
UGA stop codon readthrough to translate intergenic region of *Plautia stali* intestine virus does not require RNA structures forming internal ribosomal entry site

NOBUHIKO KAMOSHITA and SHIN-ICHI TOMINAGA¹

Department of Biochemistry, Jichi Medical University, Shimotsuke-shi, Tochigi-ken, 329-0498, Japan

ABSTRACT

The translation of capsid proteins of *Plautia stali* intestine virus (PSIV), encoded in its second open reading frame (ORF2), is directed by an internal ribosomal entry site (IRES) located in the intergenic region (IGR). Owing to the specific properties of PSIV IGR in terms of nucleotide length and frame organization, capsid proteins are also translated via stop codon readthrough in mammalian cultured cells as an extension of translation from the first ORF (ORF1) and IGR. To delineate stop codon readthrough in PSIV, we determined requirements of *cis*-acting elements through a molecular genetics approach applied in both cell-free translation systems and cultured cells. Mutants with deletions from the 3' end of IGR revealed that almost none of the sequence of IGR is necessary for readthrough, apart from the 5'-terminal codon CUA. Nucleotide replacement of this CUA trinucleotide or change of the termination codon from UGA severely impaired readthrough. Chemical mapping of the IGR region of the most active 3' deletion mutant indicated that this defined minimal element UGACUA, together with its downstream sequence, adopts a single-stranded conformation. Stimulatory activities of downstream RNA structures identified to date in gammaretrovirus, coltivirus, and alphavirus were not detected in the context of PSIV IGR, despite the presence of structures for IRES. To our knowledge, PSIV IGR is the first example of stop codon readthrough that is solely defined by the local hexamer sequence, even though the sequence is adjacent to an established region of RNA secondary/tertiary structures.

Keywords: stop codon readthrough; dicistrovirus; internal ribosomal entry site

INTRODUCTION

Stop codon readthrough, initially identified in RNA virus bacteriophage Q β coat protein translation (Weiner and Weber 1971), is a post-transcriptional mechanism to diversify the C termini of polypeptides at the level of translation. Subsequent studies revealed its prevalence in plant RNA viruses, but it has been reported in numerous eukaryotic viruses (Beier and Grimm 2001; Bertram et al. 2001) and in a wide range of host cellular genes, as in the case of rabbit β -globin (Geller and Rich 1980), *E. coli* surface antigen CS3 pili (Jalajakumari et al. 1989), large form of human myelin protein zero (MPZ) glycoprotein, L-MPZ (Yamaguchi et al. 2012), fungal peroxisomal enzymes such as *U. maydis* 3-phosphoglycerate kinase (PGK) (Freitag et al. 2012), or bovine and human VEGF-Ax (Eswarappa et al. 2014), albeit the activity of VEGF-Ax was controversial in a recent report (Loughran et al. 2017). The identification of genes accommodating readthrough has been accelerated by system-

atic analysis using bioinformatics or ribosomal profiling (Namy et al. 2002; Sato et al. 2003; Jungreis et al. 2011; Lindblad-Toh et al. 2011; Dunn et al. 2013; Loughran et al. 2014; Yordanova et al. 2018). Notable examples are peroxisomal enzymes directed via C-terminal peroxisomal targeting signal 1 (PTS1), and this targeting by stop codon readthrough was observed beyond species. For example, oxidoreductases such as human lactate dehydrogenase B (LDHB) and malate dehydrogenase 1 (MDH1), or glycolytic enzymes like *U. maydis* PGK and triosephosphate isomerase (TPI1) (Freitag et al. 2012; Schueren et al. 2014; Stiebler et al. 2014; Hoffhuis et al. 2016), are directed to peroxisomes by stop codon readthrough.

With regard to the mechanism involved, the UGACUAG heptamer was proposed as an optimal readthrough sequence through genome-wide analyses based on comparative genomics in mammals (Loughran et al. 2014) and a regression model in human cells (Schueren et al. 2014).

¹Present address: Japan Association for Development of Community Medicine (JADECOM), Tokyo, 102-0093, Japan

Corresponding author: nkamo@jichi.ac.jp

Article is online at <http://www.majournal.org/cgi/doi/10.1261/rna.065466.117>.

© 2019 Kamoshita and Tominaga This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://majournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

However, stop codon readthrough has also been identified in VEGF-Ax and rabbit β -globin, which have different contexts of UGAGCC and UGAGAU, respectively. Among the 42 human readthrough examples listed in a ribosomal profiling assay, only two use the context of UGACUA (Dunn et al. 2013). Even in cases that share the same UGACUAG heptamer sequence, the efficiency of readthrough is influenced by sequences in other regions of genes (Loughran et al. 2014; Schueren et al. 2014). Further delineation of *cis*-elements in each case would be necessary to understand how this mechanism is defined and/or controlled.

The study of RNA viruses has elucidated the *cis*-acting elements necessary for readthrough (Beier and Grimm 2001; Firth and Brierley 2012). Based on the local context of nucleotide sequences around the termination codon, Beier and Grimm (2001) classified viral examples of stop codon readthrough into three major types, in which type I uses UAGCAA (termination codon underlined), type II uses UGACGG or UGACUA in most cases, and type III typically uses UAGGGG/A/C or UAGGUA. Notably, these

classified local contexts are quite often followed by a variety of downstream RNA structures. Type III readthrough in animal gammaretroviruses at UAGGGA sequences is enhanced by conserved downstream pseudoknot structures (Wills et al. 1991; Feng et al. 1992), while UAGGUA and UAGGAA readthrough in plant luteovirus and tombusvirus, respectively, is stimulated by long-range RNA interactions (Brown et al. 1996; Cimino et al. 2011). Type II readthrough in animal alphaviruses or coltivirus is stimulated by RNA structures that extend up to 150 nucleotides (nt) downstream from the stop codon UGA, or a stem-loop structure of 9 nt downstream from UGA, respectively (Supplemental Fig. S1; Firth et al. 2011; Napthine et al. 2012). The roles of these different RNA structures in stop codon readthrough need to be clarified.

Plautia stali intestine virus (hereafter abbreviated as PSIV) is a species of the *Dicistroviridae*, a family of invertebrate RNA viruses known to use internal initiation in its intergenic region (IGR) to translate capsid proteins encoded in its second open reading frame (ORF2) (see Fig. 1A; Hellen and Sarnow 2001; Kieft 2008; Kamoshita et al.

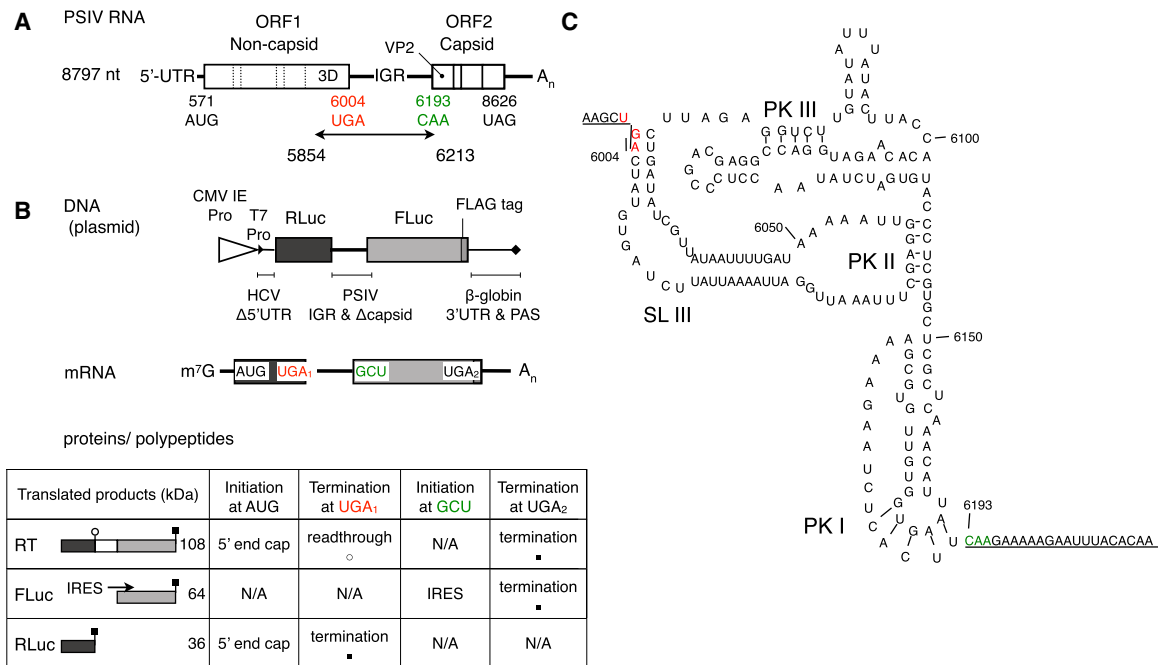


FIGURE 1. Structure of PSIV IGR and experimental system to analyze stop codon readthrough coincidental to internal initiation. (A) Schematic representation of genomic RNA of PSIV. Coding regions are boxed. Noncoding regions are shown with bold lines. The range of nucleotide sequence cloned into the expression vector is shown by the arrow. Nucleotide numbers are from registered sequence AB006531. Positions of the 3D replicase and VP2 capsid protein are shown along with the box. Dotted lines within the box represent expected cleavage sites of viral proteins. (B) Dicistronic expression unit of plasmid derived from pCI-neo with dicistronic mRNA transcribed, and with schematics of three polypeptides translated. Nucleotide and polypeptide regions for *Renilla* luciferase (RLuc) and firefly luciferase (FLuc) with a FLAG tag are shown with dark and light gray boxes, respectively. Promoters and polyadenylation site are depicted by triangles and a closed diamond, respectively. In the matrix for translated polypeptides, translation termination at UGA codon is shown by a filled square, while readthrough is shown by an open circle. Schematics of three translational products from full-length IGR, namely, readthrough polypeptides (RT), FLuc, and RLuc, are shown with numbers representing their expected molecular mass. (C) Primary nucleotide sequence of PSIV IGR presented with secondary and tertiary structural model (Kanamori and Nakashima 2001). Positions of stem-loop III (6005–6072) and three pseudoknots are shown. Termination codon of the first ORF (ORF1) and initiation codon of the second ORF (ORF2) are written in red and green, respectively. IGR, intergenic region; IRES, internal ribosomal entry site; N/A, not applicable; PK, pseudoknot; SL, stem-loop; UTR, untranslated region.

