

Structure of hepatitis C virus IRES subdomain IIa

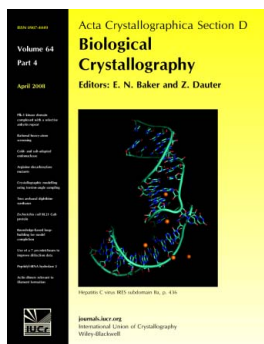
Qiang Zhao, Qing Han, Charles R. Kissinger, Thomas Hermann and Peggy A. Thompson

Acta Cryst. (2008). **D64**, 436–443

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site or institutional repository provided that this cover page is retained. Republication of this article or its storage in electronic databases other than as specified above is not permitted without prior permission in writing from the IUCr.

For further information see <http://journals.iucr.org/services/authorrights.html>



Acta Crystallographica Section D: Biological Crystallography welcomes the submission of papers covering any aspect of structural biology, with a particular emphasis on the structures of biological macromolecules and the methods used to determine them. Reports on new protein structures are particularly encouraged, as are structure–function papers that could include crystallographic binding studies, or structural analysis of mutants or other modified forms of a known protein structure. The key criterion is that such papers should present new insights into biology, chemistry or structure. Papers on crystallographic methods should be oriented towards biological crystallography, and may include new approaches to any aspect of structure determination or analysis.

Crystallography Journals Online is available from journals.iucr.org

Structure of hepatitis C virus IRES subdomain IIa

Qiang Zhao,^{a*} Qing Han,^a
Charles R. Kissinger,^a Thomas
Hermann^{a‡} and Peggy A.
Thompson^{b*}

^aDepartment of Structural Chemistry, Anadys Pharmaceuticals Inc., 3115 Merryfield Row, San Diego, CA 92121, USA, and ^bDepartment of Biology, Anadys Pharmaceuticals Inc., 3115 Merryfield Row, San Diego, CA 92121, USA

‡ Current address: Departments of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA.

Correspondence e-mail:
qiangzhao@hotmail.com,
pthompson@anadyspharma.com

The hepatitis C (HCV) internal ribosome entry site (IRES) element plays a central role in cap-independent translation of the viral genomic RNA. The unique conformation of IRES domain II is critical for 80S ribosomal assembly and initiation of viral translation. Here, the crystal structure of subdomain IIa of the HCV IRES has been determined at 2.3 Å resolution, revealing the positions of divalent metal ions and complex inter-strand interactions that stabilize the L-shaped conformation of the RNA. The presence of divalent metal ions was necessary for crystal formation. Magnesium ions occupy specific sites that appear to be critical for the formation of the folded conformation. Subdomain IIa also was crystallized in the presence of strontium, which improved the diffraction quality of the crystals and the ability to identify interactions of the RNA with metal ions and tightly bound water molecules. The hinge region and noncanonical G–U base-pair motifs are stabilized by divalent metal ions and provide unique structural features that are potential interaction sites for small-molecule ligands. The information obtained from the crystal structure provides a basis for structure-guided design of HCV translation inhibitors targeting disruption of ribosomal assembly.

Received 12 November 2007

Accepted 18 January 2008

PDB References:

Mg-subdomain IIa, 2pn3, r2pn3sf; Sr-subdomain IIa, 2pn4, r2pn4sf.

1. Introduction

Hepatitis C virus (HCV) is an enveloped single-stranded positive-sense RNA virus. The HCV genome contains an internal ribosome entry site (IRES), located within the 5′-untranslated region (5′-UTR), that is responsible for directing cap-independent translation of the viral mRNA (Tsukiyama-Kohara *et al.*, 1992; Wang *et al.*, 1993; Reynolds *et al.*, 1996). In contrast to the HCV mRNA open reading frame, the ~340 nucleotide IRES RNA sequence is highly conserved among all genotypes (Bukh *et al.*, 1992; Davidson *et al.*, 1995). Assembly of the 48S complex and subsequently the 80S ribosome is initiated when the host-cell 40S ribosomal subunit and eukaryotic initiation factor 3 (eIF3) bind to the HCV IRES RNA (Pestova *et al.*, 1998, 2001; Kieft *et al.*, 1999; Ji *et al.*, 2004; Siridechadilok *et al.*, 2005; Yu *et al.*, 2005). Formation of this complex is dependent on the IRES adopting a complex cation-dependent fold (Kieft *et al.*, 1999, 2001; Klinck *et al.*, 2000). The high sequence conservation among viral isolates and the unique structural features of the IRES, along with its critical role in translation, make it an attractive target for RNA-directed HCV therapeutics.

The IRES has been divided into four structural domains designated I, II, III and IV (Fig. 1*a*), which have been shown to

exist as independently folded regions (Brown *et al.*, 1992; Wang *et al.*, 1995; Honda *et al.*, 1996). The helical and stem-loop regions within domains II and III, along with stem loop IV, which contains the start AUG codon, all contribute to IRES-dependent translation (Brown *et al.*, 1992; Kieft *et al.*, 2001; Kikuchi *et al.*, 2003; Ji *et al.*, 2004; Otto & Puglisi, 2004; Spahn *et al.*, 2004; Boehringer *et al.*, 2005; Laletina *et al.*, 2006). The precise roles of individual IRES domains in the formation of a functional 80S complex continue to be elucidated, furthering the understanding of IRES-mediated cap-independent translation (Ji *et al.*, 2004; Lu *et al.*, 2004; Otto & Puglisi, 2004; Locker *et al.*, 2007).

The correctly folded tertiary structure of the HCV IRES has been shown to be essential for binding both the 40S ribosomal subunit and eIF3. Cryo-electron microscopy (cryo-EM) of the IRES in complex with the 40S ribosomal subunit at 20 Å resolution (Spahn *et al.*, 2004) and the 80S ribosome at 15 Å resolution (Boehringer *et al.*, 2005) confirmed that the IRES induces a significant conformational change in the 40S subunit upon binding, altering the mRNA-binding region. These studies revealed that the apical loop of domain II reaches deep into the mRNA cleft near the tRNA-exit site of the 40S ribosome complex (Spahn *et al.*, 2001). Removal of domain II does not alter the binding affinity of either the 40S ribosomal subunit or eIF3 to the IRES; however, it has been shown that this domain is responsible for inducing a large conformational rearrangement at the decoding center of the 40S subunit. This conformational change may be necessary to stabilize the viral AUG start codon in position and for assembly of the ribosomal subunits (Brown *et al.*, 1992; Kolupaeva *et al.*, 2000; Spahn *et al.*, 2001; Lafuente *et al.*, 2002;

Ji *et al.*, 2004; Otto & Puglisi, 2004; Boehringer *et al.*, 2005; Locker *et al.*, 2007). The NMR solution structure of domain II revealed that a bulge within subdomain IIa induces a bend in the helix, resulting in an overall L-shaped conformation. The structure of the isolated domain II is consistent with its conformation when bound to the 40S ribosomal subunit as determined by cryo-EM (Lukavsky *et al.*, 2003; Spahn *et al.*, 2004). Conserved structural motifs within domain II, including the unique bent conformation of subdomain IIa, play a novel functional role in 80S ribosomal assembly by promoting eIF5-induced GTP hydrolysis and mediating eIF2 release (Locker *et al.*, 2007). The connection between the uniquely folded structure of subdomain IIa and its function in 80S ribosomal formation suggests that HCV IRES IIa is a possible target for small-molecule antiviral drugs that bind and disrupt ribosomal complex formation and translation of the viral mRNA.

