PNAS
www.pnas.org

Supplementary Information for

Hatchet ribozyme structure and implications for cleavage mechanism

Luqian Zheng[#], Christoph Falschlunger[#], Kaiyi Huang, Elisabeth Mairhofer, Shuguang Yuan, Juncheng Wang, Dinshaw J. Patel*, Ronald Micura* and Aiming Ren*

[#] These authors made equal contributions

*Corresponding: aimingren@zju.edu.cn (AR), ronald.micura@uibk.ac.at (RM) and pateld@mskcc.org (DP)

This PDF file includes:

SI Materials and Methods
Figs. S1 to S12
Tables S1
References for SI reference citations

Materials and Methods

RNA preparation for crystallography

The corresponding DNA sequence of the hatchet ribozyme followed by the sequence of the hammerhead ribozyme was cloned into a pUT7 vector under the control of a T7 promoter (1). Then the plasmid was amplified in DH5α cells in large scale using the ZQZY-CF incubator shaker (Shanghai Zhichu Instrument). After linearization with the *Hind III* restriction endonuclease, the DNA can be used as templates for the transcription of hatchet ribozyme.

All the hatchet ribozyme RNA samples for crystallization were prepared by in vitro transcription with bacteriophage T7 RNA polymerase. To facilitate crystallization, the loop of stem P4 of wild-type hatchet ribozyme was replaced by either GAAA or UUCG tetraloops. The transcribed RNA was further purified by denaturing urea polyacrylamide gel electrophoresis (PAGE) followed by anion-exchange chromatography and ethanol precipitation.

Purified hatchet ribozyme RNA samples were annealed at 65 ℃ for 5 min in buffer containing 50 mM HEPES, pH 6.8, 50 mM NaCl, 5 mM $MgCl_2$, followed by cooling on ice for 30 to 60 min before setting up crystallization trials.

Crystallization

Crystallization of hatchet ribozyme RNA samples were performed by the sitting-drop vapor diffusion approach at 16 ℃ after mixing samples at a 1:1 ratio with reservoir solution. The crystal of HT-GAAA construct appeared from the crystallization high salt condition of 0.10 M HEPES, pH 6.8, 2.0 M $(NH_4)_2SO_4$ over a period of four weeks. The crystal of HT-UUCG construct appeared from the high salt crystallization condition containing 0.20 M K-formate, 2.2 M $(NH_4)_2SO_4$ over a period of more than two weeks. All the crystals were transferred quickly into cryoprotectant solution containing 3.0 M sodium malonate before flash-freezing in liquid nitrogen. For anomalous data collection, crystals of HT-UUCG constructs were soaked in the above crystallization condition with additional 100 mM $Ir(NH_3)_6^{3+}$ at 4 ℃ for 24 hours.

X-ray Data Collection and Structure Determination

Native x-Ray diffraction data of the hatchet ribozyme were collected at 100 K using the beamline BL-17U1, BL017B, BL18U1 and BL19U1 at the Shanghai Synchrotron Radiation Facility (SSRF)(2) and processed by HKL2000 (HKL Research). The $Ir(NH_3)_6^{3+}$ single-wavelength anomalous diffraction (SAD) data of HT-UUCG hatchet ribozyme were collected at NE-CAT_ID24C beamline at the Advanced Photon Source, Argonne National Laboratory, and processed with the XDS program (3). The space group was $P6_322$. The phase of the structure was solved with the SAD technique employing the anomalous signal from one iridium atom within each asymmetric unit (see the arrow in *SI Appendix,* Figure S12) by auto-solve program in the PHENIX suite (4). The initial model was built and adjusted manually in coot program (5) and refined with phenix.refine program (4). Then the structure of HT-GAAA hatchet ribozyme was solved by molecular replacement (MR) method with HT-UUCG hatchet ribozyme structure as the model. Metal ions and their coordinated waters were identified based on 2Fo-Fc and Fo-Fc maps guided by the coordination geometries. The X-ray statistics of the native HT-GAAA and $Ir(NH_3)_6^{3+}$-containing HT-UUCG crystals are listed in *SI Appendix,* Table S1.

Modeling and Energy Minimization

Modeling of the pre-catalytic state of C(-1) hatchet structure was conducted first by the cleavage site comparison of the hatchet ribozyme product structure with hammerhead ribozyme (PDB code: 2OEU) in a similar way as was done for the HDV ribozyme (6). Then the modeled C(-1) was manually rotated around the C5′-C4′ bond of U1 by 54^o to optimally fit into the predicted cleavage site pocket. The following energy minimization had been performed in Schrodinger 2018v2 software (7). The model was refined using OPLS_2005 force field. Hydrogen atoms were added to the model reflecting the physiological pH (7.4) using the PROPKA (8) tool in Protein Preparation tool in Maestro to obtain the optimized hydrogen bond network. Following that, the constrained energy minimizations

were performed on the full-atomic models, with an atom allowed motion of 0.4 Å, excluding hydrogens which were free to move.

Cleavage assays

Aliquots from aqueous stock solutions of the two strands (S, R) were mixed and lyophilized. After addition of reaction buffer (30 mM HEPES, pH 7.5, 100 mM KCl, 10 mM $MgCl_2$) to yield a final concentration of c(RNA) = 55 µM (each strand) in a total volume of 20 µL, the reaction was stopped by the addition of EDTA solution (20 µL; 3 mM) after 1, 10, 45 or 120 min, and subsequently analyzed by anion-exchange HPLC (analytical Dionex DNAPac column; (4 mm × 250 mm) at 80 °C. Flow rate: 1 mL min^{-1}, eluant A: 25 mM Tris·HCl (pH 8.0), 6 M urea; eluant B: 25 mM Tris·HCl (pH 8.0), 0.5 M $NaClO_4$, 6 M urea; gradient: 0–60% B in A within 45 min, UV detection at 260 nm).

