Self-assembling RNA square

Sergey M. Dibrov, Jaime McLean, Jerod Parsons, and Thomas Hermann[1]

Department of Chemistry and Biochemistry, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093

Edited by Dinshaw J. Patel, Memorial Sloan-Kettering Cancer Center, New York, NY, and approved February 24, 2011 (received for review December 1, 2010)

The three-dimensional structures of noncoding RNA molecules reveal recurring architectural motifs that have been exploited for the design of artificial RNA nanomaterials. Programmed assembly of RNA nanoobjects from autonomously folding tetraloop–receptor complexes as well as junction motifs has been achieved previously through sequence-directed hybridization of complex sets of long oligonucleotides. Due to size and complexity, structural characterization of artificial RNA nanoobjects has been limited to low-resolution microscopy studies. Here we present the design, construction, and crystal structure determination at 2.2 Å of the smallest yet square-shaped nanoobject made entirely of double-stranded RNA. The RNA square is comprised of 100 residues and self-assembles from four copies each of two oligonucleotides of 10 and 15 bases length. Despite the high symmetry on the level of secondary structure, the three-dimensional architecture of the square is asymmetric, with all four corners adopting distinct folding patterns. We demonstrate the programmed self-assembly of RNA squares from complex mixtures of corner units and establish a concept to exploit the RNA square as a combinatorial nanoscale platform.

crystallography | fluorescence | RNA structure

Noncoding RNA sequences can adopt intricate three-dimensional architectures whose complexity rivals those of proteins. The folding of RNA is governed by recurring structural motifs (1), the most common of which is the double helix that involves consecutively stacked pairs of complementary nucleobases interacting via hydrogen bonds. Structural motifs in RNA often form locally without the involvement of long-range tertiary interactions (2). The synthetic combination of RNA motifs has been exploited in the design of functional and architectural RNA structures (3–5), including artificial ribosensors and RNA "Lego" (6–8). Similar to well-established methods for the design and construction of DNA "origami" and nanomaterials (9, 10), approaches toward artificial RNA architectures have relied on the ability of RNA strands to hybridize via complementary base sequences (11, 12). Long oligonucleotides of considerable sequence complexity have been used to build complex RNA Lego as well as square- and cube-shaped objects (7, 13, 14). Here, we describe the design and crystal structure analysis of an RNA nanoobject that self-assembles from multiple copies of two short oligonucleotides which give rise to the smallest possible square structure that may be built from double-stranded RNA. We demonstrate sequence-dependent programmed self-assembly of the RNA square which might be exploited as a nanoscale platform for the directed combination of up to four molecular entities.

Results and Discussion

Previously, we had determined the three-dimensional structure of the domain IIa bulge in the internal ribosome entry site (IRES) of the hepatitis C virus (HCV) RNA genome (15) (Fig. 1 *A* and *B*). The domain IIa, which is a potential target of antiviral drugs that block HCV protein synthesis (16), adopts a sharply bent fold that is required for the correct spatial positioning of the HCV IRES during recruitment of host cell ribosomes (17, 18). Structural analysis of the unique 90° bend in the IIa domain led us to conclude that this viral RNA motif may constitute the most compact

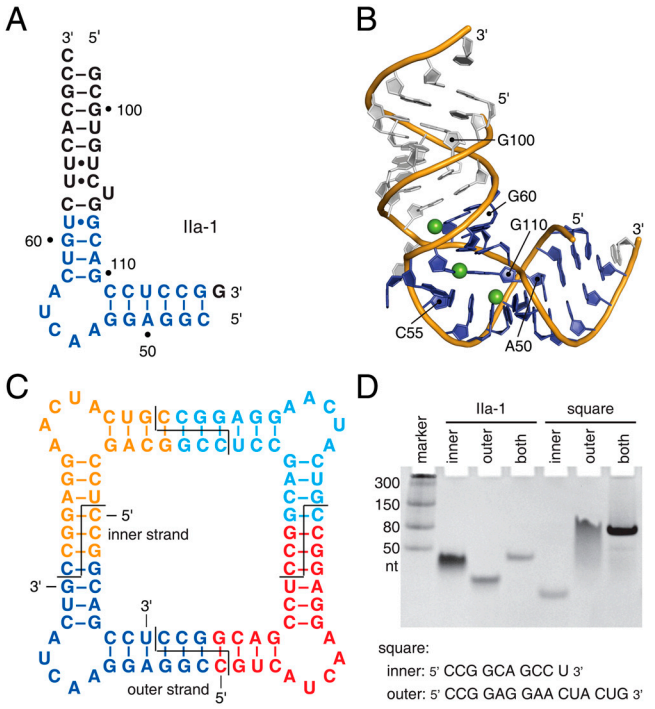

Fig. 1. Design of the self-assembling RNA square. (*A*) Secondary structure of the IIa-1 RNA representing the subdomain IIa of the IRES from HCV. Numbering was adopted from the HCV genome. (*B*) Three-dimensional structure of the IIa-1 RNA (15). Positions of magnesium ions are indicated by spheres. (*C*) Secondary structure of the RNA square. The four copies of the IIa-1 core are highlighted in different colors. Lines indicate boundaries of oligonucleotides. (*D*) Native polyacrylamide gel electrophoresis of IIa-1 and square RNA. Inner and outer oligonucleotides were loaded at higher concentrations compared to assembled IIa and square RNA.

L-shaped object, or nanocorner, that can be built from contiguous double-stranded RNA. Molecular modeling based on the IIa crystal structure suggested that combination of four such nanocorners would allow the assembly of a double-stranded RNA square (Fig. 1*C*) in which the internal loop motifs are separated by 10 base pairs, corresponding to a full turn in an A-form RNA double helix. Introduction of appropriate termini in the modeled RNA square allowed a design that required only two distinct oligonucleotides ("inner" strand of 10 nucleotides and "outer" strand of 15 nucleotides) that would self-assemble via complementary overlapping sequences of four bases. The resulting RNA square is comprised of four identical copies of each, inner and

Author contributions: T.H. designed research; S.M.D., J.M., J.P., and T.H. performed research; S.M.D. and T.H. analyzed data; 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 and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 3P59).

[1]To whom correspondence should be addressed. E-mail: tch@ucsd.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1017999108/-/DCSupplemental.

outer strand, totaling 100 nucleotides. The design of the RNA square was tested by annealing of inner and outer strand oligonucleotides which produced a single RNA species that migrated in a native polyacrylamide gel at a size compatible with the square architecture (Fig. 1*D* and *SI Appendix*, Fig. S1).

The three-dimensional structure of the 100-nucleotide RNA square was investigated by X-ray crystallography. Single wavelength anomalous diffraction data collected from crystals of inner strand oligonucleotide annealed to a 5-bromo uridine-labeled (^{5Br}U59) outer strand was used to determine the structure of the RNA square at 2.2-Å resolution (Fig. 2 and *SI Appendix*, Figs. S2–S6 and Tables S1 and S2). The crystal structure revealed a fully double-stranded RNA architecture that corresponded to the original design of the square, containing four outer strand oligonucleotides (designated A, C, E, and G) as well as four inner strands (designated B, D, F, and H) which assemble via overlapping complementary sequences and fold into the 90° bent corners of the HCV IRES domain IIa. The sides of the square measure approximately 60 Å. The distance across the interior space is about 18 Å. The termini of the outer strands are located at the interior of the square, whereas those of the inner strands are oriented at the outer circumference.