To gain further understanding of the interactions that stabilize the unique conformation of IRES domain II, we have determined the X-ray crystal structure of subdomain IIa in the presence of either Sr^{2+} or Mg^{2+} ions at resolutions of 2.3 and 2.9 Å, respectively. A 3.0 Å resolution crystal structure of subdomain IIa obtained in the presence of a mixture of Mg^{2+} and Mn^{2+} ions has also been reported (Dibrov *et al.*, 2007). We observe differences in the overall bend angle and the positions of key metal ion-binding sites. These results provide additional insight into the forces that stabilize the unique L-shaped conformation of the molecule and the possibility of molecular intervention at this step in ribosomal complex formation.

2. Materials and methods

2.1. Constructs and crystallization

Gel-purified and desalted synthetic IRES RNA oligonucleotides for subdomain IIa (5'-CGG AGG AAC UAC UGU CUU CAC GCC-3' and 5'-GCG UGU CGU GCA GCC UCC GG-3') were purchased from Dharmacon Research (Lafayette, Colorado, USA). The brominated oligonucleotide 5'-GCG (5-Br-U)GU CGU GCA GCC (5-Br-U)CC GG-3' was used for multi-wavelength anomalous dispersion (MAD) data collection.

RNA was annealed by heating in buffer (5 mM sodium cacodylate pH 6.2, 2 mM MgCl_2) at 353 K for 3 min and cooling slowly to room temperature. Crystals were obtained by the hanging-drop vapor-diffusion method, in which 1 µl 0.5 mM RNA in annealing buffer was mixed with an equal volume of a precipitant buffer containing 50 mM sodium cacodylate pH 6.2, 10 mM MgSO_4 , 25 mM NaCl, 100 mM KCl, 10% 2-methyl-2,4-pentanediol (MPD) and equilibrated over a well solution consisting of 65% MPD at 290 K or room temperature. Crystals were obtained at 290 K after 2–3 weeks. The crystals belonged to space group $P4_32_12$, with unit-cell parameters $a = b = 48.95$, $c = 120.94$ Å and one molecule per asymmetric unit.

For crystals grown in the presence of strontium, the RNA was annealed as above but without MgCl_2 . The RNA was then mixed in a 1:1 ratio with 50 mM sodium cacodylate pH 6.2,

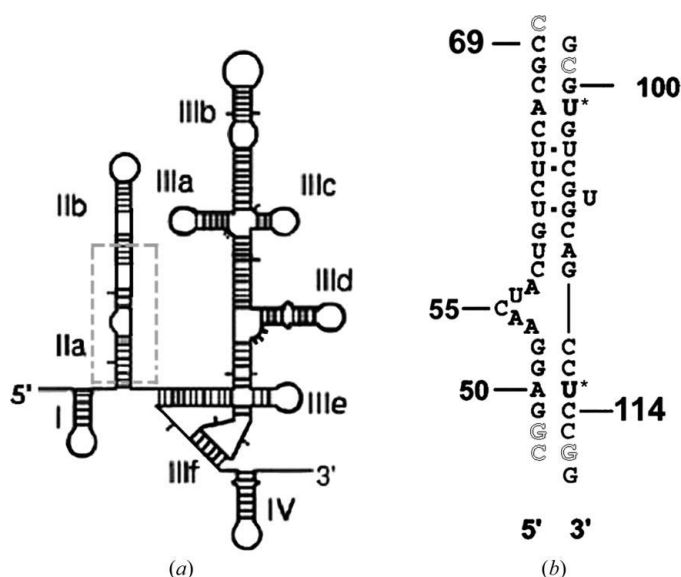


Figure 1
(a) Secondary structure of the ~340-nucleotide HCV 5'-UTR containing the four structural domains of the IRES element. Subdomain IIa is outlined with a dashed box. (b) Sequence of the subdomain IIa construct used for crystallization and structure determination. Residues identical to the HCV IRES domain II are shown in bold. Non-Watson-Crick base-pair interactions are indicated by dots. 5-Br-U-labeled uridines are marked with asterisks.

Table 1

Crystallographic and refinement data for IRES subdomain IIa.

(a) Crystallographic data. Values in parentheses are for the outer shells (3.0–2.9 and 2.43–2.35 Å for Mg-IIa and Sr-IIa, respectively).

	Mg-IIa		High remote	Sr-IIa
	Inflection point	Peak		
Wavelength (Å)	0.9203	0.9197	0.90686	0.9795
Resolution (Å)	45.4–2.9	45.4–2.9	45.4–2.9	34.3–2.32
Redundancy	11.2 (8.1)	11.0 (7.5)	11.3 (8.9)	5.5 (6.9)
$\langle I/\sigma(I) \rangle$	27.5 (4.5)	26.0 (3.8)	32.0 (5.3)	17.5 (2.4)
$R_{\text{merge}}^{\dagger}$ (%)	8.1 (34.5)	8.6 (35.4)	7.4 (31.5)	10.8 (71.4)
Completeness (%)	93.7 (75.1)	95.0 (79.7)	95.7 (81.9)	95.9 (100)
Observed reflections	162063	162376	202708	858221
Unique reflections	3420	3471	3504	12228
Mean figure of merit ‡ (acentric reflections)	0.73			

(b) Refinement statistics for the Mg-IIa and Sr-IIa structures.

	Mg-IIa	Sr-IIa
Space group	$P4_32_12$	$P2_12_12_1$
Unit-cell parameters (Å)		
<i>a</i>	48.95	48.29
<i>b</i>	48.95	48.73
<i>c</i>	120.94	120.00
No. of duplexes in ASU	1 duplex (44 nt)	2 duplexes (88 nt)
<i>R</i> factor §	0.22969	0.26141
R_{free} (10% of data)	0.28263	0.32037
RNA atoms	821	1869
Solvent atoms	—	67
Metal ions	2 (Mg $^{2+}$)	10 (Sr $^{2+}$)
R.m.s.d. from ideal ¶		
Bond lengths (Å)	0.012	0.010
Bond angles (°)	1.885	1.835
Dihedral angles (°)	16.9	16.7
Improper angles (°)	9.93	9.94
Average <i>B</i> factors (Å 2)		
RNA	44.43	51.721
Solvent	—	42.483
Cations	51.42	57.253

$^{\dagger} R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ for all data, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl) \rangle$ is the statistically weighted average intensity of multiple measurements of symmetry-related reflections. ‡ Figure of merit = $[\sum_{hkl} P(\alpha) \exp(i\alpha) / \sum_{hkl} P(\alpha)]$, where α is the phase and $P(\alpha)$ is the phase probability distribution. § *R* factor and $R_{\text{free}} = \sum_{hkl} ||F_{\text{obs}}| - k|F_{\text{calc}}|| / \sum_{hkl} |F_{\text{obs}}|$, where R_{free} includes amplitudes (10%) omitted from the refinement. ¶ Root-mean-square deviation.

