Functional conservation despite structural divergence in ligand-responsive RNA switches

Mark A. Boerneke[a], Sergey M. Dibrov[a], Jing Gu[a], David L. Wyles[b], and Thomas Hermann[a,c,1]

[a]Department of Chemistry and Biochemistry, [b]Division of Infectious Diseases, Department of Medicine, and [c]Center for Drug Discovery Innovation, University of California, San Diego, La Jolla, CA 92093

Edited by Jennifer A. Doudna, University of California, Berkeley, CA, and approved October 2, 2014 (received for review July 31, 2014)

An internal ribosome entry site (IRES) initiates protein synthesis in RNA viruses, including the hepatitis C virus (HCV). We have discovered ligand-responsive conformational switches in viral IRES elements. Modular RNA motifs of greatly distinct sequence and local secondary structure have been found to serve as functionally conserved switches involved in viral IRES-driven translation and may be captured by identical cognate ligands. The RNA motifs described here constitute a new paradigm for ligand-captured switches that differ from metabolite-sensing riboswitches with regard to their small size, as well as the intrinsic stability and structural definition of the constitutive conformational states. These viral RNA modules represent the simplest form of ligand-responsive mechanical switches in nucleic acids.

IRES elements | translation regulation | RNA viruses | hepatitis C virus

Internal ribosome entry site (IRES) elements provide an alternative mechanism for translation initiation by directing the assembly of functional ribosomes directly at the start codon in a process that does not require 5′ cap recognition or ribosomal scanning and that is independent of many host initiation factors (1–4). The genomes of *Flaviviridae* and *Picornaviridae* contain elements that share similarity with the archetypical hepatitis C virus (HCV) IRES in overall domain organization, but not sequence or details of secondary structure (5). The HCV IRES adopts a complex architecture of four independently folding domains (Fig. 1*A*) (6). Domain II is nearly 100% conserved in clinical isolates (7) and has analogous counterparts in other viral IRES elements, all of which display some secondary structure similarity, but significant sequence variation in the subdomain IIa-like internal loop (Fig. 1*B*). Domain II has been shown to promote stable entry of HCV and classic swine fever virus (CSFV) mRNA at the decoding groove of the 40S subunit (8-10) and is required for initiation factor removal before ribosomal subunit joining (11), as well as adjustment of initiator tRNA orientation (12). The transition from initiation to elongation stages of translation depends critically on domain II (13). Recently, direct interaction of HCV domain II with initiator tRNA has been demonstrated (14). In HCV, subdomain IIa folds into an L-shaped motif (15) (Fig. 1*C*) that introduces a 90° bend in domain II (16) and directs the IIb hairpin toward the E-site at the ribosomal subunit interface (17, 18).

The HCV IRES subdomain IIa is the target for viral translation inhibitors (Fig. 1*D*) that bind to the internal loop and block translation by capturing distinct conformational states of the RNA (7). Structure analysis revealed that benzimidazole inhibitors such as compound **1** (19, 20) interact with an extended architecture of IIa in which the stems flanking the internal loop are coaxially stacked on both sides of the ligand-binding pocket (Fig. 1*E*) (21). In contrast, diaminopiperidine compounds such as **2** bind and lock the IIa RNA in a bent conformation that corresponds to the ligand-free state (22). Conformational capture of the subdomain IIa switch by ligands in solution was demonstrated by FRET experiments and established as a mechanism of IRES inhibition (23). On the basis of these findings, it was proposed that subdomain IIa may be the target for a cognate biological ligand whose adaptive recognition by the RNA motif may facilitate ribosome release from the IRES-bound complex (7).

Here, we have explored potential candidates for a cognate ligand of the subdomain IIa switch and investigated the structural and functional conservation of similar ligand responsive switch motifs in other IRES RNAs.

Results

Guanine Captures an Extended Conformation of the Subdomain IIa Switch. Cryoelectron microscopy studies of the HCV IRES bound to the ribosome reveal the bent domain II interacting with the 40S subunit in a curved topology that would prevent the progression of the ribosome from initiation to elongation (17, 18). Conformational dynamics in the subdomain IIa RNA switch along with ligand capture of the extended state may facilitate removal of domain II from the ribosomal E site (21). Support for this hypothesis comes from recent cryo-EM studies that reveal major differences in the conformation of domain II in 40S-bound binary and 80S-bound initiation complexes of the HCV IRES (14). Synthetic benzimidazole translation inhibitors (compound **1**), which capture an extended conformation of subdomain IIa, appear to be fortuitous ligands of this RNA motif (7). The near-perfect conservation across clinical isolates of residues in subdomain IIa (7), along with the adaptive formation of a deep pocket that encapsulates the small molecule reminiscent of riboswitches, led us to speculate about a cognate biological ligand. Binding of compound **1** to subdomain IIa depends critically on two hydrogen bonds to the Hoogsteen edge of a G-C base pair, closely resembling isosteric hydrogen bonding patterns

Significance

RNA viruses, including the human pathogenic hepatitis C virus (HCV), use a structured untranslated region of their genome to hijack host cell ribosomes for the synthesis of viral proteins. These genome regions are termed internal ribosome entry site (IRES) elements and are encoded by distinct sequences in different viruses but share common functional RNA motifs. This study shows that viral IRES elements contain conformationally flexible RNA switches, whose state can be captured by the binding of a common ligand. Conformational switching plays a role in the function of the IRES elements. These new RNA switches are smaller than previously discovered "riboswitches" and may be the simplest form of ligand-responsive mechanical modules in nucleic acids.

Author contributions: M.A.B. and T.H. designed research; M.A.B., S.M.D., and J.G. performed research; D.L.W. contributed new reagents/analytic tools; M.A.B., S.M.D., and T.H. analyzed data; and M.A.B. and T.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 4P97 and 4PHY).

[1]To whom correspondence should be addressed. Email: tch@ucsd.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1414678111/-/DCSupplemental.

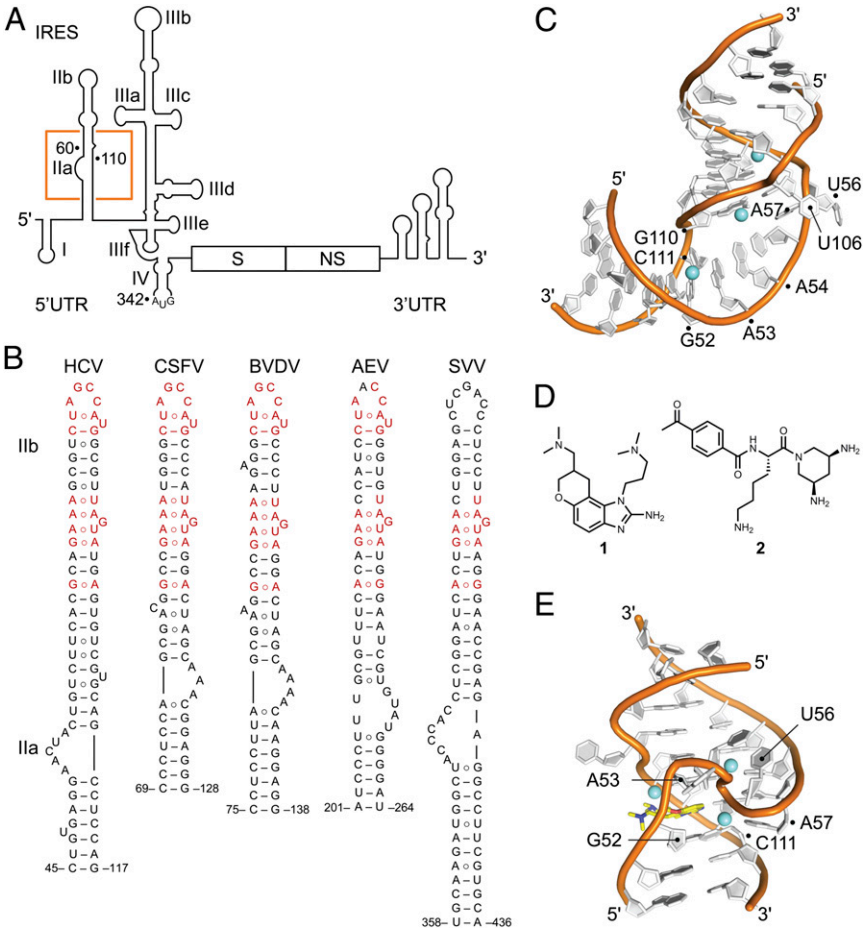

Fig. 1. Structures and ligands of viral IRES. (*A*) The IRES in the 5′ UTR of the HCV genome. The location of subdomain IIa is highlighted by an orange box. The viral genome encodes structural (S) and nonstructural (NS) proteins and contains a structured 3′ UTR. (*B*) Secondary structure predictions of domain II motifs in viral IRES elements from HCV and other flaviviruses, including CSFV and BVDV, as well as picornaviruses such as AEV and SVV. Non-Watson–Crick base pairs are indicated by the ○ symbol. Sequence conservation is indicated in red. (*C*) Crystal structure of the subdomain IIa RNA from HCV. (*D*) Benzimidazole (**1**) and diaminopiperidine (**2**) inhibitors of IRES-driven translation that target the HCV subdomain IIa. (*E*) Crystal structure of the HCV subdomain IIa RNA in complex with inhibitor **1**.

observed for G-C pairs interacting with arginine in protein–RNA complexes (24, 25) and with guanosine in C-G ∘ G triples (26, 27) (Fig. 2*A*).

Inspired by Yarus' classical study of arginine and G binding to self-splicing group I introns (28), we used a previously established FRET assay (23, 29) to test arginine, guanine, and their derivatives for binding to the subdomain IIa RNA. Conformational capture of the dye-labeled RNA switch (*SI Appendix*, Fig. S1) in the extended state leads to dose-dependent reduction of the FRET signal and concurrent increase of donor dye fluorescence resulting from diminished resonance energy transfer. The FRET assay has previously been used to demonstrate conformational capture of the IIa switch by benzimidazole ligands (23). Neither arginine nor its derivatives showed binding to the IIa RNA construct. In contrast, titration of guanine (**3**) elicited a dose-dependent FRET reduction indicative of ligand binding to the extended conformation of subdomain IIa (Fig. 2*B*). Although the binding activity of guanine was around 1 mM (EC_{50} value), which is 300 times weaker than that of the benzimidazole **1**, the concurrent increase of fluorescence from the Cy3 donor dye substantiated a selective ligand interaction, leading to conformational capture. Binding was specific for guanine, either as the base itself or in guanosine, whereas other nucleobases were inactive in the FRET assay (*SI Appendix*, Fig. S2 *A* and *B*). The 2-aminopyrimidinone ring of guanine was essential for the interaction with the subdomain IIa RNA, as indicated by the binding activity of 2-aminoquinazolinone, which shares the pyrimidinone but not the imidazole heterocycle (compound **4**, *SI Appendix*, Fig. S2*C*). The quinazolinone **4** bound the target about twofold tighter than guanine (EC_{50} value, 483 μM), perhaps as a result of more favorable stacking interactions with the benzene ring in **4**.

To test for the effect of the guanine interaction on IRES function, we measured luciferase reporter expression levels from a bicistronic construct in an in vitro translation assay (29). The bicistronic construct allows measuring effects on IRES-driven translation of Renilla luciferase while also providing an internal control of a cap-initiated firefly luciferase (*SI Appendix*, Fig. S3). In agreement with the observed binding of guanine to subdomain IIa, the nucleobase selectively inhibited expression of the IRES-driven reporter (Fig. 2*C*). Translation inhibition required a higher concentration of guanine compared with **1** (*SI Appendix*, Fig. S4), again reflecting the weaker binding of guanine relative to the synthetic inhibitor. Inhibition of reporter expression was selective for guanine, as none of the other three nucleobases showed this effect (*SI Appendix*, Fig. S4).

