nature portfolio

https://doi.org/10.1038/s41592-022-01455-w

Supplementary information

Sub-3-Å cryo-EM structure of RNA enabled by engineered homomeric self-assembly

In the format provided by the authors and unedited

Supplementary information for

Sub-3 Å cryo-EM structure of RNA enabled by engineered homomeric self-assembly

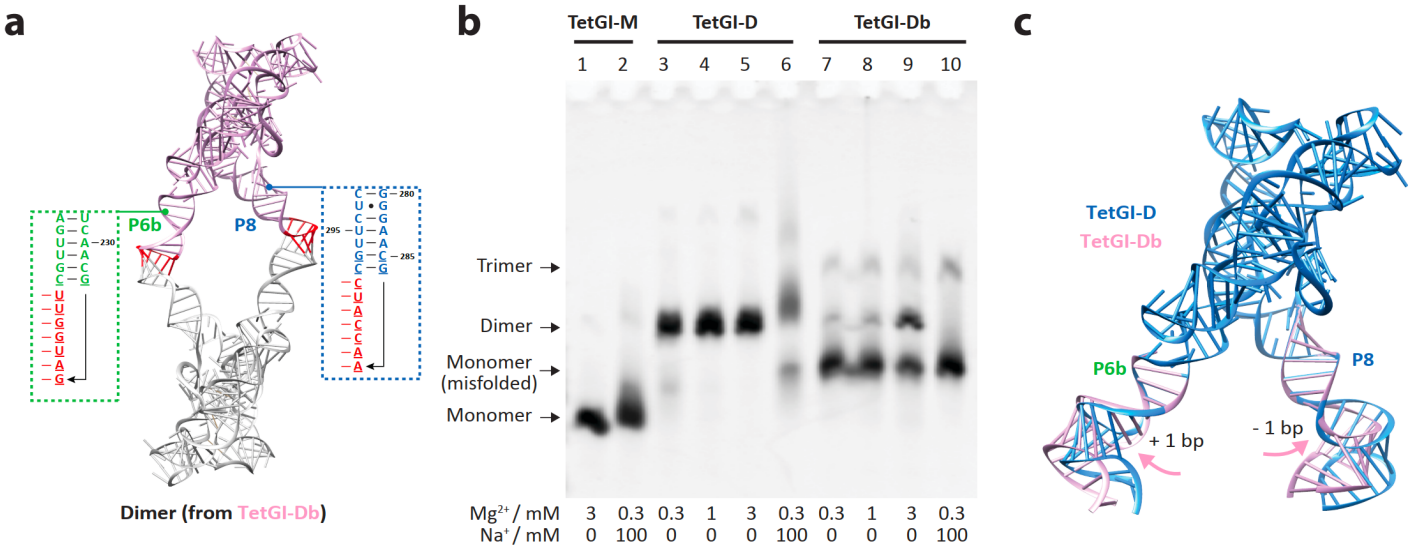
Di Liu[1,2,‡], François A. Thélot[3,‡], Joseph A. Piccirilli[4,5], Maofu Liao[3,*], Peng Yin[1,2,*]
[1]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA, USA
[2]Department of Systems Biology, Harvard Medical School, Boston, MA, USA
[3]Department of Cell Biology, Blavatnik Institute, Harvard Medical School, Boston, MA, USA
[4]Department of Chemistry, the University of Chicago, Chicago, IL, USA
[5]Department of Biochemistry and Molecular Biology, the University of Chicago, Chicago, IL, USA
[‡]These authors contribute equally to this work.
*Correspondence to: maofu_liao@hms.harvard.edu (M. L.), py@hms.harvard.edu (P. Y.)

Table of Contents

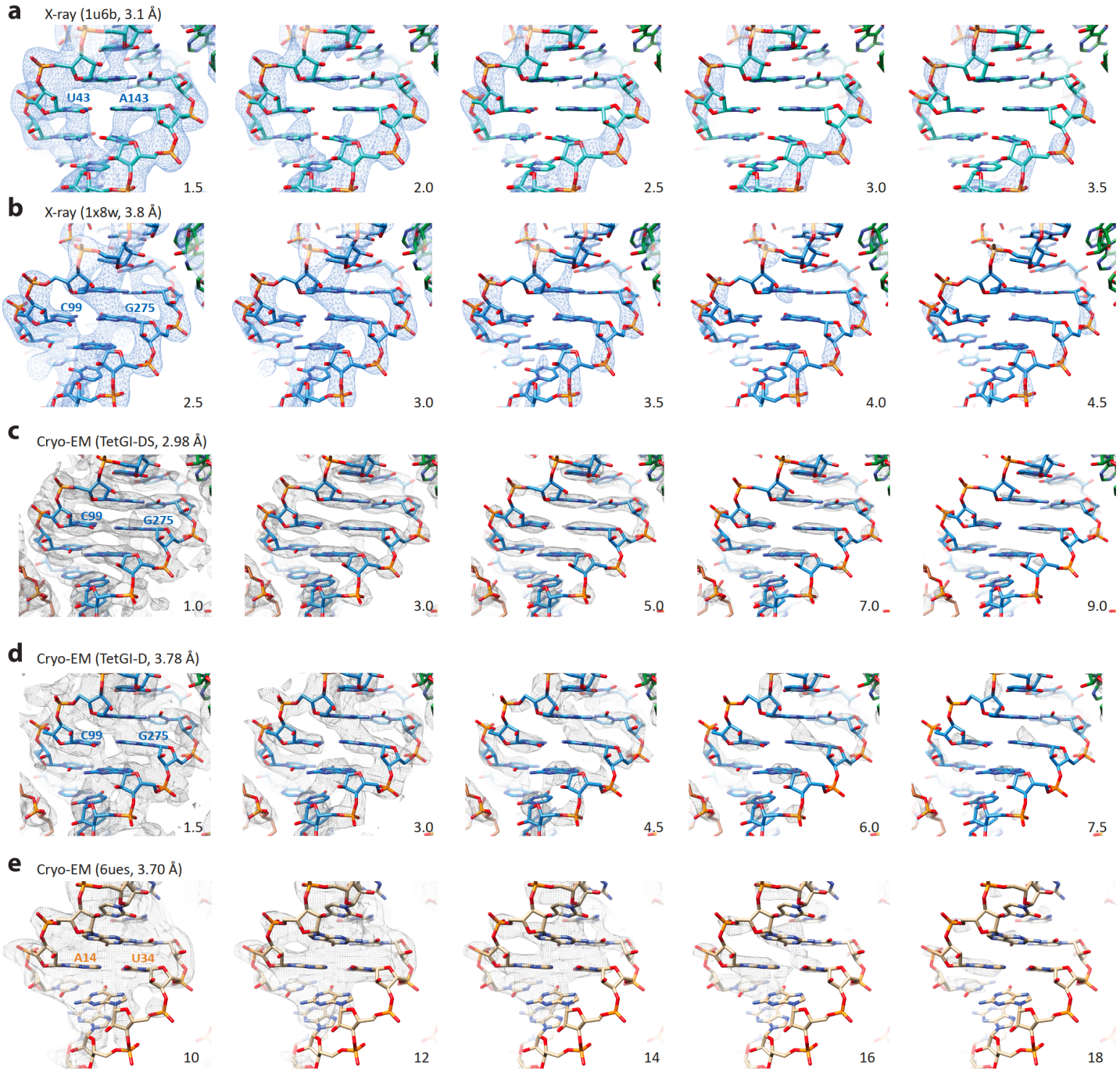
Supplementary Note 1. Choosing the 7-bp kissing-loop motif for the structural studies in this work.
Throughout this work, we employed the 7-bp kissing-loop complex (derived from the RNA I–RNA II complex of the *Escherichia coli ColE1* plasmid products[1]), which can adopt an angle from ~110° to ~130° (~120° for the average structure) based on its NMR structure[2], for mediating the inter-molecular interactions. Kissing loops, instead of sticky ends, are chosen to mediate the assembly because their paranemic characteristic[23-26] (i.e. the two interacting loops are topologically closed and can separate without the need for strand scission) minimizes the strand breaks (and thereby, the number of unique strands) and dispenses with introducing sequence permutation. Though the naturally occurring tetraloop/tetraloop receptor interactions[3] are also paranemic, they have limited specificity and programmability and their introduction is likely to interfere with the folding of a large portion of RNA molecules containing native tetraloop/tetraloop receptor interactions. A 6-bp kissing-loop complex in the HIV-1 dimerization initialization site[4], which has been employed in the crystallographic construct engineering for U1 snRNP[5], were also not chosen because of its structural variability[6] and minimal bending. In comparison, each 7-bp kissing-loop complex contributes a ~60° bending to the total curvature accumulation and therefore the ring closure of the assembly would require a smaller number of subunits than the construct designed with the 6-bp kissing-loop complex. Reducing the number of self-assembling subunits in the homooligomers can help better mitigate structural flexibility and simplify experimental procedures.