80 mM SrCl $_2$, 25 mM NaCl, 30% MPD and equilibrated over a well solution consisting of 60% MPD at 290 K. Sr-subdomain IIa crystals had unit-cell parameters that were nearly identical to those of the Mg-subdomain IIa crystals, but belonged to the lower symmetry space group $P2_12_12_1$.

2.2. Data collection and structure determination

The structure of IRES subdomain IIa with Mg $^{2+}$ bound (Mg-IIa) was determined by MAD phasing using the anomalous scattering from bromines incorporated into the RNA by replacing U101 and U113 with 5-Br-U. These two nucleotides were chosen because they form stable canonical A–U base pairs and are distant from the hinge region, minimizing the potential for altering the folded conformation.

Prior to data collection, crystals were flash-cooled in liquid nitrogen and transferred to a cold nitrogen stream. Attempts

at molecular replacement using the structure determined by NMR (Lukavsky *et al.*, 2003; PDB code 1p5m) were not successful. Therefore, MAD data sets (Table 1) were collected from a single crystal at three wavelengths using bromine as the anomalous scattering atom at 100 K on beamline 14-ID-B at the Advanced Photon Source (APS), Argonne National Laboratory. The X-ray data for the Sr-IIa crystal were collected at 100 K on beamline 8-BM also at the APS. Data were processed with *DENZO/SCALEPACK* (Otwinowski & Minor, 1997). Phases were calculated using the programs *SOLVE/RESOLVE* (Terwilliger & Berendzen, 1999; Terwilliger, 2000). The resulting MAD electron-density map at 3.2 Å resolution was used to build the initial model with the program *O* (Jones *et al.*, 1991), which gave an initial R_{free} value of 0.4798 and an *R* factor of 0.4621 after rigid-body refinement. The structure was refined by several rounds of simulated annealing with restrained individual *B*-factor refinement in the program *CNX* (Accelrys, San Diego, California, USA; Brünger *et al.*, 1998) and manual examination and rebuilding using the inflection-point data.

The Sr-IIa crystal structure was solved by the molecular-replacement method using the refined Mg-IIa model as the search model. The Sr-IIa and Mg-IIa structures were refined at 2.3 and 2.9 Å resolution to final free *R* factors of 32.0% and 28.3%, respectively (see Table 1).

3. Results and discussion

3.1. Construct design and crystallization approach

Both our own enzymatic digestion studies (data not shown) and NMR solution structures (Lukavsky *et al.*, 2003) indicate that oligonucleotide constructs comprising only the IIa subdomain adopt the same conformation as found in the full-length IRES RNA. Subdomain IIa oligonucleotide constructs of varying lengths were screened in crystallization trials (see supplementary material 1). Terminal G–C pairs were introduced to stabilize the duplex, including single-base overhangs to facilitate the formation of pseudo-continuous helices in the crystal. The flexible unpaired U48 nucleotide was also removed in most constructs in order to aid crystallization. Crystallization was carried out at both room temperature and 290 K using a variety of commercial screening kits and modified versions of an RNA-specific screen (Wahl *et al.*, 1996).

Construct length and crystallization temperature had the most impact on crystal quality. Experiments at 290 K produced higher quality crystals compared with room-temperature crystallization. Shorter constructs comprising only the hinge region produced few hits in initial screens and did not lead to well diffracting crystals. Construct lengths were progressively increased until promising crystals were obtained at 290 K under conditions containing either MgSO $_4$, NaCl and MPD or magnesium acetate, MPD and spermidine. The initial

1 Supplementary material has been deposited in the IUCr electronic archive (Reference: SX5082). Services for accessing this material are described at the back of the journal.

crystals were optimized by sampling different combinations of 10–100 mM MgCl_2 , MgSO_4 , magnesium acetate, $\text{Co}(\text{NH}_3)_6\text{Cl}_3$ with or without NaCl, KCl and spermidine in the pH range 4.2–7.4.

Further construct optimization substantially increased the resolution limit of the diffraction from 5 to 3 Å. Single-base changes in the overall length of the strands or the overhangs dramatically affected the diffraction quality of the crystals. The RNA molecule that gave the best diffracting crystals consisted of two strands with 20 and 24 bases, respectively, with single-base overhangs (Fig. 1*b*). Substituting uracil for 5-Br-U did not affect the crystal quality.

The presence of divalent metal ions was essential for crystal formation. Initial crystals were obtained using magnesium (Mg^{2+}). Unambiguous identification of bound Mg^{2+} ions and differentiation from water molecules in an electron-density map is difficult except at very high resolutions. Therefore, subdomain IIa was also crystallized in the presence of strontium ions. Although the Sr^{2+} ion has a larger ionic radius than that of Mg^{2+} , it has previously been used to substitute for Mg^{2+} in the crystallization of RNA (Mueller *et al.*, 1999; Baugh *et al.*, 2000; Deng *et al.*, 2001; Wedekind & McKay, 2003). The increased diffraction from the heavier strontium ions provided a significant advantage in differentiating the bound metal ions from water molecules. This allowed us to clearly identify additional binding sites for divalent cations in the subdomain IIa structure. The resolution of the data obtained from the crystals grown with Sr^{2+} also was higher (2.3 Å *versus* 2.9 Å; Table 1). The Sr-subdomain IIa crystals had very similar unit-cell parameters to those of the Mg-subdomain IIa crystals, but belonged to the lower symmetry space group $P2_12_12_1$ instead of $P4_32_12$.

3.2. Overall structure

Both the Sr-bound and Mg-bound subdomain IIa crystal structures show the same L-shaped conformation (Fig. 2), with individual RNA molecules aligned in a pseudo-continuous end-to-end arrangement in the crystal lattice. The two crystal forms are nearly identical; the presence of Sr led to a slight breakdown in symmetry along the fourfold axis of the tetragonal Mg form, resulting in orthorhombic crystals. In both structures, the asymmetric internal loop comprising single-stranded nucleotides A53–A57 causes the helix to bend with an angle of about 110° between the terminal vectors making up the curved axis, as calculated by the program *CURVES* (Lavery & Sklenar, 1989) with the bulged

nucleotides omitted. This compares with an angle of 129° measured by the same method for the corresponding region in the average NMR structure of isolated subdomain IIa (Lukavsky *et al.*, 2003) and 117° for the recently reported subdomain IIa crystal structure solved at 3.0 Å resolution (Dibrov *et al.*, 2007). In addition to the slightly smaller helical bend angle, the bulge region is more compact, with the distance between the ribose backbone C1 atoms of the two external bases, U56 and U106, being 3–4 Å less than previously reported (Dibrov *et al.*, 2007). In the structure determined by Dibrov and coworkers, the external U56 base re-orientates to point towards the phosphate backbone of A54 on an adjacent molecule, forming a hydrogen bond. This crystal-packing interaction is not observed in the structures reported here and could provide one explanation for the different helical bend angles.