Domain II-Like Motifs from Other IRES RNAs Are Unrelated in Sequence but Fold Similar to the HCV Archetype. IRES elements and their function in translation initiation differ fundamentally in distinct families of positive-sense RNA viruses. The IRES domains of flaviviruses such as HCV, CSFV, and bovine viral diarrhea virus (BVDV) recruit 40S subunits independent of eukaryotic initiation factors, whereas IRES elements of most picornaviruses require involvement of eukaryotic initiation factor 4 components (30). Despite the differences, RNA domains have been identified in the IRES elements of picornaviruses, including avian encephalomyelitis virus (AEV) and Seneca Valley virus (SVV) (31), which resemble analogous motifs in flaviviruses in overall secondary structure (5). RNA structures corresponding to the HCV domain II occur in the IRES elements of CSFV, BVDV, AEV, and SVV (Fig. 1*B*), as well as several other RNA viruses (5). Domain II-like RNAs share common motifs,

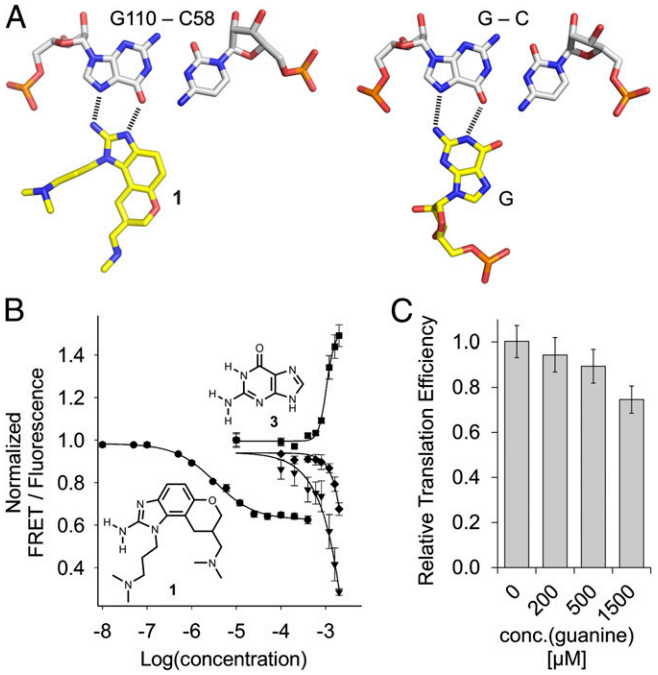

Fig. 2. Ligand binding to the HCV IRES. (*A*) Docking of the benzimidazole **1** at the C58-G110 base pair in the crystal structure of the HCV subdomain IIa target complex in comparison with the geometry of a C-G∘G triple. (*B*) Titrations of Cy3/Cy5-labeled HCV subdomain IIa RNA with benzimidazole **1** and guanine **3**. Curves show normalized FRET signal for **1** (●) and guanine (▼), as well as the normalized fluorescence signals of the donor Cy3 (■) and acceptor Cy5 (♦) for guanine. Fitting of a single-site binding curve to the benzimidazole **1** FRET response gave an EC_{50} value for ligand binding of 3.4 ± 0.3 μM. Because the FRET signal did not reach saturation in the guanine titration, affinity was estimated by fitting to the Cy3 emission, which resulted in an EC_{50} value for ligand binding of 1,050 ± 69 μM (guanine). (*C*) Effect of guanine on IRES-driven translation, as measured in an in vitro translation assay. Error bars in *B* and *C* represent ±1 SD calculated from triplicate experiments.

including a hairpin loop in subdomain IIb and a loop E motif (32), which are conserved in sequence and separation by base pairs (Fig. 3*A*). The internal loop of subdomain IIa shows no sequence conservation and differs both in size and location in the 5′ or 3′ proximal strand of the lower stem. NMR analysis of domain II from CSFV, which revealed a bent subdomain IIa (11), hinted at the structural similarity of subdomain IIa motifs despite their distinctions in sequence, size, and strand location.

To investigate whether subdomain IIa motifs from different viral IRES elements share the ability to adopt structurally similar bent architectures, we studied magnesium ion-induced folding of the RNAs by monitoring FRET. We had previously used the FRET assay to establish magnesium-dependent folding of the HCV subdomain IIa (23). We titrated dye-labeled constructs containing viral subdomain IIa motifs (*SI Appendix*, Fig. S1), which were initially free of metal ions, with increasing amounts of Mg^{2+} (Fig. 3*B*). FRET was not observed for any of the constructs in the absence of Mg^{2+} when the internal loop of subdomain IIa does not stably fold and the RNA adopts an extended conformation that places the dyes beyond the Förster radius. On the addition of Mg^{2+}, the FRET signal appeared in a dose-dependent fashion, indicating folding into a bent architecture for constructs of all viruses. Control RNA in which the internal loop was deleted did not show an increase of FRET (*SI Appendix*, Fig. S2*D*). Effective Mg^{2+} concentrations for folding were of similar magnitude for all constructs, ranging from ∼130–600 μM (Fig. 3*B*). Constructs that were more closely related with respect to overall secondary structure required similar amounts of magnesium for folding, delineating two distinct subtypes of subdomain IIa architectures (CSFV, BVDV, AEV: 131–219 μM Mg^{2+}; SVV, HCV: 579–598 μM Mg^{2+}). Separation between these subtypes did not coincide with divisions between *Flaviviridae* and *Picornaviridae*.

To further explore the ability for conformational switching that subdomain IIa analogs may share in common with the HCV archetype motif, we used the FRET assay to test other viral motifs for interaction with the benzimidazole **1**. Titration of the compound showed that ligand binding captured the subdomain IIa RNAs of CSFV, BVDV, and SVV in an extended state (*SI Appendix*, Fig. S2*E*), similar to that previously demonstrated for the HCV target (Fig. 2*B*), albeit requiring an 8–15-fold higher concentration of **1**. The motif from AEV showed weaker interaction with the compound, and saturation of the ligand bound state was not achieved. Because subdomain IIa of SVV bound the benzimidazole tightest among the viral analogs, this RNA was also tested for its interaction with guanine. Similar to the subdomain IIa motif of HCV, the SVV analog was captured in an extended state by guanine at a comparable EC_{50} value of 990 μM (*SI Appendix*, Fig. S2*F*).

Subdomain IIa RNA from the SVV IRES Element Adopts a 90° Bent Fold. Because folding studies indicated that diverse viral subdomain IIa analogs adopt bent architectures, we attempted to elucidate the 3D structure of the RNAs. Crystal structures of two different constructs of the SVV subdomain IIa were determined. A high-resolution structure (1.86 Å) was obtained for the internal loop (Fig. 4 and *SI Appendix*, Fig. S5), and a lower-resolution structure (3.2 Å) was obtained for an extended construct that included the region connecting to the loop E motif (*SI Appendix*, Fig. S6). Both structures showed identical architecture for the internal loop.

The SVV IRES subdomain IIa adopts a bent fold that arranges the flanking RNA helices at an angle of 90° (Fig. 4*A* and *SI Appendix*, Figs. S5 and S6). The overall architecture of subdomain II from SVV was nearly identical to that of the motif from HCV (*SI Appendix*, Fig. S7), despite differences in sequence and secondary structure. The folding of the bend in the SVV motif has no precedent in other RNA architectures. Except for the reverse Hoogsteen pair U369∘A424, all base pairs adopt canonical Watson–Crick geometry, and all unpaired bases are

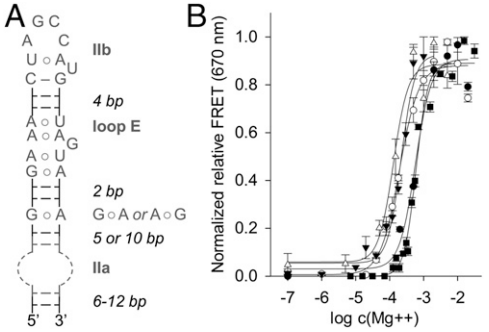

Fig. 3. Domain II-like RNAs in viral IRES elements. (*A*) Consensus secondary structure of domain II analogs from viral IRES elements. The internal loop of subdomain IIa consists of an unpaired stretch of nucleotides in either the 5′ or 3′ proximal strand of the lower stem. (*B*) Mg^{2+}-induced folding of viral subdomain IIa motifs. FRET signal from terminally Cy3/Cy5 labeled RNA constructs was monitored while increasing Mg^{2+} concentration. See *SI Appendix*, Fig. S1 for RNA construct structures. Fitting of dose–response curves resulted in EC_{50} values for Mg^{2+}-induced folding at 598 ± 25 μM (HCV, ■) (23), 196 ± 21 μM (CSFV, ○), 219 ± 19 μM (BVDV, ▼), 131 ± 21 μM (AEV, Δ), and 579 ± 86 μM (SVV, ●). Error bars represent ±1 SD calculated from triplicate experiments.

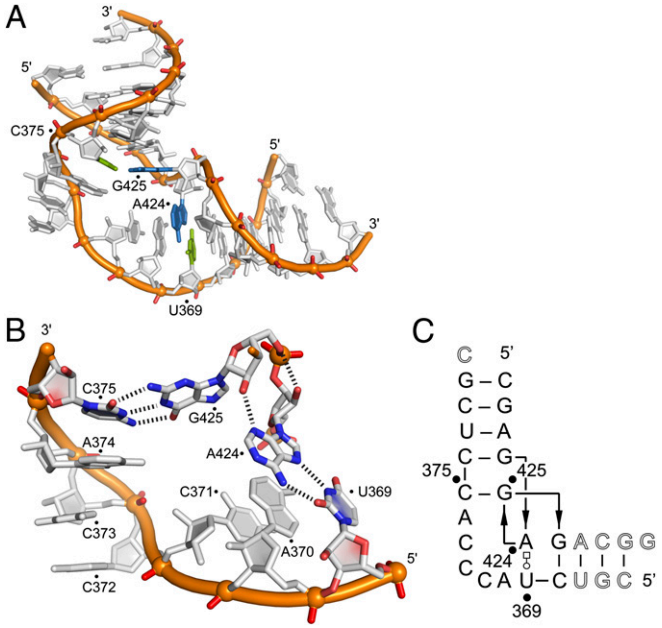

Fig. 4. Structure of the subdomain IIa RNA in the SVV IRES element. (*A*) Crystal structure highlighting the perpendicular U369 ∘ A424 trans-Watson–Crick-Hoogsteen base pair and the Watson–Crick C375-G425 base pair. (*B*) Detail view of the internal loop showing additional hydrogen bonds between A424 and G425 in the perpendicular base pairs ($N3_{A424}$...$O2'_{G425}$ and $O2'_{A424}$...$O5'_{G425}$). (*C*) Secondary structure schematic illustrating base pairing and consecutive stacking interactions. Nucleotides deviating from the SVV wild-type sequence are shown in outlined font.

continuously stacked on one of the flanking helices. The U369 ∘ A424 and C375-G425 pairs are arranged orthogonally and provide the closing ends that interface the flanking helices to the internal loop (Fig. 4*A*). The perpendicular alignment of consecutive bases A424 and G425 is stabilized by hydrogen bonds involving the sugar backbone of these residues (Fig. 4*B*). Resulting from the orthogonal orientation of A424 and G425, the U369 ∘ A424 and C375-G425 pairs form a cross-over motif, and strand directionality is locally inverted in the secondary structure (Fig. 4*C*).

Domain II-Like RNAs Are Interchangeable Functional Modules Despite Their Sequence Dissimilarity. To determine whether domain II analogs are autonomous RNA motifs that are functional as interchangeable modules, we investigated reporter expression from chimera IRES constructs in which the cognate HCV domain II was substituted by analogs from other viruses (*SI Appendix*, Fig. S3). Replacement of the HCV domain II by other viral analogs resulted in IRES proficient for translation initiation, except for the SVV domain, which does not carry a homologous hairpin sequence (Fig. 5*A*). These observations are in agreement with a previous report on translational activity for a similar HCV IRES chimera that had the domain II inserted from CSFV (33). Deletion of the subdomain IIa-like region in the chimera IRES elements, which ablates the bend in domain II, led to significant loss of function in all constructs (Fig. 5*A*, ΔIIa). Removal of two base pairs between the subdomain IIb hairpin loop and the conserved loop E motif, which affects both the distance and rotational orientation of the two motifs, resulted in a reduction of activity that was similar across different chimeras (Fig. 5*A*; Δ2bp).

The availability of a crystal structure for the SVV subdomain IIa allowed us to determine a minimal architectural module in this RNA that corresponded to the bend in the HCV motif. Guided by superposition of RNA structures from both viruses, we substituted in the HCV IRES reporter construct the internal loop comprised of residues G52...U59/A109...C111 by the corresponding module from SVV consisting of C368...C376/G423...G426 (Fig. 5*B*). Although the swapped RNA modules shared no sequence similarity, the resulting chimera IRES was fully proficient in translation initiation (Fig. 5*C*, construct X). Even a small deviation from the structure-guided module swap, the introduction of a single base pair offset (construct P, *SI Appendix*, Fig. S8), resulted in a 50% loss in IRES activity (Fig. 5*C*). The chimera construct X also conferred full translation activity of HCV replicon in human cells 4 h after transfection, when viral translation occurs, but not yet replication (34) (Fig. 5*D*). However, expression activity from the chimera replicon was greatly diminished after 24 h, when the cumulative reporter signal is expected to be dominated by replication. Apparently, the motif swap of subdomain IIa modules negatively affected communication between the 5‘and 3′ UTR, which has been implicated in viral replication (35, 36).