SUPPLEMENTARY FIGURES

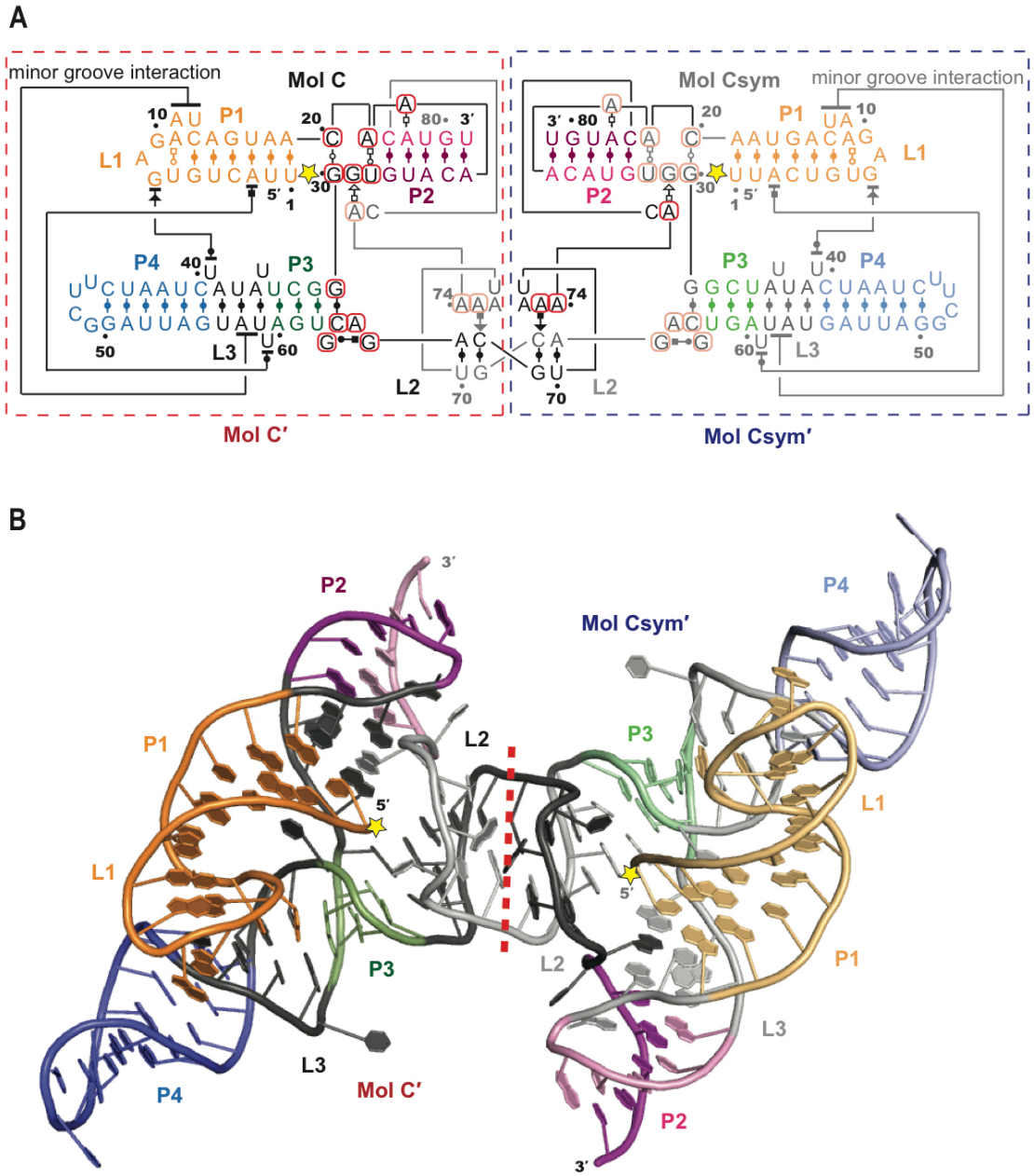

Figure S1. Schematic and tertiary fold of the HT-UUCG hatchet ribozyme product.

(A) A schematic representation of the secondary structure of the HT-UUCG hatchet ribozyme product dimer together with the long-range interactions in the

tertiary structure that are labeled with solid lines. The color-coding is the same as Figure **1C**.

(B) The tertiary structure of HT-UUCG hatchet ribozyme product. The two HT-UUCG hatchet ribozyme molecules form a dimer through swapping of the 3′-ends of the pairing strands. The cleavage site is shown with a yellow star. To simplify the structural analysis, we swapped the 3′-end of the pairing strands of the two molecules in the dimer and labeled them as the new monomer molecules C′ and Csym′ as shown by dashed rectangles.

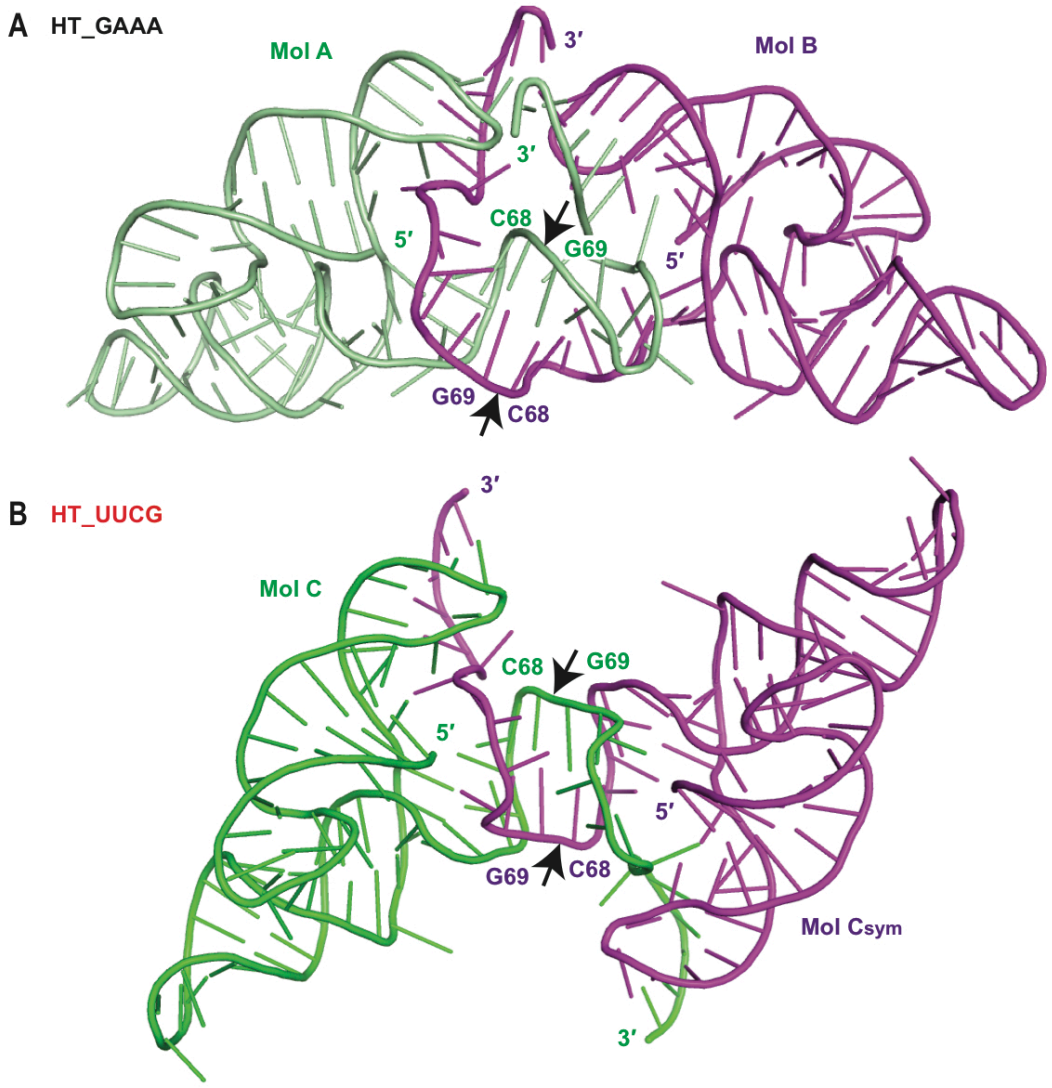

Figure S2. The dimer structure of hatchet ribozyme product observed in the crystal structure and the monomer-dimer equilibrium observed in the solution state of hatchet ribozyme product.

(A) The fold of the hatchet ribozyme product dimer formed by the HT-GAAA construct in each asymmetric unit in the crystal. The space group is $P2_12_12_1$. The pseudo-symmetric site of the dimer was located between C68 and G69.

(B) The fold of the hatchet ribozyme product dimer formed by two symmetrically-related HT-UUCG hatchet ribozyme product constructs in the crystal. The space group was $P6_322$. The symmetric site of the dimer was also located between C68 and G69.