The unpaired nucleotide A57 is stacked below the C58–G110 base pair as a result of a change in backbone direction at nucleotide U56, which projects out into the solvent. The bases

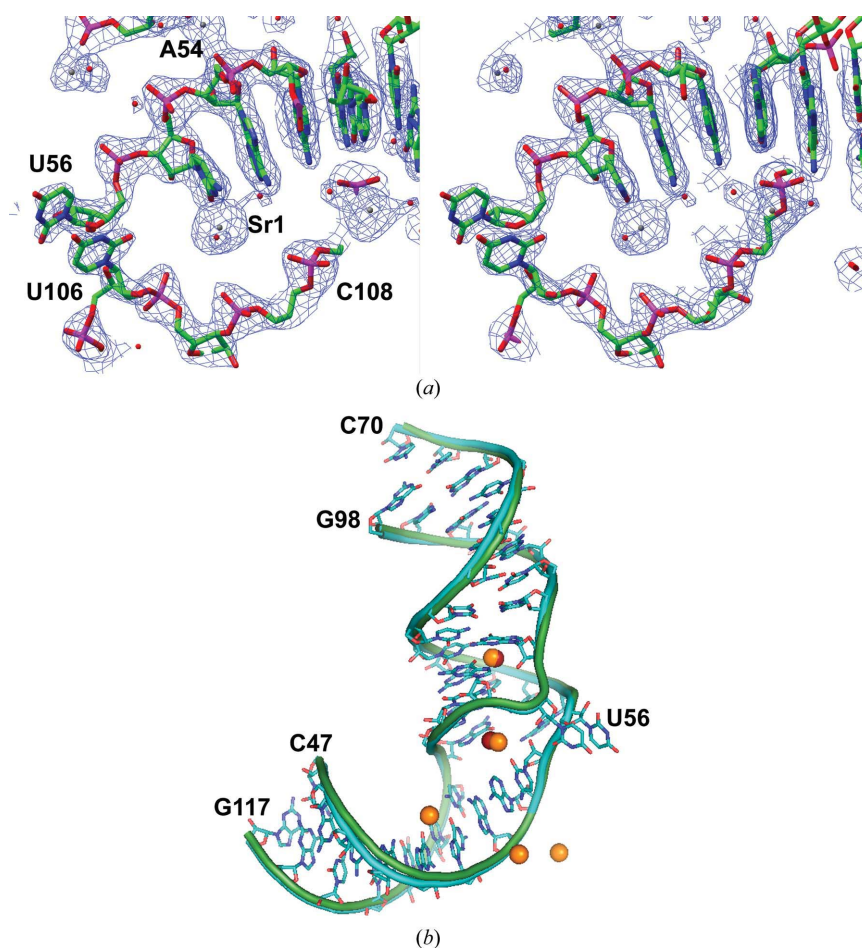


Figure 2

(*a*) Stereoview of the $2F_o - F_c$ electron-density map of the IRES subdomain IIa hinge region with the refined Sr-IIa atomic model superimposed, contoured at $0.26 \text{ e } \text{\AA}^{-3}$. (*b*) The tertiary structure of subdomain IIa shows an L-shaped helical conformation. Superposition of the Sr-IIa (blue) and Mg-IIa (green) structures demonstrates that subdomain IIa adopts the same conformation in the presence of strontium or magnesium metal ions. Two of the Sr^{2+} metal-binding sites are interchangeable with Mg^{2+} (Sr1/Mg1 and Sr2/Mg2). Mg^{2+} atoms are colored red; Sr^{2+} atoms are colored orange. Figs. 2–5 were generated using *PyMOL* (DeLano, 2002) and *CCP4MG* (Potterton *et al.*, 2002).

of nucleotides A53, A54 and C55 are also unpaired and stack continuously with the adjacent G52–C111 base pair. The pocket that is formed by the stacked unpaired bases provides a cluster of hydrogen-bond donors and acceptors that bind metal ions that stabilize the L-shaped conformation and facilitate the interaction of domain IIb with the ribosome. The L-shaped bend in the helical axis is stabilized by a hydrogen-bond interaction between the 2'-hydroxyl group of C55 and the N7 atom of A57. On the opposing strand, a hydrogen bond between the N4 amino group of C111 and the G110 phosphate group provides additional stabilization in the hinge region.

The five-nucleotide internal loop of subdomain IIa has structural features similar to the π -turn motif (Wadley & Pyle, 2004) with nucleotides 1, 2 (A53, A54) and 5 (A57) stacking on adjacent helices. A characteristic signature of π -turns is the

flipped-out nucleotide 4 (U56), which is commonly involved in RNA–RNA or RNA–protein interactions. However, an internal hydrogen bond between the 2'-hydroxyl groups of nucleotides 1 and 5 in the π -turn is absent in the subdomain IIa structure.

Stacking of C62–G105 and the noncanonical U61–G107 base pairs is facilitated by rotation of the U106 base out of the helix and into the solvent (Fig. 3). The noncanonical U61–G107 base pair adjacent to the flipped-out U106 base alters the conformation of the RNA backbone. This brings the backbone of adjacent strands into close proximity in this region. The weak electron density and resulting high atomic temperature factors for the base portions of U56 and U106 indicate that they are highly flexible. In the Sr-IIa structure these two external uridine bases are stacked parallel to each

other at a distance of 3.4–3.7 Å (Fig. 3a), while in the Mg-IIa crystal structure they form an angle of approximately 130° to each other (Fig. 3b). A hydrogen bond is formed between O2 of U56 and O4 of U106, implying tautomerization of one of the bases. These observations are consistent with a UV cross-linking study of domain II that indicated that bases U56 and U106 could be covalently linked (Lyons *et al.*, 2001), suggesting that these residues also are flipped out of the helix and in close contact in solution.

The upper and lower stem regions of loop IIa form conventional A-form helical structures. Watson–Crick base pairing resumes at G52–C111 after the bend and continues to the end of the lower stem (C47–G116). The upper stem region (C58–G110 through C69–G98) contains three noncanonical base pairs that do not significantly alter the helical structure.