Supplementary Note 2. ROCK-enabled cryo-EM is complementary to RNA crystallography.
By demonstrating that ROCK can substantially improve the utility and performance of cryo-EM in RNA structural studies, we believe that ROCK-enabled cryo-EM holds the potential as a complementary approach[7] to the currently more prevalent X-ray crystallography. First, crystallography and cryo-EM have different construct preferences due to their respective technical limitations. For crystallography, smaller RNA constructs (such as small homologs, or the experimentally identified minimal functional fragments or functional core domains) are preferred due to simpler and more compact folding and enhanced propensity for yielding good diffracting crystals. In contrast, larger constructs are better-suited for cryo-EM due to the increased signal-to-noise ratio, higher alignment accuracy and the likely extra stabilization of the structure provided by peripheral domains. As discussed in the main text, the structure of a smaller homolog or a functional fragment solved by crystallography can guide the ROCK engineering for a larger homolog or a complete structure. Second, due to the crystallization requirement, typically only small conformational motions are observed by crystallography from different sample preparations. In contrast, cryo-EM is more suitable for resolving different conformational states even from a single sample preparation. This capability is boosted by ROCK because a finer 3D classification of different functional conformations is enabled via mitigation of the RNA's nonfunctional structural flexibility that would otherwise interfere with 3D classification. Third, the maps derived from the X-ray crystallography and cryo-EM can complement each other to deliver a more comprehensive and detailed structural understanding. Whereas X-ray crystallography yields electron density maps, the cryo-EM-derived maps are electrostatic potential maps as a result of electron scattering[8]. For RNA, negatively-charged phosphates reside in the backbone, which are mostly observed as the strongest intensities in X-ray maps, but appear less prominent in cryo-EM maps (Supplementary Figure 2), especially at regions of lower local resolutions, where the electron scattering factors of negatively-charged groups disappear or even turn negative. On the other hand, the intensities for positively-charged metal ions appear more prominent in cryo-EM maps and the resolution required for locating small native metal ions (such as Mg^{2+}) for cryo-EM maps (several Mg^{2+} ions can be localized even in the 4.17-Å map of the TetGI-T construct) is substantially lower than X-ray maps. Further, the local resolution distribution for cryo-EM maps is typically nonuniform, with local resolutions descending from the core to the peripheries. The less resolved peripheral elements may be isolated and studied by crystallography, and the low-resolution structural information from cryo-EM can help guide crystallographic construct engineering[6] and phase X-ray data[9].

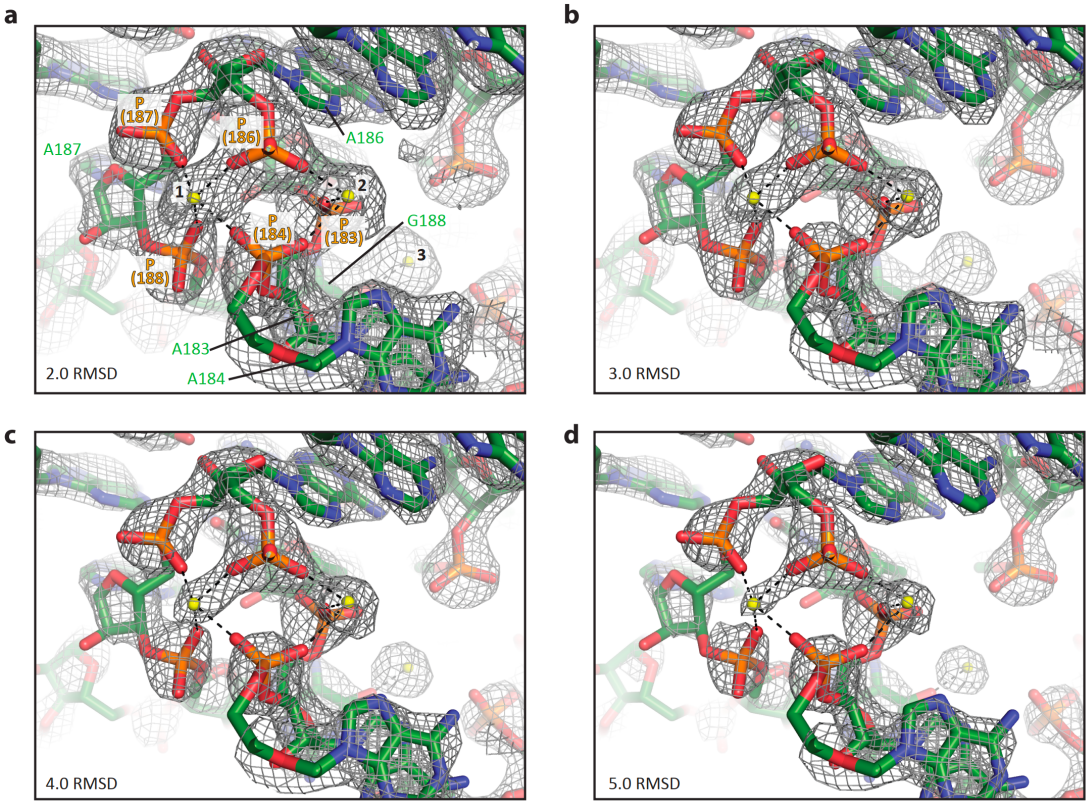
The principle of ROCK may also be applied to protein-containing systems: one exciting possibility is the direct engineering of small monomeric proteins for self-assembly to multiple the molecular weight and reduce structural flexibility, which differs from a recently devised approach of attaching small proteins to large homomeric scaffolds[10,11] aiming to increase the molecular weight. This would facilitate the cryo-EM determinations of small or flexible proteins or protein-nucleic acid complexes.