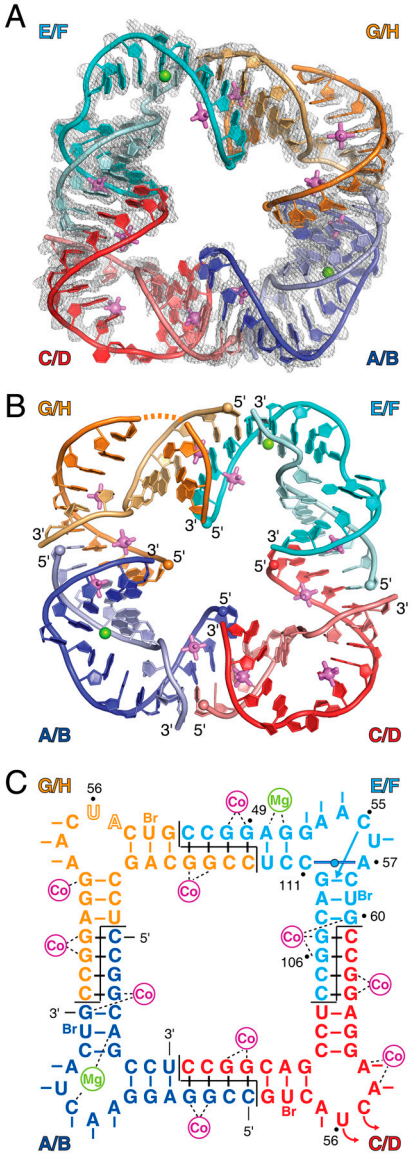

Fig. 2. Structure of the RNA square. (*A*) Minor groove side of the helical regions ($2F_o - F_c$ electron density map contoured at 1σ). (*B*) View from the major groove side. (*C*) Secondary structure revised according to the crystal structure. Dashes at nucleotides in the corner loops indicate continuous stacking of bases on an adjacent helix; arrows depict rotated-out residues that do not stack on neighboring bases. In the E/F corner, A57-C111 form a *cis*-Watson-Crick pair and C55 participates in a base triple while docking at the major groove edge of C58-G110. Binding sites of Mg^{2+} and Co^{3+} are indicated.

In agreement with the migration behavior in native polyacrylamide gel electrophoresis, the RNA square adopts a highly compact architecture in which 96 of the 100 nucleotides have their bases stacked with at least one neighboring residue. The bases of C55 and U56 in strand C are rotated out from the RNA fold and participate in stacking with a neighboring square. Nucleotides U56 and A57 in strand G were disordered in the electron density map. The overall shape of the RNA square in the crystal revealed unexpected asymmetry, attested by the fact that the four corners all adopt slightly different structures (Fig. 3). Corners formed from strands designated as A/B and G/H (Fig. 2) are most similar to each other and to the structure of the core in the HCV IRES domain IIa (*SI Appendix*, Table S3). The corner comprised of strands E/F is most dissimilar to all others, attested by the presence of a unique non-Watson-Crick A57–C111 base pair as well as a neighboring G110-C58/C55 base triple (Fig. 4), which are both absent in the original HCV IRES domain IIa. Participation of C111 in the A–C pair is at the expense of the canonical G52–C111 pair which is absent in the corner formed by strands E/F. As a consequence, the helix connecting corners E/F and G/H contains only nine base pairs, whereas the helix between corners C/D and E/F has 11 pairs. The other sides of the square, the helices between corners A/B and C/D as well as G/H and A/B, are each comprised of 10 base pairs as was designed originally.

The asymmetry observed in the four corners extends to the intermolecular contacts involved in packing of squares in the crystal (*SI Appendix*, Table S4). Within the plane, squares line up along diagonals with corners A/B in one square tightly packing against corners E/F in the adjacent molecule (*SI Appendix*, Fig. S3), forming numerous intermolecular hydrogen bonds (*SI Appendix*, Table S4). Corners C/D and G/H are pointing toward the cleavage formed between sides between corners G/H and A/B as well as E/F and G/H in neighboring squares. The tight interactions between A/B and E/F corners of adjacent squares

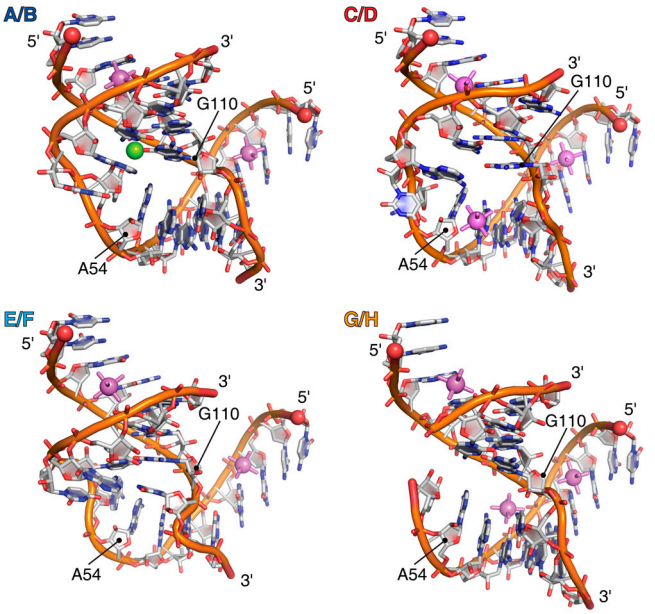

Fig. 3. Structure of the four corners within the RNA square. To facilitate comparison, the corners were rotated to show the same orientation in each case.

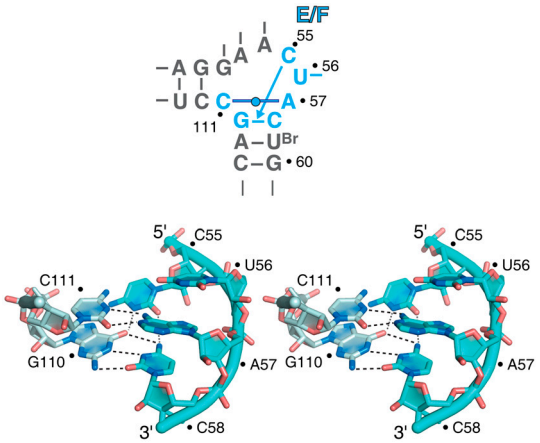

Fig. 4. Stereoview of the non-Watson-Crick A57-C111 base pair and the neighboring G110-C58/C55 base triple in the E/F corner.

is likely the cause for distortions in the secondary structure of the E/F strand pair as was discussed above. The distinct involvement of corners in the crystal packing is reflected in less surface exposure and lower thermal factors for the A/B and E/F corners as opposed to the pairs of C/D and G/H strands which show larger exposed surfaces as well as higher B factors (*SI Appendix*, Fig. S4 and Table S5).

