2362–2377 *Nucleic Acids Research, 2016, Vol. 44, No. 5* *Published online 17 January 2016*
doi: 10.1093/nar/gkw016

Attachment of ribosomal complexes and retrograde scanning during initiation on the Halastavi árva virus IRES

Irina S. Abaeva, Tatyana V. Pestova and Christopher U.T. Hellen*

Department of Cell Biology, SUNY Downstate Medical Center, Brooklyn, 11203, NY, USA

Received November 03, 2015; Revised January 06, 2016; Accepted January 07, 2016

ABSTRACT

Halastavi árva virus (HalV) has a positive-sense RNA genome, with an 827 nt-long 5′ UTR and an intergenic region separating two open reading frames. Whereas the encoded proteins are most homologous to Dicistrovirus polyproteins, its 5′ UTR is distinct. Here, we report that the HalV 5′ UTR comprises small stem-loop domains separated by long single-stranded areas and a large A-rich unstructured region surrounding the initiation codon AUG_{828}, and possesses cross-kingdom internal ribosome entry site (IRES) activity. In contrast to most viral IRESs, it does not depend on structural integrity and specific interaction of a structured element with a translational component, and is instead determined by the unstructured region flanking AUG_{828}. eIF2, eIF3, eIF1 and eIF1A promote efficient 48S initiation complex formation at AUG_{828}, which is reduced ∼5-fold on omission of eIF1 and eIF1A. Initiation involves direct attachment of 43S preinitiation complexes within a short window at or immediately downstream of AUG_{828}. 40S and eIF3 are sufficient for initial binding. After attachment, 43S complexes undergo retrograde scanning, strongly dependent on eIF1 and eIF1A. eIF4A/eIF4G stimulated initiation only at low temperatures or on mutants, in which areas surrounding AUG_{828} had been replaced by heterologous sequences. However, they strongly promoted initiation at AUG_{872}, yielding a proline-rich oligopeptide.

INTRODUCTION

Initiation of translation on most eukaryotic mRNAs is mediated by the scanning mechanism (1,2). First, initiator tRNA (Met-$tRNA_i^{Met}$) and eIF2·GTP form a ternary complex that binds cooperatively with eIF1, eIF1A and eIF3 to a 40S ribosomal subunit, yielding a 43S preinitiation complex. Attachment of 43S complexes to mRNA requires eIF4F, which comprises eIF4E (a cap-binding protein), eIF4A (a DEAD-box RNA helicase, which is activated by eIF4B and eIF4G) and eIF4G (a scaffold that binds eIF4E, eIF4A and eIF3). Group 4 eIFs promote ribosomal attachment by unwinding the cap-proximal region of mRNA, and by direct interaction of eIF4G with eIF3. The 43S complex then scans to the initiation codon, where it forms a 48S initiation complex following the establishment of codon-anticodon base pairing. Scanning involves ribosomal movement, which remains poorly characterized, and inspection of the 5′ untranslated region (5′ UTR) for an initiation codon. Ribosomal movement is also promoted by group 4 eIFs and, on mRNAs with highly structured 5′ UTRs, additionally requires DHX29, a DExH-box protein that binds to the 40S subunit near the mRNA entry site (3–5). eIF1 and eIF1A bind in the area of the ribosomal P and A sites, respectively, and induce an 'open' conformation of the mRNA-binding channel that facilitates scanning (6–10). They also ensure the fidelity of initiation codon selection, discriminating against initiation at non-AUG codons and at AUG triplets that have poor nucleotide context. Finally, eIF5 and eIF5B promote release of eIFs and joining of a 60S subunit, yielding an elongation-competent 80S ribosome.

Some viral mRNAs do not use the scanning mechanism, and initiation instead occurs by 5′-end-independent 'internal ribosomal entry'. Most internal ribosomal entry sites (IRESs) are highly structured, and are classified into a few major groups based on characteristics such as a common structural core, conserved sequence motifs and similarities in mechanisms of their function. Type I (e.g. poliovirus) and Type II (e.g. encephalomyocarditis virus) IRESs are both ∼450 nt long and contain five major domains (11), whereas the Hepatitis C virus (HCV) IRES and related IRESs in members of *Flaviviridae* and *Picornaviridae* are ∼200–450 nt long and contain two major domains (12). The intergenic region (IGR) IRESs located between ORF1 and ORF2 (which encode nonstructural and structural protein precursors, respectively) in members of *Dicistroviridae* are the shortest (∼190 nt long) and the most structured, comprising two domains formed by three pseudoknots (13). Despite these differences, the mechanisms of function of all

*To whom correspondence should be addressed. Tel: +718 270 1034; Fax: +718 270 2656; Email: christopher.hellen@downstate.edu

these IRESs are determined by specific non-canonical interactions with components of the translation apparatus. Thus, Type I and Type II IRESs bind via unrelated domains to eIF4G/eIF4A, which then promote internal attachment of 43S complexes (14–18), whereas HCV-like and IGR IRESs specifically interact with the 40S subunit (19–21). These IRESs require only a subset of initiation factors, and can thus evade regulatory mechanisms, such as those that form part of cellular innate immune responses or are triggered by viruses to suppress cellular translation. Interestingly, in addition to the IGR IRES, dicistroviruses such as Cricket paralysis virus (CrPV) (22), *Rhopalosiphum padi* virus (RhPV) (23), Triatoma virus (TrV) (24) and *Plautia stali* intestine virus (PSIV) (25), have an unrelated IRES in the 5′ UTR. The presence of two IRESs in a single mRNA would allow for differential regulation of their activities for temporal or quantitative control of synthesis of replicative and structural proteins during infection. Cellular translation is shut-off early during CrPV infection, likely due to phosphorylation of eIF2 and disruption of the eIF4E–eIF4G interaction, and initiation mediated by 5′ UTR and IGR IRESs is resistant to or even up-regulated by these changes (26). In the case of CrPV, there is no evidence for differential temporal control of the two IRESs, and the synthesis of ORF2 proteins in supramolar quantities relative to ORF1 proteins (27) likely reflects the stronger activity of the IGR IRES in these conditions (26). In contrast, the activities of the two TrV IRESs are differentially regulated by conditions that mimic viral infection (24).

An important common characteristic of the IRESs discussed above is that their function absolutely depends on structural integrity; consequently, sequence variation in picornavirus IRESs is often compensatory, and results in structural conservation (11,28). IRES activity is commonly impaired by structurally disruptive substitutions, but can be restored by compensatory second-site substitutions (e.g. 29,30). However, in contrast to these IRESs, the activity of a few viral IRESs does not depend on highly structured elements and instead appeared to be determined by the presence of specific single-stranded regions. Thus, the activity of the IRES preceding the coat protein gene of crucifer-infecting tobamovirus has been attributed to two polypurine (A-rich) sequences (PARS), and could be recapitulated by inserting PARS-like elements into heterologous RNAs (31). Another example is the 5′-terminal RhPV IRES. Internal deletions of 22–188 nt throughout the 579 nt-long RhPV 5′ UTR and long 5′- or 3′-terminal deletions had minor effects on IRES function, which is determined by the presence of U-rich, single-stranded elements (23,32). The RhPV IRES has cross-kingdom activity (being active in mammalian, insect and plant cell-free extracts) (23,33), requires eIF1, eIF2 and eIF3, is strongly stimulated by eIF1A, eIF4A and eIF4G, forms ternary IRES/eIF3/40S subunit complexes, but does not contain specific binding sites for any factors (32). The 43S complex binds directly to an unstructured region in the IRES and scans to the initiation codon; entry occurs within a large window that corresponds to a single-stranded region upstream of the initiation codon.

The identification of divergent IRESs that do not belong to established classes raises the question whether additional IRESs with unique structural and mechanistic characteristics remain to be identified. Numerous novel RNA viruses, including several with dicistronic genomes, are candidates to contain IRESs. One of them is Halastavi árva RNA virus (HalV), a positive-sense single-stranded RNA virus with a dicistronic genome that was identified in the intestinal content of freshwater carp (*Cyprinus carpio*) (34). It has not yet been placed in a virus family, but ORF1 and ORF2 proteins are most closely related to those of dicistroviruses. However, the 827 nt-long 5′ UTR, which contains 13 AUG triplets, is not related to noncoding regions from dicistroviruses or other RNA viruses. We therefore selected it for further analysis and here, determined its structure, showed that it contains an IRES and characterized the mechanism of its function.

MATERIALS AND METHODS

Plasmids

Expression vectors for eIF4A, eIF4B, eIF1, eIF1A, eIF5, eIF4A^{R362Q} mutant, eIF4G$_{736–1115}$ ('eIF4Gm'), Ligatin, *Escherichia coli* methionyl tRNA synthetase and a tRNA$_i^{Met}$ transcription vector have been described (8,14,16,35–37).

A monocistronic HalV ('MC Stem-HalV') transcription vector was made (Genscript, Piscataway, NJ, USA) by inserting DNA corresponding to a T7 promoter, a stable hairpin (GGGCCCGACCCGGTGACGGGTCGGGCCC) ($\Delta G = -32.40$ kcal/mol) and a variant of HalV nt.1–1682 (Genbank NC_016418.1) between Pst1 and EcoRV sites of pUC57. Substitutions in this insert introduced AUG triplets at positions corresponding to codons 194, 203, 209, 220, 234, 237, 246 and 254 of the 256 amino acid (a.a.)-long (30.2 kDa) coding sequence, to increase radiolabeling of the HalV translation product. The derivative MC Stem-HalV-MTN-STOP was made (NorClone Biotech, London, ON, Canada) by replacing the fourth codon (ACA) by a UAA stop codon. The derivative MC Stem-HalV containing a stem after nt 805 was made by inserting 16 nucleotides (5′GGGCCGAGAGGCCC3′) at that position (Genewiz). MC Stem-HalV(Δnt.934) was made by deleting nt. 934, thereby extending the ORF that initiates at AUG$_{872}$ from 20 to 241 codons, and truncating ORF1 (initiating at AUG$_{828}$) from 256 to 189 codons. The monocistronic transcription vector 'MC-HalV' was made by replacing a 5′-terminal Pst1-Xho1 fragment from MC Stem-HalV by a synthetic DNA fragment (Genewiz) comprising a T7 promoter, the nucleotides GG and HalV nt.1–424.

The dicistronic HalV mRNA ('DC ΔXL-HalV') transcription vector was made (Genscript) by inserting DNA between Pst1 and EcoRV sites of pUC57 that corresponded to a T7 promoter followed by a variant of nt.2–1390 of *Xenopus laevis* cyclin B2 mRNA (Genbank J03167) (XL) in which nt 1071–1218 had been replaced by the sequence CT, the linker ACTAGTCCCGGG (which contained Spe1 and Sma1 sites) and HalV nt. 1–1681 (in which the sequence AAATAATGA replaced nt. 1512–1520, and with substitutions that introduced AUG triplets at positions corresponding to codons 194, 203, 209, 214, 220, 226 and 228 of the 229 a.a.-long (26.8 kDa) coding sequence).

The dicistronic transcription vector 'DC Stem-ΔXL-HalV' was made (NorClone

Biotech) by inserting a stable hairpin (GGGCCCGACCCGGTGACGGGTCGGGCC)(ΔG = −29.10 kcal/mol) immediately downstream of the T7 promoter of DC ΔXL-HalV. This vector was used (NorClone Biotech) to generate variants with deletions (Δnt.1–568, Δnt.1–611, Δnt.1–674, Δnt.1–753, Δnt.757–824 and Δnt.614–824) and substitutions (which introduced new upstream and downstream AUGs in conditions when AUG_{828} is either retained or substituted by near- or non-cognate codons) in the HalV 5′ UTR (as shown in Figures).