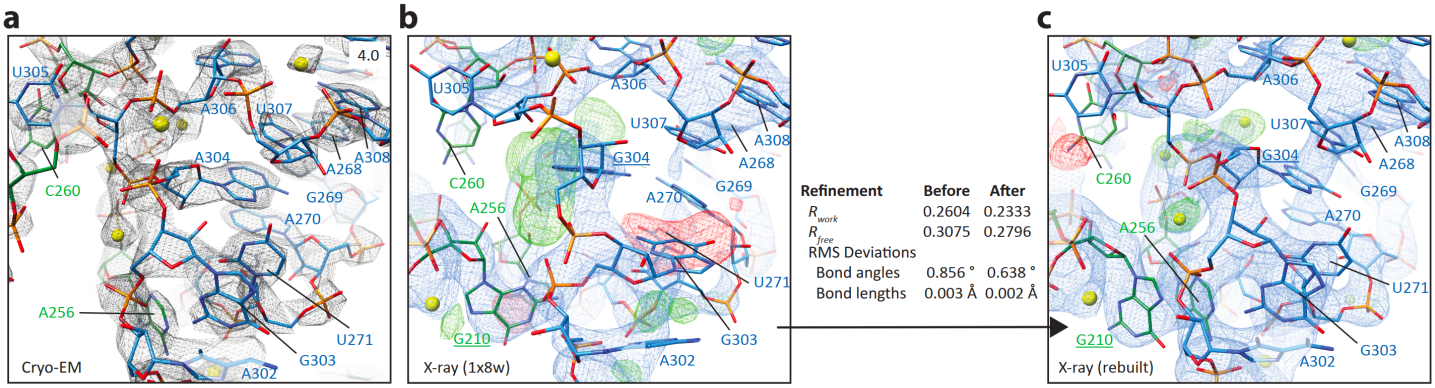
Supplementary Figure 1 | Design and assembly of a second dimeric post-2S construct of the TetGI, TetGI-Db, that fails to assembled into the desired dimer as the predominant product. a, Computer model of the assembled homodimer of TetGI-Db. Dashed boxes mark the engineering of P6b and P8a (see Fig. 2a for the secondary structure of the monomeric construct). **b**, Assembly assay of the dimeric constructs TetGI-D (lanes 3 to 6) and TetGI-Db (lanes 7 to 10), and the monomer control TetGI -M (lanes 1 and 2). Unlike TetGI-D, the desired dimer is not the dominant product under the varying assayed conditions and a faster-migrating band putatively corresponding to a misfolded monomeric species due to the undesired formation of the intra-molecular kissing loops emerges. **c**, Compared to TetGI-D, TetGI-Db has one more base pair in P6b and one less base pair in P8. These subtle differences may position the two kissing-loop regions more preferably to form the undesired intra-molecular kissing interaction instead of the desired inter-molecular interactions in the expected dimeric assembly. A potential way to avoid the interference of RNA folding caused by the presumed formation of the undesired intra-molecular kissing loops is to use a heterodimeric system. We did not experimentally explore this possibility in this work.



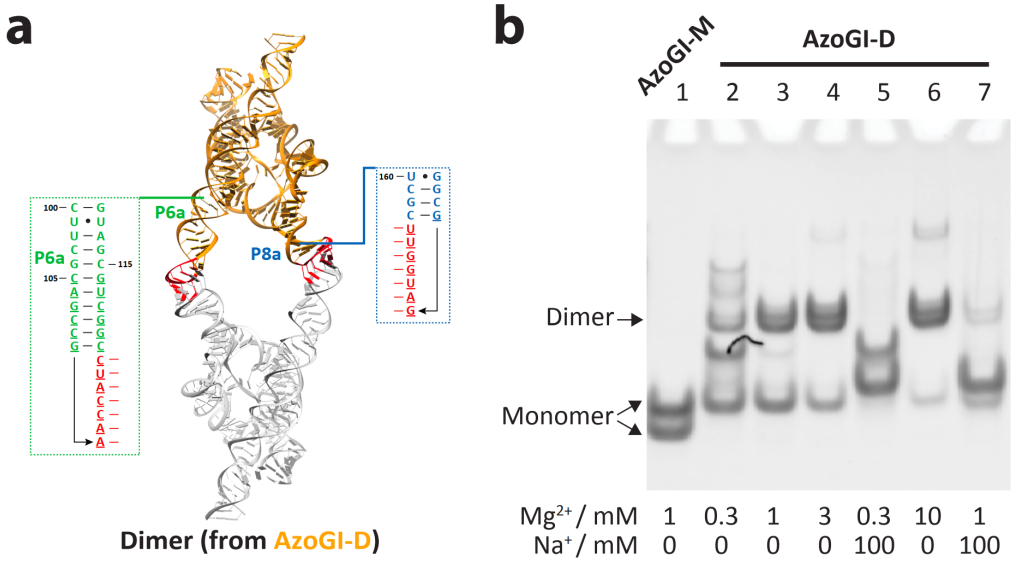
Supplementary Figure 2 | Comparing features of X-ray and cryo-EM maps at different resolutions. a, b, Structural models and X-ray maps (blue meshes) of P3 helices of the AzoGI (**a**) and the TetGI (**b**). The maps are generated from structure factors deposited at the PDB (PDB codes and resolutions are given in parentheses). **c, d**, Structural models and EM maps (grey meshes) of P3 helices of TetGI-DS (**c**) and TetGI-D (**d**). **e**, Structural model and EM map (grey meshes) of the P2a helix of the SAM-IV riboswitch in the apo state[12] (PDB code: 6ues) are shown for comparison. The contour levels (σ for X-ray maps and RMSD for cryo-EM maps) are indicated at the bottom right corner. Unlike the X-ray maps, phosphate groups are not necessarily the most prominent intensity for medium- or low-resolution cryo-EM maps of RNA due to resolution-dependent signal loss for phosphate groups in cryo-EM maps[8] (comparing **a, b** with **c, d**; also see Supplementary Note 2). This underscores the difficulty for interpreting medium- or low-resolution cryo-EM maps of RNA because the current RNA model building tools, which are mostly developed for X-ray crystallography, rely on the localization of phosphate groups as the first step to build each nucleotide. Thus, it would be more difficult to model the regions of an even lower local resolution and/or corresponding to nonhelical structures. The quality of the cryo-EM maps can be assessed on the basis that the intensities for individual bases can be well separated without breaking the backbone intensity continuity at a wide range of contour levels (comparing **c, d** and **e**).



