Science Advances
AAAS

Supplementary Materials for

Structures of artificially designed discrete RNA nanoarchitectures at near-atomic resolution

Di Liu, Yaming Shao, Joseph A. Piccirilli*, Yossi Weizmann*

*Corresponding author. Email: jpicciri@uchicago.edu (J.A.P.); yweizmann@bgu.ac.il (Y.W.)

Published 22 September 2021, *Sci. Adv.* **7**, eabf4459 (2021)
DOI: 10.1126/sciadv.abf4459

This PDF file includes:

Supplementary Text Section 1: Circularizing the RNA of BRC by enzymatic ligation.

The RNA used for the circularization was transcribed with a 5'-end hammerhead ribozyme and a 3'-end HDV ribozyme to solve the problem of 5'- and 3'-end heterogeneities of the conventional *in vitro* runoff RNA transcripts. Same as the other RNAs in this work, the HiScribeTM T7 High Yield RNA Synthesis Kit (NEB) was used to prepare the RNA by *in vitro* transcription. The DNA template was obtained by PCR amplification of a gBlocks® gene fragment (IDT) using the Q5® Hot Start High-Fidelity DNA Polymerase (NEB). The sequence of the gBlocks® gene fragment is as follows (the primer-binding regions are underlined, the sequences for promoter and both ribozymes are colored grey, and the G in bold is the transcriptional start site):

TTCTAATACGACTCACTATA**G**GGCCGTCGAATATCCCTGATGAGTCCGTGAGGACGAAACGAGCTAG
CTCGTCGGATATTCGACGGAGGCACCCAGGAACTACCGTTGAAGCTCGCACGACGGCCTGGGGTCG
AGTATCCCGGTATTTGTCGCGAGCACTGAGGAACTACTGCTGAAGCCTCCACGGCAGCCTCAGGAC
AAGTACCGGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGG
CGAATGGGAC

At the end of transcription, five thermal cycles were performed to enhance the ribozyme cleavage, with each cycle containing three steps: 70 °C for 10 s, 50 °C for 1min and 37 °C for 10 min. The target RNA was then purified by dPAGE. The resultant RNA at this stage contains 5'-hydroxyl group and 2',3'-cyclic phosphate. To make it a suitable substrate for ligase-mediated circularization, we treated the RNA (at a final concentration of 5 μM) with T4 polynucleotide kinase (NEB) at a final concentration of 0.4 U μl^{-1} in 1 × T4 DNA ligase buffer (NEB, 1 × buffer: 50mM Tris-HCl, 10mM $MgCl_2$, 10 mM DTT, 1mM ATP, pH 7.5 at 25 °C) at 37 °C for 6h. The kinase treatment is to add a phosphate to 5'-hydroxyl end and remove the 2',3'-cyclic phosphate. The RNA was then purified by phenol–chloroform extraction followed by ethanol precipitation.

T4 RNA ligase 2 (NEB) was chosen to circularize the kinase-treated RNA because it is a double-stranded RNA ligase. For each 100 μL of ligation reaction mixture containing 4 μM of kinase-treated linear RNA, 4 μl of T4 RNA ligase 2 (10 U μl^{-1}) and 10 μl of 10 × T4 DNA ligase buffer were added. The ligation was carried out at 37 °C for 30 min and then 16 °C for 3 h. The circularized RNA was purified by preparative dPAGE. The yield of circularized RNA was about 10% as estimated by analytic dPAGE, and we speculated that the low yield may be due to the poor accessibility of kinase and/or ligase caused by the complex fold of the RNA. Because we got sufficient circularized RNA for the subsequent assembly experiments, we did not try to further optimize the circularization yield.

Supplementary Text Section 2: Synthesis of the RNA of DCG with U118 substituted by 5-bromouracil.