2009; Nakashima and Uchiumi 2009; Hertz and Thompson 2011). In the case of PSIV, IGR and the coding region downstream were also translated via stop codon readthrough in a mammalian assay system (Kamoshita et al. 2009), due to its unique features of frame composition and leakiness of the first ORF (ORF1) termination codon; that is, the next in-frame downstream termination codon of ORF1 is the ORF2 termination codon.

The presence of the internal ribosomal entry site (IRES), which has been both structurally and mechanistically characterized (e.g., Fernández et al. 2014; Ruehle et al. 2015; Abeyrathne et al. 2016; Bugaud et al. 2017; Johnson et al. 2017; Yamamoto et al. 2017), raises the question of how coincident RNA structures are involved in stop codon readthrough observed in PSIV. In PSIV, the termination codon UGA is followed by the IGR nucleotide sequence 6007–6012CUAUGU (Fig. 1C), overlapping with six of the seven nucleotides of the aforementioned optimal heptamer UGACUAG. Most of this sequence in PSIV is in a P1.1 paired helical structure in the large stem-loop III (SLIII; Fig. 1C; Supplemental Fig. S1), which is involved in ribosomal 40S subunit recruitment and stabilization of the IRES–80S complex (Jan and Sarnow 2002; Nishiyama et al. 2003; Pflugstein et al. 2010). How these two components contribute to PSIV readthrough has yet to be determined.

In this study, we evaluated requirements of *cis*-acting elements in stop codon readthrough in PSIV IGR using a dicistronic dual luciferase reporter system. Of note, 3' deletion mutants of IGR revealed that trinucleotide CUA, just downstream from the UGA termination codon, was the determinant of efficient readthrough in both mammalian cultured cells and cell-free translation systems. The importance of the primary sequence of CUA after the UGA termination codon was confirmed by nucleotide replacement assays. Chemical mapping of the IGR sequence of the most active UGACUA mutant, containing only the CUA trinucleotide as an IGR, showed that the UGACUA hexamer and the downstream sequence are unlikely to form RNA structures. Thus, readthrough in PSIV, which unwinds and translates flexible IGR IRES structures, is ensured by the hexamer nucleotide sequence even without stimulatory activities of downstream RNA structures. The potential involvement of the readthrough product in viral replication is discussed at the end of this paper.

RESULTS

Expression system

To analyze stop codon readthrough of PSIV IGR, the previously described mammalian expression vector pCdEchimUAAgaCAA21LucH (Kamoshita et al. 2009) was modified so that the first cistron of RLuc and the second cistron of FLuc were positioned in-frame via the wild-type PSIV IGR sequence 6004–6213 (Fig. 1A,B). The termination

codon UGA6004–6006 of the viral first ORF1 was used as a termination codon for the first cistron of RLuc, UGA₁. The initiation codon of PSIV capsid proteins was changed from CAA/glutamine to GCU/alanine to ensure constant translation from IRES (Kamoshita et al. 2009). The AUG initiation codon of the original FLuc cDNA was deleted as previously. To improve detection through immunoblotting, the C-terminal 6× His tag of the FLuc gene was changed to a FLAG tag. A dicistronic expression system prepared in this way gives three different translational products from one mRNA (Fig. 2B, lane 2): 108 kDa RT, 64 kDa FLuc protein, and 36 kDa RLuc protein. While translation of FLuc depends on IGR IRES, RT and RLuc are translated from the 5' end structure, cap (Fig. 1B).

Evaluation of upstream viral sequence and separation of stop codon readthrough from internal initiation

It has been reported that upstream sequences influence stop codon readthrough in mouse and human cells (Cassan and Rousset 2001; Loughran et al. 2014). Thus, the effects of the sequence at the 5' end of ORF1 in PSIV were evaluated here by inserting viral sequence 5854–6003 between the penultimate CAA931–933 codon of RLuc and the UGA termination codon (Fig. 2A, sample 1). As shown in Figure 2B, dicistronic mRNA with or without insertion was translated in RRL to form 114- (lane 1) or 108-kDa (lane 2) readthrough polypeptides, respectively. In cell-free experiments, expression levels of polypeptides were determined from the radioactivity of the translational products separated on the gel. Dividing the amount of readthrough product (RT) by the sum of RT and RLuc, the proportion of readthrough products of one mRNA was calculated using the following formula:

$$\left(\frac{V_{RT}}{n_{RT}}\right) / \left\{ \left(\frac{V_{RT}}{n_{RT}}\right) + \left(\frac{V_{RLuc}}{n_{RLuc}}\right) \right\},$$

where V and n represent the radioactivity value and the number of methionine residues, respectively. In Figure 2C, the mathematical mean of the relative value to the standard sample (sample 2) is shown in a graph to indicate the relative expression level of readthrough (open bars). The inclusion of the viral upstream sequence 5854–6003 (lane 1) gave a 1.3-fold increase of readthrough level (absolute proportions in lanes 1 and 2 in Fig. 2B were 3.6% and 2.8%, respectively). In cultured cells, readthrough products RT were detected in the immunoblot against C-terminal FLAG (Fig. 2D). Chemiluminescence quantified in ImageQuant LAS4000 system was directly compared between sample 1 and sample 2. As shown in Figure 2E (open bars), a similar degree of increase was observed in HEK293 cells. While constantly higher readthrough levels of sample 1 relative to those of sample 2 were calculated, we did not consider

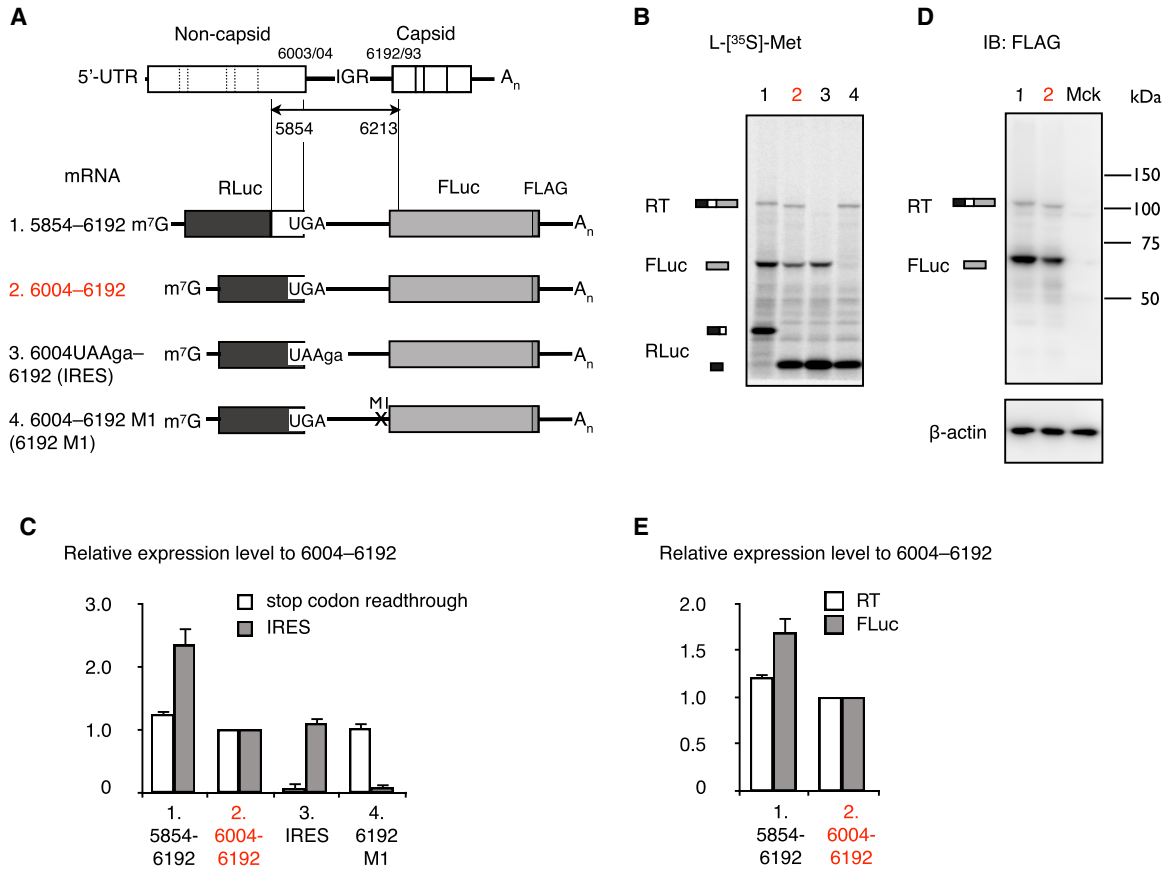


FIGURE 2. Effects of the viral sequence upstream of the UGA codon on the downstream translation of PSIV. (A) Dicistronic mRNAs with or without the viral ORF1 (open box) were used as a template for protein synthesis. Translation from dicistronic luciferase mRNA in which the wild-type full-length IGR sequence is preceded by an upstream viral sequence 5854–6003 (sample 1, 5854–6192) was compared with that of the standard sequence (sample 2, 6004–6192, written with red letters). Controls without stop codon readthrough (sample 3) or internal initiation (sample 4) were prepared by frameshift mutation after the stop codon or M1 mutation into pseudoknot I (PK I) that is essential for IRES activity, respectively. (B) Polypeptides labeled with L-[³⁵S]-methionine in rabbit reticulocyte lysate (RRL) were separated by 8% SDS-PAGE. (C) According to the radioactivity quantified from the gel, the relative expression levels of stop codon readthrough and IRES to those of sample 2 were determined as described in the text and are shown with open and shaded bars, respectively. Mean and SD of three experiments. (D) Readthrough products RT and IRES-dependent FLuc, expressed after transfection into HEK293 cells, were detected with anti-FLAG antibody. Loading control of the blot assayed with anti-β-actin antibody is shown at the bottom. (E) According to the chemiluminescence of the immunoblot obtained using anti-FLAG antibody, the expression levels of RT and FLuc products relative to those of sample 2 are expressed with open and shaded bars, respectively. Mean from the three independent experiments with SD. In B and D, schematic representations of separated polypeptides and marker positions are shown alongside the gel. Mck, mock; X in sample 4, M1 mutation to inactivate IGR IRES. Other abbreviations are as shown in Figure 1. A sample name and lane numbers acting as references are colored red.

that the difference of 1.3-fold was important and did not adopt the upstream viral sequence in the following assays.