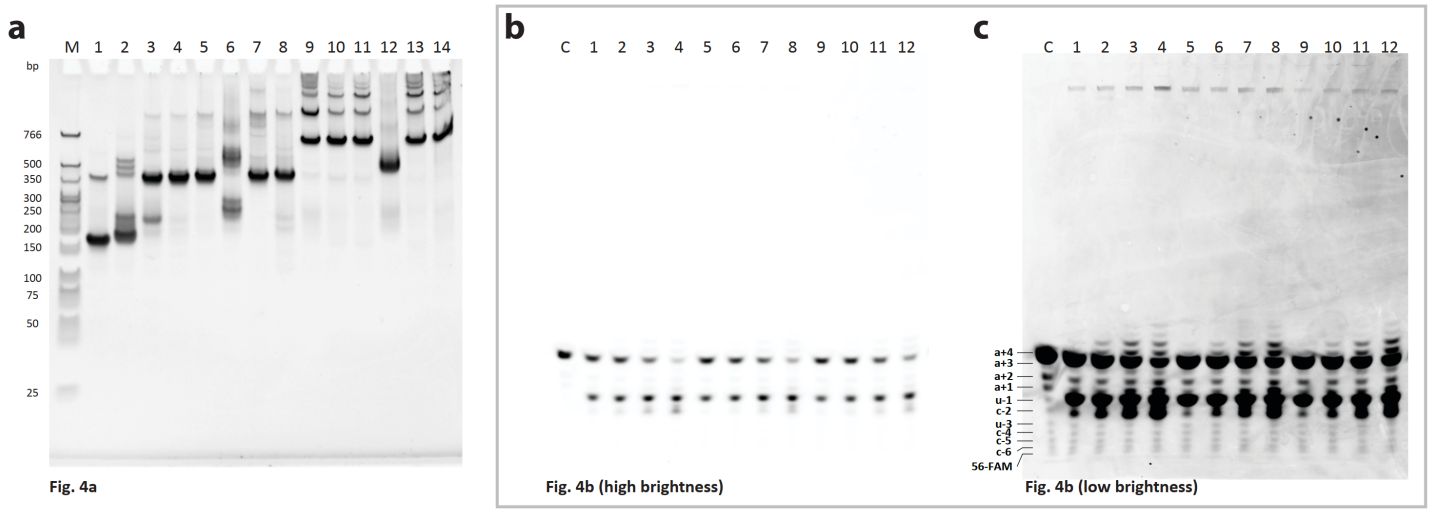
Supplementary Figure 3 | Three Mg^{2+} ions at the A-rich bulge of P4-P6 domain. The cryo-EM maps of TetGI-DS near the A-rich bulge are shown with different contour levels (indicated at the bottom-left corner of each panel of **a** to **d**). Consistent with previous studies[13-15], all the three Mg^{2+} (yellow spheres) in this region are observed in our map: #1, coordinating to the phosphate oxygens of A184, A186, A187 and G188; #2, coordinating to the phosphate oxygens of A183, A184 and A186; #3, coordinating to O6 of G188. We note that the map intensity of a Mg^{2+} ion, among other variable factors, is dependent on the number of inner-sphere coordination interactions it forms with RNA: a higher intensity is observed for a Mg^{2+} ion of a smaller number of inner-sphere coordination interactions with RNA. We attribute this to the fact that the inner-sphere coordinated water molecules, which cannot be resolved from the Mg^{2+} ions at this resolution, likely contribute to the map intensity. As is shown here, the intensity of Mg^{2+} ion #3 (with four inner-sphere ligand atoms from RNA and therefore two water molecules as inner-sphere ligands) is stronger than that of Mg^{2+} ion #2 (with three inner-sphere ligand atoms from RNA and therefore three water molecules as inner-sphere ligands), and the intensity of Mg^{2+} ion #2 is stronger than that of Mg^{2+} ion #1 (with only one inner-sphere ligand atom from RNA and therefore five water molecules as inner-sphere ligands). This feature of cryo-EM map may help determine the coordination environment of Mg^{2+} ions in RNA structures.



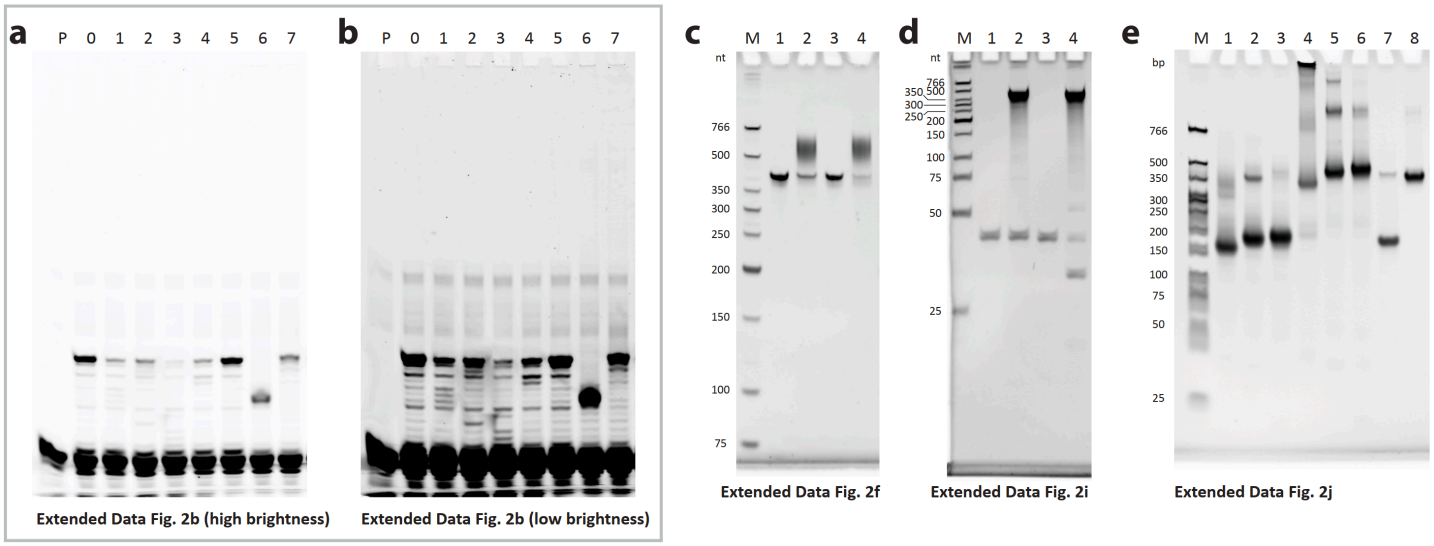
Supplementary Figure 4 | Rebuilding the crystal structure of the TetGI core using the cryo-EM model. The high-resolution cryo-EM structure of TetGI-DS (**a**) of the area around J8/7 enables the rebuilding and re-refinement of the corresponding area of the previous crystal structure of the TetGI core[16] (**b** and **c** show the maps and models before and after rebuilding, respectively; $2F_o–F_c$ map shown in blue mesh contoured at 2.5σ level, and $F_o–F_c$ map shown in green/red meshes contoured at 4σ level). Underlined nucleotides indicate the mutations introduced into the construct for crystallographic study. The crystallographic maps were generated using the factors deposited with the PDB (PDB code: 1x8w). There are four RNA molecules in the asymmetric unit of the 1x8w crystal and the shown region is from molecule B. To eliminate the potential bias of using a different refinement software, the initial maps and model before rebuilding were also re-refined using phenix.refine and the statistics also improved compared to the data originally deposited with the PDB. The improved refinement statistics may also be attributed to improvement through the RNA chains besides J8/7 region.



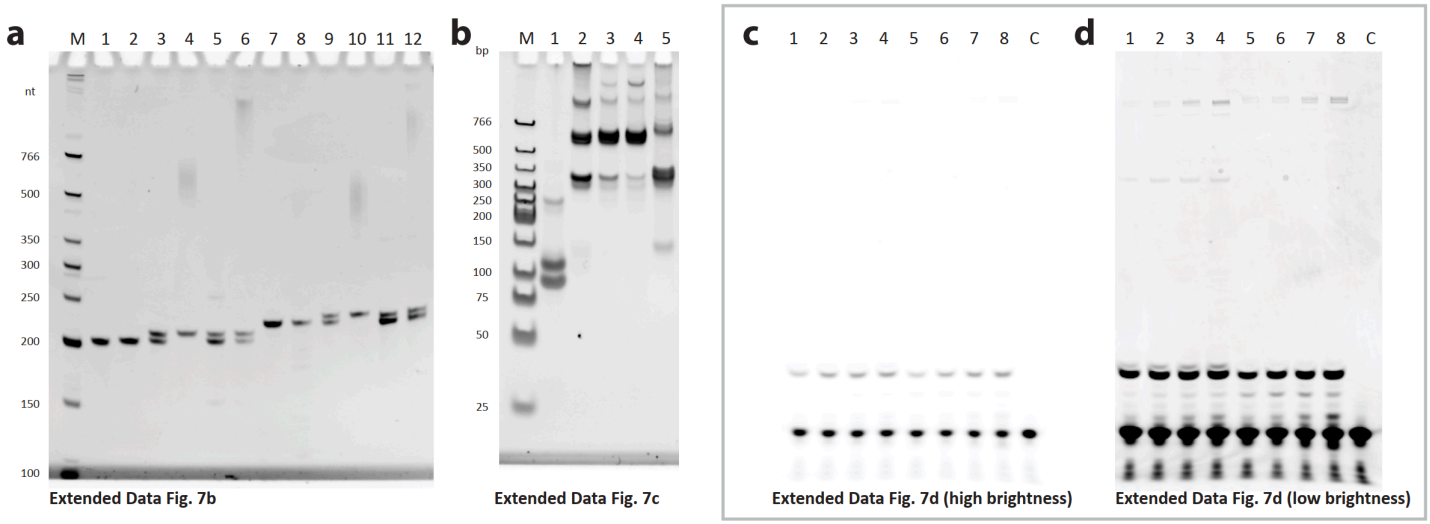
Supplementary Figure 5 | Design and assembly of a dimeric construct of the AzoGI, AzoGI-D. a, Computer model of the assembled homodimer of AzoGI-D. Dashed boxes mark the engineering of P6a and P8a (see Fig. 6a for the secondary structure of the control monomeric construct). **b**, Assembly assay of the dimeric construct AzoGI-D (lanes 2 to 7) and the monomer control AzoGI-M (lane 1). Similar to the TetGI constructs, the optimal condition for folding/assembly was determined to be 3 mM Mg^{2+} (lane 4). This dimeric construct was not further studied by cryo-EM.