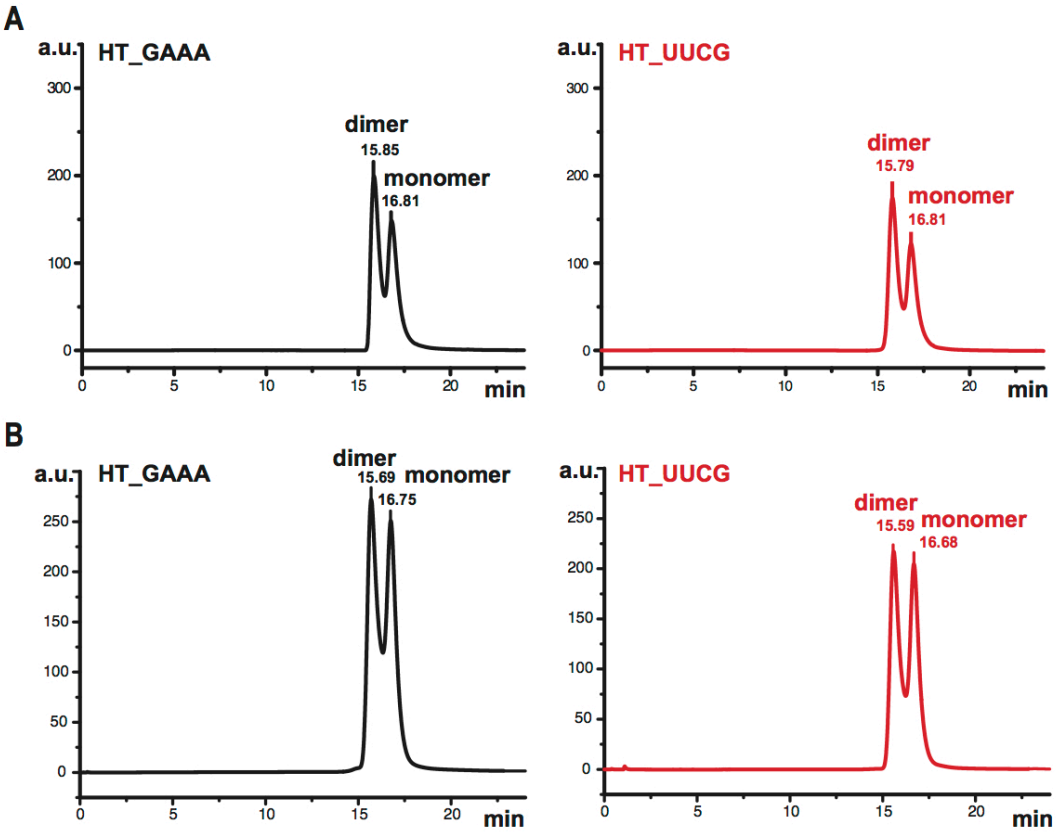

Figure S3. The solution state of HT-GAAA and HT-UUCG hatchet ribozyme . Both of the HT-GAAA and HT-UUCG hatchet ribozyme product constructs existed as an equilibrium of dimer and monomer in solution. The size-exclusion experiment was run with Superose^{TM} 6 Increase 10/300 GL column. The running buffer contained 50 mM HEPES pH 6.8, 50 mM NaCl and 5 mM $MgCl_2$ **(A)** or 50 mM HEPES pH 7.0, 140 mM KCl and 1 mM $MgCl_2$ **(B)**.

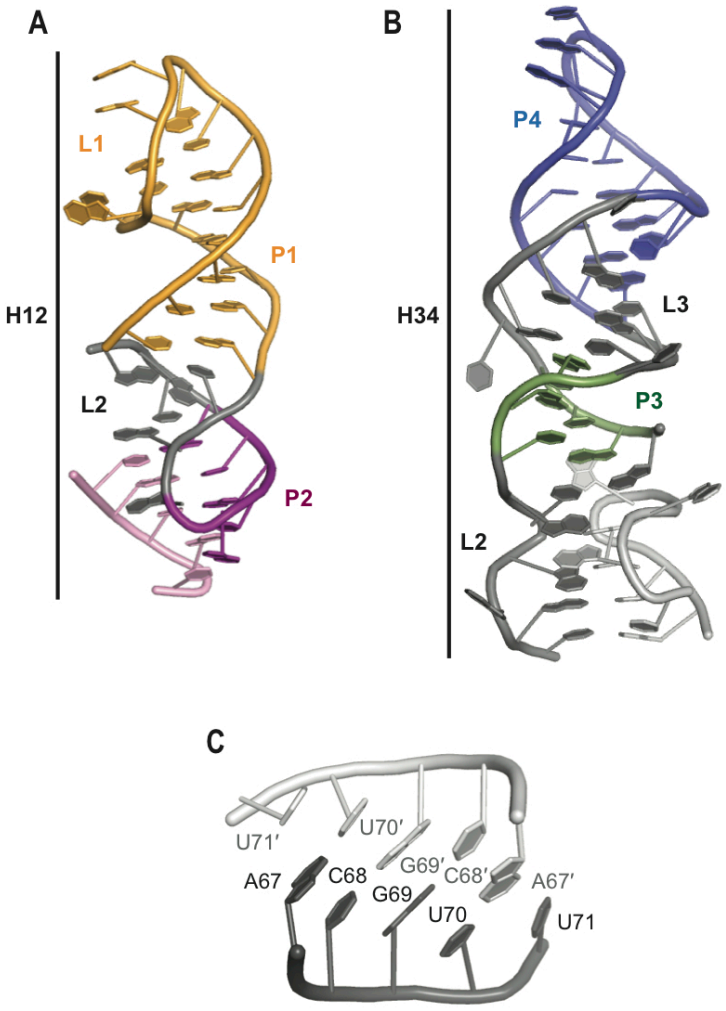

Figure S4. The stacking interaction within the two parallel long helices H12 and H34 and the pseudo-symmetric site in HT-GAAA ribozyme product structure.

(A) Pairing of some conserved residues of L2 with the three residues between stems P1 and P2 to form a continuous stacking interaction from stem P1 to stem P2, with the resulting long helix termed H12.

(B) Loop L3 zippers-up and forms a short stem between stems P3 and P4. L2 forms an intermolecular pseudo-symmetric stem with L2′ from the adjacent molecule, thereby extended the stacking interaction of stem P3. The long helix termed H34 is formed by stem P4, L3, stem P3, and part of L2.

(C) The pseudo-symmetric/symmetric intermolecular pairing stem was formed by the palindromic residues from A67 to U70 in L2 from two molecules of hatchet ribozyme product in the crystal.