The 5' fragment of the RNA was transcribed with a 3'-end HDV to generate the uniform 3'-end for ligation. The synthesis and purification processes were similar to those described in the previous section. The sequence of the gBlocks® gene fragment is as follows (the primer-binding regions are underlined, the sequences for promoter and 3'-ribozyme are colored grey and the G in bold is the transcriptional start site):

TTCTAATACGACTCACTATA**G**GACGCGGTAGCGAAGAACCATCGACTGCGACTTAGCGAGCCTGCC
CTGGCAGTGGGAGCCTGAAGCAGGCACGGGCGAAGAACACTGTCAGGTCGCTGAGTCGGCGTCCC
AGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAATGG
GAC

Three additional chemically synthesized oligonucleotides, including the 5-bromouracil-containing 3' fragment (purchased from Dharmacon), DNA splint strand and DNA block strand (both DNA strands were purchased from IDT), were required for the ligation and their sequence are shown in fig. S3.

After treatment of T4 kinase at 1 × T4 DNA ligase buffer, the 5' RNA fragment was directly mixed with 3' Br-containing RNA fragment, DNA splint and DNA block with a ratio of 1: 5: 2: 1.5 (5' RNA fragment: 3' RNA

fragment: DNA splint strand: DNA block strand). The mixture was then annealed from 70 to 16 °C over 4 h. For a typical ligation, each 100 µL of ligation reaction mixture contained 8 µM of 5' RNA fragment, 40 µM of 3' RNA fragment, 16 µM of DNA splint strand, 12 µM of DNA block strand, 4 µl of T4 RNA ligase 2 (10 U μl^{-1}), and 10 µl of 10 × T4 DNA ligase buffer. The ligation was carried out at 37 °C for 30 min and then 16 °C for 3 h. The ligated full-length RNA was purified by preparative dPAGE. The yield of the ligated RNA was about 45% as estimated by analytic dPAGE.

The statistics for SAD data-collection statistics for the crystal of Br-containing **DCG** are tabulated below.

Crystal	**DCG-Br**
Data collection	
Space group	P1
Wavelength (Å)	0.9195
Resolution (Å)	200-4.50 (4.58-4.50)
Cell dimensions	
a, b, c (Å)	60.121, 63.778, 68.986
α, β, γ (°)	85.582, 64.395, 80.204
$CC_{1/2}$	0.994 (0.922)
CC*	0.998 (0.979)
R_{merge} (%)	7.0 (106.5)
R_{meas} (%)	7.3 (38.5)
R_{pim} (%)	3.3 (18.5)
$I/\sigma I$	98.1 (1.0)
Completeness (%)	97.8 (89.4)
Redundancy	3.5 (2.5)

* Values in the parentheses are for the highest resolution shell.