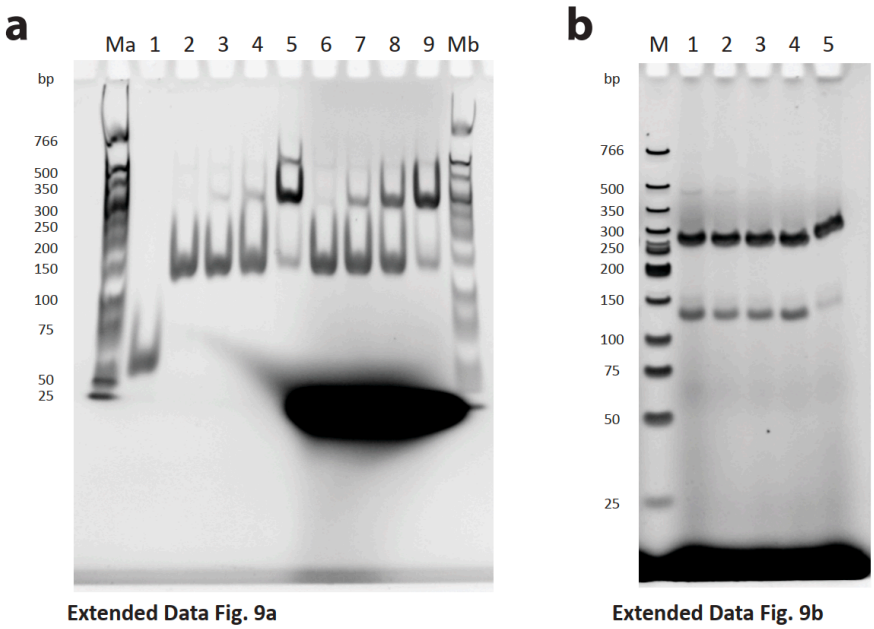
Supplementary Figure 6 | **Raw and uncropped gel images in Fig. 4.** In **a**, lane M, which is not shown in Fig. 4a, contains the low-molecular weight DNA ladder (NEB), and these dsDNA markers can only correctly indicate the sizes of linear species but not the folded RNA species or circular, assembled RNA structures. Further, the markers sometimes migrate as wide and overlapping smears in the gel under the native running conditions. **b** and **c** are the same gel image shown with different levels of brightness. Very faint bands can be visualized in **c**, and they are helpful for assigning the single nucleotides (shown on the left of the gel in **c**; numbered from the cleavage site) in the 56-FAM-labelled substrate.



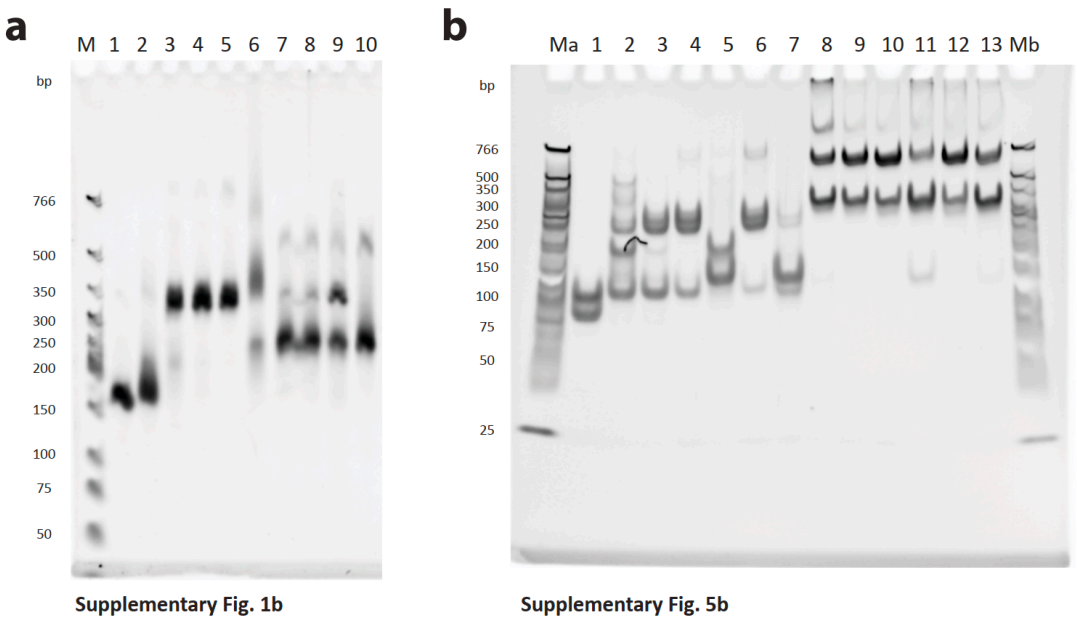
Supplementary Figure 7 | **Raw and uncropped gel images in Extended Data Fig. 2. a** and **b** are the same gel image shown with different levels of brightness. In **a** and **b**, lanes P and 0 are not shown in Extended Data Fig. 2b: lane P contains the RT primer; lane 0 contains the same sample as lane 5. Lane M in **e** is not shown in Extended Data Fig. 2j.



Supplementary Figure 8 | Raw and uncropped gel images in Extended Data Fig. 7. Lane M in **b** is not shown in Extended Data Fig. 7c. **c** and **d** are the same gel image shown with different levels of brightness.



Supplementary Figure 9 | Raw and uncropped gel images in Extended Data Fig. 9. In **a**, lanes Ma and Mb are not shown in Extended Data Fig. 9a and contain different amounts of the low-molecular weight DNA ladder with the loading in lane Mb being the half of lane Ma to facilitate the localization of the markers of different sizes. Lane M in **b** is not shown in Extended Data Fig. 9b.



Supplementary Figure 10 | Raw and uncropped gel images in Supplementary Figures. In **a**, lane M, containing the low-molecular weight DNA ladder, is not shown in Supplementary Fig. 1b. In **b**, lanes Ma and Mb are not shown in Supplementary Fig. 5b and contain different amounts of the low-molecular weight DNA ladder with the loading in lane Mb being the half of lane Ma to facilitate the localization of the markers of different sizes. Lanes 8 to 13, not shown in Supplementary Fig. 5b, contain the assembled products of AzoGI-T annealed in the buffer conditions corresponding to lanes 2 to 7, respectively.