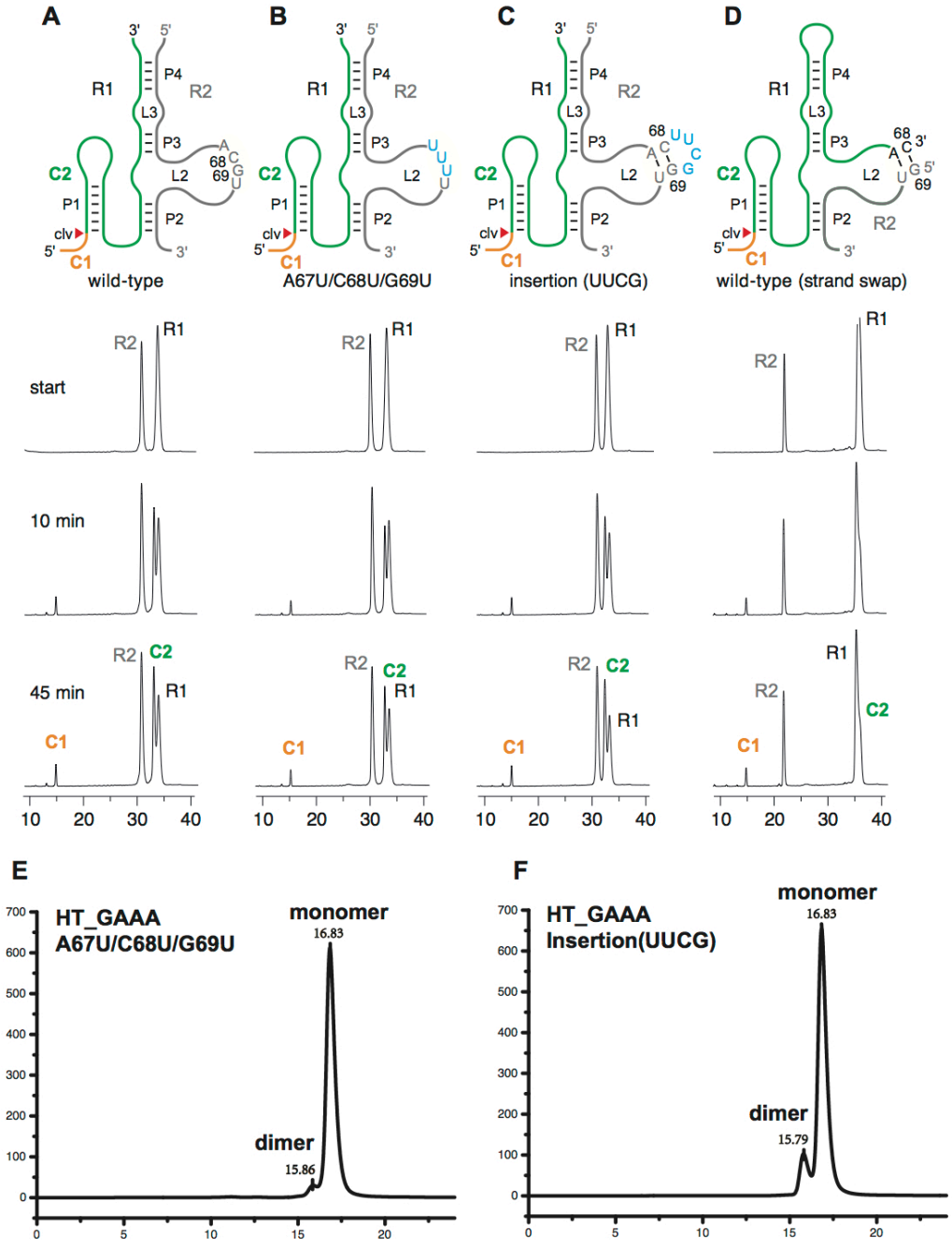

Figure S5. Cleavage activity of the *env10* hatchet ribozyme and mutants. Secondary structure cartoons of the two-stranded construct used in the cleavage assays with crucial base interactions highlighted in red and respective mutants in blue. HPLC traces following cleavage activity of wild-type ribozyme (**A**) and

mutants A67U/C68U/G69U (**B**), UUCG tetra-loop insertion (**C**), and the strand swap construct of wild-type sequence (**D**). R1 and R2 denote the substrate and ribozyme strands; C1 (orange) and C2 (green) denote cleavage products. Cleavage activity analyzed at 55 μM RNA each strand; 10 mM $MgCl_2$, 100 mM KCl, 30 mM HEPES, pH 7.5, 23 °C. HPLC conditions: Dionex DNAPac column (4 × 250 mm^2), 80 °C, 1 mL min^{-1}, 0–60% buffer B in 45 min. Buffer A: Tris–HCl (25 mM), urea (6 M), pH 8.0. Buffer B: Tris–HCl (25 mM), urea (6 M), $NaClO_4$ (0.5 M), pH 8.0. (**E, F**). Size-exclusion chromatography showed the first mutant (A67U/C68U/G69U) exclusively existed as monomer (**E**) and that the second mutant (UUCG insertion) predominantly existed as monomer (>85%) (**F**) in solution The size-exclusion experiment was run with Superose^{TM} 6 Increase 10/300 GL column. The running buffer contained 50 mM HEPES pH 6.8, 50 mM NaCl and 5 mM $MgCl_2$.

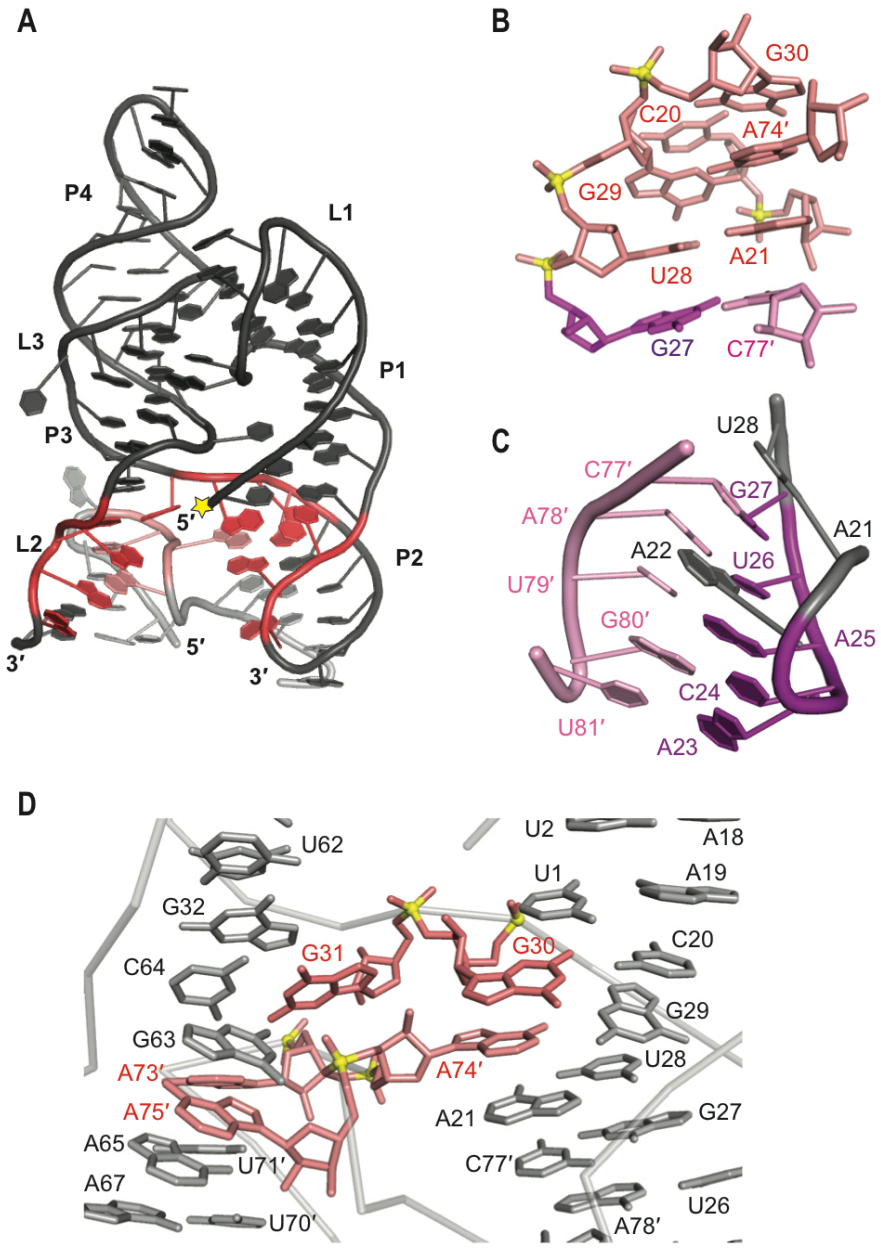

Figure S6. Structural alignment of the highly conserved residues of HT-GAAA hatchet ribozyme product.

(A) The tertiary fold of the molecule A′ of HT-GAAA structure is shown in gray with highly conserved residues shown in red.

(**B)** Two consecutive highly-conserved residues C20 and A21 between stems P1 and P2 form pairing and stacking interactions with three consecutive conserved residues U28-G29-G30 from L2. In addition, A74′ from the pseudo-symmetric molecule is intercalated between C20-G30 and U28-A21 base pairs, and forms a

non-canonical base pair with G29. The resulting three consecutive stacked base pairs formed by highly conserved residues connect stem P1 and P2.

(C) A21-A22-A23 spans along stem P2 from one end to the other with highly conserved A22 interacting with the major groove of stem P2.

(D) Some conserved residues in L2 reside in the junctional region of helix H12 and H34. The step involving G30 and G31 adopts a splayed-apart conformation, in which G30 is involved in H12 stacking and G31 participates in H34 formation. Three consecutive residues A73′, A74′ and A75′ are also involved in two parallel helix formation. A74′ is extruded from the RNA chain and participates in long helix H12 formation.