With regard to internal initiation, IRES activity in RRL was calculated using the following formula:

$$\left(\frac{V_{FLuc}}{n_{FLuc}}\right) / \left(\frac{V_{RLuc}}{n_{RLuc}}\right).$$

IRES activity in cultured cells was evaluated from the immunoblot, as described above. In the presence of the upstream sequence, monocistronic FLuc was expressed 2.4- and 1.7-fold more efficiently in RRL and in cultured cells, respectively (Fig. 2B–E, shaded bars in the graph, compare FLuc in lanes 1 and 2), consistent with the findings in a previous report (Sasaki and Nakashima 1999). In honeybee

dicistrovirus, the 5'-adjacent stem-loop structure to IGR, predicted by Firth et al. (2009), stimulates IRES activity (Ren et al. 2012; Au et al. 2018). How the PSIV upstream sequence (sample 1) confers this sort of inter-ORF translational effects in stop codon readthrough and/or internal initiation remains to be clarified.

While internal initiation and readthrough are colocalized within wild-type IGR nucleotide sequence, their separation is necessary to investigate either of these two mechanisms individually. For this purpose, key nucleotides in IGR were manipulated via a molecular genetic approach (Fig. 2, samples 3 and 4). Insertion of the dinucleotide GA to induce frameshift combined with replacement of the termination codon (sample 3) or M1 mutation to interrupt

tertiary interactions of pseudoknot I (PK I), an essential component of IRES (sample 4; Kamoshita et al. 2009), was effective to abrogate either readthrough polypeptides (sample 3) or internal initiation (sample 4), without severely impairing the other mechanism. In the following assays, these two types of mRNAs were used as controls.

Termination codon mutation

The identity of the termination codon is important for classifying readthrough. Using *Renilla* luciferase activity as an internal control of transfection, the levels of stop codon readthrough from the three different termination codons were deduced in COS-1, HEK293, and HeLa cells (Fig. 3A). Internal initiation was eliminated by the deletion of 165 nt of PSIV IGR starting from 6028A (samples 1–4) in this assay. Readthrough levels at the UAG or UAA termination codon relative to that at the UGA codon were approximately one-fifth and one-fifteenth, respectively, in the three investigated cultured cell lines. In immunoblots for the lysate from COS-1 cells (Fig. 3B), the expression of 101 kDa readthrough polypeptides from the UGA termination codon (lane 3) was prominent compared with that for the other two. The absence of a 101 kDa product in immunoblots for control samples (lanes 4 and 5) confirmed that the 101 kDa product was actually produced by readthrough. Thus, in the context of PSIV IGR, among the three termination codons, UGA gave the highest level of readthrough products, followed by UAG and then UAA.

Deletion of IGR

To clarify the relationship between stop codon readthrough and the IGR sequence in PSIV, a series of 3' deletion mutants were prepared (Fig. 4A) and readthrough

levels of mutants were compared with that from full-length IGR with M1 mutation (sample 2, see also Fig. 2). Downstream deletions up to positions 6147, 6072, and 6027 (samples 4–6) were designed to match the boundaries of expected secondary or tertiary structures in IGR, under the restriction that the numbers of deleted nucleotides were multiples of three.

As described and shown in Supplemental Results and Supplemental Figure S2, specific luciferase enzymatic activities of mutant readthrough polypeptides varied according to the length of deletion. Therefore, we used radioactivity and chemiluminescence in immunoblots to evaluate the readthrough levels of deletion mutants in cell-free assays and in cultured cells, respectively (Fig. 4B–E). In a series of assays in cultured cells, the expression of IRES-dependent FLuc (Fig. 4D, sample 1) was low, for unknown reasons, and 64 kDa Fluc polypeptides from sample 1 were not detected in immunoblot using anti-FLAG antibody (Fig. 4D, FLAG, lane 1).

The amounts of radiolabeled readthrough polypeptides, separated as molecular masses over 100 kDa in lanes 1–7 (Fig. 4B), were compared using the level obtained for the M1 mutant as a reference (Fig. 4B,C, sample 2). Assays were repeated at two different final potassium concentrations. In both conditions, the level of stop codon readthrough gradually increased as the size of the deletion in the IGR increased to more than double the reference level, until the entire IGR was deleted, except for the first three nucleotides (Fig. 4B,C, samples 2–7). However, when deletion proceeded from 6009 to 6006 (sample 8), the level of readthrough markedly decreased by 14- and 16-fold at the potassium concentration of 100 mM and 150 mM, respectively.

In immunoblots using anti-*Renilla* luciferase or FLAG antibody (Fig. 4D,E), the expression of readthrough

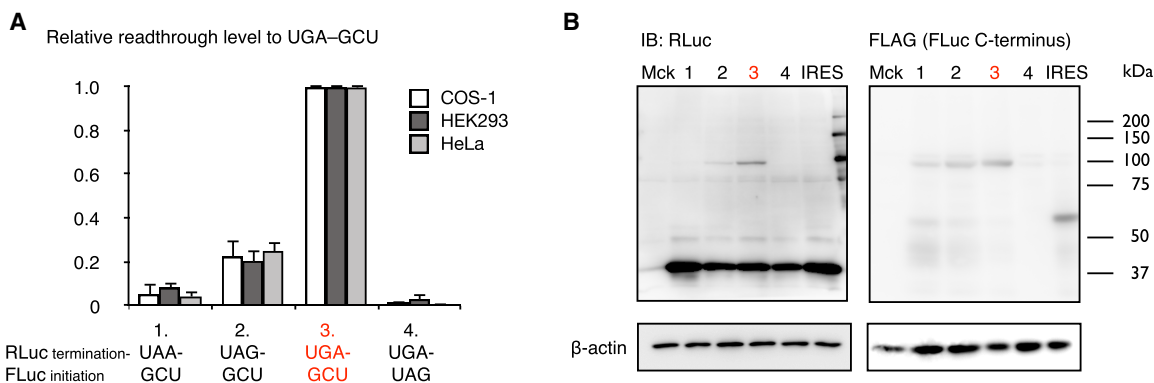


FIGURE 3. Effects of termination codon mutations on stop codon readthrough. (A) The UGA termination codon of PSIV IGR (sample 3) was changed to UAA (sample 1) or UAG (sample 2). To inactivate IRES, 165 nt of the IGR portion starting from PSIV nucleotide 6028A were deleted. In sample 4, the GCU codon following IGR was changed to UAG. The stop codon readthrough level of each sample was determined as the ratio of firefly to *Renilla* luciferase enzymatic activities after 48 h of transfection into indicated cells and is shown as the percentage relative to that of UGA-GCU (sample 3). Mean and SD from four different transfections are shown. (B) Polypeptides expressed in COS-1 cells after 60 h of transfection were immunoblotted with anti-RLuc or anti-FLAG antibody. Loading controls are shown as in Figure 2. Polypeptides expressed from IRES were run as a control in lane 5.