PDB ID	Construct description	Sub-group	Method	Resolution	Active-site metal ions	Reference (Year, Journal)
1gid	Tet P4-P6	IC1	X-ray	2.50 Å	n.a.	1996, Science
1hr2	Tet P4-P6, ΔC209 mutant	IC1	X-ray	2.25 Å	n.a.	2001, Structure
2R8S	Tet P4-P6, ΔC209 mutant, Fab-facilitated crystallization	IC1	X-ray	1.95 Å	n.a.	2007, PNAS
1k2g	Tet P7-P9.0 (G-binding site mimic)	IC1	NMR	n.a.	n.a.	2002, RNA
1grz	Tet P3-P9 & P4-P6 (ribozyme core, or Tet3-9)	IC1	X-ray	5.0 Å	n.a.	1998, Science
1x8w	Tet P3-P9 & P4-P6 (ribozyme core, or Tet3-9), 5 mutation sites	IC1	X-ray	3.8 Å	1 (M2)	2004, Mol. Cell
6wls	Tet ribozyme, truncated at U409, no exon bound	IC1	cryo-EM (Ribosolve)	6.8 Å	n.a.	2020, Nat. Meth.
7r6m	**Tet post-2SΔP10, dimer with P6b & P8 mutated, 2 deoxy mutations at -1 & +1 residues of ligated exon**	**IC1**	**cryo-EM (ROCK)**	**3.78 (3.68) Å**	**1 (M2)**	**this work**
7r6n	**Tet ribozyme, trimer with P6b & P9.2 mutated, no exon bound**	**IC1**	**cryo-EM (ROCK)**	**4.17 (4.10) Å**	**n.a.**	**this work**
7r6l	**Tet pre-2SΔ5'ex, dimer with P6b & P8 mutated, 2 deoxy mutations near scissile phosphate**	**IC1**	**cryo-EM (ROCK)**	**2.98 (2.85) Å**	**1 (M1)**	**this work**
1u6b	Azo Pre-2S, P6a mutated for U1A-facilitated crystallization, 4 deoxy mutations at exon & near scissile phosphate	IC3	X-ray	3.10 Å	2 (monovalent M2)	2004, Nature
1zzn	Azo Pre-2S, P6a mutated for U1A-facilitated crystallization, 1 deoxy mutation at -1 residue of 5'-exon	IC3	X-ray	3.37 Å	2	2005, Science
3bo2	Azo omit-P, P6a mutated for U1A-facilitated crystallization, the scissile phosphate removed	IC3	X-ray	3.31 Å	2	2008, PNAS
3bo4	Azo post-2S, P6a mutated for U1A-facilitated crystallization, 1 deoxy mutation at -1 residue ligated exon	IC3	X-ray	3.33 Å	2	2008, PNAS
3bo3	Azo pre/post-2S, P6a mutated for U1A-facilitated crystallization, a mixture of two states	IC3	X-ray	3.4 Å	2	2008, PNAS
n.a.	**Azo post-2S, trimer with P5 & P8 mutated, 2 deoxy mutations at -1 & +1 residues of ligated exon**	**IC3**	**cryo-EM (ROCK)**	**4.9 Å**	**n.a.**	**this work**
1y0q	Two ribozyme, mutations at P5a loop for crystallization, 5'-exon bound	IA2	X-ray	3.6 Å	1 (M2)	2005, Nat. SMB
2rkj	Two ribozyme, 5'-exon bound, co-crystal with CYT-18 protein	IA2	X-ray	4.5 Å	n.a.	2008, Nature
4p8z	DirLC wild-type, a mixture of two states	GIR1	X-ray	3.85 Å	n.a.	2014, PNAS
6gyv	DirLC circularly permutated	GIR1	X-ray	2.5 Å	1 (different reaction)	2014, PNAS

Supplementary Table 1 | Different group I intron structures or fragments that have been determined. The structures in this work are highlighted in bold. The resolution shown in parentheses are for the TetGI core in this work.