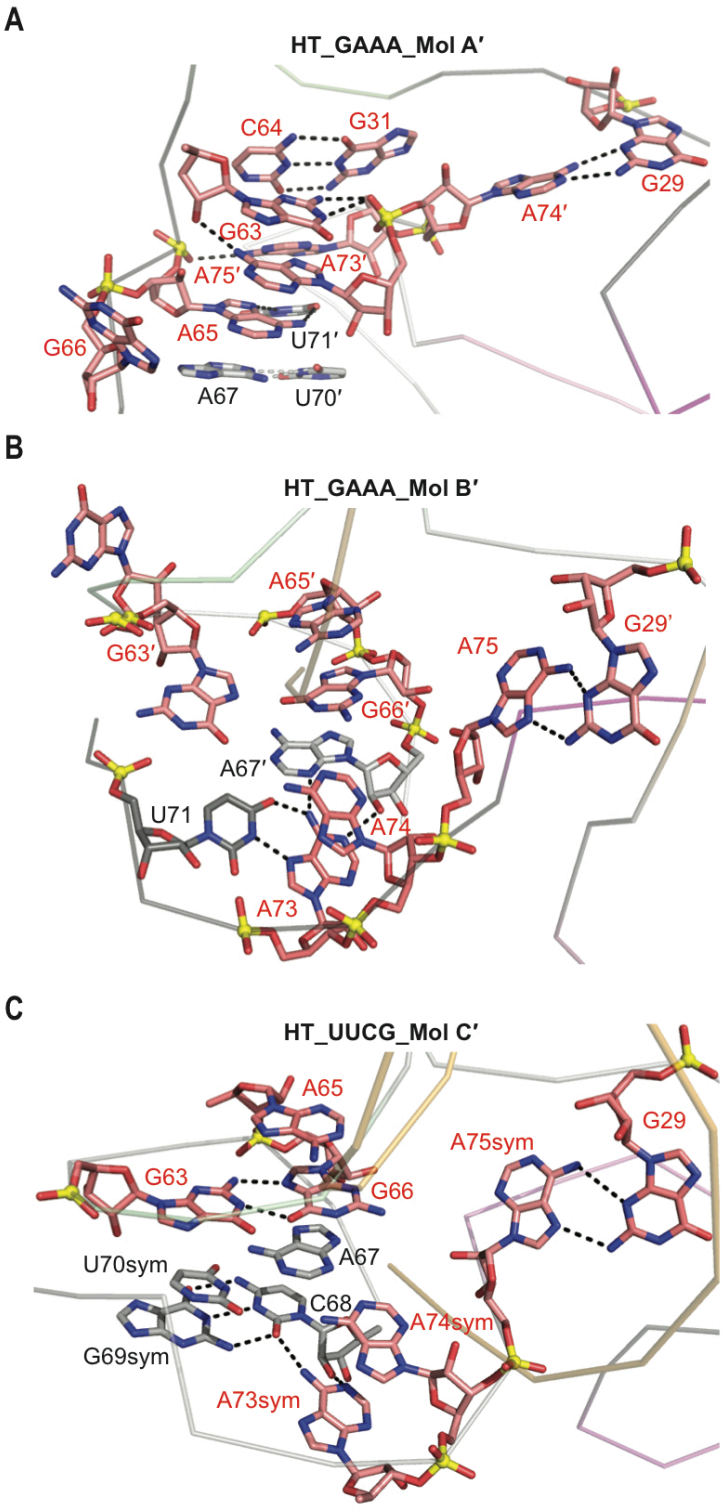
A
HT_GAAA_Mol A′
C64
G31
G29
A74′
G63
A75′
A73′
A65
U71′
G66
A67
U70′
B
HT_GAAA_Mol B′
A65′
G29′
G63′
A75
G66′
A67′
U71
A74
A73
C
HT_UUCG_Mol C′
A65
G29
A75sym
G63
G66
A67
U70sym
C68
A74sym
G69sym
A73sym

Figure S7

Figure S7. Comparison of the junctional region structure of Mol A′ and Mol B′ of HT-GAAA and Mol C′ of HT-UUCG hatchet ribozyme products.

(A) In the structure of molecule A′ of HT-GAAA construct, G63 is involved in the base stacking interaction in the loop segment and forms hydrogen bonds with A75′. Accordingly, A75′ is also involved in the loop stacking interaction of L2 and forms hydrogen bond interaction with three residues from G63 to A65. At the same time, A65 forms a *cis*-Hoogsteen-Watson-crick base pair with U71′. A73′ stacks on U71′ and hydrogen bonds with the minor groove of G31-C64 base pair. G66 projects outwards and does not participate in the intramolecular interaction. We noticed that while A74′ participated in the stacking interaction of helix H12 and formed a non-canonical base pair interaction with G29. A73′ and A75′ are involved in the stacking interaction of helix H34. Notably, it was surprising to find that the sugar pucker of these residues, namely A63, A65, A66, A73′, A74′ and A75′ in the molecule A′ all adopted C2′-*endo* conformation.

(B) In the structure of molecule B′ of HT-GAAA construct, we found G63′, A65′ and G66′ did not form any pairing interactions with the neighboring residues, but only participated in helix-stacking interactions. Different from molecule A′, it was found that the conserved residue A73 not only forms a *cis*-Hoogsteen-Watson-Crick base pair with U71, but also forms two other hydrogen bonds with A67′. A74 stacks on the A73-U71 base pair. We note that the Hoogsteen edge of A75 forms hydrogen bonds with the sugar edge of G29, which is involved in the long helix H12 staking interaction. The residue G63′ folds in two alternative conformation and partially adopts C2′-*endo* sugar pucker. The sugar pucker of other residues A65′, A66′, A73, A74 and A75 adopt a C3′-*endo* conformation.

(C) In the structure of molecule C′ of HT-UUCG construct, G63 forms a Watson-crick-Hoogsteen base pair with G66 and intercalates between A65 and A67, both of which participat in the long helix H34 stacking interaction. A73′ and A74′ interact with H34 through the A-minor interaction. A75′ forms a Hoogsteen-sugar base pair with G29, which is different with molecule A′ of HT-GAAA construct, but similar with molecule B′ of HT-GAAA construct. Notably, of these nucleotides,

only A63 and A74′ adopt C2′-endo sugar pucker conformation, the other nucleotides A65, A66, A73′ and A75′ fold in C3′-endo conformation.

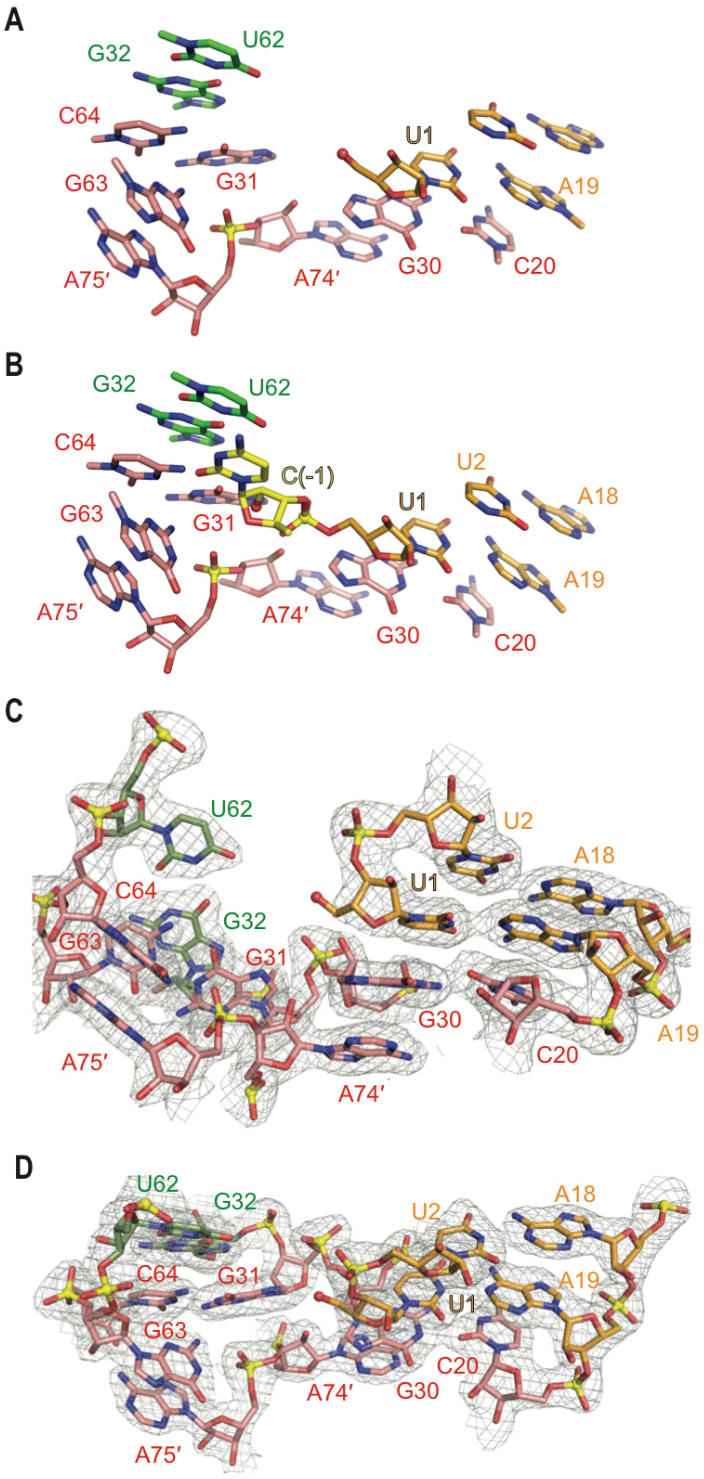
A
G32
U62
C64
U1
G63
G31
A19
A75′
A74′
G30
C20
B
G32
U62
C64
C(-1)
U2
A18
G63
G31
U1
A19
A75′
A74′
G30
C20
C
U62
U2
A18
U1
C64
G32
G63
G31
G30
A19
C20
A75′
A74′
D
U62
G32
U2
A18
C64
G31
A19
U1
G63
C20
A74′
G30
A75′

