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 MgCl₂. 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 H₂O 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 ∆II 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 ∆II 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^oC 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 ($\lambda = 1.54$ 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 $2F_0-F_c$ and F_0-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 ∆SVV-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' (antisense)

HCV Replicon Assay. SVV IRES IIa mutations were introduced into the SGR-JFH1 FEO (9) DNA plasmid as outlined above. SGR-JFH1 FEO and ∆SVV-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 ∆SVV-IRES-IIa RNAs. Transfection was performed as previously described (9, 11). Briefly, 400 μ l of a Huh-7.5.1 cell suspension (10⁷ cells/ml) was placed in a 2mm cuvette with 10μg SGR-JFH1 FEO or ∆SVV-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.

HCV	CSFV	BVDV	AEV	SVV	Control (CSFV like)
$A-U-Cy5$ $C - G$ $G - C$ $C - G$ $A-U$ $C - G$ $U \circ U$ $U \circ C$ $C - G$ $U \circ G^U$ $G - C$ $U - A$ $C_A^{\bigcup_{\substack{A \subset -G}}$ Α $G - C$ $G - C$ $A-U$ $G - C$ $G - C$ $C-G$ $Cy3-U-G$	$C - G - Cy3$ $C - G$ $G \circ A$ $G - C$ $c_{A \circ U}$ G A $C - G$ $G - C$ A Α c^{A} Α $C - G$ $C - G$ $U - A$ $C - G$ $C - G$ $G - C$ $Cy5-G-C$	$C - G - Cy3$ $C - G$ $G \circ A$ $G - C$ $A_{A} \circ U$ G A $C - G$ $G - C$ Α A c^{A} Α $U - A$ $U - A$ $C - G$ $C - G$ $U - A$ $C - G$ $Cy5-C-G$	$C - G - Cy3$ $A-U$ $C-G$ $C - G$ $U - A$ $U - A$ U U $G - C$ $C - G$ $G \circ U$ G U U U^A U $\mathsf{U}\circ\mathsf{G}$ $C - G$ $C - G$ $C - G$ $U - A$ $A-U$ $C-G$ $Cy5-C-G$	$C - G - Cy3$ $U - A$ $U - A$ $G - C$ $G - C$ $C - G$ $U - A$ $c_{c_A}^{c_A^{C-G}}$ G U $C - G$ $G - C$ $G - C$ $U - A$ $A-U$ $Cy5-G-C$	$C - G - Cy3$ $C - G$ $U - A$ $G - C$ $A-U$ $U - A$ $C - G$ $G - C$ $A - U$ $C - G$ $C - G$ $U - A$ $C - G$ $U - A$ $G - C$ $Cy5-G-C$

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

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 (∆). *(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 (\circ) 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 $483 \pm 120 \mu M$. *(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 $39 \pm 5 \mu M$ (CSFV \circ), $51 \pm 7 \mu M$ (BVDV \blacktriangledown) and $26 \pm 5 \mu M$ (SVV \bullet). 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 $990\pm170\mu$ M. In all panels, error bars represent \pm 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 \pm 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 $2F_0-F_c$ 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 2Fo-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* transcriptiontranslation 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 \pm 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

Data Collection			
Wavelength (A)	1.54		
High-resolution limit (A)	1.86		
Low-resolution limit (A)	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	H ₃		
Cell dimensions (A)			
\boldsymbol{a}	60.53		
\boldsymbol{b}	60.53		
\mathcal{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	$2 Ca2+$		
Mean <i>B</i> factors (\AA^2)			
RNA	35.3		
Solvent	40.1		
Metal	46.1		
R.m.s. deviations			
Bond lengths (A)	0.010		
Bond angles (°)	1.163		
Dihedral angles (°)	9.519		

subdomain IIa RNA structure (PDB entry 4P97).

^aNumbers in parentheses are for the highest-resolution shell.

Data Collection			
Wavelength (\AA)	1.54		
High-resolution limit (A)	3.10		
Low-resolution limit (A)	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 (A)			
a	79.61		
\boldsymbol{b}	79.61		
\overline{c}	100.94		
α	90		
β	90		
γ	120		
R_{work}/R_{free}	0.21 / 0.27		
No. atoms			
RNA atoms	1136		
Ligand	$2 \text{CH}_3\text{COO}^-$		
Metal ions	2 Mg^{2+}		
Mean <i>B</i> factors (\AA^2)			
RNA	84.1		
Ligand	81.8		
Metal	63.42		
R.m.s. deviations			
Bond lengths (A)	0.007		
Bond angles $(°)$	1.070		
Dihedral angles $(°)$	17.489		

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

^aNumbers in parentheses are for the highest-resolution shell.

*Lowercase letters indicate EcoRI restriction site.

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

*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.