Crystallization of the RNA square was critically dependent on the presence of metal ions, as was expected from the important contribution of magnesium cations in the structure of the HCV IRES domain IIa (15). In the crystallized RNA square, 10 hexamminecobalt(III) and 2 magnesium cations are associated with the RNA (Fig. 2*C* and *SI Appendix*, Table S6). The four double-helical regions that connect the corners and contain the overlapping termini of the outer and inner strands are each stabilized by two hexamminecobalt(III) ions. One such cation is bound at the major groove edges of guanines in both strands (G60, G106, G107), while another bridges the major groove edges of two guanines in the outer strands only (G48, G49). These two metal positions are identical to magnesium ion binding sites in the crystal structure of the HCV IRES domain IIa. The four internal loops that form the corners of the RNA square have additional metal ions associated, albeit each at a different position. The metals in three of the corners have magnesium counterparts in the domain IIa (one Mg^{+} each, bridging C55 and A109 in strands A/B as well as a Co^{3+} at G52 in strand G; one Mg^{2+} bridging G51 and G52 in strand E). A Co^{3+}, associated with both A53 and A54 of strand C, is unique to the square structure.

After gel shift analysis and crystal structure determination confirmed the design concept of the RNA square, proof of principle experiments were performed to establish the utility of the RNA square as a nanoscale platform for the programmed combination of molecular entities. The formation of the RNA square from noncovalently interacting corner units allows the selective assembly of distinct squares, depending on variations of the overlapping recognition sequences in the 5′ termini of outer and inner oligonucleotide strands (Fig. 5*A*). Noncomplementary sequences on both overhanging ends of a corner unit prevent the formation of the RNA square unless another corner unit with complementary sequences is provided. We have used two distinct fluorescently labeled corner units that can be discriminated by their recognition sequences (sequences A and B in a cy3-labeled unit, designated as A^B-cy3; sequences C and D in a cy5-labeled unit, designated as C^D-cy5; see *SI Appendix*, Table S1) as well as by their fluorescent properties (Fig. 5). Directed by the crystal structure, we chose to attach the fluorescent cyanine dyes at the 5′ terminus of the outer strand oligonucleotide. In a series of complementation experiments, we added unlabeled matching

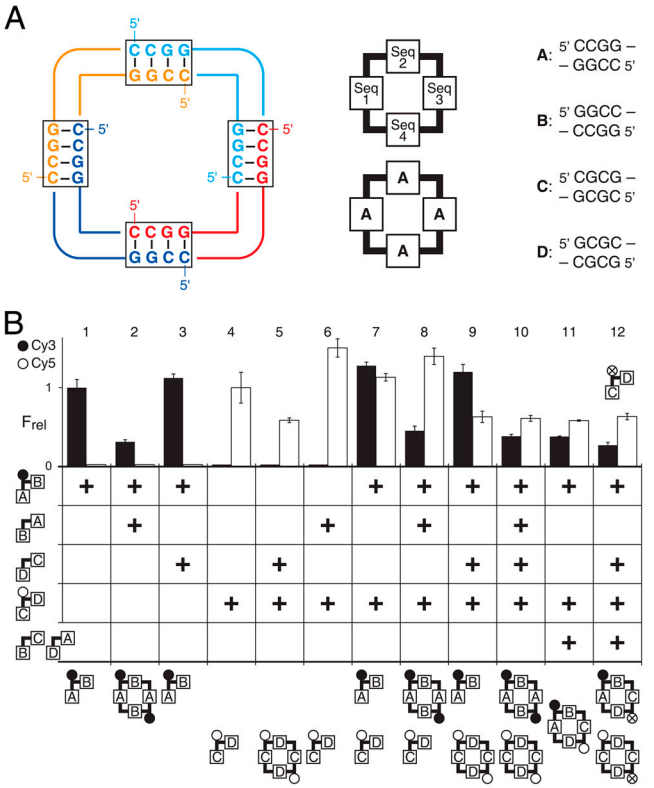

Fig. 5. Programmed self-assembly of fluorescently labeled RNA squares. (*A*) Four 5′-terminal bases in the oligonucleotides of the corner units constitute overlapping sequences for assembly via strand hybridization. In the crystallography construct, all four recognition sequences were identical (A), resulting in corners designated as A^A. Permutations were generated (B, C, D) for the programmed assembly of fluorescently labeled corners with nonidentical recognition sequences (X^Y), which are unable to self-associate. (*B*) Fluorescence of dye-labeled corners A^B-cy3 and C^D-cy5 in complementation experiments (●,cy3; ○, cy5; ⊗, quencher). A + symbol indicates the presence of a corner unit. Species that contribute to the fluorescence in each experiment are depicted.

or mismatched corners (B^A or D^C, or an equimolar mixture of B^C and D^A) and recorded the fluorescence signal of the cyanine dyes (Fig. 5, 1–12). We hypothesized that when a dye-labeled corner unit is incorporated in an RNA square, the cyanine residue at the 5′ end of the outer strand is placed inside the square where fluorescence is reduced relative to an unincorporated corner in which the dye points freely into the solvent. The corner units used in the complementation experiments were designed such that, upon square formation, two dye labels would be confined in close proximity inside the 18-Å cavity of the square, which would result in self-quenching. The complementation experiments supported these hypotheses.

Dye-labeled corner units by themselves showed high fluorescence, both as individual entities (Fig. 5, experiments 1 and 4) and as a mixture (experiment 7). Addition of unlabeled complementary corner units led to quenching (experiments 2 and 5), whereas units with mismatched recognition sequences had no effect on fluorescence (experiments 3 and 6), indicating that the formation of the RNA square is dependent on correct sequence recognition at the overhanging termini of the corner units. In mixtures of the two different dye-labeled corner units, selective assembly of only one of the labeled squares was achieved by addition of the respective complementary unit, whereas the other labeled corner was not affected (experiments 8 and 9), unless both complementary units were present (experiment 10). Hybrid squares, labeled with both cyanine dyes, could be formed by addition of a mixture of B^C and D^A units (experiment 11).

As a control, a unit carrying a high-efficiency quencher was added at 10-fold excess to a mixture of the dye-labeled corners, along with the B^C and D^A units, which led to reduction of fluorescence in both dyes comparable to self-quenching (experiment 12).

These experiments demonstrate that the selective association of distinct RNA squares from complex mixtures of corner units can be controlled by addition of sequence-complementary units. The programmed self-assembly described here establishes a concept to exploit the RNA square as a nanoscale platform for the directed combination of up to four molecular entities that are linked to the corner units. Simple modifications of the square architecture promise to provide unique RNA architectures as materials for the construction of nanostructures. For example, elongation of the square sides by one base pair (to 11 pairs; see *SI Appendix*, Fig. S7) would force the adjoining corners to rotate out of plane, which would prevent the closure of the square and perhaps allowing the association of more than four corner units to give spiral-like structures. Preliminary experiments suggest that populations of such spirals of varying number of constituting corners might indeed form as a consequence of the unhybridized recognition sequences of the terminal units (*SI Appendix*, Fig. S7).

Materials and Methods

RNA Preparation. RNA was annealed from stoichiometric amounts of HPLC-purified oligonucleotides in 10 mM sodium cacodylate buffer, pH 6.5, 5 mM $MgCl_2$ (*SI Appendix*, Table S1).

Gel Electrophoresis. RNA was analyzed on 13% native polyacrylamide gel in 40 mM MOPS buffer and 2.5 mM $MgCl_2$. Visualization was performed under UV after ethidium bromide staining.