Figure S8

Figure S8. Structural alignment of the leaving group in the hatchet ribozyme product structure and modeling the cleavage site conformation in mol A′ in HT-GAAA hatchet ribozyme.

(A) Base stacking with U1 at the cleavage site and alignment of junctional structure shown in a rotated view from Figure **3A**.

(B) The base stacking interaction of the modeled cleavage site shown in a rotated view from Figure **3C**. The base stacking interaction of the modeled cleavage site between C(-1) and U1, in which C(-1) was stacked partially between G32-U62 from stem P3 and two conserved residues G63 and A75′, whereas U1 was stacked between the conserved G30-C20 base pair and the termini of stem P1.

(C) *2Fo-Fc* electron density map contoured at 1.0 σ level was shown for the cleavage site and the junctional structure from Figure **3A**.

(D) *2Fo-Fc* electron density map contoured at 1.0 σ level was shown for the cleavage site and the junctional structure in a rotated view from Figure **S8C**.

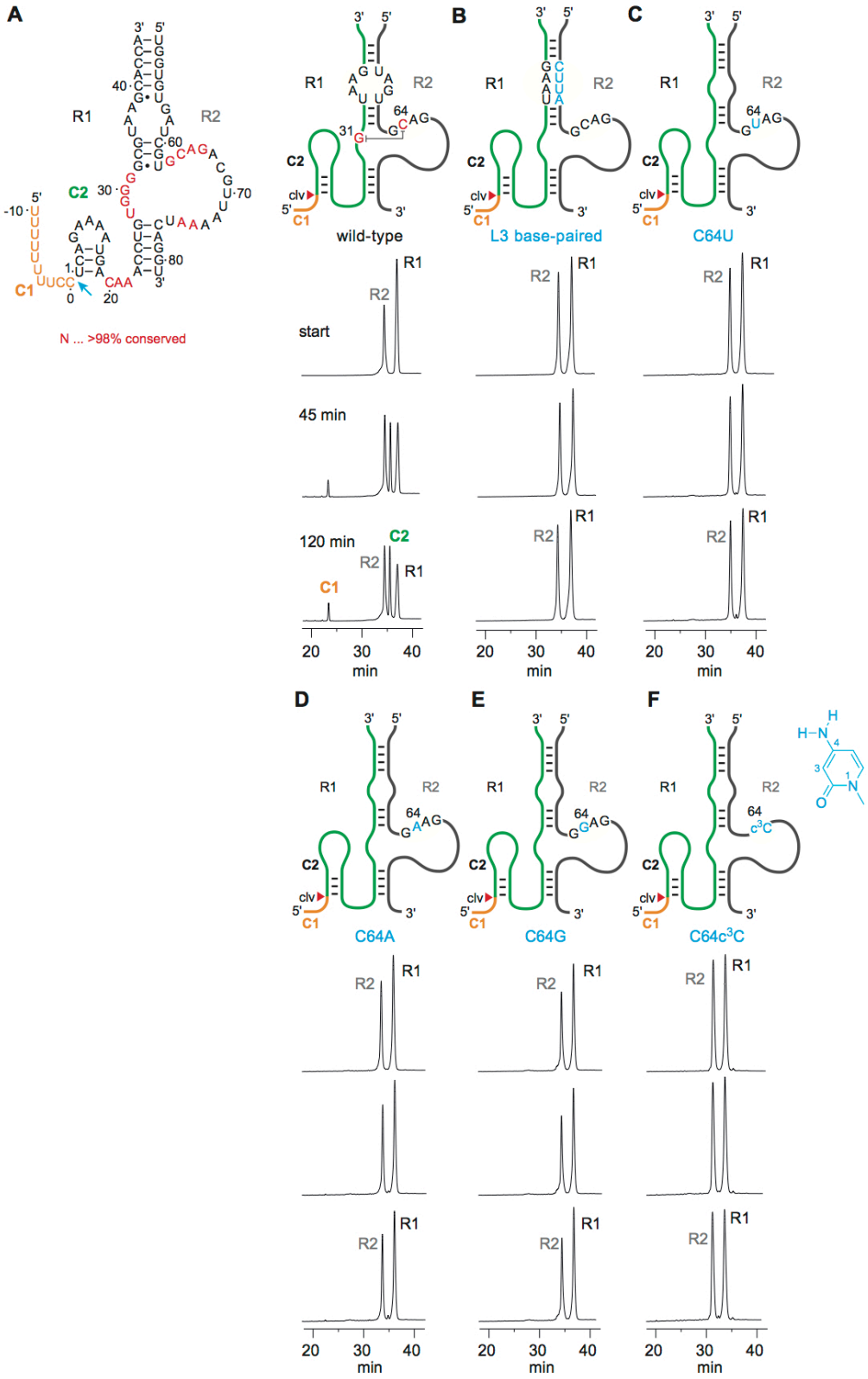
A
B
C
D
E
F
R1
R2
C1
C2
clv
wild-type
L3 base-paired
C64U
C64A
C64G
C64c3C
start
45 min
120 min
min
N ... >98% conserved

Figure S9

Figure S9. Cleavage activity of the *env214* hatchet ribozyme and mutants. Secondary structure cartoons of the two-stranded construct used in the cleavage assays with crucial base interactions highlighted in red and respective mutants in blue. HPLC traces following cleavage activity of wild-type ribozyme (**A**) and mutants with regular helix instead of L3 (**B**), C64U (**C**), C64A (**D**), C64G (**E**), and C64c^{3}C (**F**). R1 and R2 denote the substrate and ribozyme strands; C1 (orange) and C2 (green) denote cleavage products. Cleavage activity analyzed at 55 μM RNA each strand; 10 mM $MgCl_2$, 100 mM KCl, 30 mM HEPES, pH 7.5, 23 °C. HPLC conditions: Dionex DNAPac column (4 × 250 mm^2), 80 °C, 1 mL min^{-1}, 0–60% buffer B in 45 min. Buffer A: Tris–HCl (25 mM), urea (6 M), pH 8.0. Buffer B: Tris–HCl (25 mM), urea (6 M), $NaClO_4$ (0.5 M), pH 8.0.