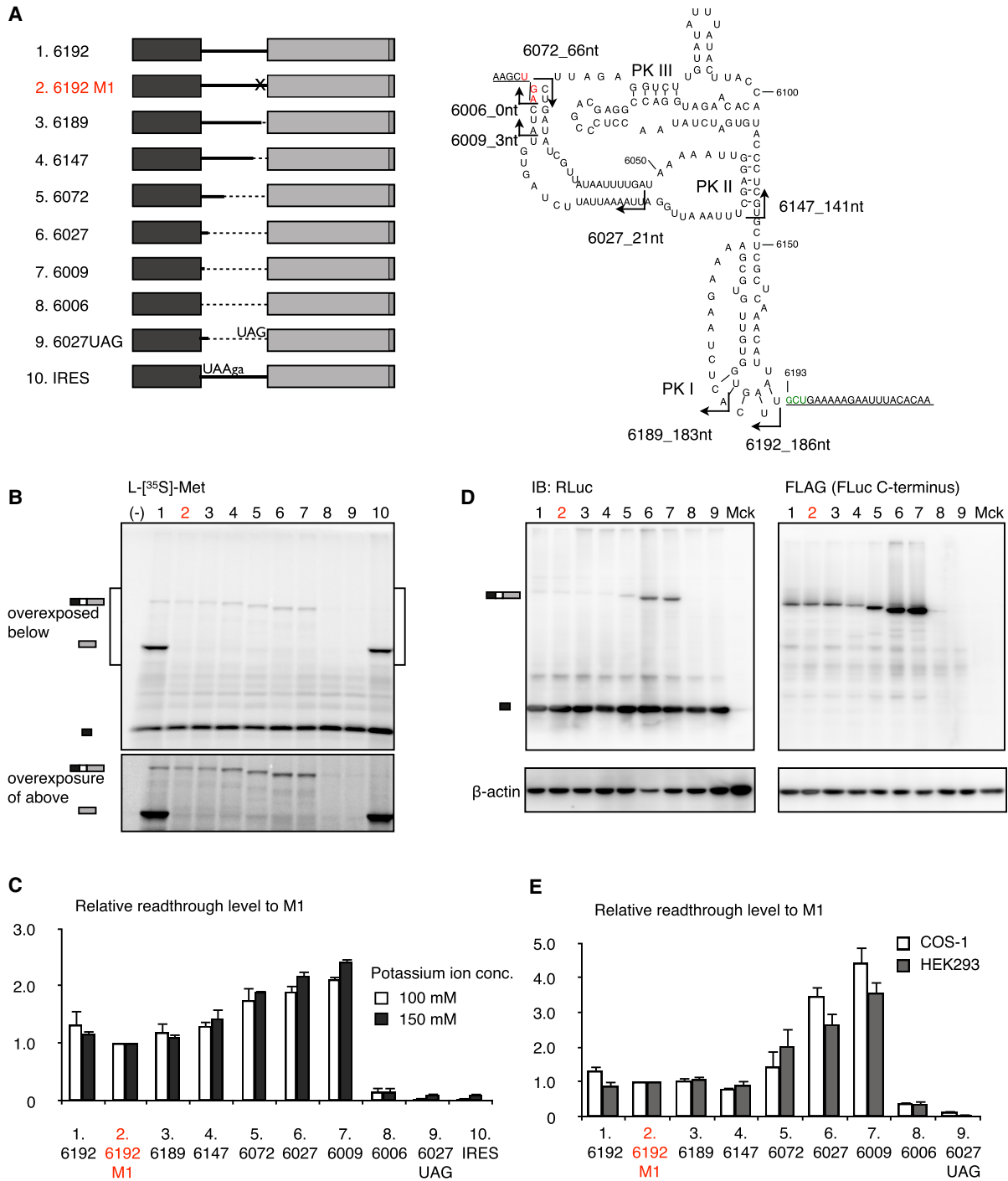


FIGURE 4. Effects of deletions from the 3' end of IGR on UGA stop codon readthrough. (A) IGR cloned into the dicistronic unit was deleted from the 3' end of full-length IGR (sample 1). Schematics of mutants (*left*) and positions of deletion alongside the nucleotide sequence in predicted structure (*right*). In the panel at *right*, numbers following an underscore symbol denote the lengths of IGR nucleotides remaining. (B) In RRL, polypeptides translated from dicistronic mRNAs with partial (lanes 3–7) to complete (lane 8) deletion of IGR were analyzed and compared with those from control mRNAs (lanes 2, 9, and 10). Shown *below* is the overexposed image of RT and FLuc products, with its corresponding part in the original image indicated with rectangles alongside. Potassium concentration was adjusted to 150 mM with KCl. Representative data among three reactions are shown. (C) Proportion of stop codon readthrough was determined from the radioactivity of each polypeptide and normalized to that of 6192M1 (sample 2). Values obtained at potassium concentrations of 100 and 150 mM are shown with open and dark gray bars, respectively. (D) Plasmids corresponding to A were transfected into COS-1 cells and expressed polypeptides were immunodetected in blots using anti-RLuc or anti-FLAG antibody. Loading controls against β-actin are shown *below*. (E) Relative expression levels determined from the chemiluminescence of readthrough polypeptides in COS-1 and HEK293 cells, normalized to that of 6192M1 mutant (lane 2), are shown with open and shaded bars, respectively. Mean from three different experiments.

polypeptides peaked again in mutant 6009 (sample 7) and expression was barely detectable in mutant 6006 (sample 8). In mutant 6009 (Fig. 4E, sample 7), normalized relative chemiluminescence increased 4.4- and 3.6-fold in COS-1 and HEK293 cells, respectively. When 6007–6009CUA was deleted and the GCU codon was directly connected to the UGA codon (sample 8, mutant 6006), the relative expression level was diminished to 0.38 and 0.36 in COS-1 and HEK293 cells, respectively.

Thus, almost none of the nucleotides in IGR were necessary, except for a trinucleotide 6007–6009CUA both in the mammalian cell-free system and in cultured cells. Compared with cell-free assays, readthrough polypeptides from mutants 6072, 6027, and 6009 (samples 5–7) were detected more in cultured cells. From this result, it is difficult to find any stimulatory contributions toward readthrough in nucleotides deleted in these mutants and we did not pursue this issue.

The result revealed by deletion mutants that the only element essential for readthrough is the 3 nt CUA strongly suggests that structures within IGR that form IRES are not necessary for maximal readthrough. Deletion of IGR up to position 6009 did not prominently abolish readthrough activity also in transdirect insect cell extract, which is derived from *Spodoptera frugiperda* 21 cells (Supplemental Fig. S3), suggesting that the observation is not an artifact of mammalian systems.

Structural probing of 3' deletion mutants

To confirm the absence of any unexpected structures in mutant 6009, structural mapping around the stop codon was performed using the base-modifying reagents and the SHAPE (selective 2'-hydroxyl acylation analyzed by primer extension) reagent, 2-methylnicotinic acid imidazole (NAI) (Fig. 5, entire gel images are shown in Supplemental Fig. S4).

If the UGACUA sequence does not form base pairs, or is unstructured, the N1 atoms of adenine A6006 and A6009 will be methylated by dimethyl sulfate (DMS) and reverse transcription from the downstream primer, labeled with ³²P, will be inhibited at these positions. Note that the products of primer extension inhibition obtained after the reaction with base-modifying reagents are 1 nt shorter than those of dideoxy sequencing shown in parallel (Ziehler and Engelke 2001). Similarly, if RNA is single-stranded at this position, N3 of U6004 and U6008 will be alkylated by *N*-cyclohexyl-*N'*-β-(4-methylmorpholinium) ethylcarbodiimide *p*-tosylate (CMCT) and detected. As shown in Figure 5A, two adenylates and two uridylates in the UGACUA hexamer sequence in mutant 6009 were reactive with DMS and CMCT, respectively, together with other bases in downstream capsid-coding nucleotides. To rule out the possibility that the reagents were modifying all of the nucleotides present, even if a structure is present, mu-

tant 6072, which contains a sequence predicted to form a large stem-loop structure called SLIII (Sasaki and Nakashima 1999; Kanamori and Nakashima 2001), was analyzed under the same conditions. As shown in the gel and schematic presentation (Fig. 5B), most of the nucleotides predicted to be located in the L1.2 loop and L1.1 bulge of SLIII in mutant 6072 were modified. In contrast, ribonucleotides A6006, U6008, and A6009 were not reactive in mutant 6072. Therefore, these three nucleotides form base pairs, as predicted in the model. The presence or the absence of the base pairs was appropriately mapped by DMS and CMCT. The UGACUA hexamer in mutant 6009 is not base-paired, while the same sequence in mutant 6072 forms a paired double helix at the base of SLIII.

To investigate the IGR structure further, the SHAPE reagent NAI was used. As shown in Figure 5A, 2'-hydroxyl groups in the IGR nucleotides CUA and the downstream capsid-coding nucleotides in mutant 6009 were modified.

Nucleotide replacement assay of 6007–6009CUA

To further evaluate the requirement for the primary nucleotide sequence 6007–6009CUA following the UGA termination codon, the GCU/Ala codon in mutant 6006 was changed to CUA/Leu and its related codons (Fig. 6; Supplemental Table S1). Readthrough levels in COS-1 or HEK293 cells were determined from the level of firefly luciferase activity relative to *Renilla* luciferase activity and expressed as a percentage of that of the CUA codon. In both cell types, any single-nucleotide replacement from CUA resulted in a decrease in the readthrough level (Supplemental Table S1), consistent with previous reports for cellular genes (Loughran et al. 2014; Schueren et al. 2014; Stiebler et al. 2014; Hofhuis et al. 2016).

Readthrough using the UGA termination codon has been reported in other contexts differing from UGACUA in 2 or 3 nt positions. The effects of mutations from CUA to GAU (rabbit β-globin) and GCC (bovine VEGF-Ax) were investigated, together with control GCU and UAG codons. The expression from these codons was <5% (Supplemental Table S1), and reported readthrough activity was not reproduced in HEK293 cells.

The significance of the CUA sequence following the UGA termination codon was further investigated using silent mutations to the other five leucine-coding codons (Fig. 6). As indicated in immunoblots for HEK293 cellular lysates (Fig. 6B), readthrough products were most prominently expressed when CUA followed UGA (sample 1). Mutations at the third position of the CUA codon (samples 2–4) diminished the expression levels of readthrough products to 26.2%–39.3%, but they were still detected in the blot using anti-*Renilla* luciferase antibody. Two mutants, in which the first base of the CUA codon was changed to uridylate, showed levels of readthrough products of <5%. Readthrough polypeptides of these two

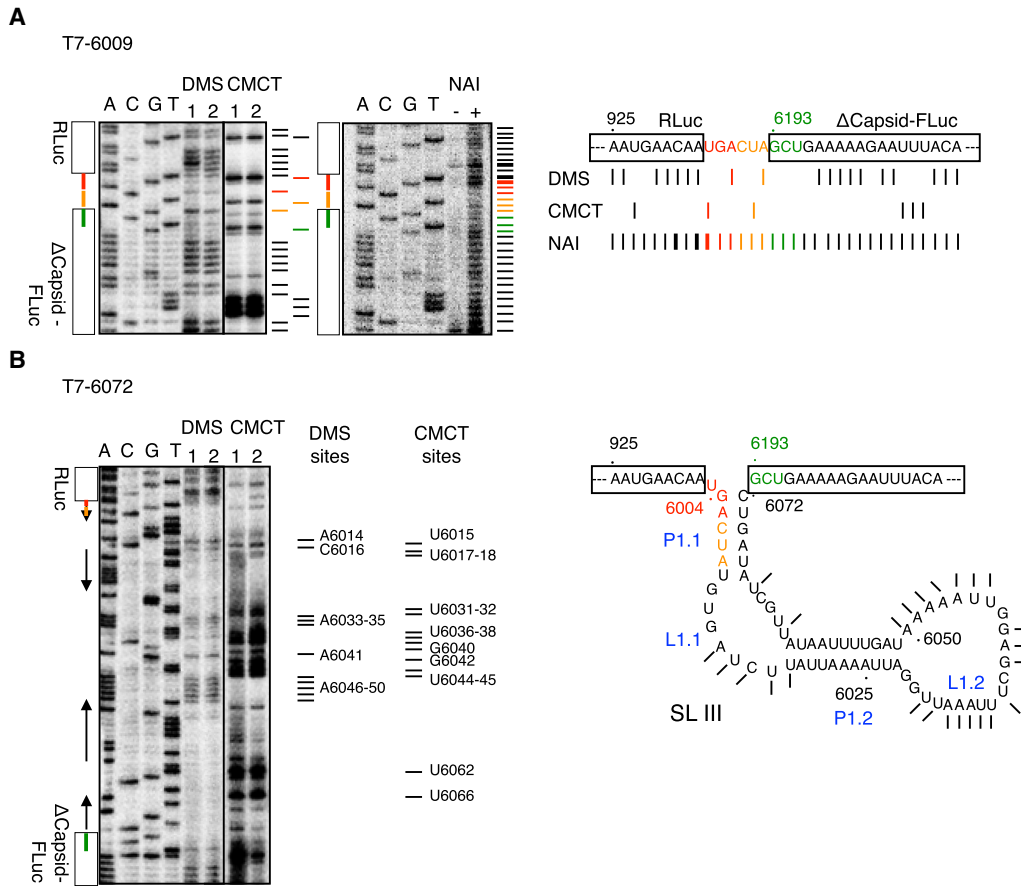


FIGURE 5. Chemical mapping of RNA structures downstream from UGA₁ codon in 3' deletion mutants 6009 and 6072. IGR RNA structures of representative deletion mutants used in Figure 4 were chemically mapped with DMS (60 mM), CMCT (25 mM), or NAI (100 mM). (Left) Chemically modified ribonucleotides in mutant 6009 (A) or 6072 (B) were determined by the inhibition of reverse transcription from the downstream ³²P-labeled primer hybridized to FLuc. Two different preparations of RNAs were modified in DMS and CMCT. The positions of modified nucleotides are indicated with lines at the side of the gel. Strong modifications observed in NAI are depicted with thick lines. (Right) In the schematics, the positions of landmark nucleotides are shown with dots and numbers. Modified nucleotides are indicated with bars (A) and lines (B). Codons 6004UGA, 6007CUA, and 6193GCU are color-coded with red, orange, and green, respectively. The positions of coding sequences for RLuc and Δcapsid-FLuc are boxed. In B (left), the positions of predicted paired regions in SLIII are indicated by arrows on the side of the gel. The positions of CMCT modification were determined from the short-exposure image (see Supplemental Fig. 54C).