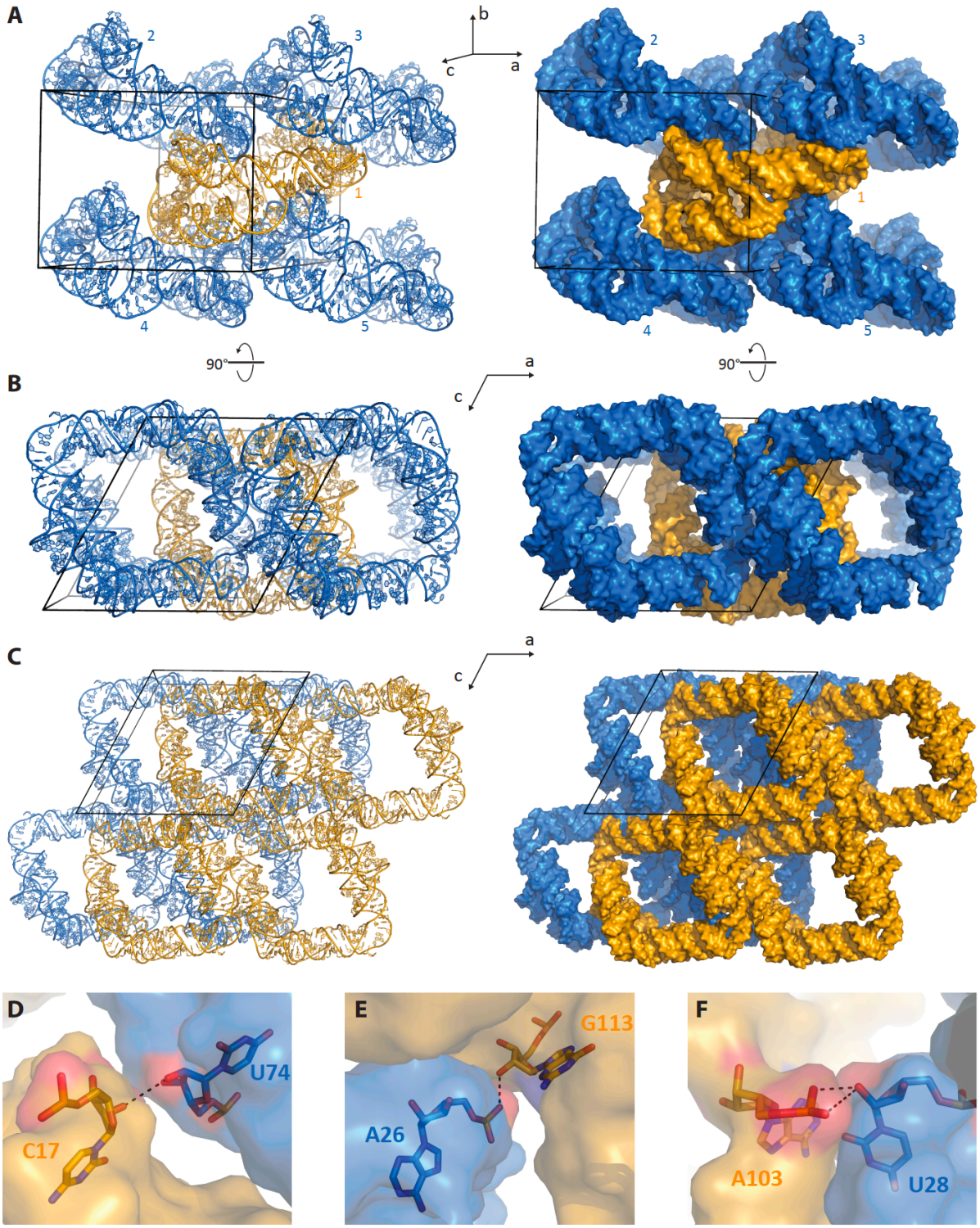

Fig. S1. Packing of RNA nanoparticles in the crystal of BRC via shape complementarity. A, The extension of lattice along the crystallographic axes a and b is mainly mediated by the shape complementarity: the cleft in the middle of a **BRC** (particle 1, colored in orange) is filled by the "ears" (described in the main text) of two neighboring particles (particles 2 and 5, colored in blue) while filling its own "ears" into the clefts of the other two particles (particles 3 and 4, colored in blue). Particle 1 is related to its four neighboring particles with the 2_1 screw-axis symmetry. Cartoon model is shown on the left and surface model on the right. **B**, A 90°-rotated view of the model presented in **A**. This view is approximately along the axis b, showing that this shape complementarity-mediated packing prevents the formation of extended channels through the openings of the bracelets and that neighboring particles' "ears" only partially occupy each cleft, leaving a cavity of cavity of ~29 nm^3 within each bracelet. **C**, The extension of lattice along the axes a and c is mediated by the contacts of RNA helices. **D-F**, Examples of crystal packing contacts mediated by hydrogen bonds involving the phosphate oxygens and 2'-hydroxyl groups.