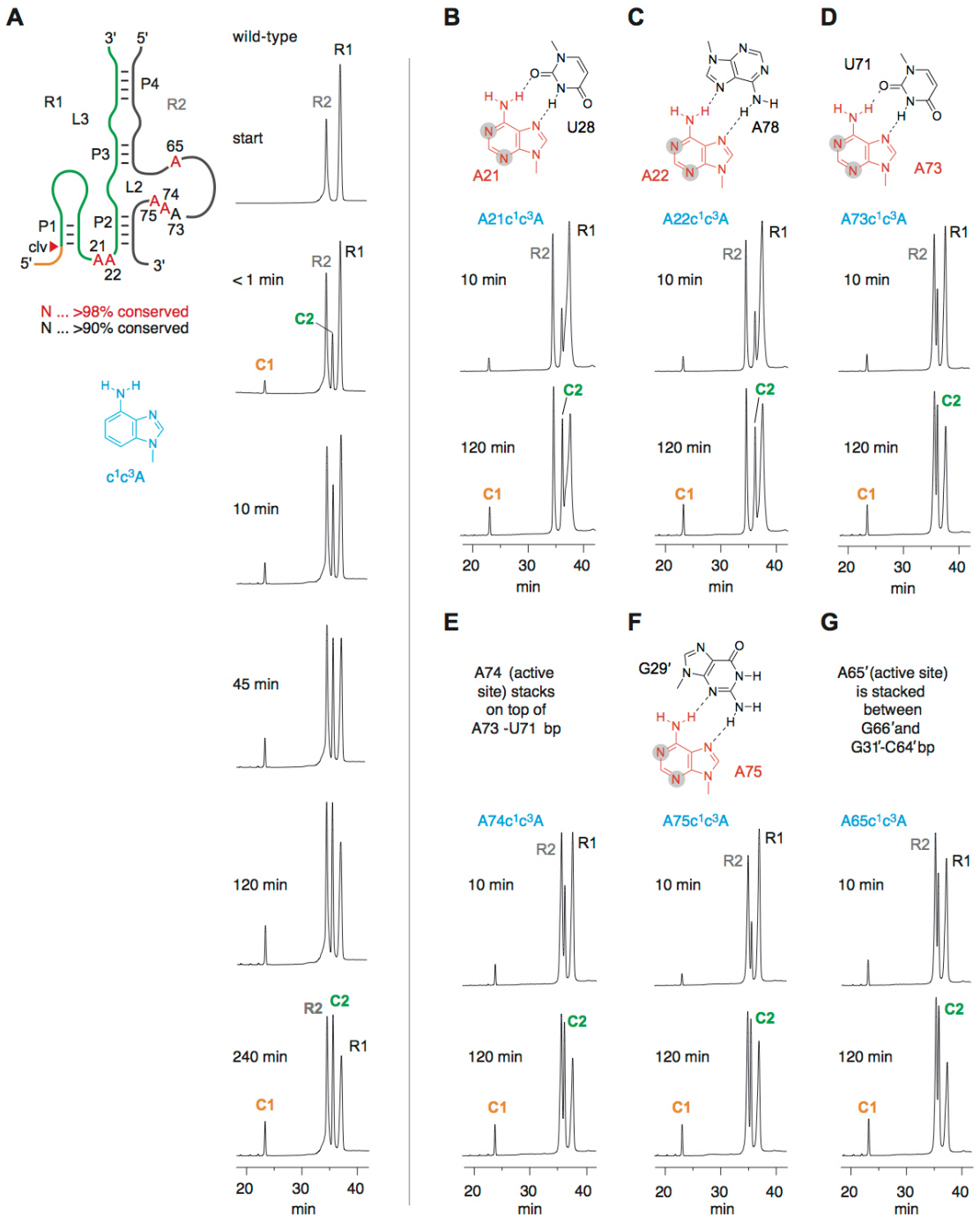

Figure S10. Cleavage activity of the *env214* hatchet ribozyme and mutants with 1,3-dideaza adenosine replacement of adenosine.

Secondary structure cartoons of the two-stranded construct used in the cleavage assays with crucial base replacement by 1,3-dideaza adenosine (c^1c^3A), an

analog that lacks Watson-Crick base-pairing propensities. HPLC traces following cleavage activity of wild-type ribozyme (**A**) and mutants with c^1c^3A replacement of A21 (**B**), A22 (**C**), A73 (**D**), A74 (**E**), A75 (**F**) and A65 (**G**). R1 and R2 denote the substrate and ribozyme strands; C1 (orange) and C2 (green) denote cleavage products. Cleavage activity analyzed at 55 μM RNA each strand; 10 mM $MgCl_2$, 100 mM KCl, 30 mM HEPES, pH 7.5, 23 °C. HPLC conditions: Dionex DNAPac column (4 × 250 mm^2), 80 °C, 1 mL min^{-1}, 0–60% buffer B in 45 min. Buffer A: Tris–HCl (25 mM), urea (6 M), pH 8.0. Buffer B: Tris–HCl (25 mM), urea (6 M), $NaClO_4$ (0.5 M), pH 8.0.

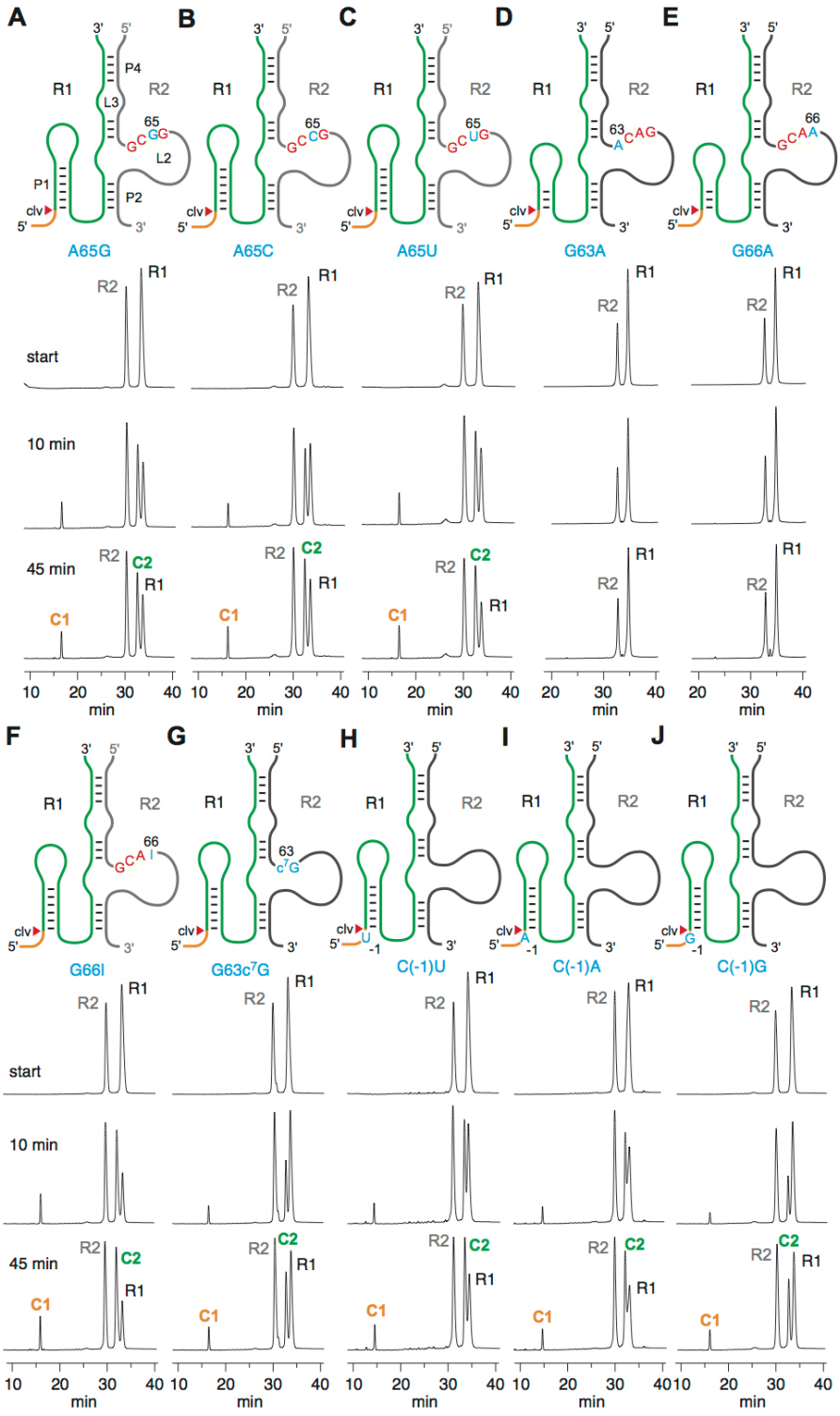
A
B
C
D
E
F
G
H
I
J
A65G
A65C
A65U
G63A
G66A
G66I
G63c7G
C(-1)U
C(-1)A
C(-1)G
start
10 min
45 min
R1
R2
C1
C2
min

Figure S11

Figure S11. Cleavage activity of *env10* and *env214* hatchet ribozyme and mutants.