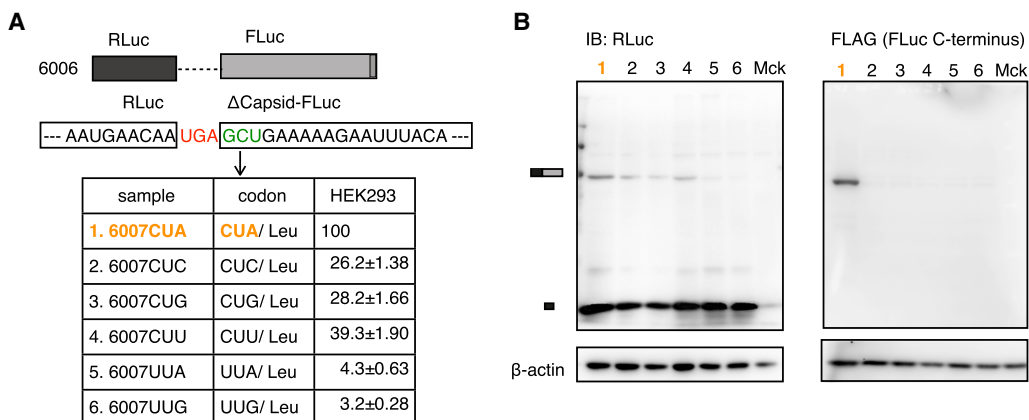


FIGURE 6. Synonymous mutations for 6007–6009CUA. (A) Synonymous mutations coding for CUA/leucine were introduced into the GCU/alanine Δcapsid initiation codon of mutant 6006. Readthrough levels in HEK293 cells normalized to the CUA wild-type sequence, indicated with orange, are shown as mean ± SEM from five different experiments. (B) Polypeptides expressed in HEK293 cells were analyzed with immunoblots using anti-RLuc or anti-FLAG antibody. Loading controls obtained in immunoblots for β-actin are shown below.

mutants were barely detectable in immunoblots (samples 5 and 6).

Together with the results shown in Supplemental Table S1, these findings show that, for maximal expression in UGA readthrough, all of the trinucleotide sequence of CUA is necessary, among which cytidylate at the fourth position of the UGACUA hexamer and uridylate at the fifth position contribute more than adenylate at the sixth position.

Efficiency of stop codon readthrough

The efficiency of readthrough in a full-length IGR was estimated from the expression level of reporters in cultured cells using M1 mutants. In these mutants, no internal initiation to produce monocistronic firefly luciferase occurs (Fig. 2, sample 4), and firefly luciferase activities are solely attributed to the amount of readthrough polypeptides (Fig. 7A). To normalize the efficiency of transfection, firefly luciferase activity was divided by cotransfected EGFP fluorescence as an internal control, instead of *Renilla* luciferase activity. This is based on our observation that specific enzymatic activity of luciferase protein became different from that of the original form when polypeptides were fused by readthrough (data not shown). As a sense codon, GGA was used to maintain base pairs in the P1.1 region of IGR IRES. Normalized firefly luciferase activity derived from the sense codon GGA was more than ten times higher than that from UGA stop codon readthrough (Fig. 7B), and readthrough efficiency was calculated to be 7.9% and 8.6% in COS-1 and HEK293 cells, respectively. Compared with the findings in previous reports, values were closer to those in assays using dual luciferase systems (Firth et al. 2011; Loughran et al. 2014, 2017) than those using other reporters (Schuere et al. 2014; Hofhuis et al.

2016). Besides the difference of reporters, this would be due to several other differences among the reported studies such as assay procedures, experimental conditions, cell types analyzed, and nucleotide sequences inserted between reporters.

An important observation from this and the previous assay using full-length IGR (Figs. 2, 7) is that readthrough polypeptides directed by the local hexamer sequence UGACUA were efficiently expressed in cultured cells, even in the presence of structured IGR IRES, which lacks stimulatory activity for readthrough (Fig. 4).

DISCUSSION

We defined *cis*-acting elements of PSIV stop codon readthrough mostly using mammalian systems and revealed that almost none of PSIV IGR, which forms IRES, was necessary for this. The determinant for readthrough in PSIV was shown to be the CUA trinucleotide immediately downstream from the UGA termination codon. This is the minimal element because one codon is the unit for the readthrough assay, and both deletion (Fig. 4) and systematic nucleotide replacement (Fig. 6; Supplemental Table S1) of this CUA sequence drastically impaired stop codon readthrough.

Deletion mutants were necessary in the investigation of *cis*-elements for stop codon readthrough. However, we encountered a problem with fused reporter genes, in which, according to the length of deletion, fused firefly luciferase genes gave different specific enzymatic activities (Supplemental Fig. S2). Here, we used physical separation and quantification of translational products in autoradiographs or immunoblots, which was helpful for evaluating the amounts of translated reporters (Figs. 2, 4).

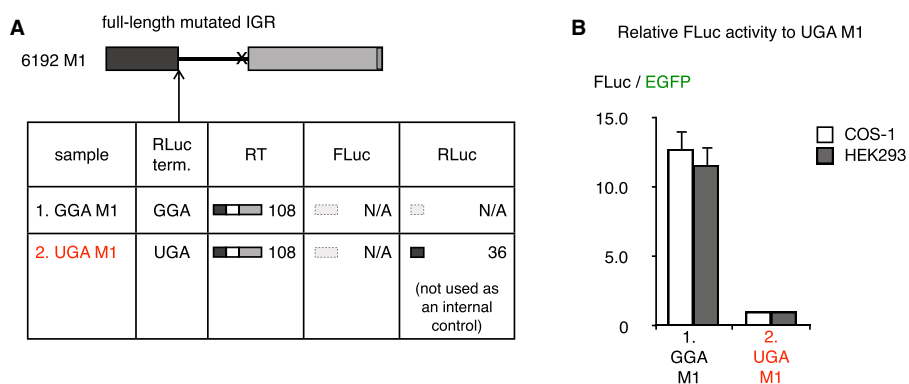


FIGURE 7. Readthrough efficiency in cultured cells. (A) Expression of 108 kDa RT polypeptides from the UGA codon (sample 2, indicated in red) was compared with that of the corresponding products from GGA sense codon control (sample 1). IRES-dependent translation was inactivated by M1 mutation, shown with an X, in pseudoknot I. In the matrix, polypeptides expressed or not expressed are shown with a solid or dotted box, respectively. (B) FLuc activity 48 h after transfection was normalized to coexpressed fluorescence of EGFP used as an internal control. Relative FLuc activities normalized to that expressed from the UGA codon (sample 2) were determined from three independent transfections. Mean and SD in COS-1 and HEK293 cells are shown with open and shaded bars, respectively. N/A, not applicable.

Role of RNA structures and sequences in PSIV readthrough

Most cases of readthrough identified in animal viruses to date are enhanced by downstream RNA structures, collectively called stimulators (Firth and Brierley 2012). Results that RNA structures in PSIV IGR, intensively studied as IGR IRES, lack stimulator activity were at first unexpected to us. The unlikely possibility that short IGR is involved in unforeseen long-range structures within dicistronic mRNA, as reported in luteovirus or alphavirus (Brown et al. 1996; Firth et al. 2011), was ruled out by chemical mapping (Fig. 5). If a stimulator were present, the level of readthrough polypeptides would decrease upon its deletion (Wills et al. 1991; Firth et al. 2011; Naphthine et al. 2012). The sole stimulatory element in our assay was 6007–6009CUA, in both cell-free systems and cultured cells (Fig. 4).

One difference between reported stimulators and full-length PSIV IGR is the complexity of the structure. During translational elongation, ribosomes need to unwind RNA structures, if present. RNA structures including pseudoknots were previously shown to decrease the rate of translation in different assay systems derived from different sources (Tu et al. 1992; Somogyi et al. 1993; Chen et al. 2013). Considering the structure of IGR IRES, namely, an abundance of base pairs with six helical portions and three pseudoknots (Kanamori and Nakashima 2001; Pflugstein et al. 2006), the results of nucleotide deletion from mutants 6147 to 6009 (Fig. 4), which showed gradual increases in the level of readthrough polypeptides, may reflect the reduction of the general inhibitory effects of complex RNA structures against ribosomal elongation. However, other possibilities cannot be ruled out. Increased readthrough in samples 5–7 can be explained by the deletion of the functional RNA region of IRES to recruit ribosomes (see Supplemental Fig. S5; Nishiyama et al. 2003), which could be a hindrance of readthrough if present. Part of the activity of the mutant 6009 can be attributed to the sequence of the VP2-coding region, which may enhance readthrough only when located in the proximity of the ORF1 stop codon, as evidenced in several mutations introduced into the spacer in type III readthrough in Moloney murine leukemia virus (Honigman et al. 1991; Feng et al. 1992; Wills et al. 1994).

Another difference of PSIV structure is its location relative to the termination codon. In stimulators for type II readthrough identified to date, the structure starts from the 12th or 16th nucleotide from the uridylylate of the UGA codon (9 or 13 nt downstream from the termination codon, Supplemental Fig. S1). In PSIV IGR, SLIII starts from the second nucleotide, or internal guanylylate, of the termination codon UGA. When the ORF1 UGA codon is decoded, base pairs within the P1.1 region of SLIII must be unpaired and the entire structure of SLIII will no longer be present at the time readthrough occurs. There will be a structure

formed with the P1.2 helix, but the stimulator activity for readthrough was not detected in the comparison between 6072 and 6009 in our assay (Fig. 4). In the case of barley yellow dwarf virus or carnation Italian ringspot virus, the viral sequence 700–750 or ~3500 nt downstream from the stop codon affects readthrough, respectively (Brown et al. 1996; Cimino et al. 2011; Xu et al. 2018). While such long-range interactions with downstream viral sequences were not investigated, our assay within the range of IGR led to the conclusion that IGR SLIII is not a stimulator.