The dicistronic transcription vector DC XL-HalV-Aichi(ORF), in which the complete HalV 5′ UTR occupies the intercistronic region and ORF2 consists of a 258 a.a.-long (28.2 kDa) fragment of the AV polyprotein was made by amplifying the HalV 5′ UTR from MC Stem-HalV by polymerase chain reaction (PCR) using the primers 5′-AATACGACTCTGGCCAGG (forward) and 5′-GTATTTCCATGGTGTTTTT (reverse) (to introduce an Nco1 site overlapping the AV initiation codon), digesting the PCR product with Msc1 and Nco1 and inserting it into Msc1 and Nco1 sites of the 'DC Aichivirus' vector (38).

Bioinformatic analysis

Secondary structure elements in the HalV 5′ UTR and proximal coding sequence (Genbank accession #JN000306) were modeled essentially as described (12,39), using probabilistic (Pfold: http://www.daimi.au.dk/~compbio/pfold/ (40); and Contrafold: http://contra.stanford.edu/contrafold/ (41)) and posterior decoding approaches (Centroidfold: http://www.ncrna.org/centroidfold (42)), and were verified and refined by free energy minimization using Mfold (http://mfold.rna.albany.edu/?q1/4mfold (43)) in all instances using default parameters. Computational analysis was done using series of overlapping RNA sequences.

Chemical and enzymatic footprinting

Five picomoles of MC Stem-HalV mRNA was enzymatically digested by incubation with RNase T1 (0.04 U/μl) or RNase V1 (0.0006 U/μl), or modified by incubation with 1-cyclohexyl-(2-morpholinoethyl) carbodiimide metho-p-toluene sulphonate (CMCT) (12.6 mg/ml) for 10 min at 30°C in 50 μl buffer A (20 mM Tris pH 7.5, 100 mM KCl, 1 mM DTT, 2.5 mM $MgCl_2$, 0.25 mM spermidine) (44) or with N-methylisatoic anhydride (NMIA) for 30 min at 37°C as described (45). Cleaved or modified sites were identified by primer extension, using avian myeloblastosis virus reverse transcriptase (AMV RT) and primers complementary to HalV nt 124–109, nt 185–169, nt 288–269, nt 387–369, nt 500–481, nt 626–610, nt 711–695, nt 811–794 and nt 911–927.

Purification of factors, ribosomal subunits and aminoacylation of tRNA

Mammalian 40S and 60S subunits, eIF2, eIF3, eIF5B, eEF1H, eEF2 and total aminoacyl-tRNA synthetases were purified from rabbit reticulocyte lysate (RRL), and insect 40S subunits were purified from *Spodoptera frugiperda* cell-free extract (Promega) (46). Recombinant eIF1, eIF1A, eIF4A, eIF4B, eIF4Gm, eIF5 and methionyl tRNA synthetase were expressed in *E. coli* (36,46). Native total tRNA (Promega) and *in vitro* transcribed $tRNA_i^{Met}$ were aminoacylated as described (46).

Assembly and toe-printing analysis of ribosomal complexes

The 48S complexes were assembled by incubating 1 pmol *wt* or mutant HalV mRNAs with 3 pmol 40S subunits, 5 pmol Met-$tRNA_i^{Met}$ and the indicated combinations of 8 pmol eIF1, 8 pmol eIF1A, 5 pmol eIF2, 4 pmol eIF3, 10 pmol eIF4A, 5 pmol eIF4B, 5 pmol eIF4Gm for 10 min at 37°C in 40 μl buffer A (20 mM Tris pH 7.5, 100 mM KCl, 1 mM DTT, 2.5 mM $MgCl_2$, 0.25 mM spermidine) supplemented with 1 mM ATP and 0.4 mM GTP. Assembled 48S complexes were analyzed by toe-printing using AMV RT and ^{32}P-labelled primers as described (46). cDNA products were resolved in 6% polyacrylamide sequencing gels.

To assay elongation on HalV-MTN-STOP mRNA, 48S complexes were supplemented with 5 pmol 60S subunits, 5 pmol eIF5, 5 pmol eIF5B, 15 μg total native aa-tRNA, 3 pmol eEF1H and 3 pmol eEF2, incubated at 37°C for an additional 15 min and analyzed by toe-printing.

Analysis of ribosomal complexes by sucrose density gradient centrifugation

40S/IRES and 40S/eIF3/IRES complexes were assembled in 100 μl reaction mixtures by incubating 10 pmol [^{32}P]UTP-labeled RNA corresponding to HalV nt.1–957 with 40 pmol 40S subunits and 45 pmol eIF3 (as indicated in Figure 3F), and analyzed by centrifugation through 10–30% sucrose density gradients prepared in buffer A (46). The HalV mRNA was transcribed from the MC-Stem-HalV vector after linearization with ApaL1. The optical density of fractionated gradients was measured at 260 nm, and the presence of [^{32}P]-labeled mRNA was monitored by Cherenkov counting.

In vitro translation

MC and DC HalV mRNAs (3 pmol) were translated in 20 μl reaction volume of Flexi RRL (at temperatures as indicated), wheat germ cell-free extract (at 25°C), or *S. frugiperda* 21 TNT T7 insect cell extract (at 30°C) (Promega) supplemented with 0.5 mCi/ml [^{35}S]methionine ((43.5 TBq/mmol); MP Biomedicals) and, where indicated, 50 pmol $eIF4A^{R362Q}$. Translation products were resolved by electrophoresis using NuPAGE 4–12% Bis-Tris precast gels (Life Technologies), followed by autoradiography and quantification on a phosphoimager.

RESULTS

The secondary structure of the HalV 5′ UTR

A model of the secondary structure of the HalV 5′-UTR (Figure 1A) was derived using complementary bioinformatic approaches, including statistical, probabilistic and free-energy minimization methods (40–43), essentially as

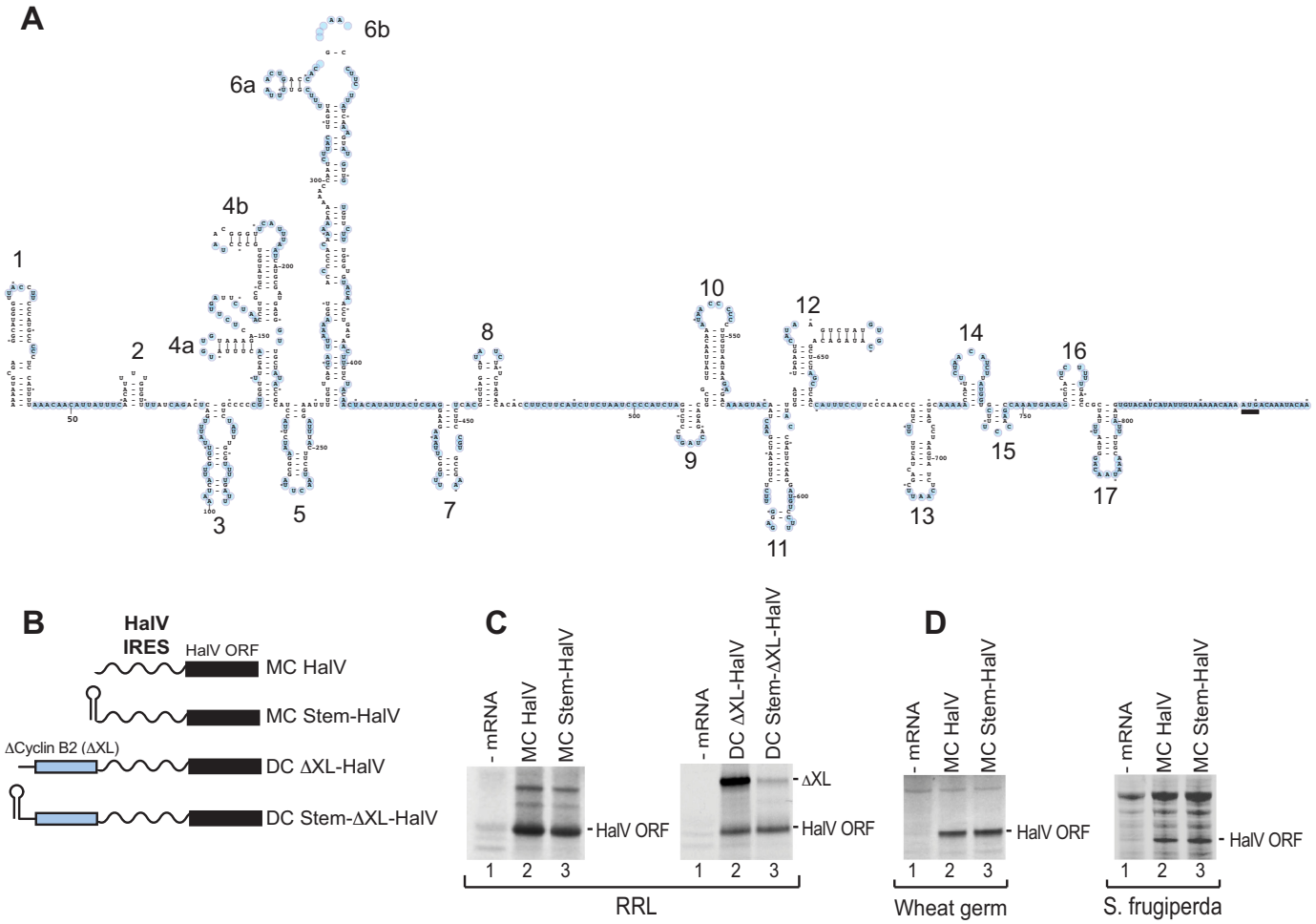

Figure 1. The HalV 5-UTR possesses the IRES activity. (**A**) Model of the secondary structure of the HalV 5′ UTR, derived as described in the text, and indicating the positions of nucleotides modified by NMIA (blue circles). (**B**) Schematic representation of monocistronic and dicistronic mRNA constructs containing cyclin B2 (XL) and HalV cistrons. (**C** and **D**) The activity of the HalV 5′ UTR in monocistronic and dicistronic constructs assayed by translation in RRL, wheat germ and *S. frugiperda* cell-free systems.

done previously for different picornavirus IRESs (12,38,39). It was largely consistent with the pattern of cleavage by RNase V1, which is specific for base-paired RNA, and of modification by CMCT (which predominantly reacts with unpaired U residues) and NMIA (which preferentially acylates conformationally flexible nucleotides at the ribose 2′-OH position) (45) (Supplementary Figures S1–S3). The HalV 5′ UTR consists of a series of stem-loop domains (designated 1–17) that (with the exception of the branched domains 4 and 6) are relatively small and separated by long single-stranded regions. The strong accessibility of the region downstream of ∼nt.710 to modification by NMIA (Supplementary Figure S3) suggested that under equilibrium conditions, the weak domains 14–17 are likely only transiently apparent, and the large area surrounding the initiation codon (AUG_{828}) is predominantly unstructured. Notably, the sequences flanking AUG_{828} are strikingly A-rich.