Replacement of the subdomain IIa motif in HCV by the corresponding analog from the CSFV IRES was performed to investigate the effect of relocating the internal loop from the 5′ to the 3′ proximal strand of the domain II lower stem. Secondary

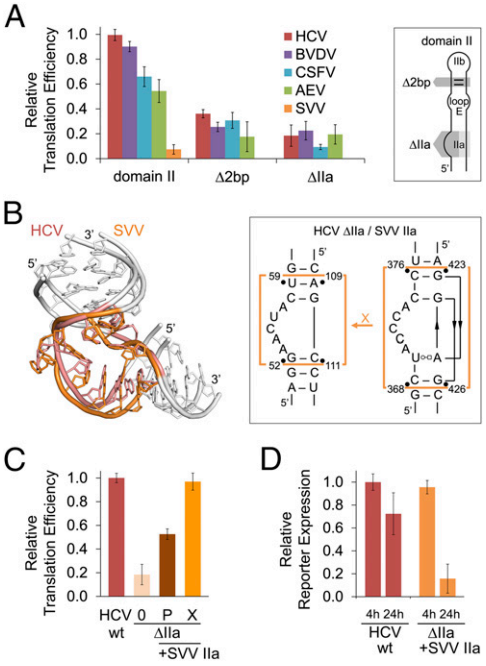

Fig. 5. Functional competence of chimera HCV IRES elements. (*A*) Effect on in vitro translation of domain II replacement in the HCV IRES by corresponding motifs from other viral IRES elements and additional deletion of two base pairs between the hairpin loop IIb and the loop E motif (Δ2bp), or deletion of the internal loop IIa (ΔIIa). See *SI Appendix*, Fig. S3 for chimera structure and sequences. (*B*) Construction of a chimera X (orange) by an exact structural motif swap based on superimposition of the subdomain IIa crystal structures from HCV (red/white) and SVV (orange/white). (*C*) Effect on in vitro translation of subdomain IIa replacement in the HCV IRES by the corresponding motif from SVV (chimera X, outlined in *B*). The chimera P (brown) is a control construct that contains a subdomain IIa swap with a one base pair offset relative to the X construct (*SI Appendix*, Fig. S8). HCV IRES wild-type and subdomain IIa deletion (ΔIIa, 0) are shown as controls. Translation efficiencies were normalized to the cap-driven expression in bicistronic dual reporter constructs. (*D*) Function of the HCV/SVV IRES chimera X in replicon-transfected human cells. Reporter expression was measured 4 and 24 h after transfection, with the subgenomic replicon RNA carrying the IRES subdomain IIa replaced by the corresponding motif from SVV. Error bars represent ±1 SD calculated from triplicate experiments, except for in *D*, where triplicates of three biological replicates (nine values) were used.

structure-guided swapping of the internal loop modules did not result in a translation-proficient IRES. An active chimera was obtained only after shortening the distance by 5 base pairs between the subdomain IIa internal loop and the conserved loop E, in combination with relocating a bulged-out pyrimidine from the 3′ to the 5′ proximal strand (*SI Appendix*, Fig. S9). These changes were motivated by the observation that in domain II analogs that carry the internal loop in the 3′ proximal strand, the distance to the loop E motif is shorter by 5 base pairs, corresponding to a half turn of an A-form RNA helix. A bulged-out pyrimidine, separated by 4 base pairs above the subdomain IIa internal loop, is found consistently in the strand opposite to the internal loop (Fig. 1*B*). These observations provided the basis for a topological model of the modular domain II architecture in viral IRES elements (Fig. 6*A*).

Discussion

The highly conserved subdomain IIa in the HCV IRES had previously been established as an RNA conformational switch that is able to convert from a ligand-free bent state to a ligand-bound extended conformation (7). The bent RNA fold is stabilized by magnesium ions and ensures the correct positioning of the IRES on the 40S ribosomal subunit and docking of domain II at the E site (17, 18). Removal of domain II from the ribosome during translation initiation may be topologically achieved by a conformational change in subdomain IIa transitioning from the bent to the extended state (21). Recent cryo-EM studies confirm the conformational flexibility of domain II through comparison of 40S-bound binary and 80S-bound initiation complexes of the HCV IRES (14). Motivated by the discovery of a riboswitch-like binding site for benzimidazole translation inhibitors (19) in the extended form of subdomain IIa (21), we hypothesized that a cognate ligand may serve as the biological actuator of the RNA switch that ultimately facilitates ribosomal release from the IRES. Inspection of the crystal structure previously determined for an inhibitor complex of subdomain IIa suggested that benzimidazole derivatives might be fortuitous ligands that exploit a recognition site for guanine. Here, we used a FRET assay to demonstrate binding and capture of the subdomain IIa RNA in an extended conformation by guanine and guanosine. Guanine binding to the subdomain IIa target led to inhibition of IRES-driven translation, albeit requiring a ∼300-fold higher concentration compared with potent benzimidazole derivatives. Structural comparison shows that the guanine base may engage in similar interactions with subdomain IIa as those observed for benzimidazole inhibitors. The resulting arrangement corresponds to the interaction of guanosine with a Watson–Crick G-C pair in a geometry that accounts for one of the most frequently occurring base triples in RNA folds (27).

Modeling of guanine in the ligand-binding pocket of the benzimidazole-RNA crystal structure revealed suboptimal stacking by the nucleobase on neighboring residues (*SI Appendix*, Fig. S10). Better stacking is achieved with the larger quinazolinone **4**, which has a ∼twofold higher affinity for the RNA target than guanine (Fig. 2*B*). The smaller size in combination with the neutral character of the ligand may explain the weaker binding affinity of guanine compared with benzimidazole inhibitors. The nucleobase as a putative biological ligand of the subdomain IIa RNA switch may function as a transient actuator that readily dissociates from the IRES, unlike the benzimidazoles, which arrest the IRES topology and thereby inhibit translation initiation. It is conceivable that a guanosine in an RNA sequence such as ribosomal RNA or the viral genome emerging from the ribosome may serve as a trigger facilitating IRES release from the ribosome. Docking of IRES fragment crystal structures to a cryo-EM map of an IRES-40S ribosomal complex revealed a cluster of guanosines from the central domain pseudoknot and domain IV residing in a single-stranded stretch upstream of the viral initiation codon and located in close proximity to subdomain IIa (37). Involvement of a guanosine from the viral genome as a trigger for the subdomain IIa switch would capitalize on a locally high concentration of the ligand as a part of the same RNA strand. In contrast, participation of cellular guanine or guanosine is unlikely because of their low physiological concentration of around 97 and 0.9 μM, respectively (38). GTP, although present in human cells at a higher level (∼300 μM), is a less suitable ligand for the IIa target because of electrostatic repulsion by the phosphate groups.

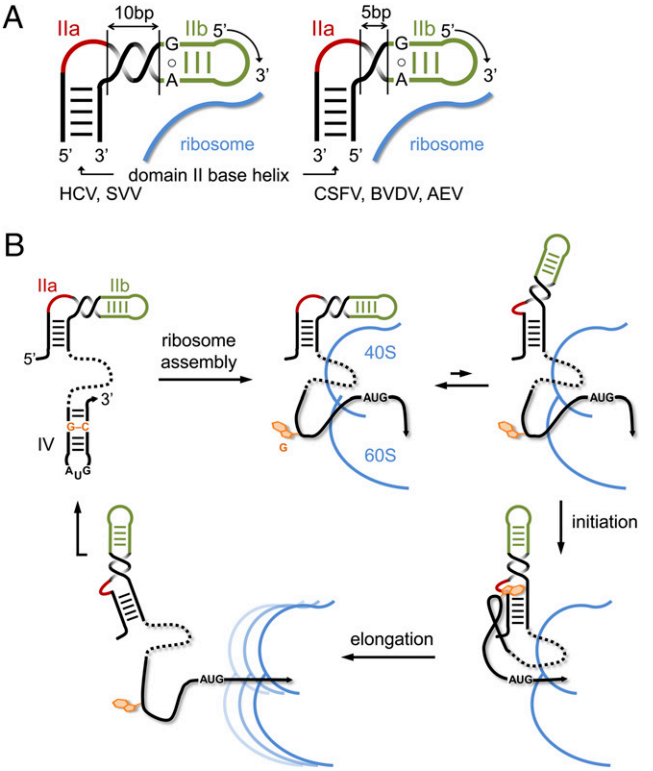

Fig. 6. Model for the modular domain II architecture in viral IRES elements and its participation in IRES regulation. (*A*) For the correct positioning of the conserved hairpin loop IIb at the 40S ribosomal subunit, the bent internal loop IIa module has to be located in the 5′ or 3′ strand of the base helix, depending on the number of base pairs in the spacer segment between IIa and IIb. A spacer of 10 base pairs, as in the IRES elements of HCV and SVV, corresponds to a full turn of an A-form RNA helix, whereas a spacer of 5 base pairs provides a half turn. (*B*) Proposed mechanism for the regulation of IRES activity by the subdomain IIa switch through an interaction with a G residue from domain IV. Ribosome binding of the IRES entails melting of the domain IV hairpin to provide access of the viral start codon to the decoding site, which in turn unmasks the trigger G residue that facilitates IRES release during initiation by capture of the IIa extended state. After the ribosome transitions to elongation, the unbound IRES may refold to the original state.

Regulation of the HCV IRES activity by the subdomain IIa switch through interaction in *cis* with a guanosine from domain IV would provide an efficient autoregulatory mechanism in which the trigger G residue is sequestered in a hairpin loop of the unbound IRES to allow ribosome assembly supported by the bent IIa RNA motif. Placement of the viral start codon at the decoding site requires melting of the domain IV hairpin, which in turn unmasks the trigger G residue that facilitates IRES release during initiation through capture of the IIa extended state (Fig. 6*B*).

The discovery of the IRES subdomain IIa as a ligand-responsive RNA switch in HCV raised the question of whether corresponding motifs in other viruses adopt analogous folds whose conformational state may be captured by ligand binding. Here, we demonstrated that RNA motifs from several flaviviruses and picornaviruses fold into magnesium-stabilized bent architectures similar to the HCV archetype, despite their sequence

dissimilarity (Fig. 3). Earlier NMR studies of domain II from CSFV performed at lower salt concentration already showed a bent structure for this RNA (11). The crystal structure of the internal loop motif from SVV (Fig. 4) revealed a fold that is overall identical to the HCV subdomain IIa despite the fact that the two motifs share little sequence or local secondary structure similarity (*SI Appendix*, Fig. S8). Analogous subdomain IIa RNAs from other viral IRESs interact with the benzimidazole **1**, as well as guanine, to be captured in an extended conformation (*SI Appendix*, Fig. S2 *E* and *F*). Binding of **1** to the subdomain IIa RNAs from other viruses was ∼10-fold weaker than the affinity measured for the HCV target, perhaps reflecting the extensive optimization of the benzimidazole derivative for inhibition of the HCV IRES (19, 39).

The analogies between viral subdomain IIa RNA motifs extend beyond the similarity of static structure and the ability to adopt two distinct conformational states. Domain swap experiments in which crystal structure information was used to precisely replace the subdomain IIa in the HCV IRES by the corresponding motif from SVV demonstrated that biological function is completely conserved between these intrinsically distinct RNA building blocks (Fig. 5 *B* and *C*). HCV-SVV chimera IRES elements were fully functional both in vitro and in replicon-infected cells. Even the structurally less accurate replacement of the whole domain II in the HCV RNA with analogous motifs from BVDV, CSFV, and AEV yielded functional chimera IRES elements (Fig. 5*A*). We conclude that subdomain IIa motifs of greatly distinct sequence and local secondary structure may serve as functionally conserved RNA conformational switches that are involved in viral IRES-driven translation and may be captured by highly similar or even identical ligands. We propose that the biological capture ligand shared between the different viral RNA motifs is likely a guanosine.

The functionally conserved subdomain IIa motifs constitute a new paradigm for ligand-captured RNA switches that differ from metabolite-sensing riboswitches with regard to their small size, as well as the intrinsic stability and structural definition of the constitutive conformational states. These viral RNA modules represent the simplest form of ligand-responsive mechanical switches in nucleic acids.

Materials and Methods

Preparation of RNA constructs is outlined in the *SI Appendix, Supporting Materials and Methods*. FRET folding and compound screening experiments were performed as described earlier (29). The in vitro transcription-translation assay was performed using HCV bicistronic luciferase constructs, as previously reported (40). HCV replicon testing followed procedures outlined earlier (23, 41). Experimental details for these methods, as well as the crystallization and structure determination of SVV subdomain IIa RNA, are described in the *SI Appendix, Supporting Materials and Methods*.

ACKNOWLEDGMENTS. M.A.B. was supported by a Graduate Assistance in Areas of National Need fellowship from the US Department of Education. Support for this project was provided by University of California, San Diego, Academic Senate Grant RM069B, for the Biomolecule Crystallography Facility by the National Institutes of Health (Grant OD011957), and for the NMR facility by the National Science Foundation (Chemistry Research Instrumentation and Facilities Program Grant CHE-0741968).