3.3. Metal-ion interactions

The presence of divalent cations is essential for HCV IRES to adopt a correctly folded structure and for formation of the IRES–40S ribosomal complex (Kieft *et al.*, 1999). Cations generally promote folding and stabilization of RNA tertiary structure by reducing the electrostatic repulsion between negatively charged phosphate groups or by coordinating directly to electronegative atoms of the nucleotide bases (Draper, 2004). While the majority of cations stabilize RNA structures through long-range diffuse interactions, X-ray crystal structures

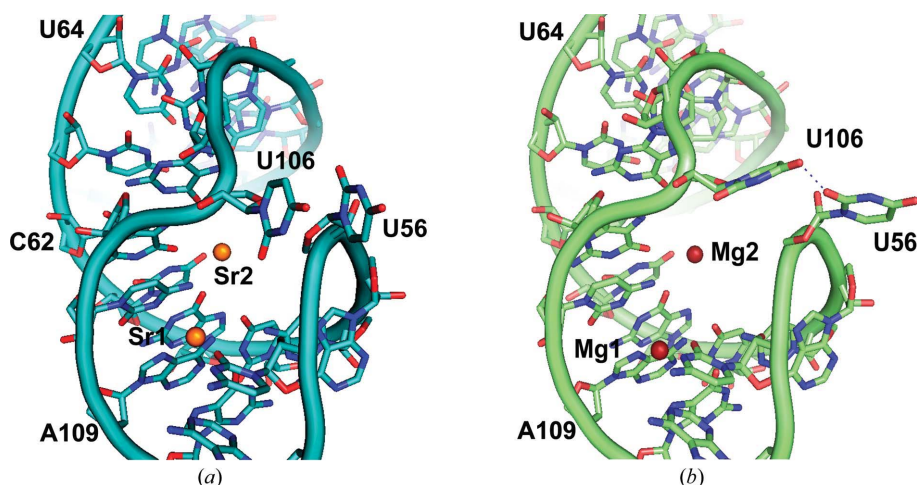


Figure 3

Hinge region of subdomain IIa showing the close proximity of the opposing backbone strands near the solvent-exposed uridines U56 and U106. The binding of Sr1/Mg1 and Sr2/Mg2 to the hinge region is shown. (a) The two looped-out uridines of the Sr-IIa model have their bases stacked parallel to each other at a distance of 3.4–3.7 Å. (b) The two bases of the looped-out uridines in the Mg-IIa model form an angle of about 130° relative to each other and participate in hydrogen-bond formation.

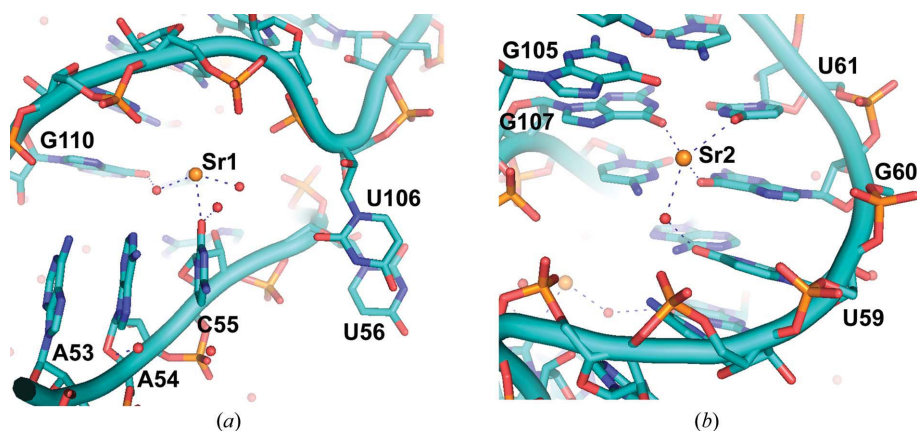


Figure 4

Bound metal ions stabilize the unpaired internal loop and non-Watson–Crick base pairs in the hinge region. (a) Sr1 binds directly to nucleotide C55 and makes a water-mediated interaction with G110. (b) Sr2 coordinates directly to nucleotides G60, U61 and G107 and indirectly to U59 through a tightly bound water molecule. The hydrogen-bond network that is formed by binding metal ions and water is depicted.

and NMR spectroscopic methods have allowed the identification of the specific binding sites of metal ions that participate in modulating RNA structure and function (Allain & Varani, 1995; Cate *et al.*, 1997; Keift & Tinoco, 1997; Hermann & Patel, 1999; Draper, 2004).

Five strontium ions were identified in association with each subdomain IIa RNA molecule in the higher resolution (2.3 Å) Sr-IIa structure (Fig. 2*b*). Two of these metal-binding sites were also clearly occupied by Mg^{2+} ions in the Mg-IIa struc-

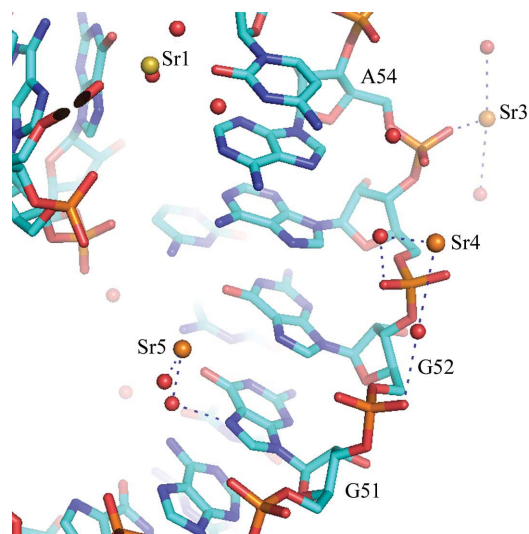


Figure 5

Metal ion-binding sites of Sr3, Sr4 and Sr5 in the Sr-IIa structure. The Sr3 metal ion is directly coordinated to the phosphate group of A54. Sr4 makes a water-mediated contact with the phosphate groups of G52 and A53. Sr5 forms an indirect interaction with the base of G51 in the major groove *via* a tightly bound water molecule. Strontium metal ions are colored orange, waters are colored red and hydrogen bonds are shown as dashed lines.