The HalV 5′ UTR contains an IRES

To test whether the HalV 5′ UTR functions as an IRES, it was placed downstream of a stable hairpin ($\Delta G = -32.40$ kcal/mol) in the monocistronic MC Stem-HalV mRNA, and also between a truncated cyclin B2 ORF (ΔXL) and part of the HalV ORF in dicistronic constructs either with or without a stable 5′-terminal hairpin of $\Delta G = -29.10$ kcal/mol (DC Stem-ΔXL-HalV and DC ΔXL-HalV mRNAs, respectively) (Figure 1B). Insertion of a strong 5′-terminal stem did not affect translation of MC HalV mRNA in RRL (Figure 1C, left panel). Moreover, the HalV 5′ UTR was equally active in the DC ΔXL-HalV mRNA, and whereas insertion of the stable hairpin upstream of the first cistron strongly reduced synthesis of ΔXL, it did not affect translation of HalV (Figure 1C, right panel). These results indicate that the HalV 5′ UTR has IRES activity. This IRES was also active in insect (*S. frugiperda*) and plant (wheat germ) cell-free extracts (Figure 1D).

Factor requirements for initiation at AUG_{828} of the HalV IRES

To characterize the mechanism of initiation on the IRES, we identified its factor requirements using the *in vitro* reconstitution approach, in which 48S complexes are assembled on MC Stem-HalV mRNA from individual 40S subunits, Met-$tRNA_i^{Met}$ and different sets of eIFs. Formation of 48S complexes is monitored by the appearance of characteristic toe-prints +15–17 nt downstream of the initiation codon.

Incubation of MC Stem-HalV mRNA with 40S subunits resulted in the appearance of weak toe-prints at nt 889–890 (Figure 2A, lane 2). Addition of eIF3 yielded new low-intensity toe-prints at positions 843–847, corresponding to 40S/eIF3/HalV IRES ternary complexes (Figure 2A, lane 4). Notably, the upper band(s) in this group would coincide with toe-prints corresponding to 48S complexes assembled at AUG_{828}. Further inclusion of eIF2-TC led to 48S complex formation at AUG_{828}, as well as at the downstream near-cognate AUA_{835} (Figure 2A, lane 5). Individually, eIF1A and eIF1 moderately enhanced initiation at AUG_{828}, concomitantly inhibiting it at AUA_{835} (Figure 2A, lanes 6 and 7). Together, eIF1 and eIF1A promoted very efficient 48S complex formation at AUG_{828} (Figure 2A, lane 8). eIF3 was essential for the process, and no 48S complexes formed in the presence of only eIFs 2, 1 and 1A (Figure 2B). Inclusion of eIFs 4A/4B/4Gm not only did not stimulate further, but instead slightly inhibited initiation at AUG_{828} (Figure 2A, lane 9). Thus, eIFs 2, 3, 1 and 1A were sufficient for initiation on the HalV IRES. 40S subunits from *S. frugiperda* could substitute for their mammalian counterparts (Figure 2C). The elongation competency of 48S complexes assembled on AUG_{828} was confirmed using MC Stem-HalV-MTN-Stop mRNA, in which the fourth codon was replaced by a UAA stop codon. Upon addition of 60S subunits, eIF5, eIF5B, eEF2, eEF1H and Σaa-tRNAs, 48S complexes underwent elongation, yielding pre-TCs at the newly introduced stop codon (Figure 2D).

Importantly, in addition to stimulating initiation at AUG_{828}, eIF1 and eIF1A induced 48S complex formation at the upstream UUG_{815}, GUG_{803} and particularly AUG_{782} (Figure 2A, lane 8). To investigate the mechanism by which 43S complexes reach AUG_{828} and the upstream AUG_{782}, we introduced a short GC-rich stem ($\Delta G = -11.90$ kCal/mol; Figure 2E, upper panel) after nt 805, which is 23 nucleotides downstream from AUG_{782} and 23 nucleotides upstream of AUG_{828}. In the absence of group 4 eIFs, such a stem would prevent ribosomal scanning, but regarding the interaction of mRNA with the 40S subunits in the mRNA-binding channel (47), it would not physically interfere with 48S complex formation on either of these codons. Strikingly, introduction of the stem abrogated 48S complex formation at AUG_{782}, but did not affect this process at AUG_{828} (Figure 2E, compare lanes 3 and 6), suggesting that these AUGs are selected by backward, rather than forward, ribosomal scanning. Consistently, initiation at the upstream UUG_{815}, GUG_{803} and particularly at AUG_{782}, was also decreased by eIFs 4A/4B/4Gm, which in addition promoted assembly of 48S complexes at the downstream AUG_{872} (Figure 2A, lane 9).

Besides stimulating 48S complex formation, eIF1 also eliminated the nt 843–847 toe-prints corresponding to 40S/eIF3/IRES ternary complexes (Figure 2F). eIF1 is thought to promote ribosomal scanning and destabilization of aberrant initiation complexes by inducing conformational changes in the mRNA-binding cleft of the 40S subunit (2). The effect of eIF1 on 40S/eIF3/HalV IRES complexes therefore suggests that the region of HalV mRNA preceding nt 843–847 might be located in the mRNA-binding channel. eIF1-induced conformational changes would weaken association of HalV mRNA with 40S/eIF3 complexes, resulting in their displacement by reverse transcriptase.

The data above are consistent with the scenario, in which initiation on the HalV IRES involves direct attachment of 43S complexes to the area at or downstream of AUG_{828}. The 40S subunits and eIF3 are likely sufficient for initial binding, which would require the start codon to be flanked by unstructured regions. After attachment, ribosomal complexes are capable of retrograde movement, the extent of which depends on eIF1 and eIF1A. Retrograde scanning would result in 48S complex formation at the upstream AUGs, and if initial attachment occurs downstream of AUG_{828}, it would also be responsible for initiation at this codon. eIFs 4A/4B/4Gm, on the other hand, impose 3′-directionality on scanning, promoting initiation at the downstream AUG_{872} and reducing it at the upstream codons.

Ribosomal attachment to multiple upstream regions in the HalV 5′ UTR

During the investigation of initiation at AUG_{828}, we noticed the appearance of additional toe-prints in the area of nt 700–750 (Figure 2A). To determine their nature, toe-printing was performed using primer complementary to the region upstream of AUG_{828} in order to avoid masking by initiation events at this codon. Incubation of MC Stem-HalV mRNA with 40S subunits and eIF3 resulted in low-intensity toe-prints at nt 750–754 and stronger ones at nt 726–728 (Figure 3A, lane 2). Like toe-prints at nt 843–847, they were susceptible to the presence of eIF1 (Figure 2F), suggesting their common nature. Inclusion of eIF2-TC yielded several additional toe-prints: (i) immediately upstream of stops at nt 750–754, which would correspond to 48S complexes at AUG_{732}, (ii) immediately upstream of stops at nt 726–728, which would correspond to 48S complexes at AUG_{708}, and (iii) further upstream, which would correspond to 48S complexes at AUG_{699} (Figure 3A, lane 3). eIF1 and eIF1A stimulated initiation at AUG_{708} and AUG_{699}, and led to low-level 48S complex formation at the further upstream AUG_{635} (Figure 3A, lane 4; Figure 2A, lane 8). Assembly of all 48S complexes was decreased by eIFs 4A/4B/4Gm (Figure 2A, lane 9).

An intriguing common feature of toe-prints at nt 843–847, 750–754 and 726–728 that appeared in the presence of 40S subunits and eIF3 (Figures 2A and 3A) was the presence in each instance of an AUG triplet located a short distance upstream of the toe-prints. To investigate whether the presence of AUG was essential, AUG_{708} was substituted to ACG. This substitution did not influence toe-prints at nt

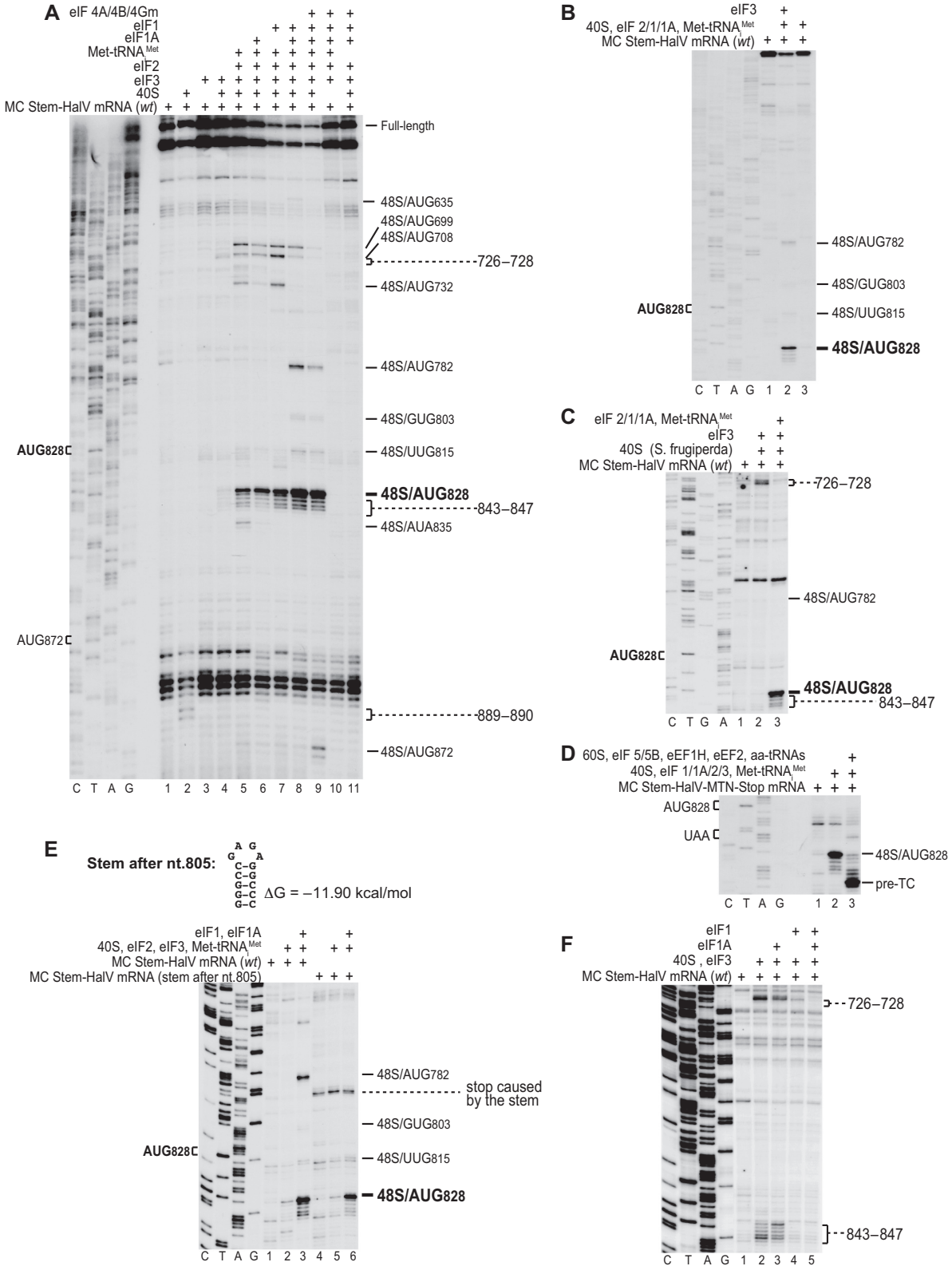

Figure 2. Factor requirements for initiation on HalV mRNA. (A–F) Toe-printing analysis of (**A** and **B**) 48S complex formation on MC Stem-HalV mRNA from mammalian 40S subunits, Met-tRNA$_i^{Met}$ and indicated combinations of eIFs, (**C**) 48S complex formation on MC Stem-HalV mRNA from *S. frugiperda* 40S subunits, and mammalian Met-tRNA$_i^{Met}$ and eIFs 2/3/1/1A, (**D**) elongation competence of 48S complexes assembled at AUG_{828}, assayed using MC Stem-HalV-MTN-Stop mRNA containing a stop codon three triplets downstream of the initiation codon, (**E**) the influence of a short stem introduced after nt 805 on 48S complex formation on MC Stem-HalV mRNA, and (**F**) the influence of eIF1 and eIF1A on 40S/eIF3/HalV IRES ternary complex formation. The positions of 48S complexes (solid lines), pre-TCs (solid line), as well as 40S/HalV IRES binary and 40S/eIF3/HalV IRES ternary complexes (dashed lines), are indicated on the right.