Sequence	Explanation
GGTTCTAATACGACTCACTATA**G**GACCTTTGGAGGGAAAAGTTATCAGGCATGCACCTGGTAGCTAGTCTTTAAACCAATAGATTGCATCGGTTTAAAAGGCAA GACCGTCAAATTGCGGGAAAGGGGTCAACAGCCGTTCAGTACCAAGTCTCAGGGGAAACTTTGAGATGGCCTTGCAAAGGGTATGGTAATAAGCTGACGGACA TGGTCCTAACCACGCAGCCAAGTCCTAAGTCAACAGATCTTCTGTTGATATGGATGCAGTTCACAGACTAAATGTCGGTCGGGGAAGATGTATTCTTCTCATAAG ATATAGTCGGACCTCTCCTTAATGGGAGCTAGCGGATGAAGTGATGCAACACTGGAGCCGCTGGGAACTAATTTGTATGCGAAAGTATATTGATTAGTTTTGGA GTACTC**G***AAGGGCGTCGTCGCCCCGAGCGGTAGTAAGCAGGGAACTCACCTCCAATTTCAGTACTGAAATTGTCGTAGCAGTTGACTACTGTTATGTGATTGGTA GAGGCTAAGTGACGGTATTGGCGTAAGTCAGTATTGCAGCACAGCA*CAAGCCCGCTTGCGAGAAT	Template for TetGI-M
GGTTCTAATACGACTCACTATA**G**GACCTTTGGAGGGAAAAGTTATCAGGCATGCACCTGGTAGCTAGTCTTTAAACCAATAGATTGCATCGGTTTAAAAGGCAA GACCGTCAAATTGCGGGAAAGGGGTCAACAGCCGTTCAGTACCAAGTCTCAGGGGAAACTTTGAGATGGCCTTGCAAAGGGTATGGTAATAAGCTGACGGACA TGGTCCTAACCACGCAGCCAAGTCCTAAGTCAAGGATGGTTCTTGATATGGATGCAGTTCACAGACTAAATGTCGGTCGGGGAAGAGAACCATCCTCTTCTCATA AGATATAGTCGGACCTCTCCTTAATGGGAGCTAGCGGATGAAGTGATGCAACACTGGAGCCGCTGGGAACTAATTTGTATGCGAAAGTATATTGATTAGTTTTG GAGTACTC**G***AAGGGCGTCGTCGCCCCGAGCGGTAGTAAGCAGGGAACTCACCTCCAATTTCAGTACTGAAATTGTCGTAGCAGTTGACTACTGTTATGTGATTGG TAGAGGCTAAGTGACGGTATTGGCGTAAGTCAGTATTGCAGCACAGCA*CAAGCCCGCTTGCGAGAAT*	Template for TetGI-D
GGTTCTAATACGACTCACTATA**G**GACCTTTGGAGGGAAAAGTTATCAGGCATGCACCTGGTAGCTAGTCTTTAAACCAATAGATTGCATCGGTTTAAAAGGCAA GACCGTCAAATTGCGGGAAAGGGGTCAACAGCCGTTCAGTACCAAGTCTCAGGGGAAACTTTGAGATGGCCTTGCAAAGGGTATGGTAATAAGCTGACGGACA TGGTCCTAACCACGCAGCCAAGTCCTAAGTCAACAGGATGGTTCTGTTGATATGGATGCAGTTCACAGACTAAATGTCGGTCGGGGAAGATGTATTCTTCTCATA AGATATAGTCGGACCTCTCCTTAATGGGAGCTAGCGGATGAAGTGATGCAACACTGGAGCCGCTGGGAACTAATCAGTGAACCATCCACTGATTAGTTTTGGAG TACTC**G***AAGGGCGTCGTCGCCCCGAGCGGTAGTAAGCAGGGAACTCACCTCCAATTTCAGTACTGAAATTGTCGTAGCAGTTGACTACTGTTATGTGATTGGTAG AGGCTAAGTGACGGTATTGGCGTAAGTCAGTATTGCAGCACAGCA*CAAGCCCGCTTGCGAGAAT*	Template for TetGI-T
GGTTCTAATACGACTCACTATA**G**GCCGGGTGGAGGGAAAAGTTATCAGGCATGCACCTGGTAGCTAGTCTTTAAACCAATAGATTGCATCGGTTTAAAAGGCAA GACCGTCAAATTGCGGGAAAGGGGTCAACAGCCGTTCAGTACCAAGTCTCAGGGGAAACTTTGAGATGGCCTTGCAAAGGGTATGGTAATAAGCTGACGGACA TGGTCCTAACCACGCAGCCAAGTCCTAAGTCAACAGATCTTCTGTTGATATGGATGCAGTTCACAGACTAAATGTCGGTCGGGGAAGATGTATTCTTCTCATAAG ATATAGTCGGACCTCTCCTTAATGGGAGCTAGCGGATGAAGTGATGCAACACTGGAGCCGCTGGGAACTAATTTGTATGCGAAAGTATATTGATTAGTTTTGGA GTACTC**G***AAGGGCGTCGTCGCCCCGAGCGGTAGTAAGCAGGGAACTCACCTCCAATTTCAGTACTGAAATTGTCGTAGCAGTTGACTACTGTTATGTGATTGGTA GAGGCTAAGTGACGGTATTGGCGTAAGTCAGTATTGCAGCACAGCA*CAAGCCCGCTTGCGAGAAT*	Template for TetGI G14-G414
GGTTCTAATACGACTCACTATA**G**GCCGGGTGGAGGGAAAAGTTATCAGGCATGCACCTGGTAGCTAGTCTTTAAACCAATAGATTGCATCGGTTTAAAAGGCAA GACCGTCAAATTGCGGGAAAGGGGTCAACAGCCGTTCAGTACCAAGTCTCAGGGGAAACTTTGAGATGGCCTTGCAAAGGGTATGGTAATAAGCTGACGGACA TGGTCCTAACCACGCAGCCAAGTCCTAAGTCAACAGATCTTCTGTTGATATGGATGCAGTTCACAGACTAAATGTCGGTCGGGGAAGATGTATTCTTCTCATAAG ATATAGTCGGACCTCTCCTTAATGGGAGCTAGCGGATGAAGTGATGCAACACTGGAGCCGCTGGGAACTAATTTGTATGCGAAAGTATATTGATTAGTTTTGGA GTACTCGACCCGGCC**A***GGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGC*TTCGGCATGGCGAATGGGAC*	Template for TetGI G14-a+9
GGTTCTAATACGACTCACTATA**G**GCCGGGTGGAGGGAAAAGTTATCAGGCATGCACCTGGTAGCTAGTCTTTAAACCAATAGATTGCATCGGTTTAAAAGGCAA GACCGTCAAATTGCGGGAAAGGGGTCAACAGCCGTTCAGTACCAAGTCTCAGGGGAAACTTTGAGATGGCCTTGCAAAGGGTATGGTAATAAGCTGACGGACA TGGTCCTAACCACGCAGCCAAGTCCTAAGTCAACAGATCTTCTGTTGATATGGATGCAGTTCACAGACTAAATGTCGGTCGGGGAAGATGTATTCTTCTCATAAG ATATAGTCGGACCTCTCCTTAATGGGAGCTAGCGGATGAAGTGATGCAACACTGGAGCCGCTGGGAACTAATTTGTATGCGA**A***GGCCGGCATGGTCCCAGCCTC CTCGCTGGCGCCGGCTGGGCAACATGC*TTCGGCATGGCGAATGGGAC*	Template for TetGI monomer G14-A386
GGTTCTAATACGACTCACTATA**G**GCCGGGTGGAGGGAAAAGTTATCAGGCATGCACCTGGTAGCTAGTCTTTAAACCAATAGATTGCATCGGTTTAAAAGGCAA GACCGTCAAATTGCGGGAAAGGGGTCAACAGCCGTTCAGTACCAAGTCTCAGGGGAAACTTTGAGATGGCCTTGCAAAGGGTATGGTAATAAGCTGACGGACA TGGTCCTAACCACGCAGCCAAGTCCTAAGTCAAGGATGGTTCTTGATATGGATGCAGTTCACAGACTAAATGTCGGTCGGGGAAGAGAACCATCCTCTTCTCATA AGATATAGTCGGACCTCTCCTTAATGGGAGCTAGCGGATGAAGTGATGCAACACTGGAGCCGCTGGGAACTAATTTGTATGCGA**A***GGCCGGCATGGTCCCAGCC CTCCTCGCTGGCGCCGGCTGGGCAACATGC*TTCGGCATGGCGAATGGGAC*	Template for TetGI dimer G14-A386 (5' IVT RNA of TetGI-DS)
GGTTCTAATACGACTCACTATA*GGGCATCAATATACTCTGATGAGTCCGTGAGGACGAAACGAGCTAGCTCGTC***A**GTATATTGATTAGTTTTGGAGTACTCGACC CGGCC**A***GGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGC*TTCGGCATGGCGAATGGGAC*	Template for rTetCIRC
GGTTCTAATACGACTCACTATA**G**GCCGTGTGCCTTGCGCCGGGAAACCACGCAAGGGATGGTGTCAAATTCGGCGAAACCTAAGCGCCCGCCCGGGCGTATGG CAACGCCGAGCCAAGCTTCGCAGCCTTCGGGCTGCGATGAAGGTGTAGAGACTAGACGGCACCCACCTAAGGCAAACGCTATGGTGAAGGCATAGTCCAGGG AGTGGCGAAAGCCACACAAACCA**G***AAGGGC*GTCGTCGCCCCGAGGATC*	Template for AzoGI-M
GGTTCTAATACGACTCACTATA**G**GCCGTGTGCCTTGCGCCGGGAAACCACGCAAGGGATGGTGTCAAATTCGGCGAAACCTAAGCGCGGATGGTTCGCGTATG GCAACGCCGAGCCAAGCTTCGCAGCCTTCGGGCTGCGATGAAGGTGTAGAGACTAGACGGCACCCACCTAAGGCGCACCGAACCATCCGGTGCGCTATGGTGA AGGCATAGTCCAGGGAGTGGCGAAAGCCACACAAACCA**G***AAGGGC*GTCGTCGCCCCGAGGATC*	Template for AzoGI-T
GGTTCTAATACGACTCACTATA**G**CGGTAGTAAGCAGGGAACTCACCTCCAATTTCAGTACTGAAATTGTCGTAGCAGTTGACTACTGTTATGTGATTGGTAGAGG CTAAGTGACGGTATTGGCGTAAGTCAGTATTGCAGCACAGCACAAGCCCGCTTGCGAGAA**T**GTCCAACCTTCATGCTTACGACG	Template for trans-acting VS ribozyme
GGTTCTAATACGACTCACTATA**G**GCCGTGTGCCTTGCGCCGGGAAACCACGCAAGGGATGGTGTCAAATTCGGCGAAACCTAAGCGCCCGCCCGGGCGTATGG CAACGCCGAGCCAAGCTTCGCAGCCTTCGGGCTGCGATGAAGGTGTAGAGACTAGACGGCACCCACCTAAGGCAAACGCTATGGTGAAGGCATAGTCCAGGG AGTGGCGAAAGCCACACAAACCAGACGGCC**A***GGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGC*TTCGGCATGGCGAATGGGAC*	Template for AzoGI*-M
GGTTCTAATACGACTCACTATA**G**GCCGTGTGCCTTGCGCCGGGAAACCACGCAAGGGATGGTGTCAAATTCGGCGAAACCTAAGCGCGGATGGTTCGCGTATG GCAACGCCGAGCCAAGCTTCGCAGCCTTCGGGCTGCGATGAAGGTGTAGAGACTAGACGGCACCCACCTAAGGCGCACCGAACCATCCGGTGCGCTATGGTGA AGGCATAGTCCAGGGAGTGGCGAAAGCCACACAAACCAGACGGCC**A***GGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGC*TTCGGCATG GCGAATGGGAC*	Template for AzoGI*-T
GTTCTAATACGACTCACTATA**G**GATCTTCGGGGCAGGGTGAAATTCCCGACCGGTGGTATAGTCCACGAAAGTATTTGCTTTGATTTGGTGAAATTCCAAAACCG ACAGTAGAGTCTGGATGAGAGAAGATT**C***GGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGC*TTCGGCATGGCGAATGGGAC*	Template for FMNrsw-M
GTTCTAATACGACTCACTATA*GGCAATCAGCGATCCCTGATGAGTCCGTGAGGACGAAACGAGCTAGCTCGTC***G**GATCGCTGATTGATCCGCTTCGGGGCAGG GTGAAATTCCCGACCGGTGGTATAGTCCACGAAAGCGCGAATCAGCCGCGCTTTGATTTGGTGAAATTCCAAAACCGACAGTAGAGTCTGGATGAGAGAAG**C***G GCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGC*TTCGGCATGGCGAATGGGAC*	Template for FMNrsw-T