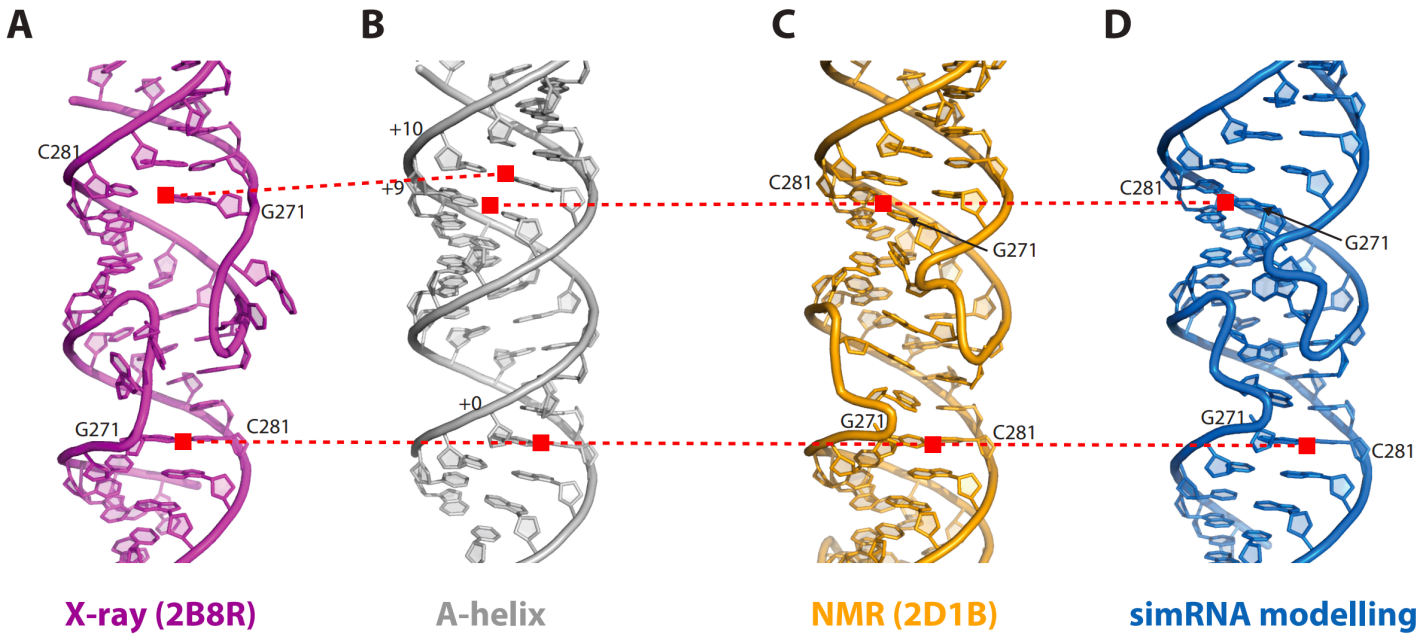

Fig. S2. Comparing different structural models of the HIV-1 DIS KL complex. Three KL complexes are shown: **A,** the 2B8R X-ray structure, **C,** the 2D1B NMR structure, and **D,** the SimRNA model based on the inter-bKL of **BRC**. An ideal A-form RNA helix (**B**) is shown for comparing the helical twist of the structures. The G271:C281 pair at the bottom loop of each KL structure is aligned with the bp+0 of the A-helix. For the 2B8R X-ray model (**A**), G271:C281 at the top loop is approximately aligned with bp +10 of the A-helix, so its KL junction contributes a twist that is equivalent to a 9-bp A-form helix. For the 2D1B NMR model (**C**) and the SimRNA model (**D**), G271:C281 at the top loop of either structure is approximately aligned with bp +9 of the A-helix, so the KL junction of either structure contributes a twist that is equivalent to an 8-bp A-form helix.