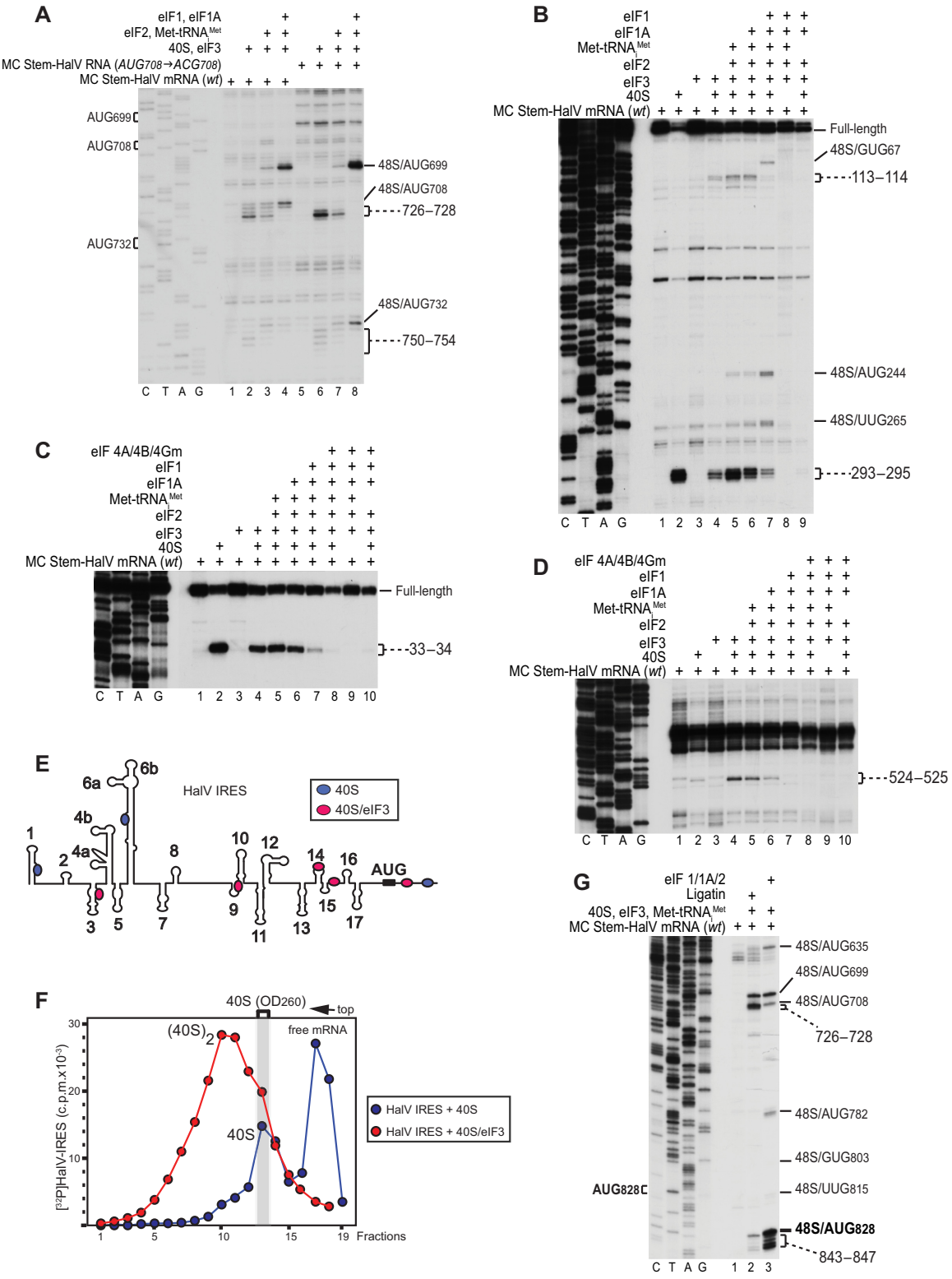

Figure 3. 40S subunits can attach to multiple regions in the HalV 5′ UTR. (**A**–**D**) Interactions of 40S subunits with different regions of the HalV 5′ UTR and assembly of 48S complexes upstream of the AUG_{828}, assayed by toe-printing. The positions of 48S complexes (solid lines), and 40S/IRES binary and 40S/eIF3/IRES ternary complexes (dashed lines) are indicated on the right. (**E**) Positions of toe-prints corresponding to 40S/IRES binary complexes (blue) and 40S/eIF3/IRES ternary complexes (red) mapped onto the secondary structure of the HalV 5′ UTR. (**F**) Association of the ^{32}P-labeled MC Stem-HalV mRNA (nt 1–957) with 40S subunits, depending on the presence of eIF3. Ribosomal complexes were separated by centrifugation in 10–30% linear sucrose density gradients, and aliquots of gradient fractions were analyzed by scintillation counting. The position of 40S subunits determined by optical density is indicated. Sedimentation was from *right* to *left*. (**G**) The activity of Ligatin in 48S complex formation on MC Stem-HalV mRNA, assayed by toe-printing. The positions of 48S complexes (solid lines) and 40S/eIF3/IRES ternary complexes (dashed lines) are indicated on the right.

726–728, and did not prevent 48S complex formation at the upstream AUG_{699} (Figure 2A, lanes 5–8), indicating that the presence of AUG is not a determinant for the 40S/eIF3 toe-prints.

Further analysis led to the identification of two other regions of HalV mRNA that supported 48S complex formation. In one, 40S and eIF3 yielded toe-prints at nt 113–114 (Figure 3B, lane 4), and supplementation of reaction mixtures with eIF2-TC, eIF1 and eIF1A led to 48S complex formation at the upstream GUG_{67} (Figure 3B, lane 7). In the other, 40S subunits alone yielded toe-prints at nt 293–295 (Figure 3B, lane 2), and addition of eIF2-TC, eIF3, eIF1 and eIF1A resulted in 48S complex formation at the upstream UUG_{265} and AUG_{244} (Figure 3B, lane 7).

The 40S subunits bound at two other sites, characterized by toe-prints at nt 33–34 (Figure 3C) and 524–525 (Figure 3D). In both cases, binding was also susceptible to dissociation by eIF1. However, supplementation of reaction mixtures with eIF1, eIF1A, eIF2, eIF3 and Met-$tRNA_i^{Met}$ did not result in 48S complex formation in the vicinity of these toe-prints.

Thus, toe-printing analysis identified several binding sites for individual 40S subunits or 40S/eIF3 complexes on the HalV IRES (Figure 3E). The 3′-terminal (not shown) and 5′-terminal deletion analysis (discussed below) indicated that these sites were independent of each other. The susceptibility of 40S and 40S/eIF3 toe-prints to eIF1 suggests that in all complexes, the IRES region was likely placed in the mRNA-binding channel. To assay how many 40S subunits or 40S/eIF3 complexes are simultaneously associated with the HalV 5′ UTR, ^{32}P-labeled HalV mRNA (nt 1–957) was incubated with 40S subunits in the presence or absence of eIF3, and resulting complexes were separated by sucrose density gradient centrifugation (Figure 3F). In the case of individual 40S subunits, mRNA was mostly bound to one 40S subunit (blue circles). eIF3 strongly stimulated binding, enforcing association of the majority of mRNA with two 40S subunits (red circles), based on the similar mobility of these complexes to that of eIF3-associated 40S dimers (48). The presence of the shoulder at the leading edge of the 40S/eIF3/IRES peak indicates that a fraction of mRNA was bound to more than two 40S subunits.

In most cases, eIFs 2, 3, 1 and 1A promoted 48S complex formation at AUGs upstream of the 40S- or 40S/eIF3-binding sites. Thus, in addition to initiation at AUG_{828}, 43S complexes were also able to initiate at several upstream AUGs, most likely by a similar mechanism, involving direct attachment of 43S complexes followed by their retrograde movement. However, initiation at all upstream AUGs was substantially lower than at AUG_{828}, and likely does not have any functional importance.

Direct attachment of 40S/eIF3 complexes in the vicinity of AUG codons also prompted us to investigate the activity of Ligatin (35,49,50) in initiation on the HalV IRES (Figure 3G). Although Ligatin was able to promote 48S complex formation at AUG_{828}, its activity was substantially lower than that of eIFs 2/1/1A. At AUG_{708} and AUG_{699}, the activities of Ligatin and eIFs 2/1/1A were more comparable, but the retrograde movement in the case of Ligatin was more limited.

The role of eIFs 4A/4B/4Gm in initiation on the HalV IRES

The influence of eIFs 4A/4B/4Gm on 48S complex formation at the upstream and downstream AUGs prompted us to investigate further their role in initiation on the IRES. At 37°C (at which all previously described assays were performed), eIFs 4A/4Gm had the same effect as eIFs 4A/4B/4Gm in reducing 48S complex formation at AUG_{828} and stimulating it at AUG_{872}, whereas eIFs 4A/4B alone did not have any influence (Figure 4A). However, there are several reasons to assume that in nature, HalV might replicate at temperatures that are lower than 37°C. Thus, HalV was isolated from the intestinal content of freshwater carp, an ectotherm, whose body temperature corresponds closely to the ambient water temperature. Moreover, although a piscine host origin for HalV was suggested on the basis of nucleotide composition analysis (34), carp are omnivorous, so that a dietary source, including insects or crustaceans, cannot be excluded. Therefore, we assayed 48S complex formation at 30°C and 25°C. Strikingly, at these temperatures, 48S complex formation at AUG_{828} was dependent on eIFs 4A/4B/4Gm (Figure 4B). Consistently, translation in RRL from AUG_{828} became progressively more sensitive to inhibition by dominant-negative $eIF4A^{R362Q}$ as the temperature decreased (Figure 4C).

The high level of utilization of AUG_{872} in the reconstituted system poses the question of whether this codon could function during translation in cell-free extracts. AUG_{872} is in a different reading frame to AUG_{828}, and initiation on it would yield a 20 amino acid-long proline-rich peptide (MLPLPPPPPPTLPLRTNTPN). To detect translation of ORF2 following initiation at AUG_{872}, a single nucleotide was deleted from the ORF2 stop codon, thereby increasing its length to 241 codons, and simultaneously reducing the length of ORF1 (initiating at AUG_{828}) from 256 to 189 codons (Figure 4D, upper panel). The resulting switch reduced the number of methionines in ORF1 from 8 to 3, and introduced 8 methionines into ORF2. Translation in RRL of MC Stem-HalV (Δnt.934) mRNA yielded two products of the expected molecular weights (Figure 4D, lower panel). Normalization of the translational efficiency for the number of methionines in each ORF indicated that initiation at AUG_{828} occurred 1.3 times more frequently than at AUG_{872}. As expected, in contrast to AUG_{828}, translation from AUG_{872} was extremely sensitive to inhibition by the $eIF4A^{R362Q}$ mutant. These findings are in good agreement with the results of toe-printng analysis obtained using the *in vitro* reconstituted system. Thus, AUG_{872} has the capacity to act as an initiation codon in infected cells, leading to expression of the short proline-rich peptide.

