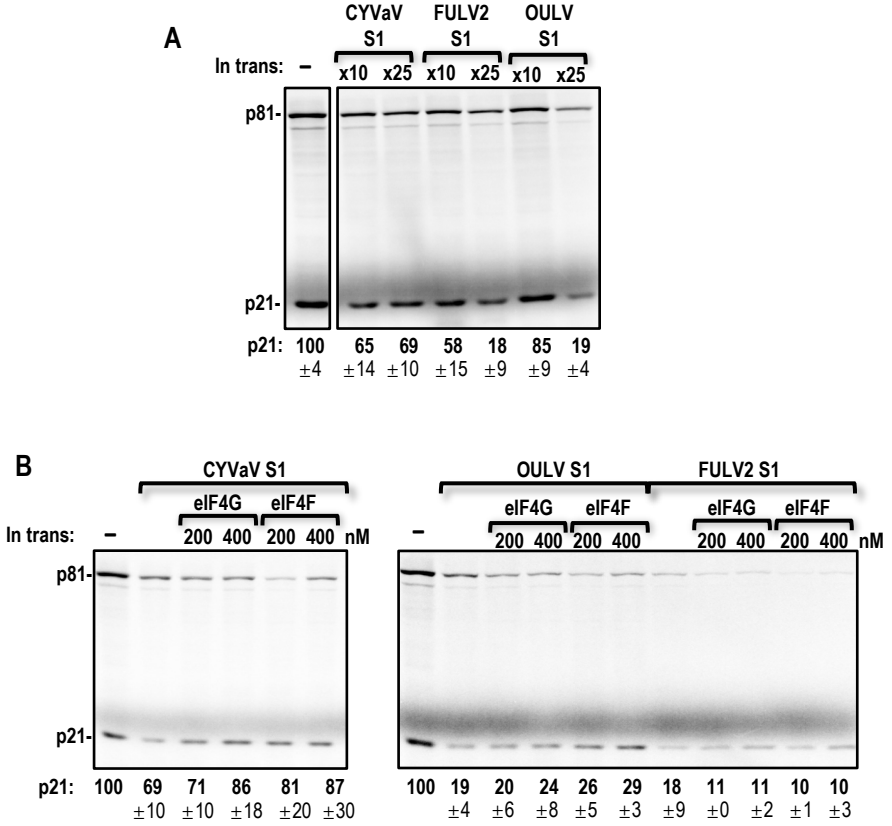
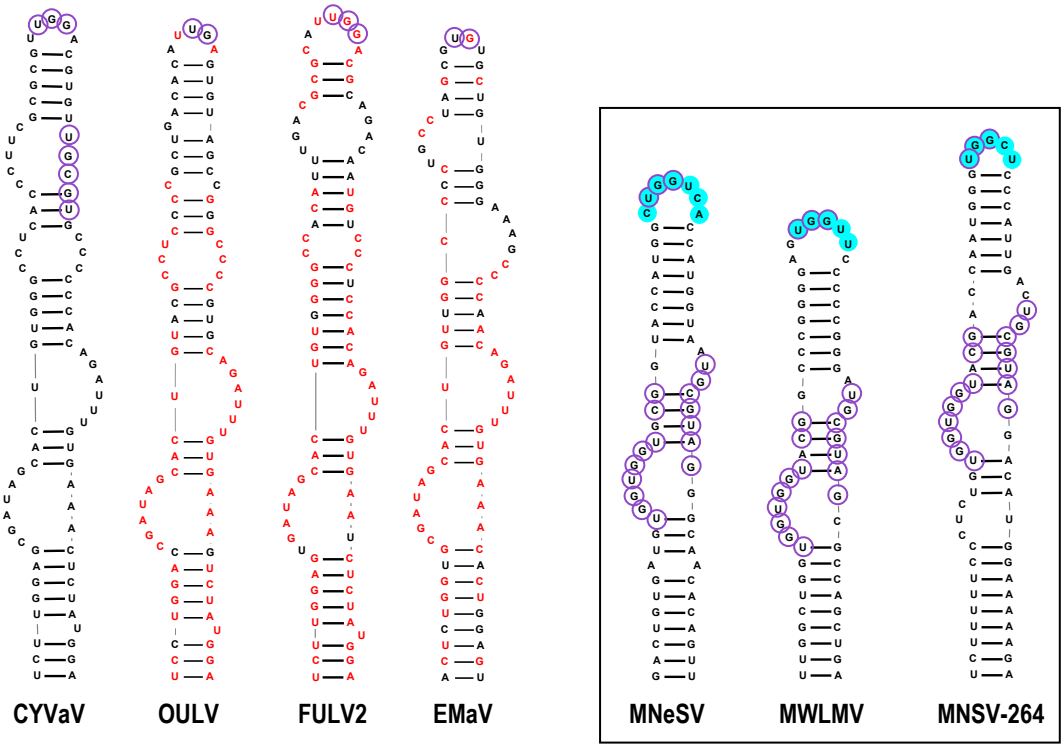

A
CYVaV S1
FULV2 S1
OULV S1
In trans: – x10 x25 x10 x25 x10 x25
p81-
p21-
p21: 100 65 69 58 18 85 19
±4 ±14 ±10 ±15 ±9 ±9 ±4
B
CYVaV S1
OULV S1
FULV2 S1
eIF4G eIF4F
In trans: – 200 400 200 400 nM
p81-
p21-
p21: 100 69 71 86 81 87
±10 ±10 ±18 ±20 ±30
100 19 20 24 26 29 18 11 11 10 10
±4 ±6 ±8 ±5 ±3 ±9 ±0 ±2 ±1 ±3

Figure 8


CYVaV
OULV
FULV2
EMaV
MNeSV
MWLMV
MNSV-264