1. Plank TD, Kieft JS (2012) The structures of nonprotein-coding RNAs that drive internal ribosome entry site function. *Wiley Interdiscip Rev RNA* 3(2):195–212.
2. Hellen CU, Pestova TV (1999) Translation of hepatitis C virus RNA. *J Viral Hepat* 6(2):79–87.
3. Ji H, Fraser CS, Yu Y, Leary J, Doudna JA (2004) Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA. *Proc Natl Acad Sci USA* 101(49):16990–16995.
4. Otto GA, Puglisi JD (2004) The pathway of HCV IRES-mediated translation initiation. *Cell* 119(3):369–380.
5. Hellen CU, de Breyne S (2007) A distinct group of hepacivirus/pestivirus-like internal ribosomal entry sites in members of diverse picornavirus genera: Evidence for modular exchange of functional noncoding RNA elements by recombination. *J Virol* 81(11):5850–5863.
6. Kieft JS, et al. (1999) The hepatitis C virus internal ribosome entry site adopts an ion-dependent tertiary fold. *J Mol Biol* 292(3):513–529.
7. Dibrov SM, et al. (2014) Hepatitis C virus translation inhibitors targeting the internal ribosomal entry site. *J Med Chem* 57(5):1694–1707.
8. Pestova TV, Shatsky IN, Fletcher SP, Jackson RJ, Hellen CU (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(1):67–83.
9. Kolupaeva VG, Pestova TV, Hellen CU (2000) Ribosomal binding to the internal ribosomal entry site of classical swine fever virus. *RNA* 6(12):1791–1807.
10. Filbin ME, Kieft JS (2011) HCV IRES domain IIb affects the configuration of coding RNA in the 40S subunit's decoding groove. *RNA* 17(7):1258–1273.
11. Locker N, Easton LE, Lukavsky PJ (2007) HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly. *EMBO J* 26(3):795–805.
12. Filbin ME, Vollmar BS, Shi D, Gonen T, Kieft JS (2013) HCV IRES manipulates the ribosome to promote the switch from translation initiation to elongation. *Nat Struct Mol Biol* 20(2):150–158.
13. Pestova TV, de Breyne S, Pisarev AV, Abaeva IS, Hellen CU (2008) eIF2-dependent and eIF2-independent modes of initiation on the CSFV IRES: A common role of domain II. *EMBO J* 27(7):1060–1072.
14. Yamamoto H, et al. (2014) Structure of the mammalian 80S initiation complex with initiation factor 5B on HCV-IRES RNA. *Nat Struct Mol Biol* 21(8):721–727.
15. Dibrov SM, Johnston-Cox H, Weng YH, Hermann T (2007) Functional architecture of HCV IRES domain II stabilized by divalent metal ions in the crystal and in solution. *Angew Chem Int Ed Engl* 46(1-2):226–229.
16. Lukavsky PJ, Kim I, Otto GA, Puglisi JD (2003) Structure of HCV IRES domain II determined by NMR. *Nat Struct Biol* 10(12):1033–1038.
17. Spahn CM, et al. (2001) Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit. *Science* 291(5510):1959–1962.
18. Boehringer D, Thermann R, Ostareck-Lederer A, Lewis JD, Stark H (2005) Structure of the hepatitis C virus IRES bound to the human 80S ribosome: Remodeling of the HCV IRES. *Structure* 13(11):1695–1706.
19. Seth PP, et al. (2005) SAR by MS: Discovery of a new class of RNA-binding small molecules for the hepatitis C virus: Internal ribosome entry site IIA subdomain. *J Med Chem* 48(23):7099–7102.
20. Parker MA, Satkiewicz E, Hermann T, Bergdahl BM (2011) An efficient new route to dihydropyranobenzimidazole inhibitors of HCV replication. *Molecules* 16(1):281–290.
21. Dibrov SM, et al. (2012) Structure of a hepatitis C virus RNA domain in complex with a translation inhibitor reveals a binding mode reminiscent of riboswitches. *Proc Natl Acad Sci USA* 109(14):5223–5228.
22. Carnevali M, Parsons J, Wyles DL, Hermann T (2010) A modular approach to synthetic RNA binders of the hepatitis C virus internal ribosome entry site. *ChemBioChem* 11(10):1364–1367.
23. Parsons J, et al. (2009) Conformational inhibition of the hepatitis C virus internal ribosome entry site RNA. *Nat Chem Biol* 5(11):823–825.
24. Puglisi JD, Chen L, Frankel AD, Williamson JR (1993) Role of RNA structure in arginine recognition of TAR RNA. *Proc Natl Acad Sci USA* 90(8):3680–3684.
25. Kondo J, Westhof E (2011) Classification of pseudo pairs between nucleotide bases and amino acids by analysis of nucleotide-protein complexes. *Nucleic Acids Res* 39(19):8628–8637.
26. Walberer BJ, Cheng AC, Frankel AD (2003) Structural diversity and isomorphism of hydrogen-bonded base interactions in nucleic acids. *J Mol Biol* 327(4):767–780.
27. Abu Almakarem AS, Petrov AI, Stombaugh J, Zirbel CL, Leontis NB (2012) Comprehensive survey and geometric classification of base triples in RNA structures. *Nucleic Acids Res* 40(4):1407–1423.
28. Yarus M (1988) A specific amino acid binding site composed of RNA. *Science* 240(4860):1751–1758.
29. Zhou S, Rynearson KD, Ding K, Brunn ND, Hermann T (2013) Screening for inhibitors of the hepatitis C virus internal ribosome entry site RNA. *Bioorg Med Chem* 21(20):6139–6144.
30. Pestova TV, et al. (2001) Molecular mechanisms of translation initiation in eukaryotes. *Proc Natl Acad Sci USA* 98(13):7029–7036.
31. Willcocks MM, et al. (2011) Structural features of the Seneca Valley virus internal ribosome entry site (IRES) element: A picornavirus with a pestivirus-like IRES. *J Virol* 85(9):4452–4461.
32. Leontis NB, Westhof E (1998) A common motif organizes the structure of multi-helix loops in 16 S and 23 S ribosomal RNAs. *J Mol Biol* 283(3):571–583.
33. Reusken CB, Dalebout TJ, Eerligh P, Bredenbeek PJ, Spaan WJ (2003) Analysis of hepatitis C virus/classical swine fever virus chimeric 5′NTRs: Sequences within the hepatitis C virus IRES are required for viral RNA replication. *J Gen Virol* 84(Pt 7):1761–1769.
34. Blight KJ, McKeating JA, Marcotrigiano J, Rice CM (2003) Efficient replication of hepatitis C virus genotype 1a RNAs in cell culture. *J Virol* 77(5):3181–3190.
35. Shi ST, Lai MMC (2006) HCV 5′ and 3′UTR: When Translation Meets Replication. *Hepatitis C Viruses: Genomes and Molecular Biology*, ed Tan SL (Horizon Bioscience, Norfolk).
36. Romero-López C, Berzal-Herranz A (2009) A long-range RNA-RNA interaction between the 5′ and 3′ ends of the HCV genome. *RNA* 15(9):1740–1752.
37. Berry KE, Waghray S, Mortimer SA, Bai Y, Doudna JA (2011) Crystal structure of the HCV IRES central domain reveals strategy for start-codon positioning. *Structure* 19(10):1456–1466.
38. Traut TW (1994) Physiological concentrations of purines and pyrimidines. *Mol Cell Biochem* 140(1):1–22.
39. Seth PP, Jefferson EA, Griffey RH, Swayze EE (2004) Benzimidazoles and analogs thereof as antivirals. World Patent Appl WO2004050035A3.
40. Brunn ND, Garcia Sega E, Kao MB, Hermann T (2012) Targeting a regulatory element in human thymidylate synthase mRNA. *ChemBioChem* 13(18):2738–2744.
41. Wyles DL, Kaihara KA, Vaida F, Schooley RT (2007) Synergy of small molecular inhibitors of hepatitis C virus replication directed at multiple viral targets. *J Virol* 81(6):3005–3008.

Supporting Information

Functional conservation despite structural divergence in ligand responsive RNA switches

Mark A. Boerneke, Sergey M. Dibrov, Jing Gu, David L. Wyles and Thomas Hermann

Supporting Materials and Methods

RNA Preparation. Cyanine dye labeled and unlabeled RNA oligonucleotides were obtained from chemical synthesis and purified by HPLC (Integrated DNA Technologies, Coralville, IA). Stock solutions were prepared by dissolving lyophilized oligonucleotides in 10mM sodium cacodylate buffer, pH 6.5.

FRET Folding Experiments. Terminally Cy3/Cy5-labeled IIa RNA constructs were annealed from single strands by heating to 65°C for 5 min followed by snap cooling in 10mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) buffer, pH 7.0. FRET folding experiments were performed as described previously (1) on a Spectra Max Gemini monochromator plate reader (Molecular Devices, Sunnyvale, CA) at 25°C. RNA was brought to a final concentration of 100nM in 10mM HEPES buffer, pH 7.0. Emission filters were set at 550 and 665 nm. The Cy3 label was excited at 520 nm and transferred fluorescence was read as Cy5 emission at 670 nm. FRET folding of labeled IIa RNA constructs was monitored while increasing Mg^{2+} concentration. Data sets were analyzed and FRET calculated as described previously (1).

FRET Compound Screening Experiments. FRET compound screening experiments were performed as described in the FRET folding experiments above, except RNA was brought to a final concentration of 100nM in 10mM sodium cacodylate buffer, pH 5.5, containing 2mM $MgCl_2$. Ligand induced FRET changes in terminally Cy3/Cy5-labelled IIa RNA constructs was monitored while increasing ligand concentration. To record fluorescence of the Cy3 dye, excitation was done at 520nm and emission read at 570nm with an emission filter at 550nm. To record fluorescence of the Cy5 dye, excitation was done at 620nm and emission read at 670nm with an emission filter at 665nm.

***In Vitro* Transcription-Translation Experiments.** The *in vitro* transcription-translation assay (IVT) was performed using the TNT Quick coupled reticulocyte lysate system (Promega, Madison, WI) and an HCV bicistronic luciferase reporter as previously described (2). The bicistronic luciferase reporter contains the sequence coding for the HCV IRES-Renilla luciferase preceded by a cap-initiated firefly luciferase internal control. Briefly, reactions were carried out according to the manufacturer's instruction at a volume of 7.5μL, containing 1.5μL reporter DNA plasmid (100ng/μL), 1.5μL H_2O or compound solution, and 4.5μL reaction buffer containing reticulocyte lysate, SP6 polymerase, RNase inhibitor, and amino acids. Detection of firefly and Renilla luciferase levels was done using the Dual-Glo Luciferase Assay System (Promega, Madison, WI) as previously described (1). Relative translation efficiencies were calculated as a ratio of IRES-driven Renilla luciferase levels to the internal control firefly luciferase levels.

Compound Testing for *In Vitro* Translation Inhibition. Compounds were dissolved in DMSO and added to the assay solution at the desired compound concentrations of 0, 200, 500, and 1500μM at a final DMSO concentration of 1.5vol%. Compound testing for *in vitro* translation inhibition was conducted using the IVT assay as described above. Relative translation efficiencies were normalized to 0μM levels.

HCV Bicistronic Reporter Mutagenesis. Mutations were introduced into the HCV bicistronic luciferase reporter DNA plasmid via standard molecular cloning techniques and site directed mutagenesis (Q5 Site-Directed Mutagenesis Kit, New England Biolabs, Ipswich, MA). Oligonucleotides for the cloning of HCV ΔII mutant bicistronic reporter plasmids are listed in the SI Appendix, Table S3.

Briefly, chemically synthesized oligonucleotides were phosphorylated and then annealed pairwise (xxx-nT with xxx-nB). Segment 2 was first ligated to segment 3 followed by the ligation of segment 1 to segment 2-3. Segment 4 was generated from a polymerase chain reaction (PCR) using the WT HCV bicistronic reporter plasmid as the template and HCV Segment 4 forward and reverse primers. The full length insert was then generated from the ligation product (segment 1-2-3) and PCR product (segment 4) using PCR by overlap extension. The resulting full length PCR product was purified with a QIAquick PCR Purification Kit and digested with EcoRI and AccI. Likewise, the WT HCV bicistronic reporter plasmid was digested with EcoRI and AccI and ligated with the EcoRI-AccI digested full length PCR product. The product of this ligation was then transformed into NEB 5-alpha competent *E. coli* cells, plated on LB-Amp medium and grown overnight at 37°C. Individual colonies with the correct insertion were confirmed by automated DNA sequencing.

Oligonucleotides for site-directed mutagenesis of HCV ΔII mutant bicistronic reporter plasmids are outlined in the SI Appendix, Table S4. Site-directed mutagenesis experiments were carried out according to the manufacturer's instructions. The sequences of all mutant bicistronic reporter plasmids were verified by DNA sequencing. Mutational studies were conducted using the IVT assay as described above. Relative translation efficiencies were normalized to HCV WT levels.