Crystallization, Data Collection, and Structure Determination. Square RNA at 0.2-mM concentration was mixed with an equal volume of precipitating solution containing 50 mM sodium cacodylate, pH 6.5, 150 mM $MgCl_2$, 8–10 mM $[Co(NH_3)_6]Cl_3$, 50 mM KCl, and 15–20% PEG400. Crystals grew at 16 °C by hanging drop vapor diffusion after equilibration against 17% PEG400 in water. Diffraction data were collected on flash-cooled crystals on beamline 17-ID-B at the Advanced Photon Source, Argonne National Laboratory. Data were processed with HKL2000 (19). Examination of diffraction data by the program Xtriage (20) revealed a twinning component with a fraction of 0.498 and following the twin law $-h$, $-k$, l. Initial phases were calculated in PHENIX (20) by the single wavelength anomalous diffraction method using the anomalous scattering from a bromine incorporated at U59. An initial model was automatically built in PHENIX followed by iterative rounds of manual building and refinement, alternating between Refmac (21) using twin refinement within CCP4 (22) and manual rebuilding in Coot (23) based on the obtained $2F_o - F_c$ and $F_o - F_c$ maps. Metal ions were assigned based on electron density and geometry of coordinating ligands. Final refinement was carried out in PHENIX with combined translation libration screw motion, individual isotropic atomic displacement parameters, and water picking.

Fluorescence Experiments. Fluorescence experiments were performed at 100-nM RNA concentration in black 96-well plates on a Spectra Max Gemini monochromator plate reader (Molecular Devices) at 25 °C. Excitation was at 540 (Cy3) or 640 nm (Cy5). Emission was read at 570 (Cy3) or 670 nm (Cy5).

ACKNOWLEDGMENTS. We thank T. Pham and J. Johnson for help during the early phase of the project. Supported in part by the National Institutes of Health, Grant R01 AI72012 (to T.H.). Use of the Advanced Photon Source for X-ray diffraction data collection was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract W-31-109-Eng-38. Use of the Industrial Macromolecular Crystallography Association–Collaborative Access Team beamline 17-ID at the Advanced Photon Source was supported by the companies of the Industrial Macromolecular Crystallography Association through a contract with the Center for Advanced Radiation Sources at the University of Chicago.

1. Holbrook SR, Cheong C, Tinoco I, Jr, Kim SH (1991) Crystal structure of an RNA double helix incorporating a track of non-Watson-Crick base pairs. *Nature* 353:579–581.
2. Woodson SA (2010) Compact intermediates in RNA folding. *Annu Rev Biophys* 39:61–77.
3. Jaeger L, Westhof E, Leontis NB (2001) TectoRNA: Modular assembly units for the construction of RNA nano-objects. *Nucleic Acids Res* 29:455–463.
4. Guo P (2005) RNA nanotechnology: Engineering, assembly and applications in detection, gene delivery and therapy. *J Nanosci Nanotechnol* 5:1964–1982.
5. Jaeger L, Chworos A (2006) The architectonics of programmable RNA and DNA nanostructures. *Curr Opin Struct Biol* 16:531–543.
6. Smalley MK, Silverman SK (2006) Fluorescence of covalently attached pyrene as a general RNA folding probe. *Nucleic Acids Res* 34:152–166.
7. Chworos A, et al. (2004) Building programmable jigsaw puzzles with RNA. *Science* 306:2068–2072.
8. Win MN, Liang JC, Smolke CD (2009) Frameworks for programming biological function through RNA parts and devices. *Chem Biol* 16:298–310.
9. Rothemund PW (2006) Folding DNA to create nanoscale shapes and patterns. *Nature* 440:297–302.
10. Seeman NC (2010) Nanomaterials based on DNA. *Annu Rev Biochem* 79:65–87.
11. Shapiro BA, Bindewald E, Kasprzak W, Yingling Y (2008) Protocols for the in silico design of RNA nanostructures. *Methods Mol Biol* 474:93–115.
12. Jaeger L (2009) Defining the syntax for self-assembling RNA tertiary architectures. *Nucleic Acids Symp Ser (Oxf)* 53:83–84.
13. Severcan I, Geary C, Verzemnieks E, Chworos A, Jaeger L (2009) Square-shaped RNA particles from different RNA folds. *Nano Lett* 9:1270–1277.
14. Afonin KA, et al. (2010) In vitro assembly of cubic RNA-based scaffolds designed in silico. *Nat Nanotechnol* 5:676–682.
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* 46:226–229.
16. Parsons J, et al. (2009) Conformational inhibition of the hepatitis C virus internal ribosome entry site RNA. *Nat Chem Biol* 5:823–825.
17. Lukavsky PJ, Kim I, Otto GA, Puglisi JD (2003) Structure of HCV IRES domain II determined by NMR. *Nat Struct Biol* 10:1033–1038.
18. Spahn CM, et al. (2001) Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit. *Science* 291:1959–1962.
19. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
20. Adams PD, et al. (2002) PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* 58:1948–1954.
21. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53:240–255.
22. Collaborative Computational Project, N (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50:760–763.
23. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.

Supporting Information for

Self-assembling RNA square

Sergey M. Dibrov, Jaime McLean, Jerod Parsons, and Thomas Hermann[1]

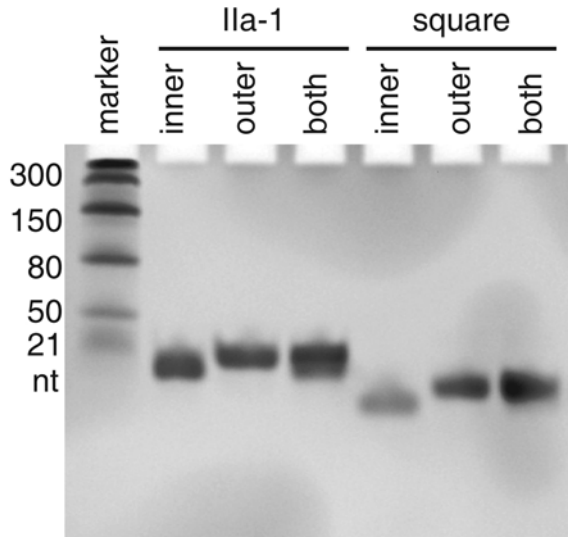

Fig. S1. Denaturing polyacrylamide gel electrophoresis analysis of IIa-1 and square RNA. The same samples as shown in Figure 1*D* were analyzed. Marker oligonucleotides are single stranded. Individual inner and outer oligonucleotides were loaded at higher concentrations compared to assembled IIa-1 and square RNA ("both").