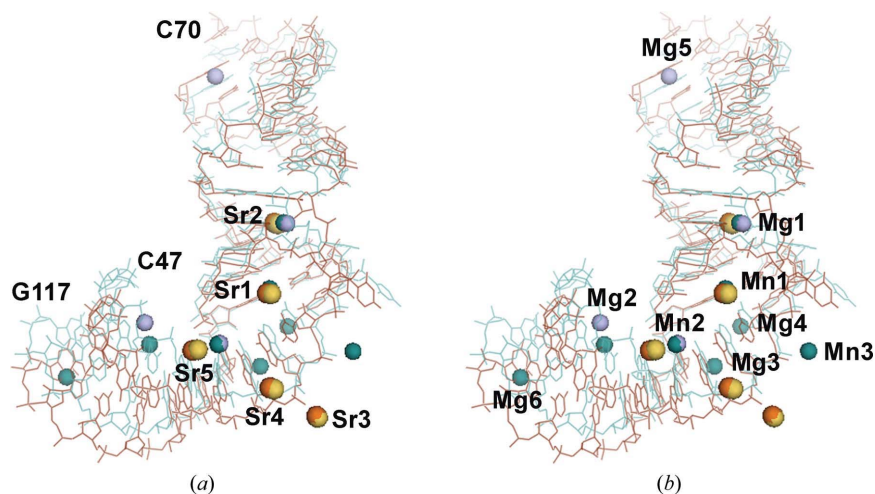


Figure 6

Stereo superposition of the Sr-subdomain IIa model (RNA is colored brown, strontium ions are orange and yellow) with Mg-Mn subdomain IIa model (Dibrov *et al.*, 2007; RNA is colored teal, metal ions are light-blue and teal). For each model, metal ions from both copies in the asymmetric unit are superimposed and represented as spheres. For clarity, water molecules have been removed and only one RNA molecule from each model is shown. The Sr1 metal ion-binding site corresponds to the Mn1 site in the Mg-Mn IIa model. The remaining metal ion-binding sites differ.

ture, demonstrating the ability to interchange Mg^{2+} with Sr^{2+} without altering the RNA fold. The presence of Mg^{2+} or Sr^{2+} was essential for the generation of well diffracting crystals. Interestingly, from our experiments, the divalent cations could be added after annealing the RNA, suggesting that metal ions may not be necessary in the initial RNA-annealing step but are required for the formation of a stable folded structure. Direct binding studies indicated that Mg^{2+} binds with a K_d of $\sim 30 \mu\text{M}$ to the highest affinity sites on subdomain IIa (manuscript in preparation).

In both crystal structures, a metal ion (Mg1, Sr1) is directly coordinated to the O2 carbonyl O atom of nucleotide C55 and makes a water-mediated contact to O6 of G110 (Fig. 4*a*). The single-stranded stacked nucleotides in the hinge region form a metal ion-binding pocket that provides direct coordination with unpaired bases. The bound metal ion acts to further stabilize the bend at the RNA internal loop.

A second metal ion (Mg2, Sr2) binds to the mismatched G–U base pair in the major groove through direct coordination to O4 of U61, O6 of G107 and O6 of G60 (Fig. 4*b*). A hydrogen-bond network in this region, which includes a water molecule that coordinates the metal ion and forms hydrogen bonds to the O6 atoms of U59 and G60, contributes to stabilization of the non-Watson–Crick base pair as well as the unique conformation of the RNA backbone that results in U106 being rotated out into the solvent. The potential instability caused by the close proximity of the backbone strands at U56 and U106 may be mitigated by the binding of Sr2 (Mg2), along with the presence of a tightly bound water molecule that bridges the metal ion and bases. This metal-binding site may also serve a functional role, as base-pair mismatches are often found in ribosomal RNAs, where they provide recognition sites for proteins and small molecules (Gautheret *et al.*, 1995; Chandrasekhar & Malathi, 2003).

The crystal structure shows that the hinge region is stabilized by coordinated metal ions, π -stacking of single-stranded nucleotides and a hydrogen-bonding network that involves tightly bound water molecules as well as nucleotide bases. The perpendicular orientation of bases C55 and G110 is fixed though mutual coordination of Mg1/Sr1 and appears to be critical for stabilizing the conformation of subdomain IIa. The locations of the two metal-binding sites are consistent with the NMR study of subdomain IIa, which suggested the presence of one or two metal-binding sites near nucleotides A54 and C55 (Lukavsky *et al.*, 2003). The existence of an Mg^{2+} -binding site in the hinge region near A109 and G110 has also been suggested by Fe^{II} -EDTA footprinting (Kieft *et al.*, 1999).

Three additional Sr^{2+} -binding sites were observed in the subdomain IIa crystal structure. Sr3 binds directly to the phosphate group of A54, whereas Sr4 is

indirectly coordinated to the phosphate groups of G52 and A53 *via* a water-mediated interaction (Fig. 5). Sr5 is bound in the major groove of the lower stem. This metal ion does not form direct interactions with the RNA, but a coordinating water molecule is hydrogen bonded to N7 of G51. These three Sr^{2+} ions have relatively high temperature factors and appear to be less tightly bound than Sr1 and Sr2 in the hinge region.

The metal-binding sites that we observe differ significantly from those reported by Dibrov and coworkers for the same subdomain IIa construct determined at 3.0 Å resolution in the presence of a mixture of Mg^{2+} and Mn^{2+} ions (Dibrov *et al.*, 2007). Only one metal ion, Mg1/Sr1, coordinated to C55 (see Fig. 6), corresponds closely in position to a metal ion (Mn1) in the Mg-Mn subdomain IIa structure. Mg1 in the Mg-Mn model is in the same proximity as Mg2/Sr2; however, Mg2/Sr2 is significantly closer to the base of U61 (2.5 Å *versus* 5.3–6 Å) and stabilizes the wobble base pair through direct contacts with U61 and G107. In the Mg-Mn subdomain IIa structure, the Mg1 metal ion makes only one contact with N7 of G107 rather than coordinating directly to both base pairs to stabilize their conformation. The metal ion Mn2 in the Mg-Mn model is near Sr5 in our structure; however, it is in contact with the N7 atom of G52 rather than N7 of G51 as seen for Sr5.

We observed no convincing electron density in either of our crystal structures for the Mg5 and Mg6 metal ions located at both ends of the duplex or for Mg2 near G48. Also, no density was observed at sites corresponding to the Mg3, Mg4 and Mn3 ions found between the two RNA molecules in the Mg-Mn subdomain IIa structure (Dibrov *et al.*, 2007).

The 2.3 Å resolution data obtained from the Sr-subdomain IIa crystals allowed us to accurately identify metal ion-binding sites in the electron-density map. In addition, we were able to assign the positions of bound water molecules (see Figs. 4 and 5) which could not be reliably modeled in the lower resolution Mg-subdomain IIa structure and also are absent from previously reported structures (Lukavsky *et al.*, 2003; Dibrov *et al.*, 2007). This has provided an initial view of the role that these water molecules play in coordinating the bound metal ions and stabilizing the RNA conformation.

3.4. Potential for small-molecule interactions

The uniquely folded conformation of IRES subdomain IIa and its critical role in 80S ribosomal assembly (Locker *et al.*, 2007) suggest that it is a possible target for small molecules that bind to subdomain IIa, alter the conformation of the IRES element and thereby inhibit translation of viral mRNA. Recently, a novel class of benzimidazole compounds was identified that bind to HCV IRES subdomain IIa with sub-micromolar affinity (Seth *et al.*, 2005). The binding of these compounds was mapped to the hinge region by RNA-foot-printing studies, which showed that nucleotides C55 and U56 had the greatest protection from enzymatic cleavage in the presence of the compound (Seth *et al.*, 2005).