Ribosomal attachment window and the mechanism for selection of AUG_{828}

To determine the site of eIF4A/4B/4G-independent attachment of 43S complexes and the mechanism by which they subsequently reach AUG_{828}, we investigated 48S complex formation at newly introduced upstream and downstream AUGs in conditions when AUG_{828} is either retained or substituted by near- or non-cognate codons (Figure 5A).

Substitution of AUG_{828} by UUG or GUG strongly reduced, whereas substitution by AUA or AAA abrogated

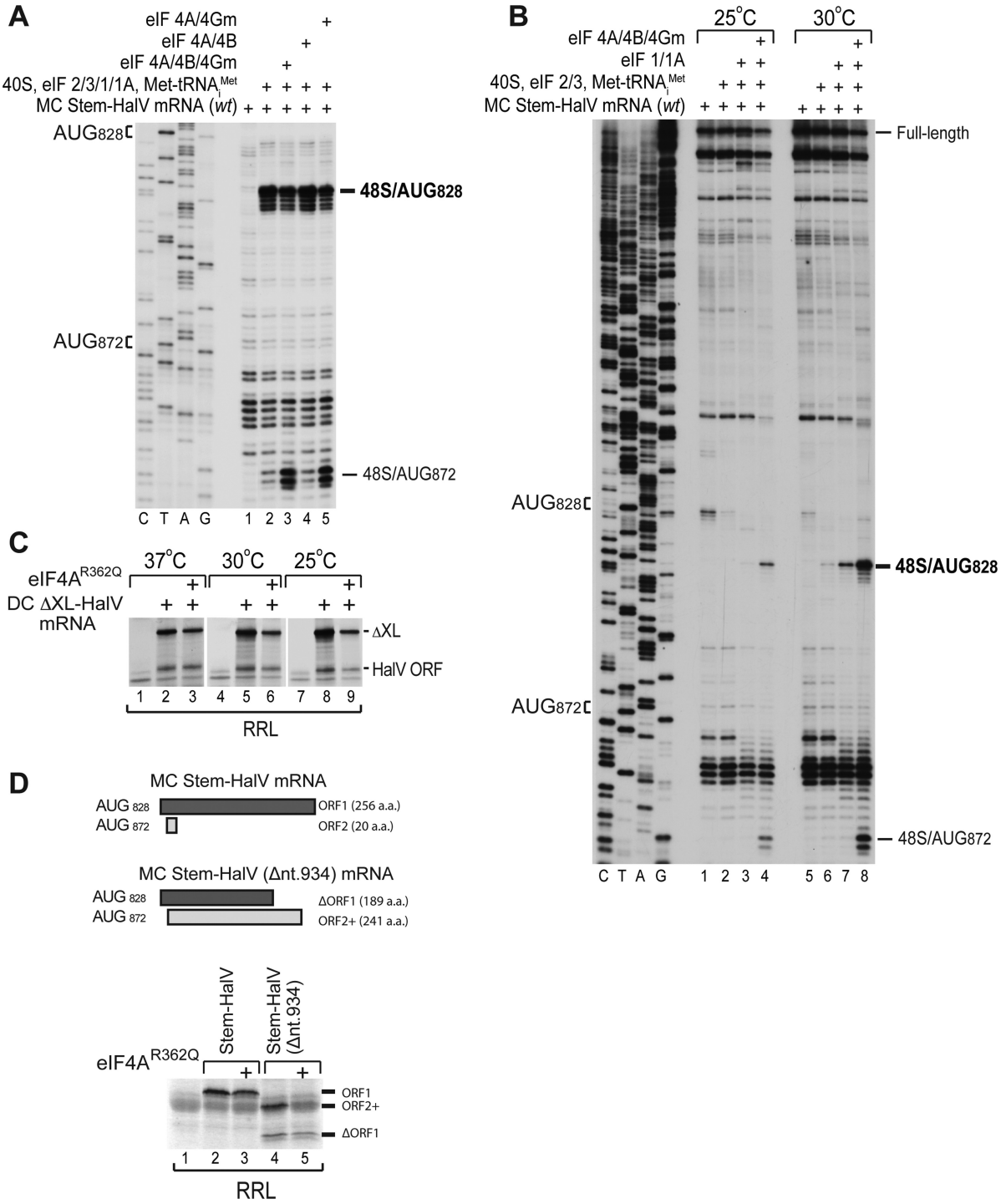

Figure 4. The influence of the group 4 eIFs on initiation on the HalV mRNA. (A, B) The influence of (**A**) eIF4A, eIF4B and eIF4Gm and (**B**) the reaction temperature on 48S complex formation at AUG_{828} and AUG_{872} of MC Stem-HalV mRNA, assayed by toe-printing. The positions of 48S complexes are indicated on the right. (**C**) The influence of the eIF4A^{R362Q} mutant on translation of DC ΔXL-HalV mRNA in RRL at different temperatures. (**D**) *Upper panel* - schematic representation of the ORFs starting at AUG_{828} or AUG_{872} in the *wt* and (Δnt.934) mutant MC Stem-HalV mRNAs. *Lower panel* – the influence of the eIF4A^{R362Q} mutant on translation in RRL from AUG_{828} and AUG_{872} of the *wt* and (Δnt.934) mutant MC Stem-HalV mRNAs.

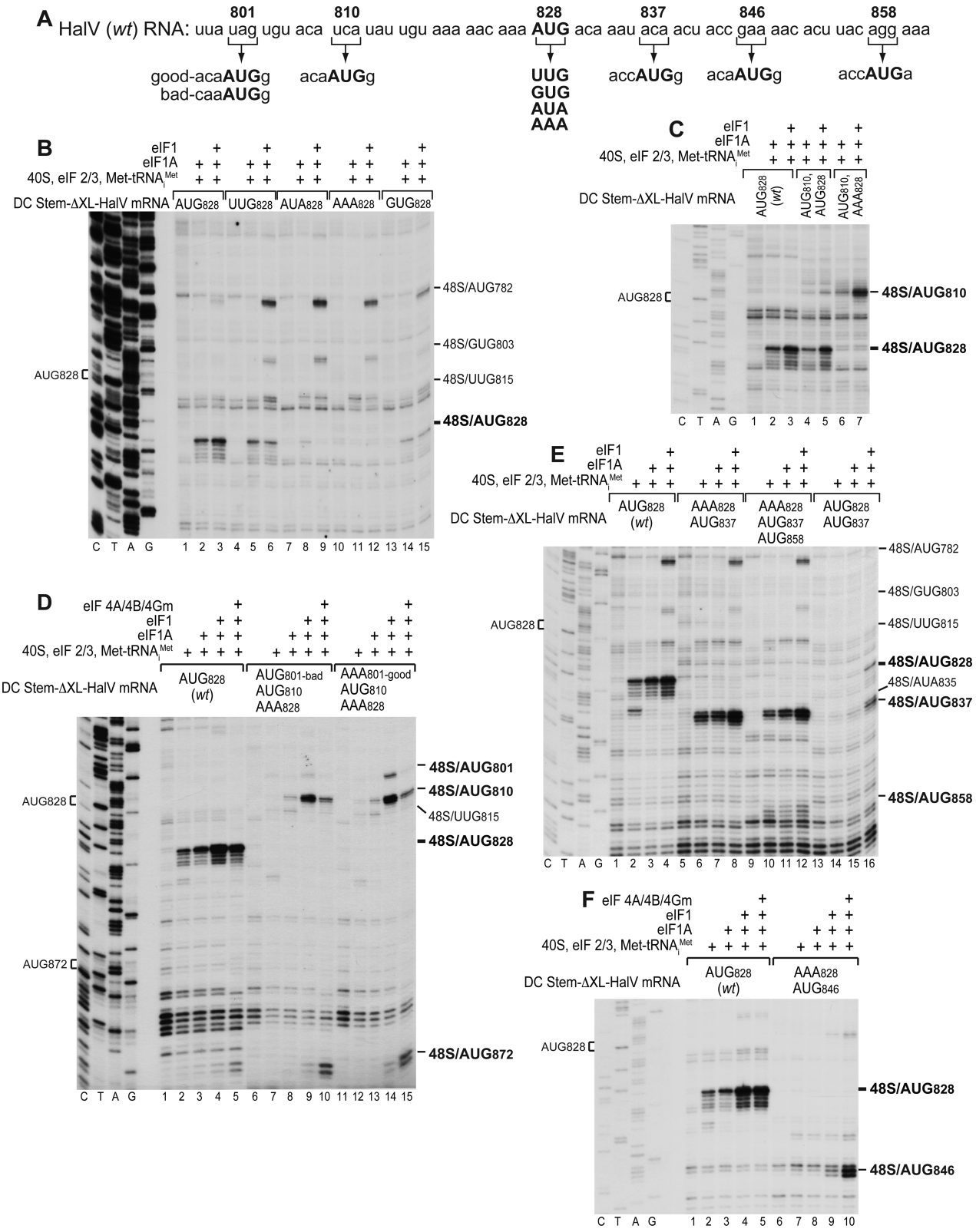

Figure 5. The mechanism of selection of the HalV initiation codon AUG_{828}. (**A**) Mutations in AUG_{828} and positions of new AUGs introduced into the DC Stem-ΔXL-HalV construct. (**B**) Toe-printing analysis of 48S complex formation on mutant DC Stem-ΔXL-HalV mRNAs with substitutions in AUG_{828}, depending on the presence of different sets of eIFs. (**C**–**F**) Toe-printing analysis of 48S complex formation on mutant DC Stem-ΔXL-HalV mRNAs containing additional AUGs introduced (**C**, **D**) upstream and (**E**, **F**) downstream of the AUG_{828}, depending on the presence of different sets of eIFs. The positions of 48S complexes are indicated on the right (B–F).

48S complex formation at this position (Figure 5B). Strikingly, in the presence of eIF1 and eIF1A, all mutations in AUG_{828} dramatically increased 48S complex formation at the upstream native UUG_{815}, GUG_{803} and AUG_{782}. When the new AUG was introduced at position 810, efficient initiation on it also occurred only after substitution of AUG_{828} by AAA (Figure 5C). When AUG_{810} was introduced simultaneously with the upstream AUG_{801}, either in optimum or suboptimum nucleotide context, efficient 48S complex formation occurred only at AUG_{810} (Figure 5D, lanes 9 and 14). Taken together, these results are consistent with the scenario, in which 43S complexes initially attach either directly at AUG_{828}, or downstream from it, in which case AUG_{828} is reached by backward scanning. If AUG_{828} is not present, 43S complexes can scan in the 5′ direction in search of an alternative start codon. Interestingly, 48S complex formation at AUG_{810} was strongly inhibited by eIFs 4A/4B/4Gm (Figure 5D, lanes 10 and 15), which would be consistent with their greater inhibitory influence on more extensive retrograde scanning.

To determine the 3′ border of the attachment window, new AUGs were introduced downstream of AUG_{828}. Insertion of AUG_{837} resulted in more efficient 48S complex formation at it than at AUG_{828} (Figure 5E, lane 16), which suggests that at least a large proportion of 43S complexes attach at or downstream of AUG_{837}. This would also be consistent with 48S complex formation at AUA_{835} on the *wt* mRNA in the presence of only eIF3 and eIF2-TC (Figure 2A). Introduction of AUG_{837} also reduced the overall efficiency of 48S complex formation, likely due to formation of some secondary structure in this area. For this reason we created another mutant in which AUG_{837} was introduced simultaneously with substitution of AUG_{828} by AAA. In this construct, 48S complex formation at AUG_{837} was as efficient as at AUG_{828} on the *wt* mRNA, and the efficiencies of 48S complex formation at the upstream UUG_{815}, GUG_{803} and AUG_{782} were also similar on the *wt* and AUG_{837} mutant mRNAs (Figure 5E, lanes 4 and 8). These results confirm the conclusion that attachment of 43S complexes occurs at or downstream of AUG_{837}. Introduction of AUG_{858} resulted in only low-efficiency 48S complex formation at this codon, and did not affect initiation at AUG_{837} and the upstream codons UUG_{815}, GUG_{803} and AUG_{782} (Figure 5E, lane 12), which suggests that ribosomal attachment occurs upstream of AUG_{858}. To narrow the window, AUG was placed at position 846. Efficient initiation at AUG_{846} occurred only in the presence of eIFs 4A/4B/4Gm (Figure 5F), indicating that 43S complexes mostly attach upstream of AUG_{846}. Taken together, our results suggest that the window for eIF4A/4B/4G-independent attachment of 43S complexes on the IRES is located between nt 828–846.