Within the UGACUA sequence identified in PSIV readthrough, the importance of cytidylate at the fourth position on readthrough was previously pointed out (Li and Rice 1993). This would be the reflection of the weak termination signal formed by the UGAC sequence, as pointed out by others (Brown et al. 1990; McCaughan et al. 1995; Loughran et al. 2014). Recent cryo-EM-based studies on the mammalian termination complex (Brown et al. 2015; Mathiesl et al. 2015), showing the presence of four nucleotides in the ribosomal A site, would be structural proofs of this phenomenon. The effect of uridylylate at the fifth position is enigmatic. The fifth uridylylate was almost as unexchangeable as the fourth cytidylate in our nucleotide replacement assays (Supplemental Table S1). Compared with these two positions, the effect of adenylate at the sixth position was mild.

In PSIV IGR, as long as UGACUA hexamer is present, a certain level of readthrough occurred, even in the presence of the large RNA structure of IRES (Figs. 2, 7), which can intrinsically inhibit translational elongation, as discussed above. The number of examples of readthrough using the UGACUA sequence is increasing, as identified by Loughran and colleagues (Loughran et al. 2014, 2018), including metabolic enzymes targeted to peroxisomes (Schueren et al. 2014; Stiebler et al. 2014). It is unlikely that the 3'-UTRs of all of these mRNAs are equipped with stimulatory RNA structures or are completely free from RNA structures. We propose that, for one or some of these cellular mRNAs, an essential level of readthrough would be secured by the hexamer sequence even in the presence of downstream RNA structures lacking stimulator activities, as observed in our assay.

Conversely, the previous reports of readthrough in rabbit β -globin (local hexamer UGAGAU) and bovine VEGF-Ax (UGAGCC) could not be reproduced in our cell-based system (Supplemental Table S1). In both cases, hexamer sequences differ from UGACUA by three nucleotides. The readthrough activity of these two genes does not need to be ubiquitous or essential and would be up-regulated by *trans*-acting factors (Eswarappa et al. 2014).

Role of PSIV readthrough in viral replication

Within 15 dicistroviruses, the UGACUA motif identified in this study is only present in PSIV (Supplemental Fig. S6),

while an in-frame termination codon, which abolished the extension of readthrough products (Fig. 3, sample 4; Fig. 4, sample 9), is present within the downstream three codons in five viruses (HiPV, BQCV, HoCV-1, TSV, and KBV). Therefore, readthrough to translate capsid proteins will never be conserved. The observation of readthrough in an insect cell-free system (Supplemental Fig. S3), however, suggests that the translation of amino acid sequence encoded in IGR (IGR polypeptides, hereafter abbreviated as IGRp) does occur in infected cells.

IGRp will be composed of 63 amino acids (Supplemental Figs. S5, S7), with its expected molecular mass and pI of 7.5 kDa and 9.6, respectively. Besides methionine, all other 19 amino acids are present. Overlapping with the predicted transmembrane α -helix in the N-terminal region, it has homology with the prokaryotic transmembrane protein CydD thiol exporter (Supplemental Fig. S5; Supplemental Table S2). One possible function of readthrough is to increase the affinity of 3D replicase (Supplemental Fig. S8) to the membrane structure where viral RNAs are synthesized (Flint et al. 2015).

Since IGRp and ORF2 capsid proteins (ORF2p) share the same reading frame, the following three polypeptides will be translated in PSIV: ORF1p, ORF2p, and ORF1p-IGRp-ORF2p derived from 5' IRES, IGR IRES, and 5' IRES followed by readthrough, respectively. The uniqueness of PSIV readthrough resides in that C-terminally diversified ORF1p is N-terminally fused to the otherwise functional ORF2p via IGRp. Dicistoviral polypeptides are thought to be proteolytically processed to each protein (Bonning and Miller 2010; Nakashima and Ishibashi 2010). As a result, the unique readthrough in PSIV may produce a unique viral protein, 3D-IGRp-VP2 (Fig. 1A), which has no counterparts in any other viruses in the order *Picornavirales*. The structures of virions are similar between picornavirus and dicistovirus (Tate et al. 1999). The N-terminal attachment of 3D-IGRp to VP2 will be too large (594 vs. 254 amino acids), if present, to be assembled into the virion. Localization to the membrane evoked by the predicted membrane affinity in IGRp may alleviate the interference with assembly but disturb the supply of viral proteins. Cleavage within or around IGRp, therefore, will be the most effective way for PSIV to replicate in the presence of readthrough.

Dicistoviral VP2 is the most upstream capsid protein translated by IGR IRES. In contrast to the picornaviral VP2, which is encoded downstream from VP4 and the N terminus of which is strictly determined by the self-cleavage inside the virion, there is remarkable heterogeneity of the amino acid sequence and length at the N termini (Supplemental Fig. S10). Virion crystal structures of dicistovirus revealed the extension of the VP2 N-terminal region to the outside of the pentamer unit (Tate et al. 1999), where chances of cleavage will be present. According to the homology search, six candidate sites, which in-

clude unique initiator glutamine of PSIV, are present (Gln1-Glu2; for details, see Supplemental Discussion and Supplemental Fig. S9). Whether and how these sites are cleaved needs an experimental proof. In case neither of the candidate sites is cleaved, readthrough will be a bystander. Note that 3D-IGRp-VP2 will be localized to the membrane and other detrimental effects of readthrough will be tolerated.

The development of a cultivation system together with an infectious cDNA clone, which is currently unavailable in PSIV but has been reported in another dicistovirus (Kerr et al. 2016), will advance the virological understanding of PSIV readthrough. The roles of PSIV readthrough should be addressed via a multidisciplinary approach using virological, biochemical, bioinformatic, and any other methods available.

MATERIALS AND METHODS

Plasmids

Plasmids used in this study were prepared from the previously described mammalian expression vector pCdEchimUAAgaCAA21 LucH (Kamoshita et al. 2009) and were purified using JetStar 2.0 Plasmid Midiprep Kit (Genomed/Veritas). Mutations were introduced by standard primer-based PCR using KOD DNA polymerase (ToYoBo) and sequences were verified with ABI PRISM 310 Genetic Analyzer (Applied Biosystems). A more detailed description is provided in the Supplemental Material.

Transfection

COS-1, HEK293, or HeLa cells, plated in a 24-well plate the day before, were transfected with a mixture of 0.4 μ g of plasmid and 0.375 μ L of Lipofectamine 2000 (Life Technologies) assembled in 50 μ L of Opti-MEM (Life Technologies). Medium was changed to complete medium and further incubated for 48 h at 33°C. For immunoblot, cells plated in a six-well plate were transfected with an increased amount of DNA/Lipofectamine 2000 complex in proportion to surface area of the well. The duration of incubation was extended to 60 h.

Immunoblotting

Cells were lysed with TNE buffer and cleared by centrifugation (Kamoshita et al. 2009). After quantification of protein concentration of the cleared lysate using Pierce BCA Protein Assay Kit (Thermo 23225), 8 μ g of lysate was separated by 8% SDS-PAGE and transferred to an Immobilon-P PVDF membrane (Millipore IPVH0001). Expressed polypeptides were detected by rabbit anti-*Renilla* luciferase antibody (MBL PM047) or mouse monoclonal anti-FLAG M2 antibody (SIGMA F1804-20), together with appropriate secondary antibodies. As a loading control, β -actin was stained with mouse monoclonal antibody (Santa Cruz sc-47778). Chemiluminescence from the blot was recorded on an ImageQuant LAS4000 Luminescent Image Analyzer (GE Healthcare). When necessary, the luminescence of the bands of

interest was quantified as volume values using the Analysis Toolbox module in ImageQuant TL software (GE Healthcare).

RNA

Plasmids, linearized with the SpeI site, were used as a template for the T7 RNA polymerase reaction. Template DNA was digested with DNase I (Promega) and transcripts were purified using a spun column packed with Sephadex G-50 (GE Healthcare). Capped RNA was prepared with anti-reverse cap analog (NEB S1411S).

In vitro translation

Rabbit Reticulocyte Lysate (RRL, Promega L416A) or Transdirect Insect Cell Extract (Shimadzu 292-30000-91) from *Spodoptera frugiperda* ovary Sf21 cells was used, in accordance with each manufacturer's protocol. The final reaction concentrations for potassium, magnesium, and template RNA were adjusted to 150 or 100 mM, 1.1 mM, and 0.03 µg/µL (RRL) or 0.3 µg/µL (Sf21), respectively. Products were labeled with L-[³⁵S]-methionine (Perkin-Elmer, NEG709A) and separated by SDS-PAGE, as described previously (Kamoshita et al. 1997). Radioactivities of polypeptides on the images scanned with a Typhoon 9410 Imager (GE Healthcare) were measured as volume values V of photo-stimulated luminescence (PSL) in ImageQuant TL software (GE Healthcare). When polypeptides with different lengths were compared, the PSL value was divided by the number of methionine residues in each polypeptide. Readthrough efficiency of certain mRNA was calculated as $(V_{RT}/n_{RT})/(V_{RT}/n_{RT} + V_{RLuc}/n_{RLuc})$, where V_X and n_X represent the PSL V value and the number of methionine residues, respectively, for translated polypeptide X , where polypeptide X is either RT (stop codon readthrough) or RLuc (*Renilla* luciferase). IRES activity was calculated as $(V_{FLuc}/n_{FLuc})/(V_{RLuc}/n_{RLuc})$.

Reporter assay

Luciferase activities in 5–10 µL of cellular lysate were quantified using Dual Luciferase Reporter Assay (Promega). Firefly and *Renilla* luciferase activities were measured by Lumat LB9507 Tube Luminometer or LB96V Microplate Luminometer (Berthold).

Fluorescence intensity from EGFP was quantified using ARVomx Plate Reader (PerkinElmer). After 10-fold dilution of 10 µL of lysate with PBS, fluorescence of 485 nm excitation and 535 nm emission was measured for 1 sec. Samples with mock transfection were used as blank controls.

Determination of the expression level of stop codon readthrough polypeptides

Calculation in cell-free systems was performed as described above. In cultured cells, the expression level of stop codon readthrough polypeptides was compared among mutants by one of the following three procedures. In all three procedures, values obtained in one series of transfection were normalized to that of a reference sample, shown in red or the indicated color in each figure, and averaged from at least three independent experiments.