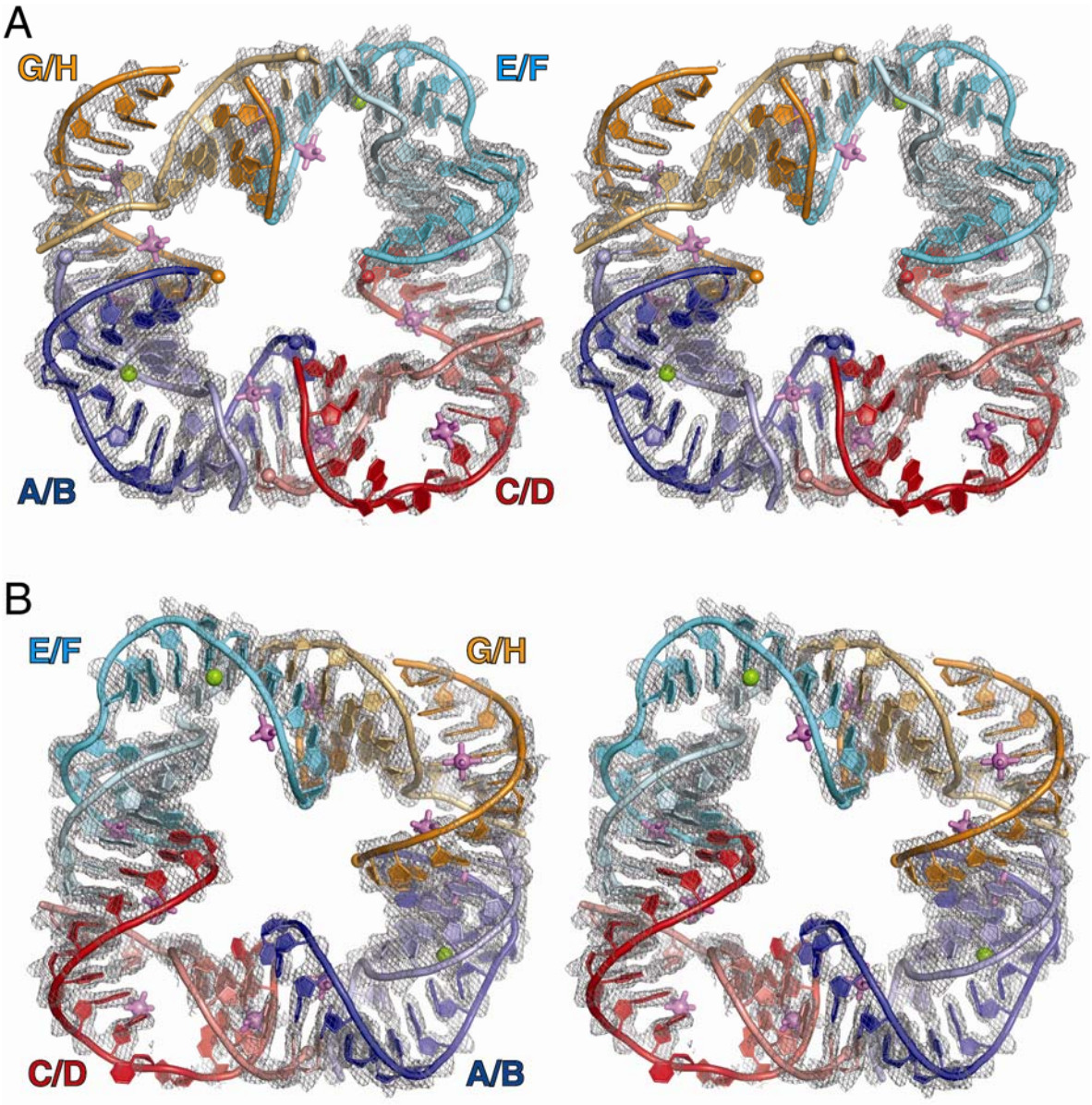

Fig. S2. Stereo views of the self-assembling RNA square overlaid with a 2Fo-Fc electron density map contoured at 1σ. (*A*) Front view. (*B*) Back view. The color scheme of the strands corresponds to the colors in Fig. 1*C*.

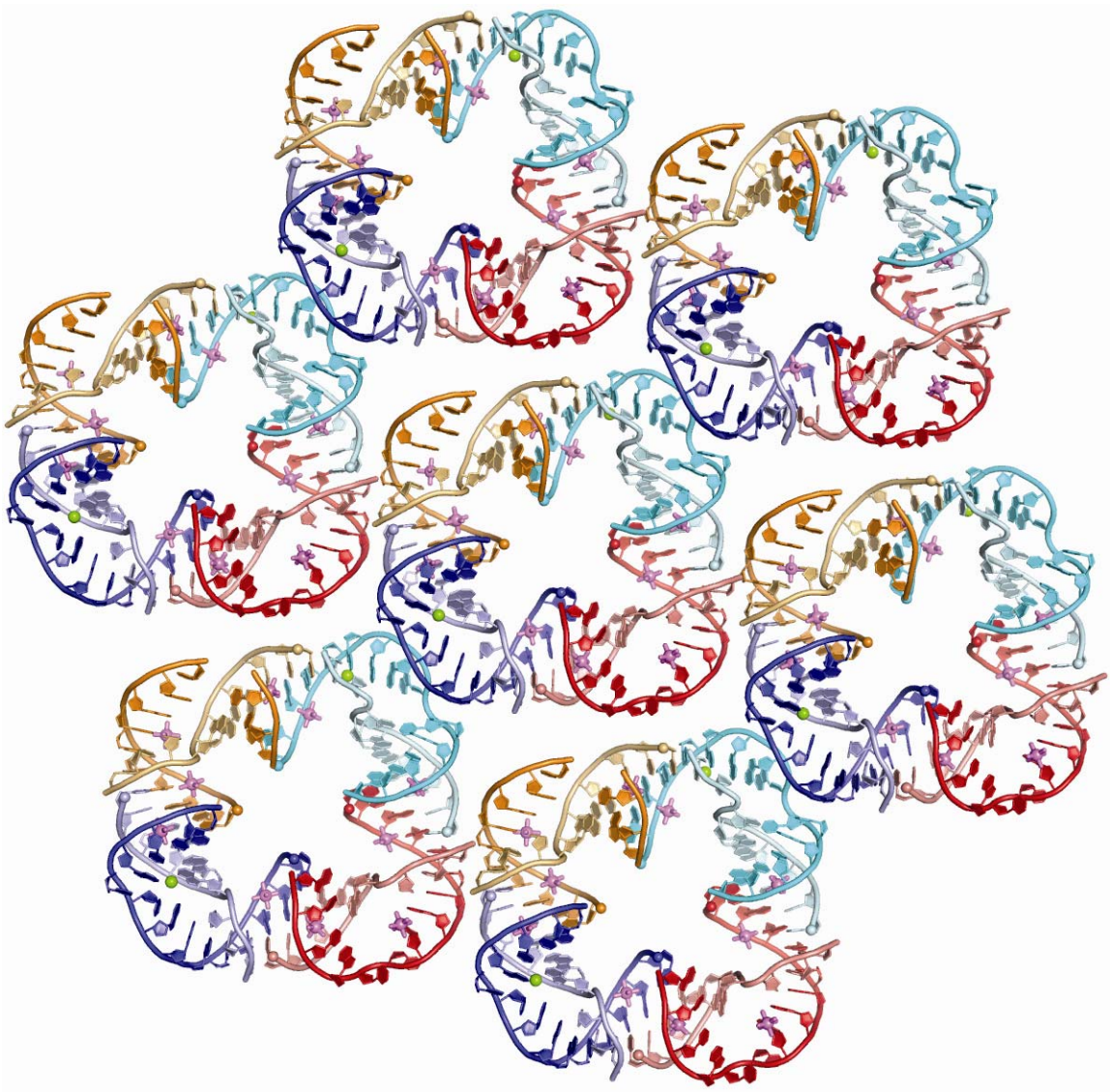

Fig. S3. Crystal packing of the self-assembling RNA square. View of next neighbors within a plane of squares. The color scheme of the strands corresponds to the colors in Figure 1*C*.