Secondary structure cartoons of the two-stranded construct used in the cleavage assays with crucial base interactions highlighted in red and respective mutants in blue. HPLC traces following cleavage activity of hatchet ribozyme mutants A65G (**A**), A65C (**B**), A65U (**C**), G63A (**D**), G66A (**E**), C66I (**F**), G63c^{7}G (**G**), C(-1)U (**H**), C(-1)A (**I**), and C(-1)G (**J**). R1 and R2 denote the substrate and ribozyme strands; C1 (orange) and C2 (green) denote cleavage products. Cleavage activity analyzed at 55 μM RNA each strand; 10 mM $MgCl_2$, 100 mM KCl, 30 mM HEPES, pH 7.5, 23 °C. HPLC conditions: Dionex DNAPac column (4 × 250 mm^2), 80 °C, 1 mL min^{-1}, 0–60% buffer B in 45 min. Buffer A: Tris–HCl (25 mM), urea (6 M), pH 8. Buffer B: Tris–HCl (25 mM), urea (6 M), $NaClO_4$ (0.5 M), pH 8.

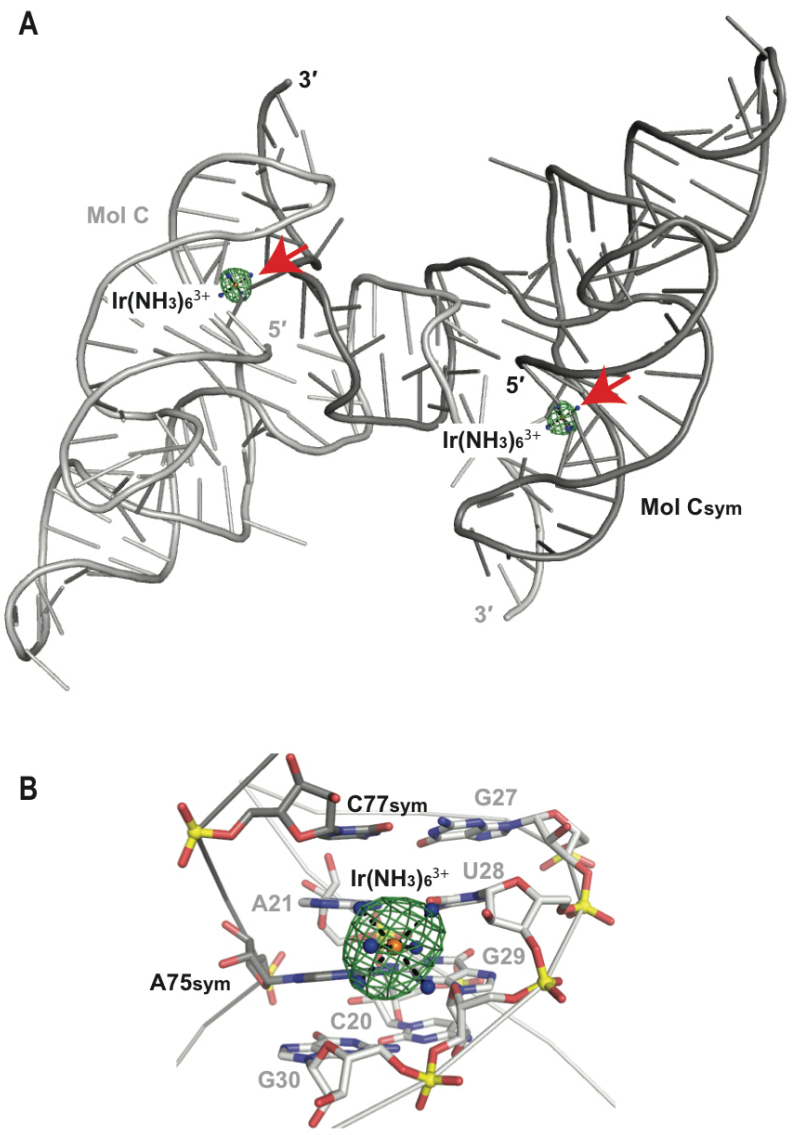

Figure S12. Packing alignments in the crystal lattice and bound $Ir(NH_3)_6^{3+}$ sites in structure of HT-UUCG hatchet ribozyme product.

(A) Anomalous map for $Ir(NH_3)_6^{3+}$ sites (shown by red arrows) was contoured at 4.0 σ level involving the molecule C of the HT-UUCG ribozyme in one asymmetrical unit and the symmetry-related molecule Csym in another asymmetrical unit in the same crystal lattice.

(B) Expanded view of panel highlighting $Ir(NH_3)_6^{3+}$ metal ion in molecule C of the HT-UUCG ribozyme.

Table S1. Crystallographic Statistics for the hatchet Ribozyme

Crystal	**HT-GAAA Native**	**HT-UUCG $[Ir(NH_3)_6]^{3+}$ soaked**
Data collection	BL17U (SSRF)	NE-CAT_ID24C (APS)
Space group	$P2_12_12$	$P6_322$
Cell dimensions		
a, *b*, *c* (Å)	40.6, 104.3, 126.9	91.8, 91.8, 140.6
α, β, γ (º)	90, 90, 90	90, 90, 120
Wavelength (Å)	0.9792	1.102 (peak)
Resolution (Å)	50-2.10(2.18-2.10)*	79.5-2.61 (2.72-2.61)
R_{pim}	0.036(0.90)	0.041 (2.29)
I/sI	31.9 (1.1)	20.2 (0.6)
Completeness (%)	99.6 (98.9)	99.0 (92.2)
Redundancy	12.1 (12.0)	36.9 (37.1)
Refinement		
Resolution (Å)	33.5-2.06	79.5-2.63
No. reflections	33267	19049
R_{work}/R_{free}	0.19/0.23	0.20/0.23
No. atoms		
RNA	3495	1742
Cations	1	7
Water	79	9
B-factors		
RNA	57.1	70.6
Cations	83.1	80.6
Water	54.0	61.3
R.m.s deviations		
Bond lengths (Å)	0.007	0.006
Bond angles (º)	1.034	1.236

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

References

1. Pikovskaya O, Serganov AA, Polonskaia A, Serganov A, & Patel DJ (2009) Preparation and crystallization of riboswitch-ligand complexes. *Methods in molecular biology* 540:115-128.
2. Wang Q-S*, et al.* (2018) Upgrade of macromolecular crystallography beamline BL17U1 at SSRF. *Nucl. Sci. Tech.* 29(5):68.
3. Kabsch W (2010) XDS. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 2):125-132.
4. Adams PD*, et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 2):213-221.
5. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 4):486-501.
6. Chen JH*, et al.* (2010) A 1.9 A crystal structure of the HDV ribozyme precleavage suggests both Lewis acid and general acid mechanisms contribute to phosphodiester cleavage. *Biochemistry* 49(31):6508-6518.
7. Anonymous (Schrödinger Release 2018-2: Maestro, Schrödinger, LLC, New York, NY, 2018.
8. Sondergaard CR, Olsson MHM, Rostkowski M, & Jensen JH (2011) Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pK(a) Values. *J Chem Theory Comput* 7(7):2284-2295.