Mutational analysis of the HalV 5′ UTR

To determine regions and structural elements that are essential for initiation on the IRES, 5′-terminal truncations and internal deletions were introduced in the context of the DC Stem-ΔXL-HalV construct (Figure 6A). Strikingly, progressive 5′ terminal deletions up to nt 674 did not affect the efficiency or factor requirements of 48S complex formation at AUG_{828}, and also did not influence initiation at AUG_{872} (Figure 6B). The 48S complex formation at AUG_{828}, as well as at AUG_{872}, was also not affected by internal deletion of nt 757–824, located immediately upstream of AUG_{828} (Figure 6C, lanes 6–10). Notably, in this deletion mutant AUG_{828} would again be linked to an unstructured upstream sequence (Figure 1A). All these mutations also did not prevent formation of 40S/eIF3/IRES complexes characterized by toe-prints at nt 726–728, or of 48S complexes at the upstream AUG_{708} and AUG_{699}.

A larger 5′-terminal deletion of 753 nt strongly reduced 48S complex formation at AUG_{828} and made it dependent on eIFs 4A/4B/4Gm (Figure 6C, lanes 11–15). This deletion would fuse the residual ∼75 nt-long segment of the 5′ UTR to a completely unrelated sequence that could potentially base pair with some regions in the vicinity of AUG_{828}, thereby reducing the unstructured nature of the mRNA surrounding the initiation codon. However, the biggest effect on initiation at AUG_{828} resulted from deletion of nt 616–824, which very strongly reduced 48S complex formation in a manner that could not be rescued by eIFs 4A/4B/4Gm (Figure 6C, lanes 16–20). At the same time, this deletion did not affect 48S complex formation at AUG_{872} in the presence of eIFs 4A/4B/4Gm, and even increased it in their absence. In contrast to deletion of nt 757–824, this deletion would link AUG_{828} to the structured domain 11 (Figure 1A). This would likely shorten the 43S attachment window at the 5′-end and also prevent retrograde movement of 43S complexes to AUG_{828}. At the same time, it would not interfere with initiation at AUG_{872}. Analysis of 48S complex formation at 30°C showed the same pattern of activity for all mutants, albeit with a greater dependence on eIF4A/4B/4Gm, and the level of 48S complex formation on mutant mRNAs was generally consistent with their translation in RRL (data not shown).

To characterize the importance of the coding sequence immediately downstream of the initiation codon on internal ribosomal entry on the IRES, we replaced the HalV ORF in DC ΔXL-HalV mRNA by a heterologous 5′-terminal segment of the Aichivirus polyprotein coding region (Figure 6D, upper panel). Introduction of this heterologous sequence rendered initiation at AUG_{828} highly dependent on eIF4A/4B/4Gm (Figure 6D, lower panel). In their presence, the levels of 48S complex formation on DC ΔXL-Aichi(ORF) and DC ΔXL-HalV mRNAs were similar (Figure 6D), and consistently, translation of Aichi and HalV ORFs from these two mRNAs occurred at comparable levels in RRL (Figure 6E).

Taken together, these results suggest that the HalV IRES does not contain functionally important structural elements, and that efficient initiation depends only on the presence of an unstructured region immediately upstream and downstream of the initiation codon. The activity of some HalV mutants with heterologous sequences around AUG_{828} could be rescued by addition of eIF4A/4B/4Gm.

DISCUSSION

The HalV 5′ UTR contains 13 AUG triplets upstream of the viral ORF1 initiation codon AUG_{828}, and comprises 17 mostly small stem-loop domains interspersed between long single-stranded elements. Extensive 5′-terminal dele-

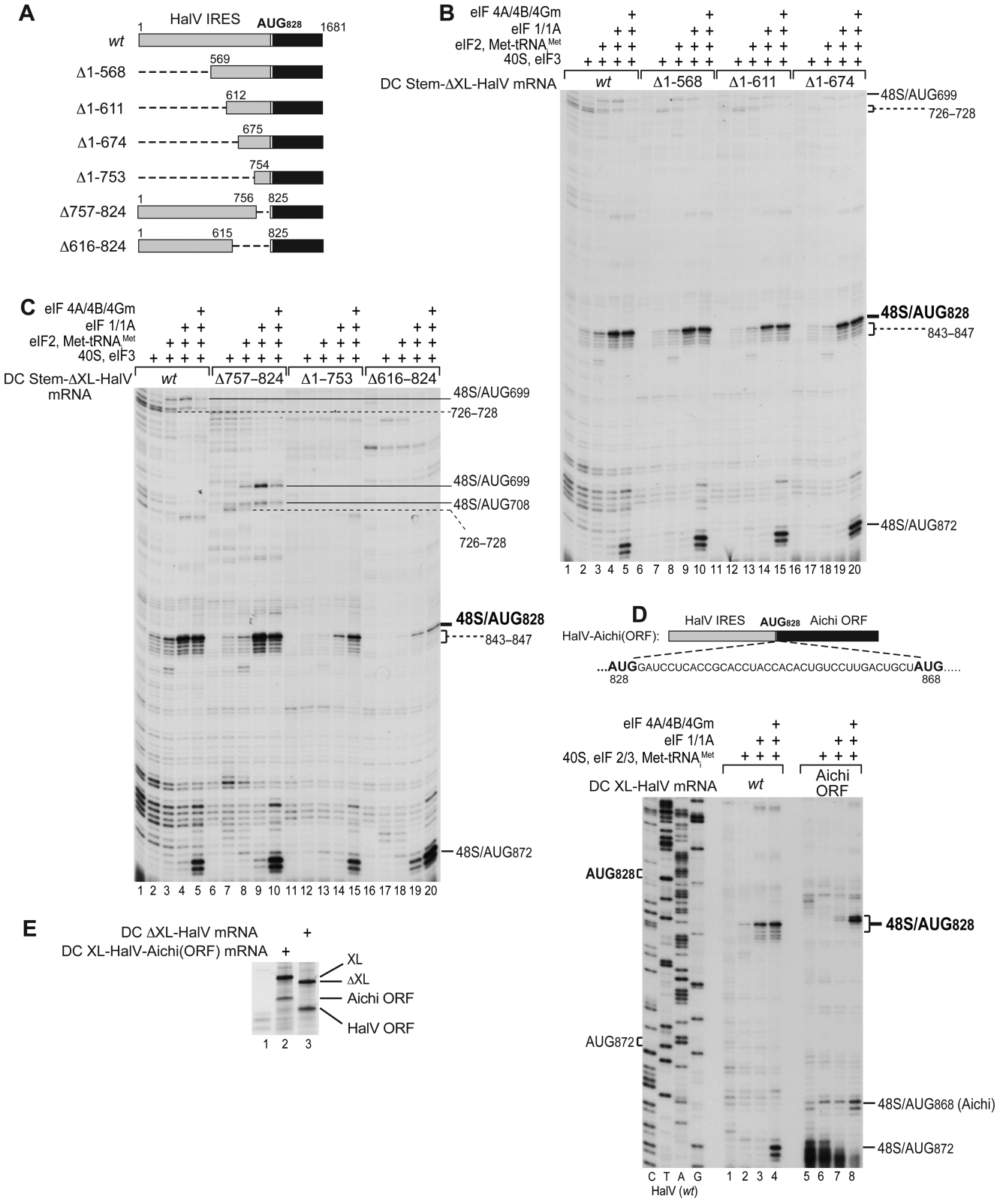

Figure 6. Mutational analysis of the HalV 5′ UTR. (**A**) Schematic representation of deletions introduced into the HalV 5′ UTR in the context of the DC Stem-ΔXL-HalV construct. (**B** and **C**) Toe-printing analysis of 48S complex formation on mutant DC Stem-ΔXL-HalV mRNAs containing deletions in the HalV 5′ UTR, depending on the presence of different sets of eIFs. The positions of 48S complexes (solid lines) and 40S/eIF3/HalV IRES ternary complexes (dashed lines) are indicated on the right. (**D**) *Upper panel* - schematic representation of the second cistron of DC XL-HalV-Aichi(ORF) mRNA, showing the sequence of the first 43 coding nucleotides in the insert. *Lower panel* – toe-printing analysis of 48S complex formation on DC ΔXL-HalV and DC XL-HalV-Aichi(ORF) mRNAs, depending on the presence of different sets of eIFs. The positions of 48S complexes are indicated on the right. (**E**) Relative activities of DC ΔXL-HalV and DC XL-HalV-Aichi(ORF) mRNAs assayed by translation in RRL.

tions (e.g. of nt 1–674) do not affect the 5′ UTR's activity as an IRES, whereas the predominantly unstructured A-rich region flanking AUG_{828} is critical. Initiation on the HalV IRES therefore differs from that on almost all other IRESs characterized to date, in that it does not require distinct highly-structured domains that interact specifically with translational components, and instead appears to depend on the single-stranded nature of the segment of mRNA comprising the initiation codon and flanking regions.

Efficient assembly of 48S complexes on the HalV IRES required four initiation factors: eIF2, eIF3, eIF1 and eIF1A. However, only eIF2 and eIF3 were absolutely essential, and could promote initiation with ∼20% of maximal efficiency even in the absence of eIF1 and eIF1A. Analysis of initiation codon selection on mutant mRNAs containing additional AUGs upstream and/or downstream of AUG_{828} indicated that 43S complexes attach directly at or immediately downstream of AUG_{828}, within a short window located between nt 828–846. After attachment, 43S complexes can scan in the 5′-direction. The ability of 43S complexes to scan in a retrograde manner strongly depends on the presence of eIFs 1 and 1A, which induce a scanning-competent conformation of the 43S complex (e.g. 9,10). The processivity of retrograde scanning was very high, and if AUG_{828} was substituted by a non-cognate codon, 48S complexes were able to form with the same efficiency at the upstream AUGs (e.g. AUG_{810} and AUG_{782}). Most likely, 43S complexes attach not to the precise nucleotides, but to any point within the attachment window, and if ribosomal binding occurs downstream of AUG_{828}, 43S complexes reach this codon by retrograde scanning. Notably, a similar efficient 5′-directional scanning was observed during eIF4A/4B/4G-independent reinitiation of translation on a derivative of β-globin mRNA (49).