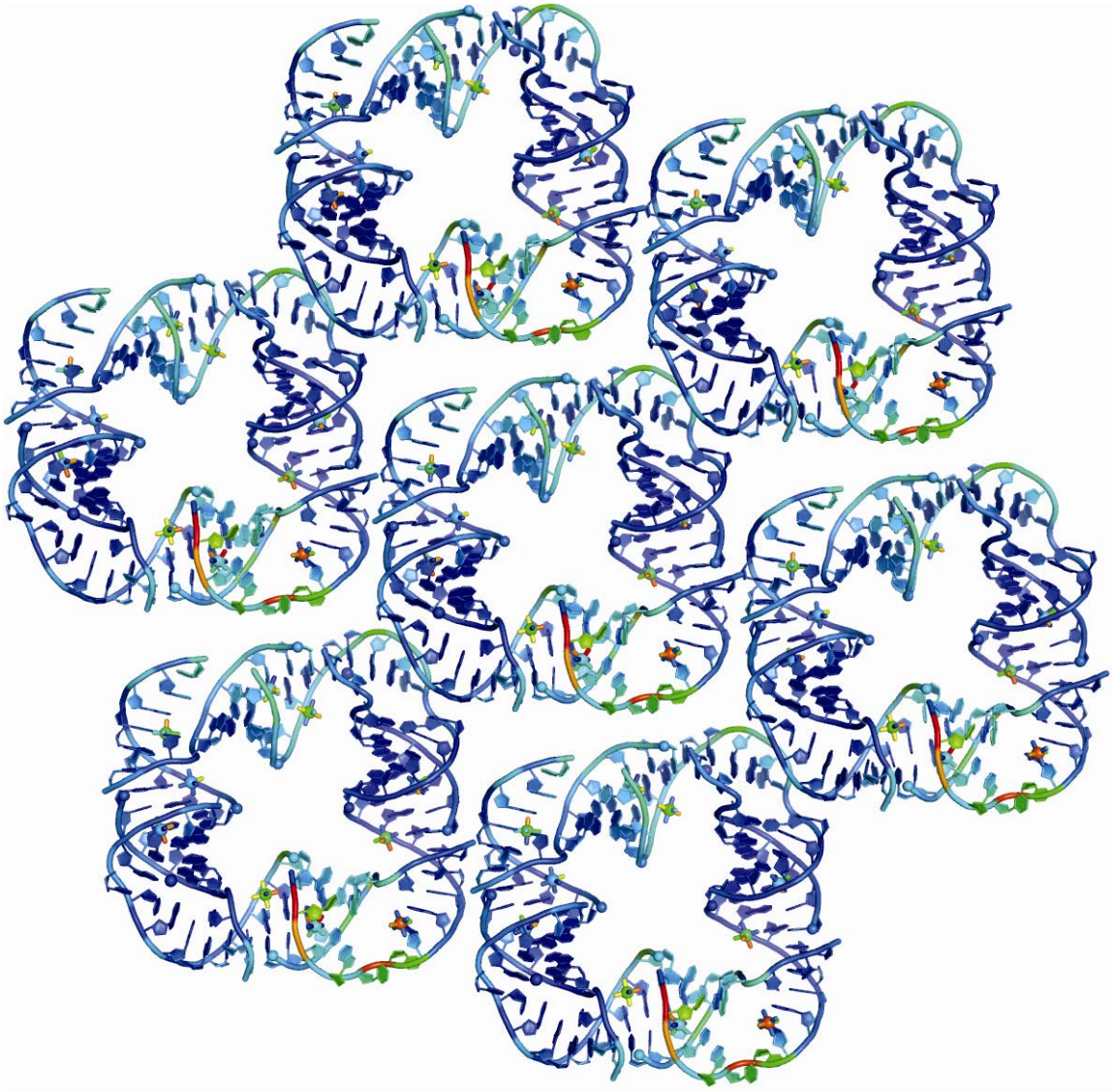

Fig. S4. Crystal packing of the self-assembling RNA square. View of next neighbors within a plane of squares. Strands are colored according to B factors (blue = low, red = high B factors).

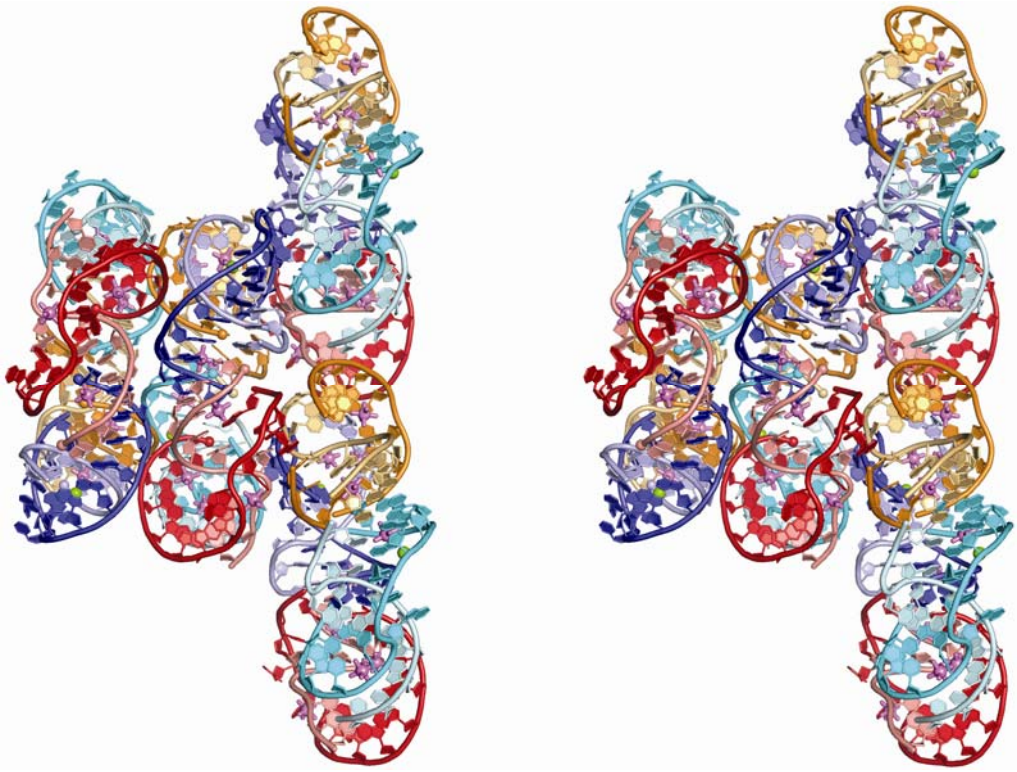

Fig. S5. Stereo view of the crystal packing of the self-assembling RNA square. View of next neighbors in adjacent planes of squares. The color scheme of the strands corresponds to the colors in Fig. 1*C*.