The crystal structure of IRES subdomain IIa, coupled with the identification of small molecules that bind to this region,

will allow a new structure-guided approach for designing small-molecule inhibitors of HCV IRES-mediated translation.

We thank the staff at APS beamlines 14-ID and 8-BM for providing access to their synchrotron-radiation facility and for their help with data collection and we are grateful to Dr Ke Shi at UMN for helpful discussions during the initial Mg-IIa structure determination.

References

- Allain, F. H. & Varani, G. (1995). *Nucleic Acids Res.* **23**, 341–350.
- Baugh, C., Grate, D. & Wilson, C. (2000). *J. Mol. Biol.* **301**, 117–128.
- Boehringer, D., Theremann, R., Ostareck-Lederer, A., Lewis, J. D. & Stark, H. (2005). *Structure*, **13**, 1695–1706.
- Brown, E. A., Zhang, H., Ping, L. H. & Lemon, S. M. (1992). *Nucleic Acids Res.* **20**, 5041–5045.
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst. D* **54**, 905–921.
- Bukh, J., Purcell, R. H. & Miller, R. H. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 4942–4946.
- Cate, J. H., Hanna, R. L. & Doudna, J. A. (1997). *Nature Struct. Biol.* **4**, 553–558.
- Chandrasekhar, K. & Malathi, R. (2003). *J. Biosci.* **28**, 547–555.
- Davidson, F., Simmonds, P., Ferguson, J. C., Jarvis, L. M., Dow, B. C., Follett, E. A., Seed, C. R., Krusius, T., Lin, C. & Medgyesi, G. A. (1995). *J. Gen. Virol.* **76**, 1197–1204.
- DeLano, W. L. (2002). *The PyMOL Molecular Graphics System*. Palo Alto, California, USA: DeLano Scientific.
- Deng, J., Xiong, Y. & Sundaralingam, M. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 13665–13670.
- Dibrov, S. M., Johnston-Cox, H., Weng, Y. H. & Hermann, T. (2007). *Angew. Chem. Int. Ed. Engl.* **46**, 226–229.
- Draper, D. E. (2004). *RNA*, **10**, 335–343.
- Gautheret, D., Konings, D. & Gutell, R. R. (1995). *RNA*, **1**, 807–814.
- Hermann, T. & Patel, D. J. (1999). *J. Mol. Biol.* **294**, 829–849.
- Honda, M., Brown, E. A. & Lemon, S. M. (1996). *RNA*, **2**, 955–968.
- Ji, H., Fraser, C. S., Yu, Y., Leary, J. & Doudna, J. A. (2004). *Proc. Natl. Acad. Sci. USA*, **101**, 16990–16995.
- Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst. A* **47**, 110–119.
- Keft, J. S. & Tinoco, I. J. (1997). *Structure*, **5**, 713–721.
- Kieft, J. S., Zhou, K., Jubin, R. & Doudna, J. A. (2001). *RNA*, **7**, 194–206.
- Kieft, J. S., Zhou, K., Jubin, R., Murray, M. G., Lau, J. Y. & Doudna, J. A. (1999). *J. Mol. Biol.* **292**, 513–529.
- Kikuchi, K., Umehara, T., Fukuda, K., Hwang, J., Kuno, A., Hasegawa, T. & Nishikawa, S. (2003). *J. Biochem. (Tokyo)*, **133**, 263–270.
- Klinck, R., Westhof, E., Walker, S., Afshar, M., Collier, A. & Aboul-Ela, F. (2000). *RNA*, **6**, 1423–1431.
- Kolupaeva, V. G., Pestova, T. V. & Hellen, C. U. (2000). *J. Virol.* **74**, 6242–6250.
- Lafuente, E., Ramos, R. & Martinez-Salas, E. (2002). *J. Gen. Virol.* **83**, 1113–1121.
- Laletina, E., Graifer, D., Malygin, A., Ivanov, A., Shatsky, I. & Karpova, G. (2006). *Nucleic Acids Res.* **34**, 2027–2036.
- Lavery, R. & Sklenar, H. (1989). *J. Biomol. Struct. Dyn.* **6**, 655–667.
- Locker, N., Easton, L. E. & Lukavsky, P. J. (2007). *EMBO J.* **26**, 795–805.
- Lu, H., Li, W., Noble, W. S., Payan, D. & Anderson, D. C. (2004). *J. Proteome Res.* **3**, 949–957.
- Lukavsky, P. J., Kim, I., Otto, G. A. & Puglisi, J. D. (2003). *Nature Struct. Biol.* **10**, 1033–1038.

- Lyons, A. J., Lytle, J. R., Gomez, J. & Robertson, H. D. (2001). *Nucleic Acids Res.* **29**, 2535–2541.
- Mueller, U., Schubel, H., Sprinzl, M. & Heinemann, U. (1999). *RNA*, **5**, 670–677.
- Otto, G. A. & Puglisi, J. D. (2004). *Cell*, **119**, 369–380.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pestova, T. V., Kolupaeva, V. G., Lomakin, I. B., Pilipenko, E. V., Shatsky, I. N., Agol, V. I. & Hellen, C. U. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 7029–7036.
- Pestova, T. V., Shatsky, I. N., Fletcher, S. P., Jackson, R. J. & Hellen, C. U. (1998). *Genes Dev.* **12**, 67–83.
- Potterton, E., McNicholas, S., Krissinel, E., Cowtan, K. & Noble, M. (2002). *Acta Cryst. D* **58**, 1955–1957.
- Reynolds, J. E., Kaminski, A., Carroll, A. R., Clarke, B. E., Rowlands, D. J. & Jackson, R. J. (1996). *RNA*, **2**, 867–878.
- Seth, P. P., Miyaji, A., Jefferson, E. A., Sannes-Lowery, K. A., Osgood, S. A., Propp, S. S., Ranken, R., Massire, C., Sampath, R., Ecker, D. J., Swayze, E. E. & Griffey, R. H. (2005). *J. Med. Chem.* **48**, 7099–7102.
- Siridechadilok, B., Fraser, C. S., Hall, R. J., Doudna, J. A. & Nogales, E. (2005). *Science*, **310**, 1513–1515.
- Spahn, C. M., Jan, E., Mulder, A., Grassucci, R. A., Sarnow, P. & Frank, J. (2004). *Cell*, **118**, 465–475.
- Spahn, C. M., Kieft, J. S., Grassucci, R. A., Penczek, P. A., Zhou, K., Doudna, J. A. & Frank, J. (2001). *Science*, **291**, 1959–1962.
- Terwilliger, T. C. (2000). *Acta Cryst. D* **56**, 965–972.
- Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst. D* **55**, 849–861.
- Tsukiyama-Kohara, K., Iizuka, N., Kohara, M. & Nomoto, A. (1992). *J. Virol.* **66**, 1476–1483.
- Wadley, L. M. & Pyle, A. M. (2004). *Nucleic Acids Res.* **32**, 6650–6659.
- Wahl, M. C., Ramakrishnan, B., Ban, C., Chen, X. & Sundaralingam, M. (1996). *Acta Cryst. D* **52**, 668–675.
- Wang, C., Le, S. Y., Ali, N. & Siddiqui, A. (1995). *RNA*, **1**, 526–537.
- Wang, C., Sarnow, P. & Siddiqui, A. (1993). *J. Virol.* **67**, 3338–3344.
- Wedekind, J. E. & McKay, D. B. (2003). *Biochemistry*, **42**, 9554–9563.
- Yu, Y., Ji, H., Doudna, J. A. & Leary, J. A. (2005). *Protein Sci.* **14**, 1438–1446.