Crystallization and Data Collection. SVV subdomain IIa RNA (SI Appendix, Figure S5*A*) was annealed from stoichiometric amounts of the single strands by heating to 65°C for 4min followed by slow cooling to room temperature. After cooling, the RNA was crystallized at 16 °C by hanging drop vapor diffusion. For crystallization, 1μL of 0.2mM RNA was mixed with an equal volume of precipitating solution containing 10mM calcium chloride, 200mM ammonium

chloride, 50mM Tris hydrochloride buffer, pH 8.5, and 25% w/v polyethylene glycol 4,000. Cube-shaped crystals appeared and grew to full size over 2-4 days of equilibration against 700μL of well solution containing precipitating solution.

Extended SVV Subdomain IIa RNA (SI Appendix, Fig. S6*A*) was crystallized using the same methods as above after mixing 1uL of 0.2mM RNA with an equal volume of precipitating solution containing 100mM magnesium acetate, 200mM potassium chloride, 50mM sodium cacodylate buffer, pH 6.5, and 10% w/v polyethylene glycol 8,000. Plate-shaped crystals grew over 2 months of equilibration against 700μL of well solution containing precipitating solution.

Crystals were flash-cooled in liquid nitrogen. X-ray diffraction data were collected at 110K on a Rigaku rotating anode X-ray generator (λ = 1.54 Å) equipped with a MAR345 imaging plate detector system. Datasets were processed, integrated, and scaled with the HKL2000 package (3).

Structure Solution and Refinement. The three-dimensional structure of the SVV subdomain IIa RNA was solved by molecular replacement with the program Phaser (4) using A-form RNA duplexes as search models and refined by the program Refmac (5) both within the CCP4 package (6). Subsequent iterative rounds of manual building and refinement, alternating between Refmac and manual rebuilding in Coot (7), were based on the obtained $2F_o$-F_c and F_o-F_c maps. Final refinement was carried out in PHENIX (8) with individual isotropic atomic displacement parameters and water picking (SI Appendix, Tables S1, S2). Coordinates and structure factors for both SVV subdomain IIa structures have been deposited in the RCSB Protein Data Bank under accession codes 4P97 and 4PHY.

HCV Replicon Mutagenesis. SVV IRES IIa mutations were introduced into the SGR-JFH1 FEO (9) DNA plasmid via site directed mutagenesis (Q5 Site-Directed Mutagenesis Kit, New

England Biolabs, Ipswich, MA). The sequence of the ΔSVV-IRES-IIa SGR-JFH1 FEO plasmid was verified by automated DNA sequencing. Sequences of the mutagenic oligonucleotides for SGR-JFH1 FEO were as follows, with lowercase letters indicating mutated residues:

5'- GCCATGGCGTTAGTATGAGTGTCGTACgaggCTCCAGGCCCCCCCCTCC-3' (sense)

5'- TAGGCGCTTTCTGCGTGAAGACggtgggtagCTCACAGGGGAGTGATTCATGGCG-3' (anti-sense)

HCV Replicon Assay. SVV IRES IIa mutations were introduced into the SGR-JFH1 FEO (9) DNA plasmid as outlined above. SGR-JFH1 FEO and ΔSVV-IRES-IIa RNAs were generated from the corresponding DNA plasmid using T7 RNA polymerase as previously described (9, 10). Briefly, plasmids were linearized with XbaI and then digested with mung bean nuclease to generate an authentic 3' end. *In vitro* transcription (T7 RiboMAX Express Large Scale RNA Production System, Promega, Madison, WI) was carried out according to the manufacturer's instructions to yield SGR-JFH1 FEO and ΔSVV-IRES-IIa RNAs. Transfection was performed as previously described (9, 11). Briefly, 400µl of a Huh-7.5.1 cell suspension (10^7 cells/ml) was placed in a 2mm cuvette with 10µg SGR-JFH1 FEO or ΔSVV-IRES-IIa RNA. The mixture was electroporated (Bio-Rad Gene Pulser, Hercules, CA) with an exponential protocol (140V, 950µF). Cells were then seeded into 96-well plates at a density of 20,000 cells/well and resuspended in 100µL complete media. All conditions were run in triplicate. Luciferase activity was determined at 4h and 24h post-transfection using a combined lysis buffer and luciferin reagent (OneGlo, Promega, Madison, WI) according to the manufacturer's instructions. Luciferase activity was determined using a microplate luminometer (Veritas microplate luminometer; Turner Biosystems, Sunnyvale, CA). 4h and 24h luciferase levels were normalized to WT luciferase level at 4h.

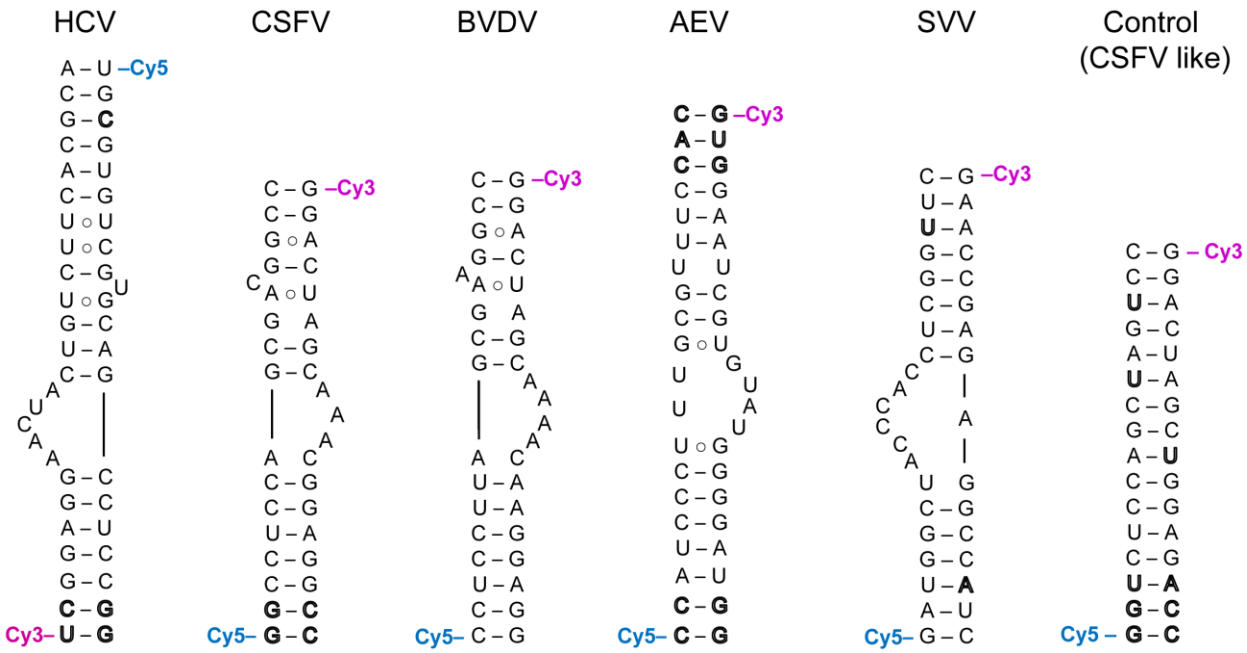

Fig. S1. Structures and sequences of dye-labeled viral IRES subdomain IIa constructs for FRET experiments.

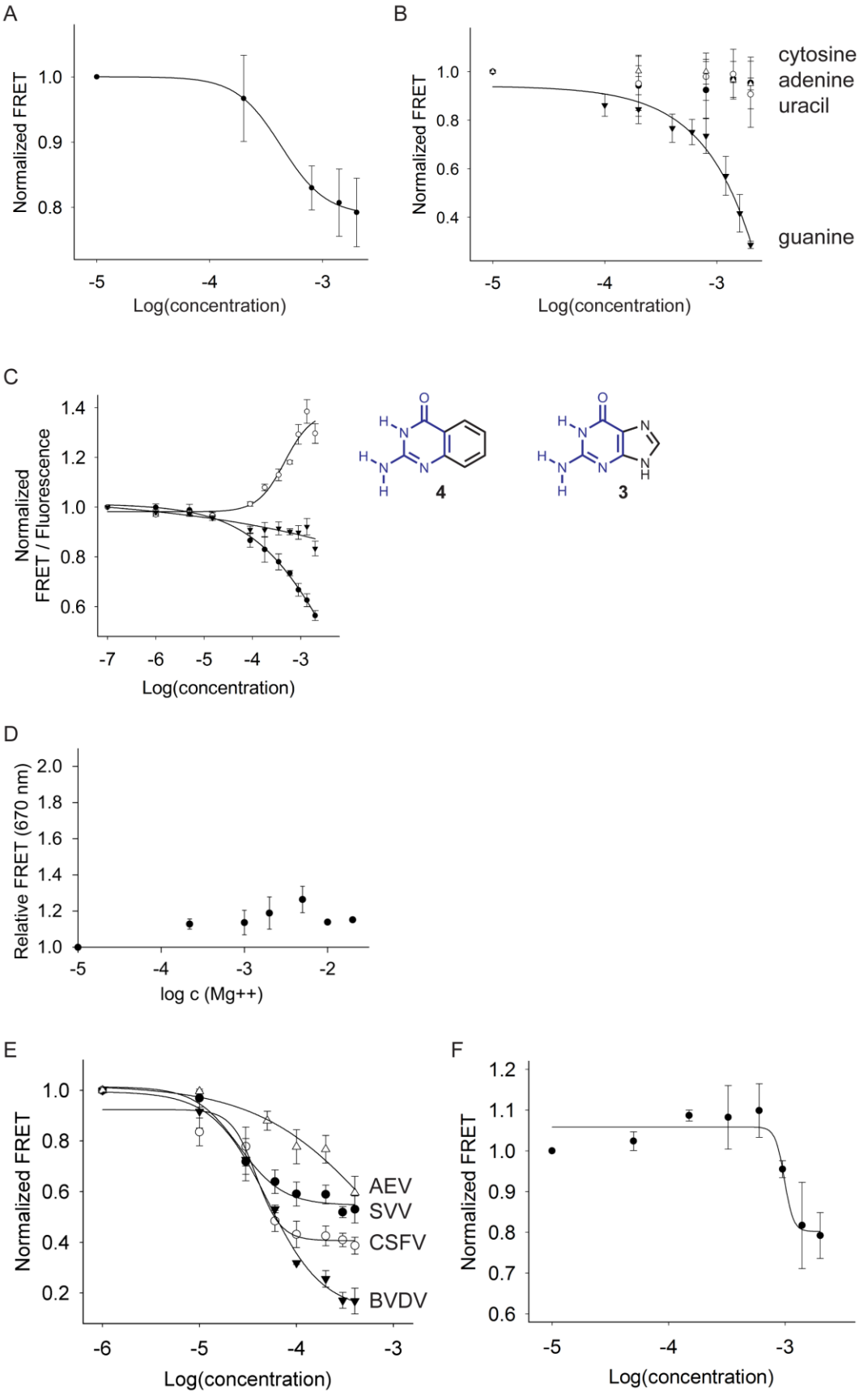
A
Normalized FRET
1.0
0.9
0.8
-5
-4
-3
Log(concentration)
B
Normalized FRET
1.0
0.8
0.6
0.4
cytosine
adenine
uracil
guanine
-5
-4
-3
Log(concentration)
C
Normalized
FRET / Fluorescence
1.4
1.2
1.0
0.8
0.6
-7
-6
-5
-4
-3
Log(concentration)
4
3
D
Relative FRET (670 nm)
2.0
1.8
1.6
1.4
1.2
1.0
-5
-4
-3
-2
log c (Mg++)
E
Normalized FRET
1.0
0.8
0.6
0.4
0.2
AEV
SVV
CSFV
BVDV
-6
-5
-4
-3
Log(concentration)
F
Normalized FRET
1.2
1.1
1.0
0.9
0.8
0.7
0.6
-5
-4
-3
Log(concentration)

Fig. S2. FRET experiments with constructs shown in SI Appendix, Fig. S1. *(A)* Titration of Cy3/Cy5-labelled HCV subdomain IIa with guanosine (●). *(B)* Titrations of Cy3/Cy5-labelled HCV subdomain IIa with guanine (▼), cytosine (●), adenine (○) and uracil (Δ). *(C)* Titration of Cy3/Cy5-labelled HCV subdomain IIa with 2-aminoquinazolin-4(3H)-one **4** which contains the 2-aminopyrimidinone heterocycle (blue) of guanine **3** but has the imidazole ring replaced by a benzene. Curves show normalized FRET signal for **4** (●) as well as the normalized fluorescence signals of the donor Cy3 (○) and acceptor Cy5 (▼). Because the FRET signal did not reach saturation, affinity was estimated by fitting single-site binding to the Cy3 emission which resulted in an EC_{50} value for ligand binding of 483±120μM. *(D)* Titration of terminally Cy3/Cy5 labeled RNA Control construct with Mg^{2+}. *(E)* Titrations of a Cy3/Cy5-labelled viral subdomain IIa RNA constructs with benzimidazole **1**. Fitting of single-site binding curves resulted in EC_{50} values for ligand binding of 39±5μM (CSFV ○), 51±7μM (BVDV ▼) and 26±5μM (SVV ●). Titration of the AEV RNA did not result in a saturating binding curve (Δ). *(F)* Titration of the SVV RNA with guanine. Fitting of a single-site binding curve resulted in an EC_{50} value for ligand binding of 990±170μM. In all panels, error bars represent ± 1s.d. calculated from triplicate experiments.