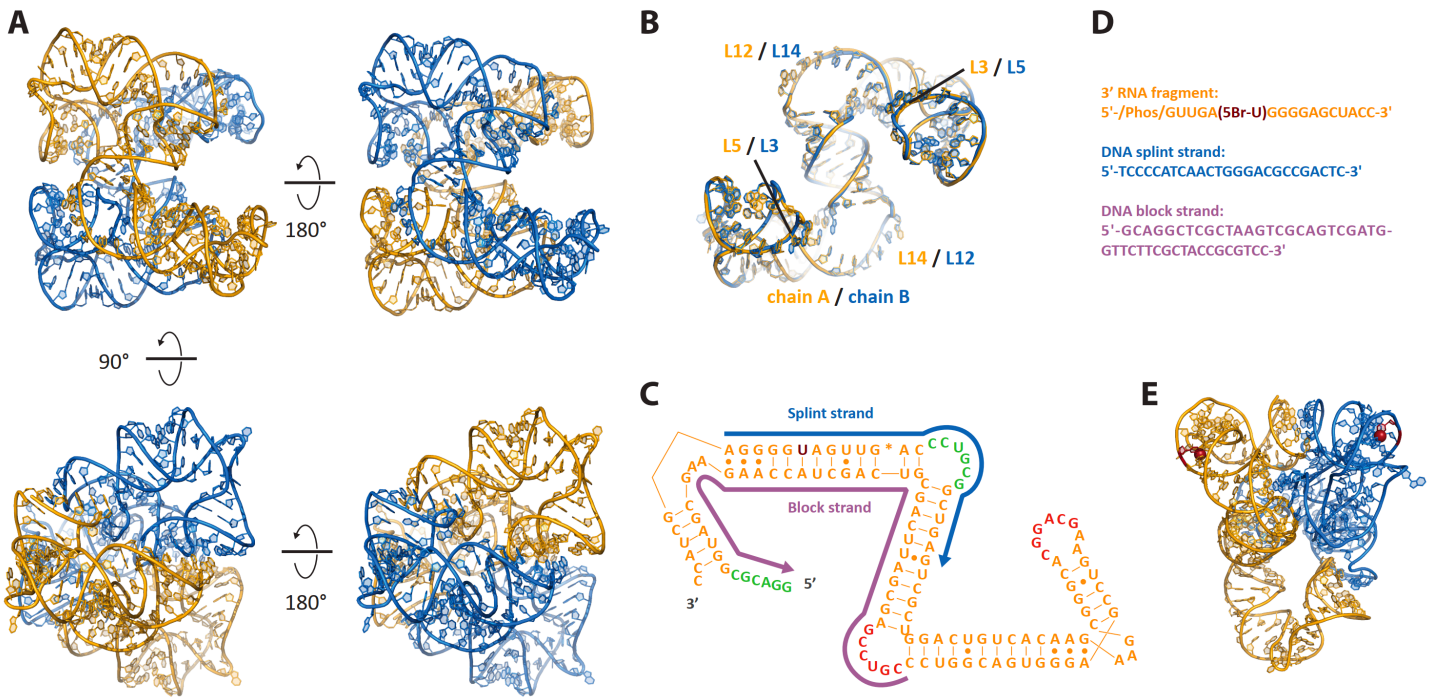

Fig. S3. RNA containing 5-bromouracil is used to unambiguously determine the orientation of the DCG nanoparticle in the lattice. A, The shapes of the top and bottom of **DCG** closely resemble each other so that the **DCG** cage has a 2-fold rotational pseudosymmetry. The rotation axis of this pseudosymmetry passes through the center of the cage and is perpendicular to the crystallographic axes a and b. **B**, A superposition of chain A with chain B of the flipped orientation achieves near-perfect alignment. **C**, The heavy-atom modified RNA was prepared through splinted enzymatic ligation using the *in vitro* transcribed 5' fragment and a chemically synthesized 3' fragment containing the 5-bromouracil. Splinted RNA ligation is conducted in the presence of a DNA splint strand (blue) and a DNA block strand (purple), the latter of which is to inhibit the secondary structure formation of the RNA. The 5' fragment of the RNA is prepared by *in vitro* transcription with an HDV ribozyme at the 3'-end to generate a uniform 3'-end. The 3' fragment of the RNA containing the 5-bromouracil is chemically synthesized. An asterisk (between A112 and G113) indicates the site of ligation. The 5-bromouracil (U118) is claret. **D**, Sequences of the 3' fragment of RNA, the DNA splint strand and the DNA block strand. **E**, The locations of bromine atoms (shown as claret spheres) in the crystal structure of **DCG** as determined by single-wavelength anomalous diffraction (SAD).