Crystal Structure of Hepatitis C Virus IRES Subdomain IIa

QIANG ZHAO^{1*}, QING HAN¹, CHARLES R. KISSINGER¹, THOMAS HERMANN^{1,3} and PEGGY A. THOMPSON^{2*}

Departments of ¹Structural Chemistry and ²Biology
Anadys Pharmaceuticals, Inc.
3115 Merryfield Row, San Diego, CA 92121

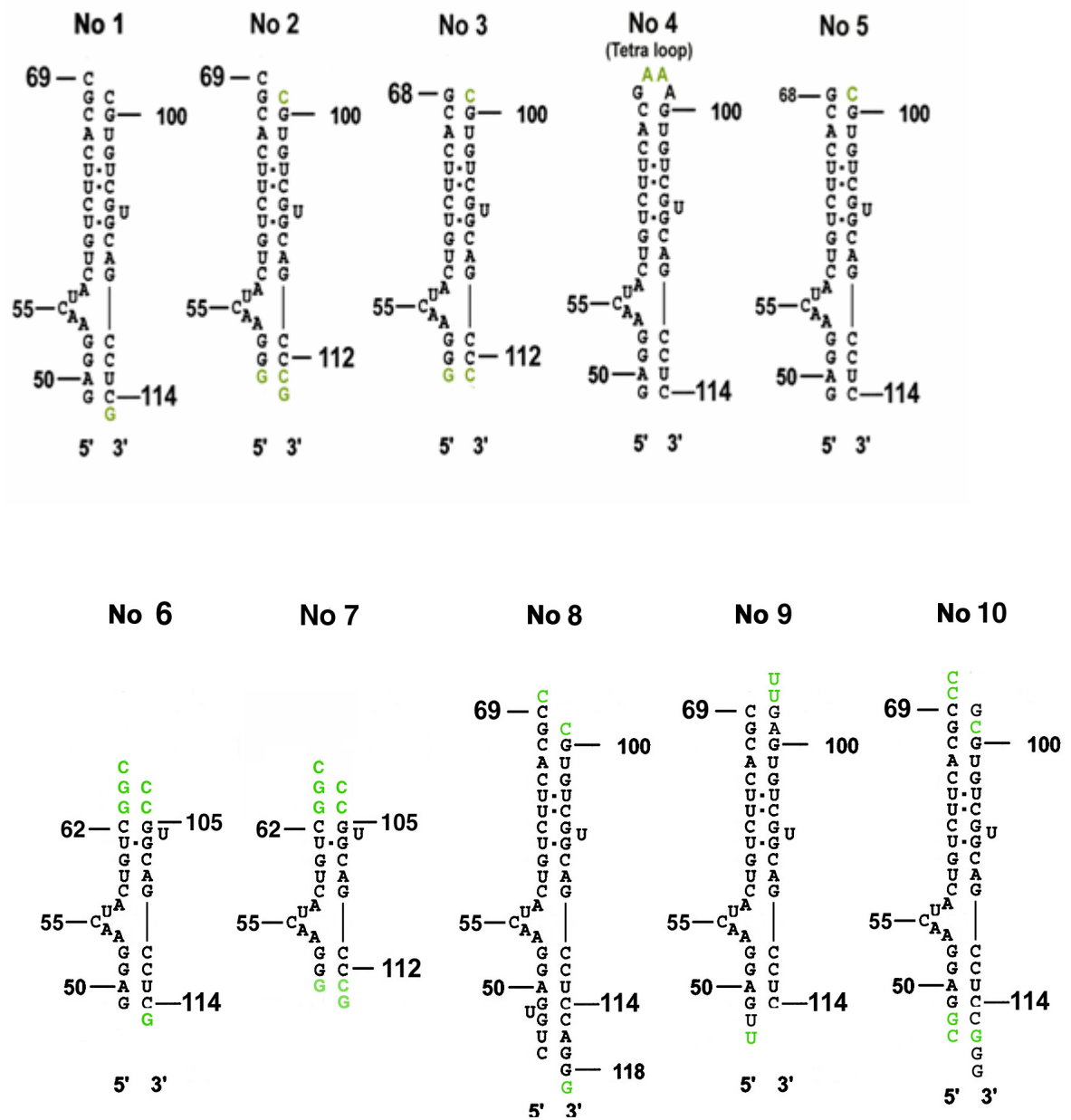
³Current address:
Departments of Chemistry and Biochemistry
University of California, San Diego
9500 Gilman Drive, La Jolla, CA 92093

*Corresponding authors

E-mail addresses of corresponding authors:
qiangzhao@anadyspharma.com; pthompson@anadyspharma.com

Running title: Crystal Structure of IRES Subdomain IIa

Supplementary Material



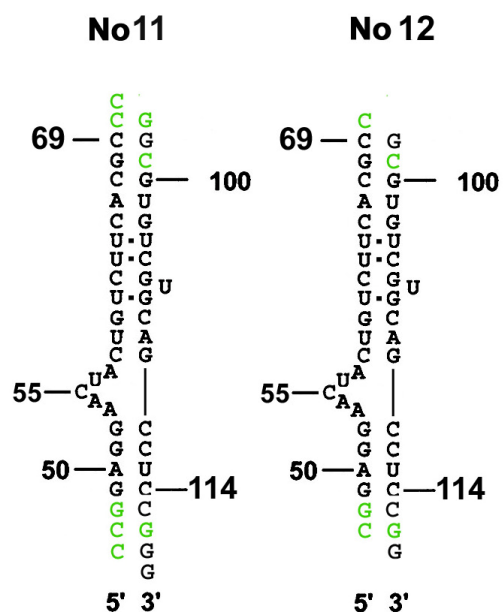


Figure S1. Sequences of the subdomain IIa constructs used for crystallization.

Residues identical to the HCV IRES domain II are shown in black. U-101 and U-113 were also substituted with 5-Br-U in constructs No. 10 and 12. The RNA molecule that gave the best diffracting crystals was construct No. 12. Construct No. 10 also produced nice crystals but diffracted poorly.