1. When the lengths in amino acids of the readthrough products to be compared were equal (Figs. 3, 6; Supplemental Table S1), the ratio of firefly luciferase activity to *Renilla* luciferase activity was normalized to that of a reference sample.
2. In 3' deletion mutant assays (Fig. 4), chemiluminescence from readthrough polypeptides obtained with anti-FLAG antibody and 36 kDa *Renilla* luciferase proteins obtained with anti-RLuc antibody (Fig. 4D) was quantified as volume values using ImageQuant TL software. Relative chemiluminescence values of readthrough polypeptides divided by those of 36 kDa *Renilla* luciferase were normalized to that of a reference sample 6192M1 (Fig. 4E). In upstream sequence assay (Fig. 2), chemiluminescence from readthrough polypeptides obtained with anti-FLAG antibody was directly compared between sample 1 and sample 2.
3. When readthrough was compared with a sense control (Fig. 7), in which no 36 kDa *Renilla* luciferase was expressed, *Renilla* luciferase activity values were no longer used as an internal control. Firefly luciferase activity was divided by co-transfected fluorescence of EGFP and this relative value was normalized to that of a reference sample UGA M1 (Fig. 7B).

Structural mapping of RNA

Chemical reagents including DMS and CMCT were purchased from SIGMA. SHAPE reagent NAI was prepared from 2-methylpyridine-3-carboxylic acid (2MNA), and 1,1'-carbonyldiimidazole (CDI), in accordance with the procedures described by Spitale et al. (2013).

Dicistronic RNAs to be mapped were renatured after 1 min of heat denaturation at 80°C. At a potassium concentration of 150 mM, 0.08 µg/µL RNA was reacted with 60 mM DMS, 25 mM CMCT, or 100 mM NAI at 33°C for 15 min. The positions of modifications were determined by reverse transcription inhibition from ³²P-labeled oligodeoxynucleotide (ODN) annealed to downstream firefly luciferase sequence (Kamoshita et al. 2009).

Bioinformatic analysis

Homology of PSIV polypeptides was analyzed in NCBI BLASTP. In the search using a short query (length 50–85), BLOSUM80 was used as a substitution matrix. Secondary structure prediction was based on PSIPRED and JPRED. The alignment of viral proteins (Supplemental Figs. S8–S10) was prepared in the Web version of ESPript 3.0 (Robert and Gouet 2014). Six candidate sites for 3C or 3CD cleavage were chosen according to the presence of P1 glutamine, or P1 glutamate followed by serine or glycine. Details are described in the Supplemental Material.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

The earliest part of this work was carried out in the RajBhandary Laboratory at MIT. We are grateful to Dr. Nobuhiko Nakashima

for his generous gift of cDNA of PSIV and Dr. Shusuke Kuge and Dr. Izumu Saito for plasmids. We are also grateful to laboratory members for discussions and to Dr. Tsukasa Ohmori for comments on the manuscript and support for revisions of experiments. This work was supported by Grants-in-Aid for Scientific Research (C) from the Japan Society for the Promotion of Science (JSPS KAKENHI grant numbers JP22590413 to N.K. and JP25460393 to S.T.) and the MEXT-Supported Program for the Strategic Research Foundation at Private Universities, 2013–2017.

NOTE ADDED IN PROOF

After this paper was submitted, investigations into the *AMD1* gene and all human genes containing UGA_CUAG were reported in Yordanova et al. (2018) and Loughran et al. (2018), respectively. Also, the long-range RNA interaction of another member of luteovirus was reported in Xu et al. (2018).

Received December 25, 2017; accepted October 8, 2018.

REFERENCES

- Aberathne PD, Koh CS, Grant T, Grigorieff N, Korostelev AA. 2016. Ensemble cryo-EM uncovers inchworm-like translocation of a viral IRES through the ribosome. *Elife* **5**: e14874. doi:10.7554/eLife.14874
- Au HHT, Elspass VM, Jan E. 2018. Functional insights into the adjacent stem-loop in honey bee dicistroviruses that promotes IRES-mediated translation and viral infection. *J Virol* **92**: e01725-17. doi:10.1128/JVI.01725-17
- Beier H, Grimm M. 2001. Misreading of termination codons in eukaryotes by natural nonsense suppressor tRNAs. *Nucleic Acids Res* **29**: 4767–4782. doi:10.1093/nar/29.23.4767
- Bertram G, Innes S, Minella O, Richardson J, Stansfield I. 2001. Endless possibilities: translation termination and stop codon recognition. *Microbiology* **147**: 255–269. doi:10.1099/00221287-147-2-255
- Bonning BC, Miller WA. 2010. Dicistroviruses. *Annu Rev Entomol* **55**: 129–150. doi:10.1146/annurev-ento-112408-085457
- Brown CM, Stockwell PA, Trotman CN, Tate WP. 1990. Sequence analysis suggests that tetra-nucleotides signal the termination of protein synthesis in eukaryotes. *Nucleic Acids Res* **18**: 6339–6345. doi:10.1093/nar/18.21.6339
- Brown CM, Dinesh-Kumar SP, Miller WA. 1996. Local and distant sequences are required for efficient readthrough of the barley yellow dwarf virus PAV coat protein gene stop codon. *J Virol* **70**: 5884–5892.
- Brown A, Shao S, Murray J, Hegde RS, Ramakrishnan V. 2015. Structural basis for stop codon recognition in eukaryotes. *Nature* **524**: 493–496. doi:10.1038/nature14896
- Bugaud O, Barbier N, Chommy H, Fiszman N, Le Gall A, Dulin D, Saguy M, Westbrook N, Perronet K, Namy O. 2017. Kinetics of CrPV and HCV IRES-mediated eukaryotic translation using single-molecule fluorescence microscopy. *RNA* **23**: 1626–1635. doi:10.1261/ma.061523.117
- Cassan M, Rousset JP. 2001. UAG readthrough in mammalian cells: effect of upstream and downstream stop codon contexts reveal different signals. *BMC Mol Biol* **2**: 3. doi:10.1186/1471-2199-2-3
- Chen C, Zhang H, Broitman SL, Reiche M, Farrell I, Cooperman BS, Goldman YE. 2013. Dynamics of translation by single ribosomes through mRNA secondary structures. *Nat Struct Mol Biol* **20**: 582–588. doi:10.1038/nsmb.2544
- Cimino PA, Nicholson BL, Wu B, Xu W, White KA. 2011. Multifaceted regulation of translational readthrough by RNA replication elements in a tombusvirus. *PLoS Pathog* **7**: e1002423. doi:10.1371/journal.ppat.1002423
- Dunn JG, Foo CK, Belletier NG, Gavis ER, Weissman JS. 2013. Ribosome profiling reveals pervasive and regulated stop codon readthrough in *Drosophila melanogaster*. *Elife* **2**: e01179. doi:10.7554/eLife.01179
- Eswarappa SM, Potdar AA, Koch WJ, Fan Y, Vasu K, Lindner D, Willard B, Graham LM, DiCorleto PE, Fox PL. 2014. Programmed translational readthrough generates antiangiogenic VEGF-Ax. *Cell* **157**: 1605–1618. doi:10.1016/j.cell.2014.04.033
- Feng YX, Yuan H, Rein A, Levin JG. 1992. Bipartite signal for readthrough suppression in murine leukemia virus mRNA: an eight-nucleotide purine-rich sequence immediately downstream of the gag termination codon followed by an RNA pseudoknot. *J Virol* **66**: 5127–5132.
- Fernández IS, Bai XC, Murshudov G, Scheres SH, Ramakrishnan V. 2014. Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. *Cell* **157**: 823–831. doi:10.1016/j.cell.2014.04.015
- Firth AE, Brierley I. 2012. Non-canonical translation in RNA viruses. *J Gen Virol* **93**: 1385–1409. doi:10.1099/vir.0.042499-0
- Firth AE, Wang QS, Jan E, Atkins JF. 2009. Bioinformatic evidence for a stem-loop structure 5'-adjacent to the IGR-IRES and for an overlapping gene in the bee paralysis dicistroviruses. *Virology* **6**: 193. doi:10.1186/1743-422X-6-193
- Firth AE, Wills NM, Gesteland RF, Atkins JF. 2011. Stimulation of stop codon readthrough: frequent presence of an extended 3' RNA structural element. *Nucleic Acids Res* **39**: 6679–6691. doi:10.1093/nar/gkr224
- Flint J, Racaniello VR, Rall GF, Skalka AM, Enquist LW. 2015. *Principles of virology. Volume I: molecular biology*, 4th ed. ASM Press, Washington, DC.
- Freitag J, Ast J, Böcker M. 2012. Cryptic peroxisomal targeting via alternative splicing and stop codon read-through in fungi. *Nature* **485**: 522–525. doi:10.1038/nature11051
- Geller AI, Rich A. 1980. A UGA termination suppression tRNA^{Trp} active in rabbit reticulocytes. *Nature* **283**: 41–46. doi:10.1038/283041a0
- Hellen CU, Sarnow P. 2001. Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev* **15**: 1593–1612. doi:10.1101/gad.891101
- Hertz MI, Thompson SR. 2011. Mechanism of translation initiation by *Dicistroviridae* IGR IRESs. *Virology* **411**: 355–361. doi:10.1016/j.virol.2011.01.005
- Hofhuis J, Schuere F, Nötzel C, Lingner T, Gärtner J, Jahn O, Thoms S. 2016. The functional readthrough extension of malate dehydrogenase reveals a modification of the genetic code. *Open Biol* **6**: 160246. doi:10.1098/rsob.160246
- Honigman A, Wolf D, Yaish S, Falk H, Panet A. 1991. *cis* acting RNA sequences control the Gag–Pol translation readthrough in murine leukemia virus. *Virology* **183**: 313–319. doi:10.1016/0042-6822(91)90144-Z
- Jalajakumari MB, Thomas CJ, Halter R, Manning PA. 1989. Genes for biosynthesis and assembly of CS3 pili of CFA/II enterotoxigenic *Escherichia coli*: novel regulation of pilus production by bypassing an amber codon. *Mol Microbiol* **3**: 1685–1695. doi:10.1111/j.1365-2958.1989.tb00154.x
- Jan E, Sarnow P. 2002. Factorless ribosome assembly on the internal ribosome entry site of cricket paralysis virus. *J Mol Biol* **324**: 889–902. doi:10.1016/S0022-2836(02)01099-9
- Johnson AG, Grosely R, Petrov AN, Puglisi JD. 2017. Dynamics of IRES-mediated translation. *Philos Trans R Soc Lond B Biol Sci* **372**: 20160177. doi:10.1098/rstb.2016.0177
- Jungreis I, Lin MF, Spokony R, Chan CS, Negre N, Victorsen A, White KP, Kellis M. 2011. Evidence of abundant stop codon