Supplementary Table 2 | Sequences of synthesized genes for PCR amplification to produce the IVT templates. The 5'- and 3'ends of the target RNA are in bold. Sequences removed by ribozyme(s) are in italics. Primer-binding regions for PCR amplification are underlined. For the preparation of AzoGI-M and AzoGI-T, the IVT was performed with the trans-acting VS ribozyme, which was introduced by adding a 0.1 equivalent of the PCR-amplified template for trans-acting VS ribozyme.

	TetGI-DS	TetGI-D	TetGI-T	AzoGI-T	FMNrw-T
	PDB code: 7R6L EMDB code: EMD-24281	PDB code: 7R6M EMDB code: EMD-24282	PDB code: 7R6N EMDB code: EMD-24283	PDB code: N/A EMDB code: EMD-24284	PDB code: N/A EMDB code: EMD-24285
Data collection and processing					
Microscope	Titan Krios	Titan Krios	Titan Krios	Polara	Polara
Detector	K3	K3	K3	K2 Summit	K2 Summit
Magnification	105,000	105,000	105,000	31,000	31,000
Voltage (kV)	300	300	300	300	300
Electron Exposure ($e^-/Å^2$)	47	47	47	52	52
Defocus Range (μm)	0.8-2.0	0.8-2.0	0.8-2.0	1.0-2.5	1.0-2.5
Acquisition Pixel Size (Å)	0.825	0.825	0.825	1.23	1.23
Final symmetry imposed	C1	C1	C1	C1	C1
Initial particle images (#)	550,754	384,170	124,353	474,349	471,007
Final particle images (#)	82,575 S.E. monomers	113,548 S.E. monomers	85,596 S.E. monomers	486,860 S.E. monomers	266,623 S.E. monomers
Reconstruction Pixel Size (Å)	0.825	1.03	1.24	1.23	1.23
Map resolution (Å)	2.85 core / 2.98 overall	3.68 core / 3.78 overall	4.10 core / 4.17 overall	4.9	5.9
FSC threshold	0.143	0.143	0.143	0.143	0.143
Map resolution range (Å)	2.5-6.0	3.5-8.0	3.0-7.0	4.0-8.0	4.0-8.0
Refinement					
Model resolution (Å)	2.97	3.82	4.20		
FSC threshold	0.5	0.5	0.5		
Map sharpening B factor ($Å^2$)	N/A	N/A	N/A	N/A	N/A
Model composition					
Non-hydrogen atoms	7857	7755	7588		
RNA bases	366	362	354		
Ligand (Mg^{2+} ions)	20	13	12		
B factor ($Å^2$)					
RNA bases	53.79	100.66	89.86		
Ligand (Mg^{2+} ions)	19.61	55.48	46.19		
R.m.s. deviations					
Bond lengths (Å)	0.008	0.006	0.008		
Bond angles (°)	0.893	0.943	1.069		

Supplementary Table 3 | Data collection and refinement statistics.

Supplementary references

1. Tomizawa, J.-i. Control of ColE1 plasmid replication: The process of binding of RNA I to the primer transcript. *Cell* 1984, **38**(3)**:** 861-870.
2. Lee, A. J. & Crothers, D. M. The solution structure of an RNA loop–loop complex: the ColE1 inverted loop sequence. *Structure* 1998, **6**(8)**:** 993-1007.
3. Fiore, J. L. & Nesbitt, D. J. An RNA folding motif: GNRA tetraloop-receptor interactions. *Q. Rev. Biophys.* 2013, **46**(3)**:** 223-264.
4. Ennifar, E., Walter, P., Ehresmann, B., Ehresmann, C. & Dumas, P. Crystal structures of coaxially stacked kissing complexes of the HIV-1 RNA dimerization initiation site. *Nature Struct. Biol.* 2001, **8**(12)**:** 1064-1068.
5. Kondo, Y., Oubridge, C., van Roon, A. M. & Nagai, K. Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5' splice site recognition. *Elife* 2015, **4**.
6. Liu, D., Shao, Y., Piccirilli, J. A. & Weizmann, Y. Structures of artificially designed discrete RNA nanoarchitectures at near-atomic resolution. *Sci Adv* 2021, **7**(39)**:** eabf4459.
7. Wang, H. W. & Wang, J. W. How cryo-electron microscopy and X-ray crystallography complement each other. *Protein Sci.* 2017, **26**(1)**:** 32-39.
8. Wang, J. & Moore, P. B. On the interpretation of electron microscopic maps of biological macromolecules. *Protein Sci.* 2017, **26**(1)**:** 122-129.
9. Stuart, D. I. & Abrescia, N. G. From lows to highs: using low-resolution models to phase X-ray data. *Acta Crystallogr D Biol Crystallogr* 2013, **69**(Pt 11)**:** 2257-2265.
10. Liu, Y., Gonen, S., Gonen, T. & Yeates, T. O. Near-atomic cryo-EM imaging of a small protein displayed on a designed scaffolding system. *Proc. Natl. Acad. Sci. USA* 2018, **115**(13)**:** 3362-3367.
11. Liu, Y., Huynh, D. T. & Yeates, T. O. A 3.8 A resolution cryo-EM structure of a small protein bound to an imaging scaffold. *Nat Commun* 2019, **10**(1)**:** 1864.
12. Zhang, K. *et al.* Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 A resolution. *Nat Commun* 2019, **10**(1)**:** 5511.
13. Cate, J. H. *et al.* Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* 1996, **273**(5282)**:** 1678-1685.
14. Cate, J. H., Hanna, R. L. & Doudna, J. A. A magnesium ion core at the heart of a ribozyme domain. *Nat Struct Biol* 1997, **4**(7)**:** 553-558.
15. Ye, J. D. *et al.* Synthetic antibodies for specific recognition and crystallization of structured RNA. *Proc. Natl. Acad. Sci. USA* 2008, **105**(1)**:** 82-87.
16. Guo, F., Gooding, A. R. & Cech, T. R. Structure of the Tetrahymena ribozyme: base triple sandwich and metal ion at the active site. *Mol Cell* 2004, **16**(3)**:** 351-362.