The appearance of toe-prints downstream of AUG_{828} at nt 843–847 in the presence of only 40S subunits and eIF3 suggests that the initial binding of 43S complexes to the IRES is determined by these two components. This is consistent with the earlier finding that eIF3, 40S subunits and single-stranded RNA can form ternary complexes, in which the RNA is located in the mRNA-binding cleft of the 40S subunit (51). Several observations also indicate that eIF3 forms an extension of the mRNA-binding channel, including eIF3's position on the 40S subunit (5,52), the interaction of eIF3 with mRNA in 48S complexes detected by site-directed UV cross-linking (47), and the mutual stabilization of binding of eIF3 and mRNA on the 40S subunit following hydrolysis of GTP and consequent release of eIF2 from 48S complexes during subunit joining (48). However, it is currently not known whether the single-stranded nature of the mRNA flanking AUG_{828} is the only feature that determines attachment of ribosomal complexes, or if its sequence-specific characteristics, such as its (A)-rich nature, also contribute to recognition. Notably, it has recently been shown that mRNA with the 5′ UTR composed of a poly(A) tract can promote initiation independently of eIF3 and eIF4F, suggesting that poly(A) has the capacity to recruit 40S subunits, likely due to its specific affinity to the mRNA-binding channel (53). An obvious question that arises from the proposed mechanism of initiation on the HalV IRES is why 43S complexes attach at or downstream, but not upstream, of the initiation codon. One possibility is that 43S complexes have a higher capacity to scan in the 3′-5′ than in the 5′-3′ direction. Alternatively, initial recognition of the IRES might involve specific binding of one of eIF3's subunits to mRNA downstream of AUG_{828}, in which case downstream scanning would be inhibited by the anchored eIF3, whereas upstream movement would be permitted.

Strikingly, at 37°C, group 4 eIFs not only did not increase, but instead reduced 48S complex formation at AUG_{828}, concomitantly strongly promoting it at the downstream AUG_{872}. The activity of eIFs 4A/4B/4Gm was identical to that of eIFs 4A/4Gm, indicating that eIF4B was dispensable. The effect of group 4 eIFs on initiation on the HalV IRES is consistent with their ability to impose 3′-directionality on scanning (49) and with the intrinsic 5′-3′ directionality of RNA unwinding by yeast eIF4F (54). Thus, if attachment occurs downstream of AUG_{828}, these factors would inhibit retrograde movement of 43S complexes, and if 43S complexes were to bind directly to AUG_{828}, eIF4A/eIF4G would increase the leakiness of selection of this codon. However, initiation at AUG_{828} became strongly dependent on eIF4A/eIF4G at lower temperatures and on some IRES mutants, in which upstream or downstream areas adjacent to the initiation codon had been replaced by heterologous sequences. The reduced level of initiation by 43S complexes alone in these instances was likely caused by an increase in the degree of secondary structure at the attachment site. But if eIF4A/4G-independent ribosomal attachment mostly occurs downstream of AUG_{828}, then how could group 4 eIFs stimulate initiation at this codon? One possibility is that these factors shift the attachment toward the 5′-border of the attachment window, or could even extend this border beyond that observed during initiation by 43S complexes alone. This would be possible if the RNA-binding activity of eIF4G (55,56) enabled it to interact directly with some RNA sequences in or near the attachment window.

The high efficiency of utilization of AUG_{872} in an *in vitro* reconstituted system as well as in RRL suggests that it could also act as an initiation codon in infected cells. Initiation in an alternate reading frame downstream of the principal initiation site has been reported for various classes of IRES, including the Type II Theiler's murine encephalomyocarditis virus (TMEV) IRES (57), the HCV IRES (58) and the Israeli acute paralysis virus IGR IRES (59). In the case of TMEV, the alternate ORF encodes a 156 amino acid (a.a)-long protein that impairs an arm of the innate immune response and thus promotes virus persistence (60). Initiation at HalV AUG_{872} would yield a 20 a.a.-long proline-rich oligopeptide, and although polyproline motifs are commonly involved in protein–protein interactions (61) and this translation product could thus potentially have a physiological impact on infected cells, an alternative possibility is that the coding sequence of this ORF has a regulatory function in translation of the HalV genome. Peptide bond formation at consecutive proline residues is slow, because proline is a both a poor donor and a poor acceptor in the peptidyl transferase reaction (62); the consequent stalling of ribosomes in eukaryotes is alleviated by eIF5A (63) but the presence of consecutive proline codons nev-

ertheless remains associated with slowed elongation in ribosome profiling experiments (64,65). Thus, an accessible short ORF encoding six consecutive prolines located a short distance downstream of the polyprotein initiation codon would likely induce ribosomal stalling and attenuate initiation at that codon, consequently downregulating synthesis of nonstructural proteins and potentially promoting a switch to translation of capsid proteins.

The HalV IRES is distinct from most of the characterized viral IRESs, but there are interesting parallels between its activity and that of the IRES in the RhPV 5′ UTR (32). They both exhibit cross-kingdom activity, form stable ternary complexes with 40S subunits and eIF3 (albeit less efficiently on the RhPV IRES), and rather than depending on specific structural elements that interact specifically with components of the translation apparatus, have an activity that depends on the presence of extensive single-stranded elements in the region of the 5′ UTR adjacent to the initiation codon. However, there are important differences between the two mechanisms. Thus, 43S complexes appear to be able to bind productively to a region of the RhPV IRES that is much larger than the narrow 'window' on the HalV IRES, and this 'entry' region is located upstream of the RhPV initiation codon, whereas on the HalV IRES, attachment occurs at or immediately downstream of the initiation codon. These differences in the characteristics of the attachment region on the two IRESs determine the factor requirements for initiation and the mechanisms employed for initiation complexes to reach the initiation codon. Thus, on the RhPV IRES, 43S complexes reach the initiation codon by forward scanning and formation of 48S complexes absolutely requires eIF1 and is strongly dependent on eIF4A/eIF4G, whereas on the HalV IRES, 43S complexes either bind directly to the initiation codon or reach it by backward movement, initiation occurs efficiently in the absence of group 4 eIFs, and eIF4A/eIF4G are needed only at low temperatures or when unstructured mRNA elements flanking the initiation codon are replaced by heterologous sequences. Retrograde scanning could also be involved in initiation on Turnip yellow mosaic virus (66) and although the process has not been well characterized, could account for initiation at AUG1 on Human immunodeficiency virus 2 genomic mRNA following recruitment of a ribosomal preinitiation complex to the IRES located in the Gag coding region (67,68).

The mechanism of function of both HalV and RhPV IRESs raises an important question, which is why they retain sequences in their 5′ UTRs that do not seem to be essential for their activity in translation, and that are much larger than conventional viral replication elements. This issue remains to be resolved, although the functional redundancy within both 5′ UTRs could confer a competitive advantage in infected cells, by enhancing the IRES capacity to recruit 43S complexes in the cytoplasmic milieu. However, an initiation mechanism that relies on the presence of a largely single-stranded window for entry of 43S complexes is likely not used by several other dicistrovirus 5′ UTRs, including those of CrPV and PSIV, because they do not exhibit cross-kingdom activity and because the activity of the PSIV IRES is impaired by deletions that impinge on a 350 nt-long 3′-terminal fragment and is strongly dependent on structural integrity (22,25). It remains to be determined whether this mechanism is involved in initiation on the 5′ UTR of Kakugo virus (family *Iflaviridae*), which shares some characteristics with the RhPV 5′ UTR (69).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

National Institutes of Health (NIH) [AI51340 to C.U.T.H. and GM59660 to T.V.P.). Funding for open access charge: NIH [AI51340].
Conflict of interest statement. None declared.

REFERENCES

1. Jackson,R.J., Hellen,C.U. and Pestova,T.V. (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell. Biol.*, **11**, 113–127.
2. Hinnebusch,A.G. (2014) The scanning mechanism of eukaryotic translation initiation. *Annu. Rev. Biochem.*, **83**, 779–812.
3. Pisareva,V.P., Pisarev,A.V., Komar,A.A., Hellen,C.U. and Pestova,T.V. (2008) Translation initiation on mammalian mRNAs with structured 5′ UTRs requires DExH-box protein DHX29. *Cell*, **135**, 1237–1250.
4. Abaeva,I.S., Marintchev,A., Pisareva,V.P., Hellen,C.U. and Pestova,T.V. (2011) Bypassing of stems versus linear base-by-base inspection of mammalian mRNAs during ribosomal scanning. *EMBO J.*, **30**, 115–129.
5. Hashem,Y., des Georges,A., Dhote,V., Langlois,R., Liao,H.Y., Grassucci,R.A., Hellen,C.U., Pestova,T.V. and Frank,J. (2013) Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. *Cell*, **153**, 1108–1119.
6. Yu,Y., Marintchev,A., Kolupaeva,V.G., Unbehaun,A., Veryasova,T., Lai,S.C., Hong,P., Wagner,G., Hellen,C.U. and Pestova,T.V. (2009) Position of eukaryotic translation initiation factor eIF1A on the 40S ribosomal subunit mapped by directed hydroxyl radical probing. *Nucleic Acids Res.*, **37**, 5167–5182.
7. Weisser,M., Voigts-Hoffmann,F., Rabl,J., Leibundgut,M. and Ban,N. (2013) The crystal structure of the eukaryotic 40S ribosomal subunit in complex with eIF1 and eIF1A. *Nat. Struct. Mol. Biol.*, **20**, 1015–1017.
8. Pestova,T.V., Borukhov,S.I. and Hellen,C.U. (1998) Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. *Nature*, **394**, 854–859.
9. Pestova,T.V. and Kolupaeva,V.G. (2002) The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev.*, **16**, 2906–2922.
10. Llácer,J.L., Hussain,T., Marler,L., Aitken,C.E., Thakur,A., Lorsch,J.R., Hinnebusch,A.G. and Ramakrishnan,V. (2015) Conformational differences between open and closed states of the eukaryotic translation initiation complex. *Mol. Cell*, **59**, 399–412.
11. Jackson,R.J. and Kaminski,A. (1995) Internal initiation of translation in eukaryotes: the picornavirus paradigm and beyond. *RNA*, **1**, 985–1000.
12. Asnani,M., Kumar,P. and Hellen,C.U. (2015) Widespread distribution and structural diversity of Type IV IRESs in members of Picornaviridae. *Virology*, **478**, 61–74.
13. Nakashima,N. and Uchiumi,T. (2009) Functional analysis of structural motifs in dicistroviruses. *Virus Res.*, **139**, 137–147.
14. Pestova,T.V., Hellen,C.U. and Shatsky,I.N. (1996) Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Mol. Cell. Biol.*, **16**, 6859–6869.
15. Pestova,T.V., Shatsky,I.N. and Hellen,C.U. (1996) Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. *Mol. Cell. Biol.*, **16**, 6870–6878.
16. Kolupaeva,V.G., Lomakin,I.B., Pestova,T.V. and Hellen,C.U. (2003) Eukaryotic initiation factors 4G and 4A mediate conformational