- readthrough in *Drosophila* and other metazoa. *Genome Res* **21**: 2096–2113. doi:10.1101/gr.119974.110
- Kamoshita N, Tsukiyama-Kohara K, Kohara M, Nomoto A. 1997. Genetic analysis of internal ribosomal entry site on hepatitis C virus RNA: implication for involvement of the highly ordered structure and cell type-specific transacting factors. *Virology* **233**: 9–18. doi:10.1006/viro.1997.8600
- Kamoshita N, Nomoto A, RajBhandary UL. 2009. Translation initiation from the ribosomal A site or the P site, dependent on the conformation of RNA pseudoknot I in dicistrovirus RNAs. *Mol Cell* **35**: 181–190. doi:10.1016/j.molcel.2009.05.024
- Kanamori Y, Nakashima N. 2001. A tertiary structure model of the internal ribosome entry site (IRES) for methionine-independent initiation of translation. *RNA* **7**: 266–274. doi:10.1017/S1355838201001741
- Kerr CH, Ma ZW, Jang CJ, Thompson SR, Jan E. 2016. Molecular analysis of the factorless internal ribosome entry site in *Cricket Paralysis virus* infection. *Sci Rep* **6**: 37319. doi:10.1038/srep37319
- Kieft JS. 2008. Viral IRES RNA structures and ribosome interactions. *Trends Biochem Sci* **33**: 274–283. doi:10.1016/j.tibs.2008.04.007
- Li G, Rice CM. 1993. The signal for translational readthrough of a UGA codon in Sindbis virus RNA involves a single cytidine residue immediately downstream of the termination codon. *J Virol* **67**: 5062–5067.
- Lindblad-Toh K, Garber M, Zuk O, Lin MF, Parker BJ, Washietl S, Kheradpour P, Ernst J, Jordan G, Maudsli E, et al. 2011. A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478**: 476–482. doi:10.1038/nature10530
- Loughran G, Chou MY, Ivanov IP, Jungreis I, Kellis M, Kiran AM, Baranov PV, Atkins JF. 2014. Evidence of efficient stop codon readthrough in four mammalian genes. *Nucleic Acids Res* **42**: 8928–8938. doi:10.1093/nar/gku608
- Loughran G, Howard MT, Firth AE, Atkins JF. 2017. Avoidance of reporter assay distortions from fused dual reporters. *RNA* **23**: 1285–1289. doi:10.1261/rna.061051.117
- Loughran G, Jungreis I, Tzani I, Power M, Dmitriev RI, Ivanov IP, Kellis M, Atkins JF. 2018. Stop codon readthrough generates a C-terminally extended variant of the human vitamin D receptor with reduced calcitriol response. *J Biol Chem* **293**: 4434–4444. doi:10.1074/jbc.M117.818526
- Matheisl S, Bermingham O, Becker T, Beckmann R. 2015. Structure of a human translation termination complex. *Nucleic Acids Res* **43**: 8615–8626. doi:10.1093/nar/gkv909
- McCaughan KK, Brown CM, Dalphin ME, Berry MJ, Tate WP. 1995. Translational termination efficiency in mammals is influenced by the base following the stop codon. *Proc Natl Acad Sci* **92**: 5431–5435. doi:10.1073/pnas.92.12.5431
- Nakashima N, Ishibashi J. 2010. Identification of the 3C-protease-mediated 2A/2B and 2B/2C cleavage sites in the nonstructural polyprotein precursor of a dicistrovirus lacking the NPGP motif. *Arch Virol* **155**: 1477–1482. doi:10.1007/s00705-010-0723-z
- Nakashima N, Uchiumi T. 2009. Functional analysis of structural motifs in dicistroviruses. *Virus Res* **139**: 137–147. doi:10.1016/j.virusres.2008.06.006
- Namy O, Duchateau-Nguyen G, Rousset JP. 2002. Translational readthrough of the PDE2 stop codon modulates cAMP levels in *Saccharomyces cerevisiae*. *Mol Microbiol* **43**: 641–652. doi:10.1046/j.1365-2958.2002.02770.x
- Naphtine S, Yek C, Powell ML, Brown TD, Brierley I. 2012. Characterization of the stop codon readthrough signal of Colorado tick fever virus segment 9 RNA. *RNA* **18**: 241–252. doi:10.1261/rna.030338.111
- Nishiyama T, Yamamoto H, Shibuya N, Hatakeyama Y, Hachimori A, Uchiumi T, Nakashima N. 2003. Structural elements in the internal ribosome entry site of *Plautia stali* intestine virus responsible for binding with ribosomes. *Nucleic Acids Res* **31**: 2434–2442. doi:10.1093/nar/gkg336
- Pfingsten JS, Costantino DA, Kieft JS. 2006. Structural basis for ribosome recruitment and manipulation by a viral IRES RNA. *Science* **314**: 1450–1454. doi:10.1126/science.1133281
- Pfingsten JS, Castile AE, Kieft JS. 2010. Mechanistic role of structurally dynamic regions in *Dicistroviridae* IGR IRESs. *J Mol Biol* **395**: 205–217. doi:10.1016/j.jmb.2009.10.047
- Ren Q, Wang QS, Firth AE, Chan MM, Gouw JW, Guarna MM, Foster LJ, Atkins JF, Jan E. 2012. Alternative reading frame selection mediated by a tRNA-like domain of an internal ribosome entry site. *Proc Natl Acad Sci* **109**: E630–E639. doi:10.1073/pnas.1111303109
- Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**: W320–W324. doi:10.1093/nar/gku316
- Ruehle MD, Zhang H, Sheridan RM, Mitra S, Chen Y, Gonzalez RL, Cooperman BS, Kieft JS. 2015. A dynamic RNA loop in an IRES affects multiple steps of elongation factor-mediated translation initiation. *Elife* **4**: e08146. doi:10.7554/eLife.08146
- Sasaki J, Nakashima N. 1999. Translation initiation at the CUU codon is mediated by the internal ribosome entry site of an insect picorna-like virus in vitro. *J Virol* **73**: 1219–1226.
- Sato M, Umeki H, Saito R, Kanai A, Tomita M. 2003. Computational analysis of stop codon readthrough in *D. melanogaster*. *Bioinformatics* **19**: 1371–1380. doi:10.1093/bioinformatics/btg183
- Schueren F, Lingner T, George R, Hofhuis J, Dickel C, Gärtner J, Thoms S. 2014. Peroxisomal lactate dehydrogenase is generated by translational readthrough in mammals. *Elife* **3**: e03640. doi:10.7554/eLife.03640
- Somogyi P, Jenner AJ, Brierley I, Inglis SC. 1993. Ribosomal pausing during translation of an RNA pseudoknot. *Mol Cell Biol* **13**: 6931–6940. doi:10.1128/MCB.13.11.6931
- Spitale RC, Crisalli P, Flynn RA, Torre EA, Kool ET, Chang HY. 2013. RNA SHAPE analysis in living cells. *Nat Chem Biol* **9**: 18–20. doi:10.1038/nchembio.1131
- Stiebler AC, Freitag J, Schink KO, Stehlik T, Tillmann BA, Ast J, Bötker M. 2014. Ribosomal readthrough at a short UGA stop codon context triggers dual localization of metabolic enzymes in fungi and animals. *PLoS Genet* **10**: e1004685. doi:10.1371/journal.pgen.1004685
- Tate J, Liljas L, Scotti P, Christian P, Lin T, Johnson JE. 1999. The crystal structure of cricket paralysis virus: the first view of a new virus family. *Nat Struct Biol* **6**: 765–774. doi:10.1038/11543
- Tu C, Tzeng TH, Bruenn JA. 1992. Ribosomal movement impeded at a pseudoknot required for frameshifting. *Proc Natl Acad Sci* **89**: 8636–8640. doi:10.1073/pnas.89.18.8636
- Weiner AM, Weber K. 1971. Natural read-through at the UGA termination signal of Q β coat protein cistron. *Nat New Biol* **234**: 206–209. doi:10.1038/newbio234206a0
- Wills NM, Gesteland RF, Atkins JF. 1991. Evidence that a downstream pseudoknot is required for translational read-through of the Moloney murine leukemia virus gag stop codon. *Proc Natl Acad Sci* **88**: 6991–6995. doi:10.1073/pnas.88.16.6991
- Wills NM, Gesteland RF, Atkins JF. 1994. Pseudoknot-dependent read-through of retroviral gag termination codons: importance of sequences in the spacer and loop 2. *EMBO J* **13**: 4137–4144. doi:10.1002/j.1460-2075.1994.tb06731.x
- Xu Y, Ju HJ, DeBlasio S, Carino EJ, Johnson R, MacCoss MJ, Heck M, Miller WA, Gray SM. 2018. A stem-loop structure in *Potato leafroll virus* open reading frame 5 (ORF5) is essential for readthrough translation of the coat protein ORF stop codon 700 bases upstream. *J Virol* **92**: e01544-17. doi:10.1128/JVI.01544-17

- Yamaguchi Y, Hayashi A, Campagnoni CW, Kimura A, Inuzuka T, Baba H. 2012. L-MPZ, a novel isoform of myelin P0, is produced by stop codon readthrough. *J Biol Chem* **287**: 17765–17776. doi:10.1074/jbc.M111.314468
- Yamamoto H, Unbehaun A, Spahn CMT. 2017. Ribosomal chamber music: toward an understanding of IRES mechanisms. *Trends Biochem Sci* **42**: 655–668. doi:10.1016/j.tibs.2017.06.002
- Yordanova MM, Loughran G, Zhdanov AV, Mariotti M, Kiniry SJ, O'Connor PBF, Andreev DE, Tzani I, Saffert P, Michel AM, et al. 2018. AMD1 mRNA employs ribosome stalling as a mechanism for molecular memory formation. *Nature* **553**: 356–360. doi:10.1038/nature25174
- Ziehler WA, Engelke DR. 2001. Probing RNA structure with chemical reagents and enzymes. *Curr Protoc Nucleic Acid Chem* **Chapter 6**: Unit 6.1. doi:10.1002/0471142700.nc0601s00