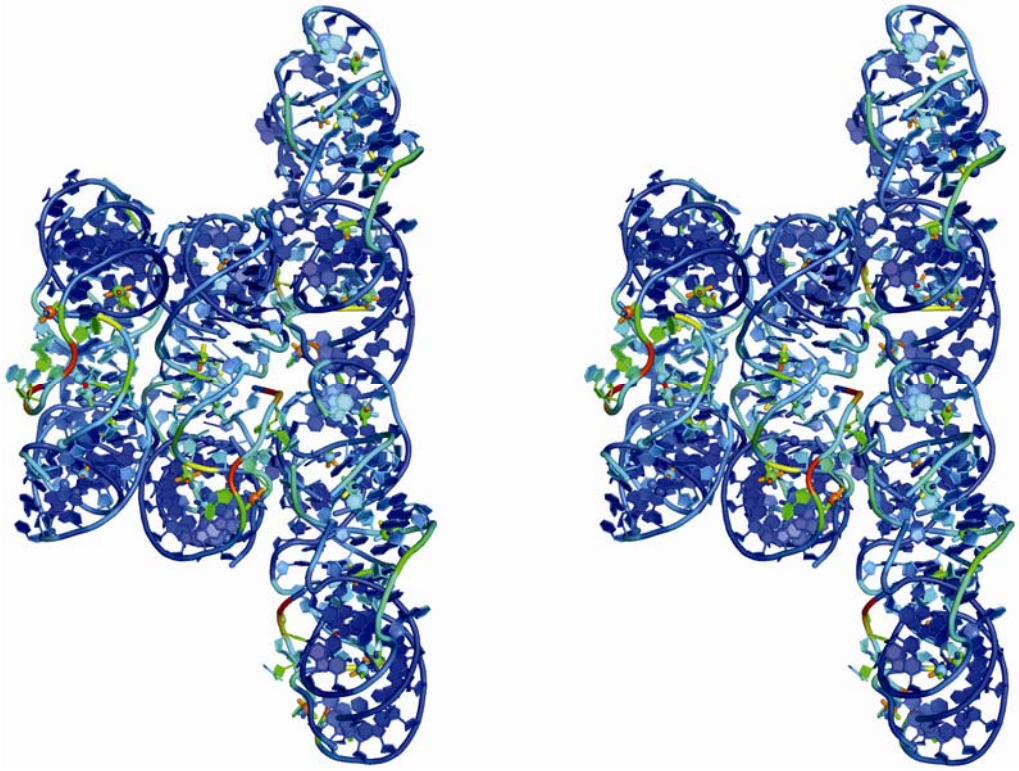

Fig. S6. Stereo view of the crystal packing of the self-assembling RNA square. View of next neighbors in adjacent planes of squares. Strands are colored according to B factors (blue = low, red = high B factors).

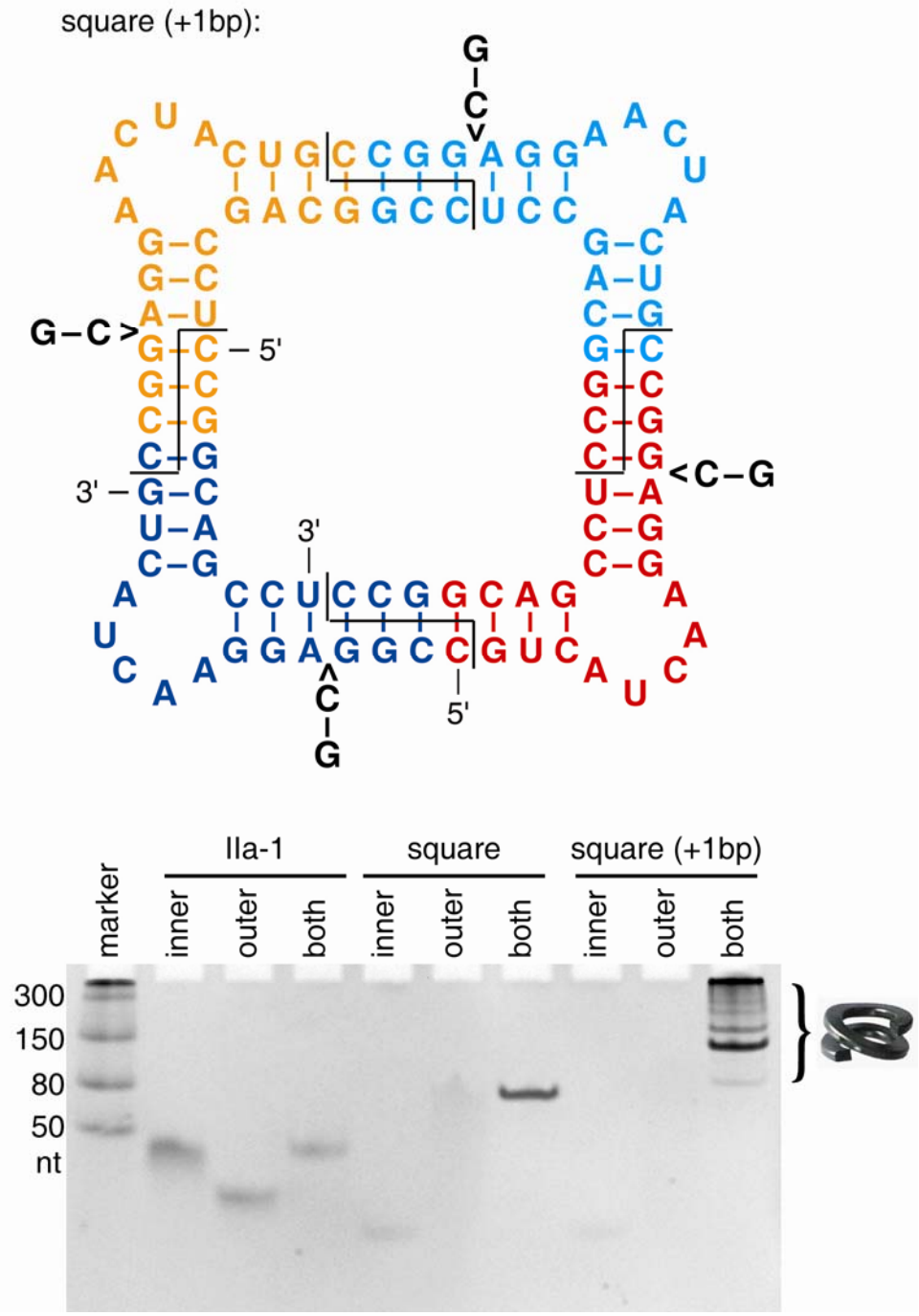

Fig. S7. Native polyacrylamide gel electrophoresis analysis of IIa-1, square RNA and an RNA that contained one additional base pair in each side of the square (square +1bp). Marker oligonucleotides are single stranded.

Table S1. Oligonucleotides used for experiments.