changes downstream of the initiation codon of the encephalomyocarditis virus internal ribosomal entry site. *Mol. Cell. Biol.*, **23**, 687–698.
17. de Breyne,S., Yu,Y., Unbehaun,A., Pestova,T.V. and Hellen,C.U. (2009) Direct functional interaction of initiation factor eIF4G with type 1 internal ribosomal entry sites. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 9197–9202.
18. Sweeney,T.R., Abaeva,I.S., Pestova,T.V. and Hellen,C.U. (2014) The mechanism of translation initiation on Type 1 picornavirus IRESs. *EMBO J.*, **33**, 76–92.
19. Pestova,T.V., Shatsky,I.N., Fletcher,SP, Jackson,R.J. and Hellen,C.U. (1998) A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev.*, **12**, 67–83.
20. Hashem,Y., des Georges,A., Dhote,V., Langlois,R., Liao,H.Y., Grassucci,R.A., Pestova,T.V., Hellen,C.U. and Frank,J. (2013) Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature*, **503**, 539–543.
21. Wilson,J.E., Pestova,T.V., Hellen,C.U. and Sarnow,P. (2000) Initiation of protein synthesis from the A site of the ribosome. *Cell*, **102**, 511–520.
22. Wilson,J.E., Powell,M.J., Hoover,S.E. and Sarnow,P. (2000) Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites. *Mol. Cell. Biol.*, **20**, 4990–4999.
23. Woolaway,K.E., Lazaridis,K., Belsham,G.J., Carter,M.J. and Roberts,L.O. (2001) The 5′ untranslated region of *Rhopalosiphum padi* virus contains an internal ribosome entry site which functions efficiently in mammalian, plant, and insect translation systems. *J. Virol.*, **75**, 10244–10249.
24. Czibener,C., Alvarez,D., Scodeller,E. and Gamarnik,A.V. (2005) Characterization of internal ribosomal entry sites of Triatoma virus. *J. Gen. Virol.*, **86**, 2275–2280.
25. Shibuya,N. and Nakashima,N. (2006) Characterization of the 5′ internal ribosome entry site of *Plautia stali* intestine virus. *J. Gen. Virol.*, **87**, 3679–3686.
26. Garrey,J.L., Lee,Y.Y., Au,H.H., Bushell,M. and Jan,E. (2010) Host and viral translational mechanisms during cricket paralysis virus infection. *J. Virol.*, **84**, 1124–1138.
27. Moore,N.F., Kearns,A. and Pullin,J.S. (1980) Characterization of cricket paralysis virus-induced polypeptides in Drosophila cells. *J. Virol.*, **33**, 1–9.
28. Pöyry,T., Kinnunen,L. and Hovi,T. (1992) Genetic variation in vivo and proposed functional domains of the 5′ noncoding region of poliovirus RNA. *J. Virol.* **66**, 5313–5319.
29. Wang,C., Sarnow,P. and Siddiqui,A. (1994) A conserved helical element is essential for internal initiation of translation of hepatitis C virus RNA. *J. Virol.*, **68**, 7301–7307.
30. Kanamori,Y. and Nakashima,N. (2001) A tertiary structure model of the internal ribosome entry site (IRES) for methionine-independent initiation of translation. *RNA*, **7**, 266–274.
31. Dorokhov,Y.L., Skulachev,M.V., Ivanov,P.A., Zvereva,S.D., Tjulkina,L.G., Merits,A., Gleba,Y.Y., Hohn,T. and Atabekov,J.G. (2002) Polypurine (A)-rich sequences promote cross-kingdom conservation of internal ribosome entry. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 5301–5306.
32. Terenin,I.M., Dmitriev,S.E., Andreev,D.E., Royall,E., Belsham,G.J., Roberts,L.O. and Shatsky,I.N. (2005) A cross-kingdom internal ribosome entry site reveals a simplified mode of internal ribosome entry. *Mol. Cell. Biol.*, **25**, 7879–7888.
33. Groppelli,E., Belsham,G.J. and Roberts,L.O. (2007) Identification of minimal sequences of the *Rhopalosiphum padi* virus 5′ untranslated region required for internal initiation of protein synthesis in mammalian, plant and insect translation systems. *J. Gen. Virol.*, **88**, 1583–1588.
34. Boros,Á., Pankovics,P., Simmonds,P. and Reuter,G. (2011) Novel positive-sense, single-stranded RNA (+ssRNA) virus with di-cistronic genome from intestinal content of freshwater carp (*Cyprinus carpio*). *PLoS One*, **6**, e29145.
35. Skabkin,M.A., Skabkina,O.V., Dhote,V., Komar,A.A., Hellen,C.U. and Pestova,T.V. (2010) Activities of Ligatin and MCT-1/DENR in eukaryotic translation initiation and ribosomal recycling. *Genes Dev.*, **24**, 1787–1801.
36. Lomakin,I.B., Shirokikh,N.E., Yusupov,M.M., Hellen,C.U. and Pestova,T.V. (2006) The fidelity of translation initiation: reciprocal activities of eIF1, IF3 and YciH. *EMBO J.*, **25**, 196–210.
37. Pestova,T.V. and Hellen,C.U. (2001) Preparation and activity of synthetic unmodified mammalian tRNAi(Met) in initiation of transl.ation in vitro. *RNA*, **7**, 1496–1505.
38. Yu,Y., Sweeney,T.R., Kafasla,P., Jackson,R.J., Pestova,T.V. and Hellen,C.U. (2011) The mechanism of translation initiation on Aichivirus RNA mediated by a novel type of picornavirus IRES. *EMBO J.*, **30**, 4423–4436.
39. Sweeney,T.R., Dhote,V., Yu,Y. and Hellen,C.U. (2012) A distinct class of internal ribosomal entry site in members of the Kobuvirus and proposed Salivirus and Paraturdivirus genera of the Picornaviridae. *J. Virol.*, **86**, 1468–1486.
40. Knudsen,B. and Hein,J. (2003) Pfold: RNA secondary structure prediction using stochastic context-free grammars. *Nucleic Acids Res.*, **31**, 3423–3428.
41. Do,C.B., Woods,D.A. and Batzoglou,S. (2006) CONTRAfold: RNA secondary structure prediction without physics-based models. *Bioinformatics*, **22**, e90–e98.
42. Sato,K., Hamada,M., Asai,K. and Mituyama,T. (2009) CENTROIDFOLD: a web server for RNA secondary structure prediction. *Nucleic Acids Res.*, **37**, W277–W280.
43. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
44. Kolupaeva,V.G., de Breyne,S., Pestova,T.V. and Hellen,C.U. (2007) In vitro reconstitution and biochemical characterization of translation initiation by internal ribosomal entry. *Methods Enzymol.*, **430**, 409–439.
45. Wilkinson,K.A., Merino,E.J. and Weeks,K.M. (2006) Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat. Protoc.*, **1**, 1610–1616.
46. Pisarev,A.V., Unbehaun,A., Hellen,C.U. and Pestova,T.V. (2007) Assembly and analysis of eukaryotic translation initiation complexes. *Methods Enzymol.*, **430**, 147–177.
47. Pisarev,A.V., Kolupaeva,V.G., Yusupov,M.M., Hellen,C.U. and Pestova,T.V. (2008) Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes. *EMBO J.*, **27**, 1609–1621.
48. Unbehaun,A., Borukhov,S.I., Hellen,C.U. and Pestova,T.V. (2004) Release of initiation factors from 48S complexes during ribosomal subunit joining and the link between establishment of codon-anticodon base-pairing and hydrolysis of eIF2-bound GTP. *Genes Dev.*, **18**, 3078–3093.
49. Skabkin,M.A., Skabkina,O.V., Hellen,C.U. and Pestova,T.V. (2013) Reinitiation and other unconventional posttermination events during eukaryotic translation. *Mol. Cell*, **51**, 249–264.
50. Zinoviev,A., Hellen,C.U. and Pestova,T.V. (2015) Multiple mechanisms of reinitiation on bicistronic calicivirus mRNAs. *Mol. Cell*, **57**, 1059–1073.
51. Kolupaeva,V.G., Unbehaun,A., Lomakin,I.B., Hellen,C.U. and Pestova,T.V. (2005) Binding of eukaryotic initiation factor 3 to ribosomal 40S subunits and its role in ribosomal dissociation and anti-association. *RNA*, **11**, 470–486.
52. des Georges,A., Dhote,V., Kuhn,L., Hellen,C.U., Pestova,T.V., Frank,J. and Hashem,Y. (2015) Structure of mammalian eIF3 in the context of the 43S preinitiation complex. *Nature*, **525**, 491–495.
53. Shirokikh,N.E. and Spirin,A.S. (2008) Poly(A) leader of eukaryotic mRNA bypasses the dependence of translation on initiation factors. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 10738–10743.
54. Rajagopal,V., Park,E.H., Hinnebusch,A.G. and Lorsch,J.R. (2012) Specific domains in yeast translation initiation factor eIF4G strongly bias RNA unwinding activity of the eIF4F complex toward duplexes with 5′-overhangs. *J. Biol. Chem.*, **287**, 20301–20312.
55. Lomakin,I.B., Hellen,C.U. and Pestova,T.V. (2000) Physical association of eukaryotic initiation factor 4G (eIF4G) with eIF4A strongly enhances binding of eIF4G to the internal ribosomal entry site of encephalomyocarditis virus and is required for internal initiation of translation. *Mol. Cell. Biol.*, **20**, 6019–6029.
56. Prévôt,D., Décimo,D., Herbreteau,C.H., Roux,F., Garin,J., Darlix,J.L. and Ohlmann,T. (2003) Characterization of a novel RNA-binding region of eIF4GI critical for ribosomal scanning. *EMBO J.*, **22**, 1909–1921.

57. Kong,W.P. and Roos,R.P. (1991) Alternative translation initiation site in the DA strain of Theiler's murine encephalomyelitis virus. *J. Virol.*, **65**, 3395–3399.
58. Walewski,J.L., Keller,T.R., Stump,D.D. and Branch,A.D. (2001) Evidence for a new hepatitis C virus antigen encoded in an overlapping reading frame. *RNA*, **7**, 710–721.
59. Ren,Q., Wang,Q.S., Firth,A.E., Chan,M.M., Gouw,J.W., Guarna,M.M., Foster,L.J., Atkins,J.F. and Jan,E. (2012) Alternative reading frame selection mediated by a tRNA-like domain of an internal ribosome entry site. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, E630–E639.
60. Sorgeloos,F., Jha,B.K., Silverman,R.H. and Michiels,T. (2013) Evasion of antiviral innate immunity by Theiler's virus L* protein through direct inhibition of RNase L. *PLoS Pathog.*, **9**, e1003474.
61. Mandal,A., Mandal,S. and Park,M.H. (2014) Genome-Wide analyses and functional classification of proline repeat-Rich proteins: potential role of eIF5A in eukaryotic evolution. *PLoS One*, **9**, e111800.
62. Doerfel,L.K., Wohlgemuth,I., Kubyshkin,V., Starosta,A.L., Wilson,D.N., Budisa,N. and Rodnina,M.V. (2015) Entropic contribution of elongation Factor P to proline positioning at the catalytic center of the ribosome. *J. Am. Chem. Soc.*, **137**, 12997–13006.
63. Gutierrez,E., Shin,B.S., Woolstenhulme,C.J., Kim,J.R., Saini,P., Buskirk,A.R. and Dever,T.E. (2013) eIF5A promotes translation of polyproline motifs. *Mol. Cell*, **51**, 35–45.
64. Stadler,M. and Fire,A. (2011) Wobble base-pairing slows in vivo translation elongation in metazoans. *RNA*, **17**, 2063–2073.
65. Artieri,C.G. and Fraser,H.B. (2014) Accounting for biases in riboprofiling data indicates a major role for proline in stalling translation. *Genome Res.*, **24**, 2011–2021.
66. Matsuda,D. and Dreher,T.W. (2006) Close spacing of AUG initiation codons confers dicistronic character on a eukaryotic mRNA. *RNA*, **12**, 1338–1349.
67. Herbreteau,C.H., Weill,L., Décimo,D., Prévôt,D., Darlix,J.L., Sargueil,B. and Ohlmann,T. (2005) HIV-2 genomic RNA contains a novel type of IRES located downstream of its initiation codon. *Nat. Struct. Mol. Biol.*, **12**, 1001–1007.
68. Locker,N., Chamond,N. and Sargueil,B. (2011) A conserved structure within the HIV gag open reading frame that controls translation initiation directly recruits the 40S subunit and eIF3. *Nucleic Acids Res.*, **39**, 2367–2377.
69. Roberts,L.O. and Groppelli,E. (2009) An atypical IRES within the 5′ UTR of a dicistrovirus genome. *Virus Res.*, **139**, 157–165.