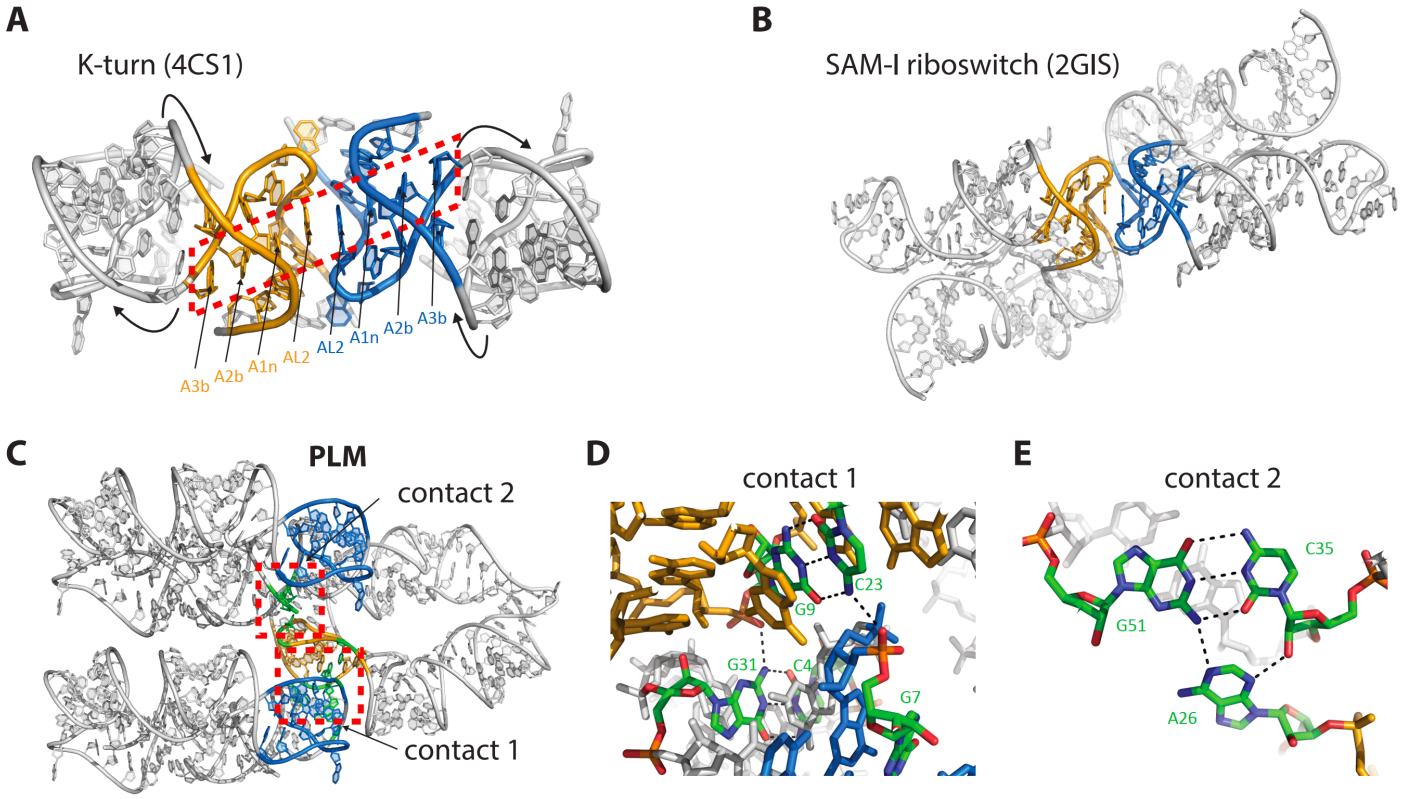

Fig. S4. A recurrent contact of K-turns plays an important role in mediating the crystal packing of some RNAs. A, Packing of RNA molecules in the crystal of K-turn of *H. marismortui* ribosomal Kt-7. For clarity, the two interacting K-turns are colored in orange and blue while the other parts are in gray. The red dashed box highlights the continuous base stacking to connect the NC helices of the two contacting K-turns. **B**, A similar K-turn contact is also in the crystal structure of a SAM-I riboswitch. **C-E**, This recurrent K-turn contact is, however, not observed in the crystal of **PLM**, where two important interactions (highlighted with red dashed boxes in **C** with details shown in **D** and **E**, respectively) mediate the crystal packing. For clarity, the interacting K-turns are colored in orange and blue with the relevant nucleotides involved in inter-particle interactions highlighted in green, while the other parts are in gray. Contact 1 (**D**) involves two phosphate-base pair interactions. Contact 2 (**E**) involves an undocumented A-minor interaction (though bearing some resemblance to the type II A-minor interaction) formed by A26 (at the L3 site of K-turn) and a G:C pair from another **PLM** particle.