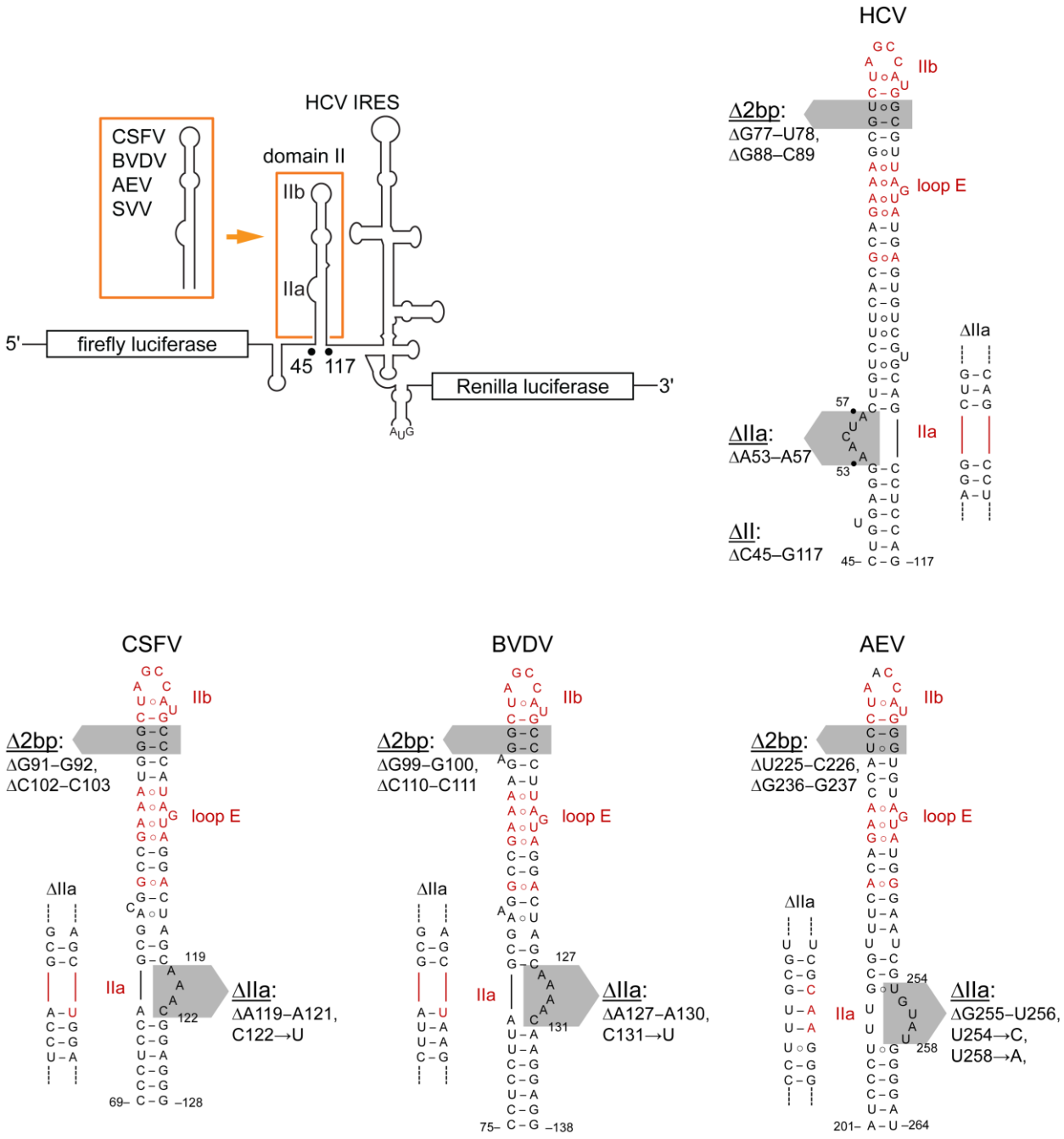

Fig. S3. Structures and sequences of HCV IRES domain II chimera constructs. Bicistronic (dual) luficerase reporter constructs were used for the IVT experiments. Domain II was mutated or replaced by corresponding RNA motifs from other viruses as indicated. NCBI reference sequences: HCV (NC_004102), CSFV (NC_002657.1), BVDV (NC_001461.1), AEV (NC_003990.1), SVV (NC_011349.1).

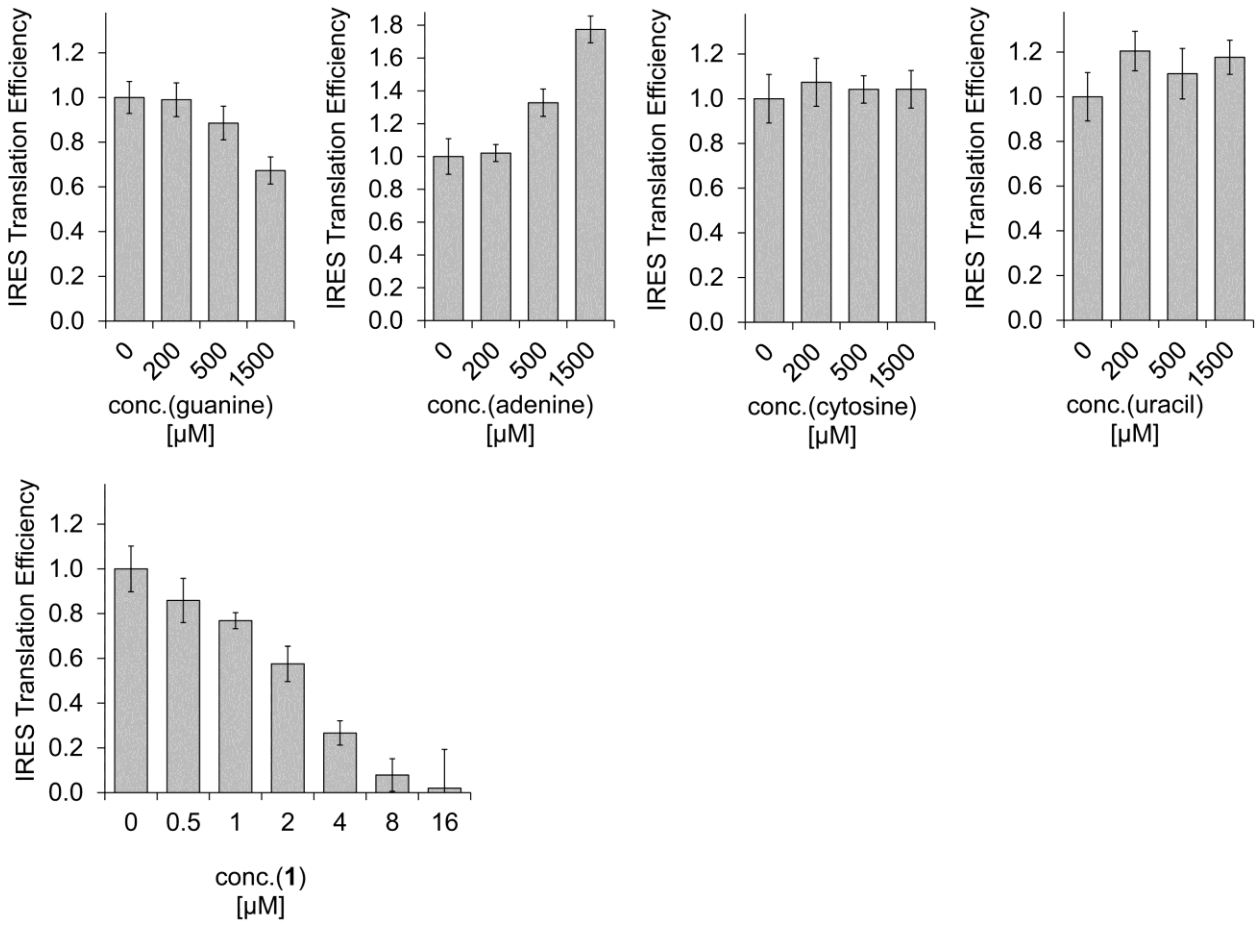

Fig. S4. Impact of guanine, adenine, cytosine, uracil and the benzimidazole **1** on IRES-driven translation as measured in an *in vitro* translation assay. Error bars represent ± 1s.d. calculated from triplicate experiments.

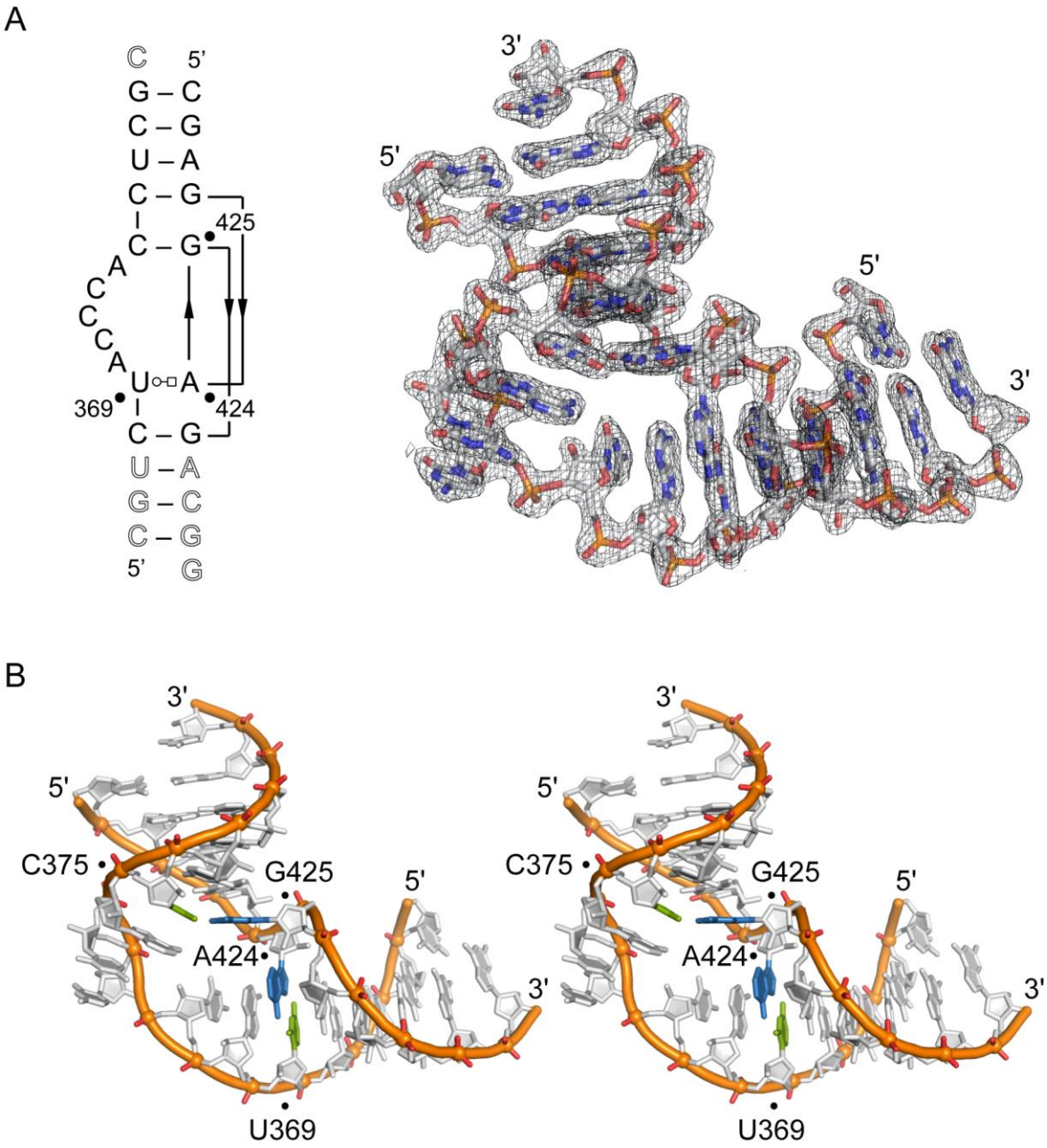

Fig. S5. Crystal structure of the SVV subdomain IIa RNA, determined at 1.86Å resolution (SI Appendix, Table S1; PDB entry 4P97). *(A)* The RNA construct that was used for crystallization is shown on the left, with secondary structure indicated as observed in the crystal. A $2F_o$-F_c electron density map is shown contoured at 1σ. *(B)* Stereo view of the crystal structure.