Oligonucleotide	Name	Sequence
Gel electrophoresis:		
IIA1-inner strand	IIA1I	5’ GCG UGU CUG GCA GCC UCC G
IIA1-outer strand	IIA1O	5’ CGG AGG AAC UAC UGU CUU CAC GCC
square inner strand	AAI	5’ CCG GCA GCC U
square outer strand	AAO	5’ CCG GAG GAA CUA CUG
Crystallization (square):		
inner strand	AAI	5’ CCG GCA GCC U
outer strand	AAO5Br	5’ CCG GAG GAA CUA C^{5Br}UG[1]
Fluorescence Experiments:		
inner strand for A^B-cy3	ABI	5’ GGC CCA GCC U
outer strand for A^B-cy3	ABOCy3	5’ cy3-CCG GAG GAA CUA CUG[2]
inner strand for C^D-cy5	CDI	5’ GCG CCA GCC U
outer strand for C^D-cy5	CDOCy5	5’ cy5-CGC GAG GAA CUA CUG[2]
inner strand for C^D-Q	CDI	5’ GCG CCA GCC U
outer strand for C^D-Q	CDOIBRQ	5’ IAbRQ-CGC GAG GAA CUA CUG[3]
inner strand for B^A	BAI	5’ CCG GCA GCC U
outer strand for B^A	BAO	5’ GGC CAG GAA CUA CUG
inner strand for D^C	DCI	5’ CGC GCA GCC U
outer strand for D^C	DCO	5’ GCG CAG GAA CUA CUG
inner strand for B^C	DCI	5’ CGC GCA GCC U
outer strand for B^C	BAO	5’ GGC CAG GAA CUA CUG
inner strand for D^A	BAI	5’ CCG GCA GCC U
outer strand for D^A	DCO	5’ GCG CAG GAA CUA CUG

[1] ^{5Br}U = 5-bromo uridine

[2] cy3, cy5 = fluorescent cyanine dyes attached to the 5’ O atom via a $(CH_2)_4$- linker.

[3] IAbRQ = Iowa Black RQ Dark quencher attached to the 5’ O atom via a $(CH_2)_4$- linker.

Table S2. Crystallographic data collection and refinement statistics for the square RNA.

	Native	Crystal 1
Data collection		
Space group		$P3_121$
Cell dimensions		
a, *b*, *c* (Å)		62.38
		62.38
		126.27
α, β, γ (º)		90
		120
		90
		Peak
Wavelength		0.920275
Resolution (Å)		2.16
R_{sym} or R_{merge}		6.5 (33.0)
$I/\sigma I$		49.5 (2.03)
Completeness (%)		86.7 (30.0)
Redundancy		30.2 (2.0)
Refinement		
Resolution (Å)		2.16
No. reflections		428297
R_{work}/ R_{free}		0.156/0.192
No. atoms		
Protein		2074
Ligand/ion		72
Water		118
B-factors		
Protein		47.9
Ligand/ion		97.1
Water		36.6
R.m.s deviations		
Bond lengths (Å)		0.008
Bond angles (º)		1.11

*Highest resolution shell is shown in parenthesis.

Table S3. Root mean square deviation [Å] within strands of the square RNA and compared to IIa-1 RNA.

	CD	EF	GH	IIa-1
AB	2.27	3.47	1.25	0.83
CD		3.93	2.15	2.27
EF			3.02	3.54
GH				1.26

*Nucleotides 107-113 of the inner strand and 50-55/58-60 of the outer strand were used for superposition.

Table S4. Potential hydrogen bonds and stacking interactions in crystal contacts of the square RNA.

Strand	Residue	Atom	Strand*	Residue	Atom	Distance
A	G52	O2'	E'	U56	O4'	2.80
A	U56	O4	F'	G110	N2	3.06
A	U56	base	F'	C111	sugar	stacking
A	A57	N1	E'	C58	O2'	2.31
A	C58	O2'	E'	G60	O1P	2.72
B	C104	O2'	C'	G52	O2'	2.77
B	G107	O3'	C'	G48	O3'	2.43
B	G107	O2'	C'	G48	O4'	3.00
B	C111	O4'	E'	A57	O2'	2.78
C	G48	O2'	B'	G107	O3'	2.43
C	G48	O4'	B'	G107	O2'	3.00
C	G51	O1P	E'	A54	O2'	3.05
C	A50	sugar	E'	A54	base	stacking
C	G52	O2'	B'	C104	O2'	2.77
C	U56	O4	G'	C58	O2'	2.29
C	U56	O4	G'	C58	O3'	2.57
C	A57	base	G'	C58	base	stacking
D	G110	O2'	H'	C111	O1P	2.99
D	C112	O2'	H'	U113	O2'	2.46
D	U113	O2'	G'	G51	N3	2.53
D	U113	O3'	G'	G51	N2	2.93
E	G48	O2'	H'	G107	O2'	2.42
E	A54	O2'	C'	G51	O1P	3.05
E	A54	base	C'	A50	sugar	stacking
E	U56	O4'	A'	G52	O2'	2.80
E	A57	O2'	B'	C111	O4'	2.78
E	C58	O2'	A'	A57	N1	2.31
E	G60	O1P	A'	C58	O2'	2.72
F	C104	O2'	G'	G51	O2'	2.46
F	C104	O2'	G'	G51	O3'	3.02
F	C108	O2'	G'	G48	O2'	2.57
F	C108	O4'	G'	G48	O2'	2.94
F	G110	N2	A'	U56	O4	3.06
F	G110	sugar	A'	U56	base	stacking
G	G48	O2'	F'	C108	O2'	2.57
G	G51	N3	D'	U113	O2'	2.53
G	G51	N2	D'	U113	O3'	2.93
G	G51	O2'	F'	C104	O2'	2.46
G	C58	O2'	C'	U56	O4	2.29
G	C58	O3'	C'	U56	O4	2.57
G	C58	base	C'	A57	base	stacking
H	G107	O2'	E'	G48	O2'	2.42
H	C111	O1P	D'	G110	O2'	2.99
H	U113	O2'	D'	C112	O2'	2.46

*Primed strand designations indicate a neighboring square RNA.

Table S5. Exposed surface area and B-factors in the crystal structure of the square RNA.

Chain	**A/B**	**C/D**	**E/F**	**G/H**
Residues	25	25	25	23
Exposed Surface Area [Å^2]	4434	4819	4510	4356
Exposed Surface Area / Residue [Å^2]	350	379	354	375
Average B-factor [Å^2]	40	47	41	46

Table S6. Metal ions and their potential ligands in the square RNA crystal structure.

Metal	Chain	Residue	Atom	Distance [Å]
Co1	A	G48	N7	3.63
			O6	3.94
		G49	N7	3.59
			O6	4.20
Co2	D	G106	N7	4.04
			O6	4.20
		G107	N7	4.20
			O6	3.93
Co3	C	A53	N3	4.63
		A54	O2'	4.18
			O4'	3.88
Co4	C	G48	N7	4.14
			O6	4.66
		G49	N7	4.29
			O6	4.37
Co5	F	G106	N7	4.38
			O6	4.55
		G107	O6	4.53
	E	G60	O6	4.34
Co6	E	G48	N7	3.72
		G49	N7	4.62
			O6	4.16
Co7	H	G106	N7	4.27
			O6	3.71
		G107	O6	4.46
Co8	G	G52	N7	4.68
			O6	4.41
Co9	G	G48	N7	4.28
			O6	4.46
		G49	N7	4.40
			O6	4.69
Co10	B	G107	O6	4.35
	A	G60	O6	4.36
Mg1	A	C55	O2	2.37
	B	A109	N7	3.96
		G110	O6	4.02
Mg2	E	A50	N7	3.84
		G51	N7	2.96
			O6	3.08
			O2P	3.59

*Hexamminecobalt(III) ions are listed as "Co", hydrated magnesium ions as "Mg".