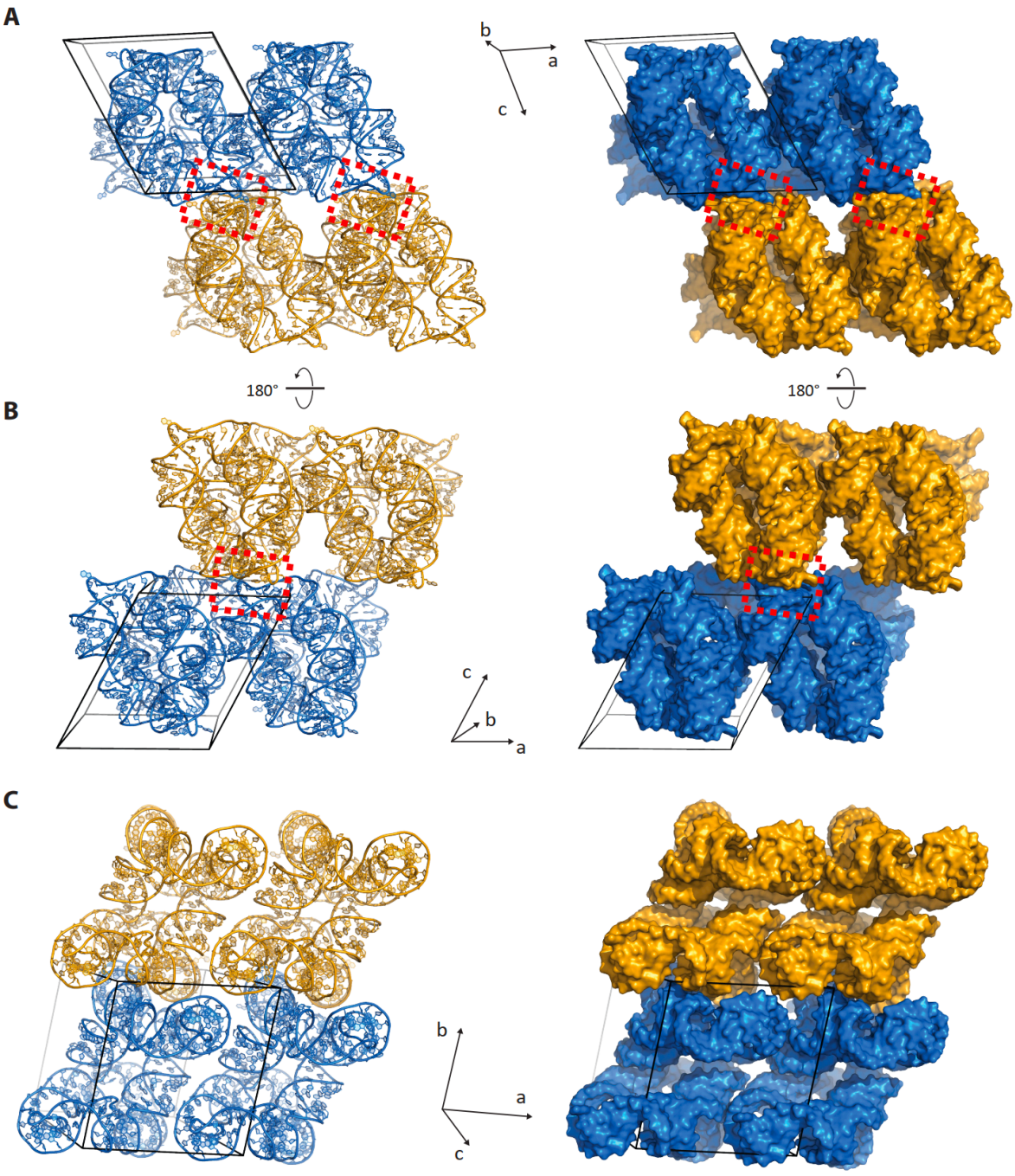

Fig. S5. Packing of RNA nanoparticles in the crystal of DCG. A, The extension of lattice along the crystallographic axes a and c is mainly mediated by the recurrent contacting interface of K-turns. The K-turn contacts are highlighted with red dashed boxes. **B**, A 180°-rotated view of the model presented in **A**. **C**, The extension of lattice along the axis b is mediated by shape complementarity of the corrugated surfaces involving the helices at the top and base of the nanoparticles. Note that the coloring in **A** and **B** is different from that in **C**. In **A** and **B**, the differently colored particles are in the same plane ac, while the particles in different colors are in the different planes ac.

Crystal	PLM	BRC	DCG
PDB code	**7JRR**	**7JRT**	**7JRS**
Data collection			
Space group	$P4_32_12$	$P2_1$	P1
Resolution (Å)	28.16-2.16 (2.24-2.16)	45.53-3.07 (3.18-3.07)	50.00-3.21 (3.27-3.21)
Cell dimensions			
a, b, c (Å)	40.22, 40.22, 202.43	86.44, 71.99, 89.44	60.62, 64.80, 69.58
α, β, γ (°)	90, 90, 90	90, 117.61, 90	85.70, 64.25, 79.91
$CC_{1/2}$	0.998 (0.661)	0.998 (0.976)	0.996 (0.686)
CC*	0.999 (0.892)	0.999 (0.994)	0.999 (0.902)
R_{merge} (%)	5.3 (75.9)	10.0 (51.0)	6.9 (80.4)
R_{meas} (%)	5.7 (85.2)	10.3 (52.8)	8.2 (95.6)
R_{pim} (%)	2.0 (37.8)	2.5 (13.2)	4.2 (51.1)