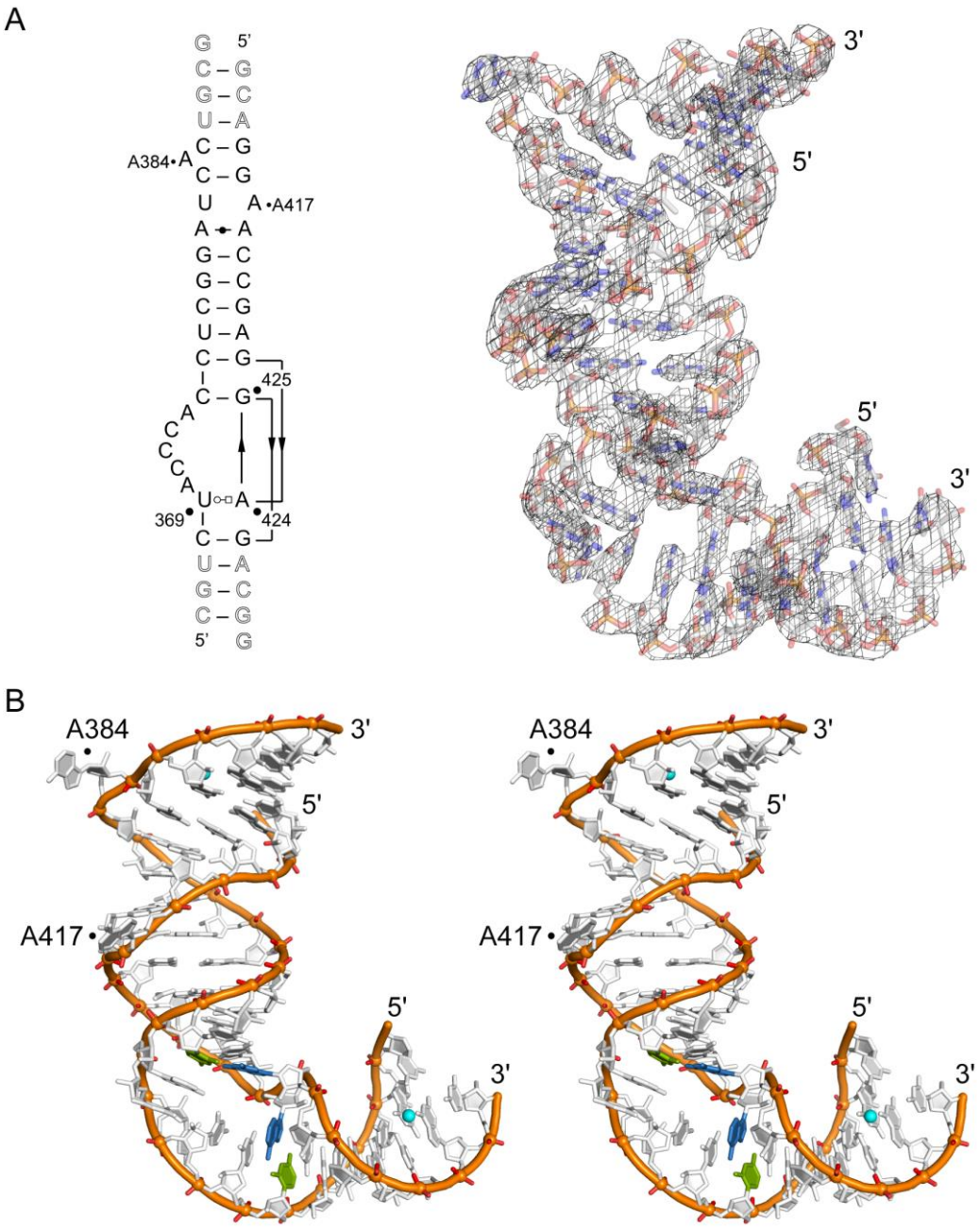

Fig. S6. Crystal structure of an extended SVV subdomain IIa RNA construct, determined at 3.2Å resolution (SI Appendix, Table S2; PDB entry 4PHY). *(A)* The RNA construct that was used for crystallization is shown on the left, with secondary structure indicated as observed in the crystal. A $2F_o$-F_c electron density map is shown contoured at 1.5σ. *(B)* Stereo view of the crystal structure.

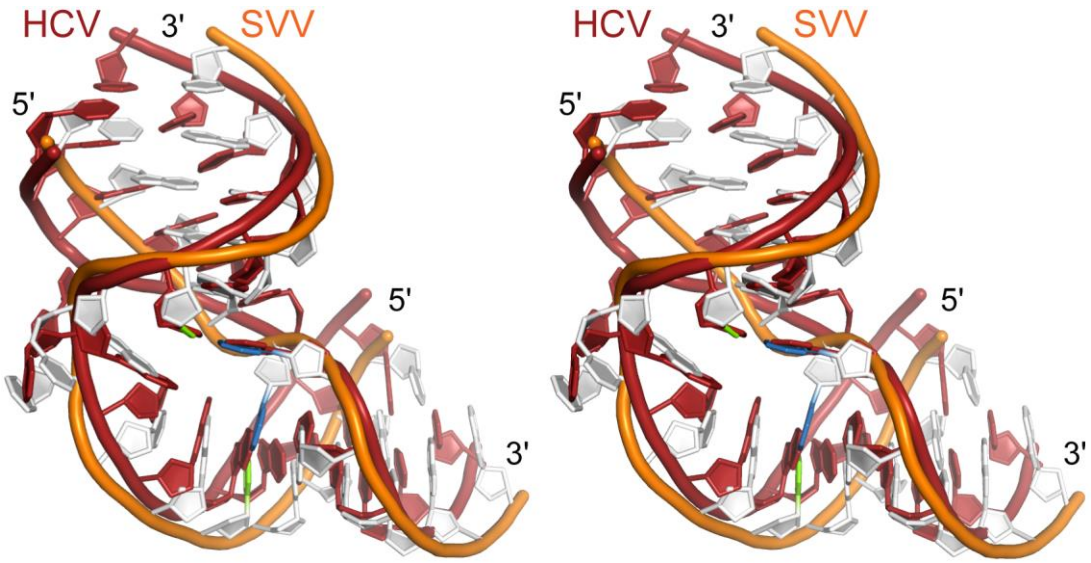

Fig. S7. Stereo view of a superposition of the subdomain IIa crystal structures from the HCV (red) and SVV (orange/white) IRES elements. The cross-stacked U369○A424 reverse Hoogsten base pair and the Watson-Crick C375-G425 base pair in the SVV RNA structure are highlighted in green and blue (see Fig. 4*A*).

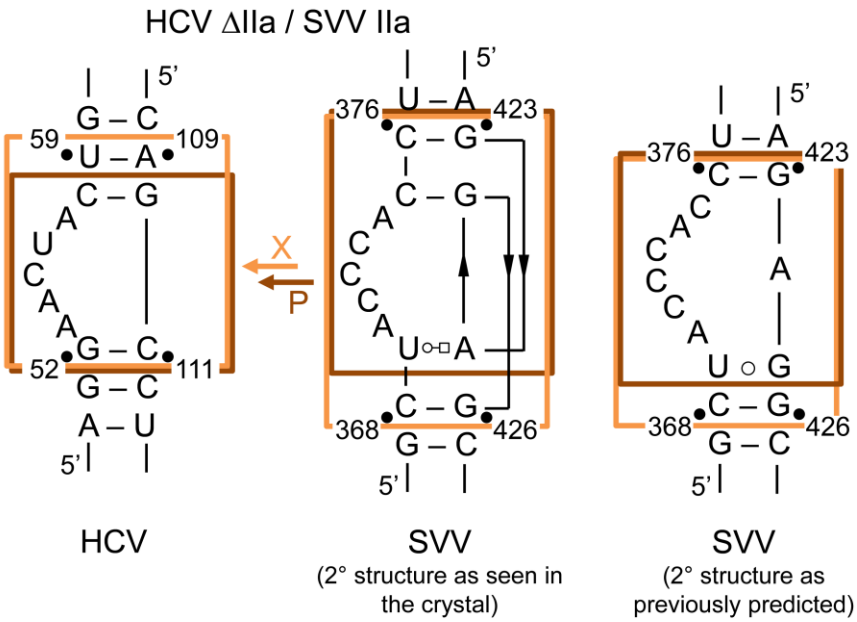

Fig. S8. Construction of HCV IRES chimeras. The chimera construct X (orange) is an exact structural motif swap based on superimposition of the subdomain IIa crystal structures from HCV (see Fig. 1*C*) and SVV (see Fig. 4*A*). The chimera P (brown) is a control construct that contains a subdomain IIa swap with a one base pair offset that is guided by the secondary structure of the RNAs without regard of the 3D structure information from crystallography. Two representations of the subdomain IIa secondary structure in SVV are shown.

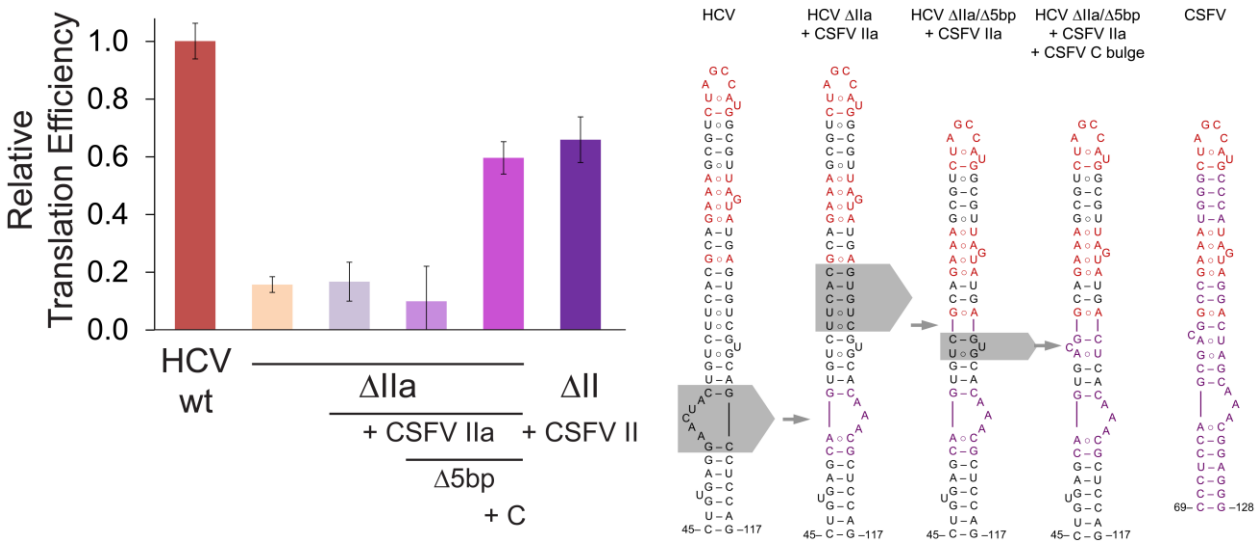

Fig. S9. Effect on translation efficiency of subdomain IIa replacement in the HCV IRES by the corresponding motif from CSFV, and mutants thereof, as measured in an *in vitro* transcription-translation assay. The structure of CSFV chimera constructs is outlined on the right. The HCV wt and CSFV wt chimera are shown as controls. Translation efficiencies were normalized to the cap driven expression in bicistronic dual reporter constructs. Error bars represent ± 1s.d. calculated from triplicate experiments.

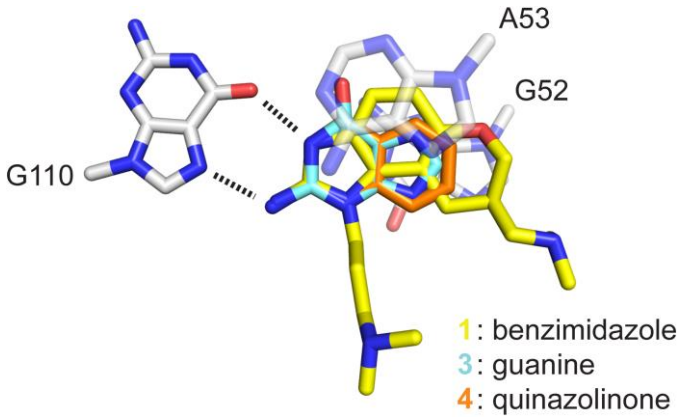

Fig. S10. Modeling of guanine **3** and quinazolinone **4** by superimposition on the benzimidazole **1** in the ligand binding site of the subdomain IIa complex crystal structure. Neighboring stacking bases are shown along with the docking site at G110.

Table S1. Crystallographic data collection and refinement statistics for the SVV IRES subdomain IIa RNA structure (PDB entry 4P97).