$I/\sigma I$	25.8 (1.9)	39.2 (5.15)	29.6 (1.6)
Completeness (%)	98.3 (91.0)	99.3 (99.6)	98.2 (95.9)
Redundancy	7.6 (4.6)	16.2 (14.2)	3.7 (3.2)
Refinement			
No. of reflections	9,320	18,129	15,105
R_{work}/R_{free}	0.2261/0.2598	0.1886/0.2268	0.2745/0.2966
r.m.s deviations			
Bond angles (°)	1.237	1.220	0.318
Bond length (Å)	0.005	0.006	0.001
B-factors	57.76	86.65	151.85

* Values in the parentheses are for the highest resolution shell.

Table S1. Data collection and refinement statistics.

Construct	DNA Template Sequence
PLM	TTCTAATACGACTCACTATA**G**GACGGGAGCTGAACCATCCAGCGAAGAACGTCC CGACGGATGGTTCGTCGGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGG CTGGGCAACATGCTTCGGCATGGCGAATGGGAC
BRC	TTCTAATACGACTCACTATA**G**GATATTCGACGGAGGCACCCAGGAACTACCGTTG AAGCTCGCACGACGGCCTGGGGTCGAGTATCCCGGTATTTGTCGCGAGCACTG AGGAACTACTGCTGAAGCCTCCACGGCAGCCTCAGGACAAGTACCG
DCG	TTCTAATACGACTCACTATA**G**GACGCGGTAGCGAAGAACCATCGACTGCGACTTA GCGAGCCTGCCCTGGCAGTGGGAGCCTGAAGCAGGCACGGGCGAAGAACACT GTCAGGTCGCTGAGTCGGCGTCCCAGTTGATGGGGAGCTACC

Table S2. Sequences of gBlocks® gene fragments used for the templates of RNA transcription. Sequences have been colored to match the secondary structures presented in the figures of the main text. In each sequence, a G in bold marks the transcriptional start site. Primer-binding regions for PCR amplification are underlined.