Data Collection	
Wavelength (Å)	1.54
High-resolution limit (Å)	1.86
Low-resolution limit (Å)	19.56
Redundancy[a]	8.2 (3.2)
Completeness (%)[a]	99.12 (90.3)
$I/\sigma(I)$[a]	10.78 (3.67)
Total reflections	13932
Unique reflections	1699
Refinement	
Space group	H3
Cell dimensions (Å)	
a	60.53
b	60.53
c	122.68
α	90
β	90
γ	120
R_{work}/R_{free}	0.19 / 0.23
No. atoms	
RNA atoms	1136
Solvent atoms	111
Metal ions	2 Ca^{2+}
Mean *B* factors (Å^2)	
RNA	35.3
Solvent	40.1
Metal	46.1
R.m.s. deviations	
Bond lengths (Å)	0.010
Bond angles (°)	1.163
Dihedral angles (°)	9.519

[a]Numbers in parentheses are for the highest-resolution shell.

Table S2. Crystallographic data collection and refinement statistics for the SVV IRES subdomain IIa extended RNA structure (PDB entry 4PHY).

Data Collection	
Wavelength (Å)	1.54
High-resolution limit (Å)	3.10
Low-resolution limit (Å)	19.90
Redundancy[a]	28.0 (15.5)
Completeness (%)[a]	88.4 (56.4)
$I/\sigma(I)$[a]	22.28 (2.48)
Total reflections	92932
Unique reflections	3319
Refinement	
Space group	P6322
Cell dimensions (Å)	
a	79.61
b	79.61
c	100.94
α	90
β	90
γ	120
R_{work}/R_{free}	0.21 / 0.27
No. atoms	
RNA atoms	1136
Ligand	2 CH_3COO^-
Metal ions	2 Mg^{2+}
Mean *B* factors (Å^2)	
RNA	84.1
Ligand	81.8
Metal	63.42
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.070
Dihedral angles (°)	17.489

[a]Numbers in parentheses are for the highest-resolution shell.

Table S3. Oligonucleotide sequences for the cloning of HCV ΔII mutants in bicistronic reporter plasmids.

Oligonucleotide Name	**Sequence (5' → 3')***
HCV Segment 1T (sense)	CTAGAGaattCCCAGCCCCCGATTGGGGGCGACACTCCACCATAGATC
HCV Segment 1B (antisense)	GGAGTGATCTATGGTGGAGTGTCGCCCCCAATCGGGGGCTGGGaattCTCT AG
HCV Segment 4 Forward Primer (sense)	CCCTCCCGGGAGAGCCATA
HCV Segment 4 Reverse Primer (antisense)	CGAAGGATTCGTGCTCATGG
HCV ΔII + BVDV II-2T (sense)	ACTCC CCTCCTTAGCGAAGGCCGAAAAGAGGCTAGCCATGCCCT
HCV ΔII + BVDV II-2B (antisense)	TACTAAGGGCATGGCTAGCCTCTTTTCGGCCTTCGCTAAGGAGG
HCV ΔII + BVDV II-3T (sense)	TAGTAGGACTAGCAAAACAAGGAGGACCCCCCCTCCCGGGAGAGC
HCV ΔII + BVDV II-3B (antisense)	GCTCTCCCGGGAGGGGGGGTCCTCCTTGTTTTGCTAGTCC
HCV ΔII + CSFV II-2T (sense)	ACTCCCCTCCAGCGACGGCCGAAATGGGCTAGCCATGCCCA
HCV ΔII + CSFV II-2B (antisense)	TACTATGGGCATGGCTAGCCCATTTCGGCCGTCGCTGGAGG
HCV ΔII + CSFV II-3T (sense)	TAGTAGGACTAGCAAACGGAGGACCCCCCCTCCCGGGAGAGC
HCV ΔII + CSFV II-3B (antisense)	GCTCTCCCGGGAGGGGGGGTCCTCCGTTTGCTAGTCC
HCV ΔII + SVV II-2T (sense)	ACTCCGATGGCTACCCACCTCGGATCACTGAACTGGAGCTCGACCCT
HCV ΔII + SVV II-2B (antisense)	TAAGGAGGGTCGAGCTCCAGTTCAGTGATCCGAGGTGGGTAGCCATC
HCV ΔII + SVV II-3T (sense)	CCTTAGTAAGGGAACCGAGAGGCCTTCACCCCCCCTCCCGGGAGAGC
HCV ΔII + SVV II-3B (antisense)	GCTCTCCCGGGAGGGGGGGTGAAGGCCTCTCGGTTCCCTTAC
HCV ΔII + BVDV II/ΔIIa-2T (sense)	ACTCCCCTCCTTAGCGAAGGCCGAAAAGAGGCTAGCCATGCCCT
HCV ΔII + BVDV II/ΔIIa-2B (antisense)	TACTAAGGGCATGGCTAGCCTCTTTTCGGCCTTCGCTAAGGAGG
HCV ΔII + BVDV II/ΔIIa-3T (sense)	TAGTAGGACTAGCTAAGGAGGACCCCCCCTCCCGGGAGAGCCATA
HCV ΔII + BVDV II/ΔIIa-3B (antisense)	TATGGCTCTCCCGGGAGGGGGGGTCCTCCTTAGCTAGTCC
HCV ΔII + CSFV II/ΔIIa-2T (sense)	ACTCCCCTCCAGCGACGGCCGAAATGGGCTAGCCATGCCCA
HCV ΔII + CSFV II/ΔIIa-2B (antisense)	TACTATGGGCATGGCTAGCCCATTTCGGCCGTCGCTGGAGG
HCV ΔII + CSFV II/ΔIIa-3T (sense)	TAGTAGGACTAGCTGGAGGACCCCCCCTCCCGGGAGAGCCATA
HCV ΔII + CSFV II/ΔIIa-3B (antisense)	TATGGCTCTCCCGGGAGGGGGGGTCCTCCAGCTAGTCC
HCV Δ2bp-2T (sense)	ACTCCCCTGTGAGGAACTACTGTCTTCACGCAGAAAGCCTAGCCATGGT
HCV Δ2bp-2B (antisense)	TACTAACCATGGCTAGGCTTTCTGCGTGAAGACAGTAGTTCCTCACAGG

Table S3.	*continued*
HCV Δ2bp-3T (sense)	TAGTATGAGTGTCGTGCAGCCTCCAGGACCCCCCCTCCCGGGAGAGCCATA
HCV Δ2bp-3B (antisense)	TATGGCTCTCCCGGGAGGGGGGGTCCTGGAGGCTGCACGACACTCA
HCV ΔII + BVDV II/Δ2bp-2T (sense)	ACTCCCCTCCTTAGCGAAGGCCGAAAAGACTAGCCATGCT
HCV ΔII + BVDV II/Δ2bp-2B (antisense)	TACTAAGCATGGCTAGTCTTTTCGGCCTTCGCTAAGGAGG
HCV ΔII + BVDV II/Δ2bp-3T (sense)	TAGTAGGACTAGCAAAACAAGGAGGACCCCCCCTCCCGGGAGAGC
HCV ΔII + BVDV II/Δ2bp-3B (antisense)	GCTCTCCCGGGAGGGGGGGTCCTCCTTGTTTTGCTAGTCC
HCV ΔII + CSFV II/Δ2bp-2T (sense)	ACTCCCCTCCAGCGACGGCCGAAATGCTAGCCATGCA
HCV ΔII + CSFV II/Δ2bp-2B (antisense)	TACTATGCATGGCTAGCATTTCGGCCGTCGCTGGAGG
HCV ΔII + CSFV II/Δ2bp-3T (sense)	TAGTAGGACTAGCAAACGGAGGACCCCCCCTCCCGGGAGAGC
HCV ΔII + CSFV II/Δ2bp-3B (antisense)	GCTCTCCCGGGAGGGGGGGTCCTCCGTTTGCTAGTCC

*Lowercase letters indicate EcoRI restriction site.

Table S4. Oligonucleotide sequences for site-directed mutagenesis of bicistronic reporter constructs.

Parent Construct	Oligonucleotide Name	Sequence (5' → 3')*
HCV WT	HCV ΔII sense (sense)	GACCCCCCCTCCCGGGAG
HCV WT	HCV ΔII antisense (antisense)	GGGAGTGATCTATGGTGGAGTGTCG
HCV WT	HCV ΔII + AEV II (sense)	atgggtgtagtatgggaatcgtgtatggggatACCCC CCCTCCCGGGAGA
HCV WT	HCV ΔII + AEV II (antisense)	ggttaggatggttctgtgaaacgcaaagggatGGAGT GATCTATGGTGGAGTGTCGC
HCV WT	HCV ΔIIa (sense)	CTGTCTTCACGCAGAAAG
HCV WT	HCV ΔIIa (antisense)	CCTCACAGGGGAGTGATC
HCV ΔII + AEV II	HCV ΔII + AEV II/ΔIIa (sense)	ATGGGAATCGcaaGGGGATACCCCC
HCV ΔII + AEV II	HCV ΔII + AEV II/ΔIIa (antisense)	ACTACACCCATGGTTAGG
HCV ΔII + AEV II	HCV ΔII + AEV II/Δ2bp (sense)	catgTGTAGTATGGGAATCGTG
HCV ΔII + AEV II	HCV ΔII + AEV II/Δ2bp (antisense)	gttagTGGTTCTGTGAAACGCAAAG
HCV WT	HCV ΔIIa + SVV IIa Predicted (sense)	atggcgttagtatgagtgtcgtgcAGAGGCCTTCACC CCCCC
HCV WT	HCV ΔIIa + SVV IIa Predicted (antisense)	ggctagacgctttctgcgtgaagacAGGTGGGTAGCC ATCGGAG
HCV WT	HCV ΔIIa + SVV IIa X-Ray (sense)	gccatggcgttagtatgagtgtcgtgcgaggCTCCAG GACCCCCCCTCC
HCV WT	HCV ΔIIa + SVV IIa X-Ray (antisense)	tagacgctttctgcgtgaagacggtgggtagCTCACA GGGGAGTGATCTATGGTG
HCV WT	HCV ΔIIa + CSFV IIa (sense)	tggcgttagtatgagtgtcgtgcacaaacgCTCCAGG ACCCCCCCTCC
HCV WT	HCV ΔIIa + CSFV IIa (antisense)	tggctagacgctttctgcgtgaagacactgCTCACAG GGGAGTGATCTATGGTG
HCV WT	HCV ΔIIa + CSFV IIa/Δ5bp (sense)	tggcgttagtatgagtgcacaaacgCTCCAGGACCCC CCCTCC
HCV WT	HCV ΔIIa + CSFV IIa/Δ5bp (antisense)	tggctagacgctttctgcgacactgCTCACAGGGGAG TGATCTATGGTG

Table S4.	*continued*	
HCV ΔIIa + CSFV IIa	HCV ΔIIa + CSFV IIa/Δ5bp/+C (sense)	catggcgttagtatgactagCAAACGCTCCAGGACCCC
HCV ΔIIa + CSFV IIa	HCV ΔIIa + CSFV IIa/Δ5bp/+C (antisense)	gctagacgctttctgccgtcgCTGCTCACAGGGGAGTGATC

*Lowercase letters indicate a substitution from the parent construct. The absence of lowercase letters indicates a simple deletion.

Supporting Information References

1. Zhou S, Rynearson KD, Ding K, Brunn ND, & Hermann T (2013) Screening for inhibitors of the hepatitis C virus internal ribosome entry site RNA. *Bioorg. Med. Chem.* 21:6139-6144.
2. Brunn ND, Garcia Sega E, Kao MB, & Hermann T (2012) Targeting a regulatory element in human thymidylate synthase mRNA. *ChemBiochem* 13(18):2738-2744.
3. Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276:307-326.
4. McCoy AJ, *et al.* (2007) Phaser crystallographic software. *J. Appl. Cryst.* 40:658-674.
5. Murshudov GN, Vagin AA, & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* 53(Pt 3):240-255.
6. Collaborative Computational Project N (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 50(Pt 5):760-763.
7. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60(Pt 12 Pt 1):2126-2132.
8. Adams PD, *et al.* (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* 58(Pt 11):1948-1954.
9. Wyles DL, Kaihara KA, Vaida F, & Schooley RT (2007) Synergy of small molecular inhibitors of hepatitis C virus replication directed at multiple viral targets. *J. Virol.* 81(6):3005-3008.
10. Wyles DL, *et al.* (2009) The octadecyloxyethyl ester of (S)-9-[3-hydroxy-2-(phosphonomethoxy) propyl]adenine is a potent and selective inhibitor of hepatitis C virus replication in genotype 1A, 1B, and 2A replicons. *Antimicrob. Agents Chemother.* 53(6):2660-2662.
11. Parsons J, *et al.* (2009) Conformational inhibition of the hepatitis C virus internal ribosome entry site RNA. *Nat. Chem. Biol.* 5(11):823-825.