A distinct pseudoknot fold defines the hovlinc self-cleaving ribozyme in human lncRNA

Luqian Zheng[1,2,*], Wenfei Li[1], Wenkai Wang[3], Heqin Zhu[4], Zejia Hu[1], Guiqi Zhi[2], Yuting Tang[2], Xinyue Bao[1], Huiqin You[1], Anna Rázková[5], Abhishek Suman[6], Zhenghao Qiao[7], Ronald Micura[5], Dinshaw Patel[6], Jianyi Yang[3], Peng Xiong[4] and Aiming Ren[1,*]

[1]Department of Cardiology of The Second Affiliated Hospital and Life Sciences Institute and School of Medicine and Liangzhu Laboratory, Zhejiang University, Hangzhou 310058, China

[2]College of Life Sciences, Anhui Provincial Key Laboratory of Molecular Enzymology and Mechanism of Major Metabolic Diseases, Anhui Normal University, Wuhu 241000, Anhui, China

[3]MOE Frontiers Science Center for Nonlinear Expectations, Shandong Province Key Laboratory of Financial Risk, Research Center for Mathematics and Interdisciplinary Sciences, Shandong University, Qingdao, 266237, China.

[4]Suzhou Institute for Advanced Research, University of Science and Technology of China, Suzhou, Jiangsu 215123, China

[5]Institute of Organic Chemistry, Center for Molecular Biosciences Innsbruck, Leopold Franzens University, Innsbruck A6020, Austria

[6]Structural Biology Program, Memorial Sloan-Kettering Center, New York, New York 10065, USA.

[7]Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou 510623, Guangdong, China

*To whom correspondence should be addressed. Tel: 86-571-88981228; Fax: 86-571-88981336;

Email: aimingren@zju.edu.cn (AR) and luqianzheng@ahnu.edu.cn (LZ)

Abstract

The hovlinc ribozyme is the first ribozyme identified within a human very long intergenic noncoding RNA (vlincRNA), uncovering a previously unexplored layer of functional RNA biology in the human genome. To elucidate its tertiary organization, we determined the crystal structure of a catalytically active truncated hovlinc ribozyme, which adopts an unanticipated homodimeric architecture composed of two functional units. Within each functional unit, the catalytic core is organized and stabilized by a previously unrecognized Rainbow-type (R-type) pseudoknot, defining a new RNA topological class distinct from all known pseudoknot architectures. At the cleavage site, A6–C7 adopts a splayed-apart conformation precisely organized by base-pairing and stacking interactions. Structure-based comparative analysis of 22 homologs, including an inactive gorilla variant, provides a structural rationale for widespread evolutionary loss of activity. Upon retaining the catalytic domain, we performed systematic truncation, sequence engineering, and SELEX-based selection to probe the variable regions, uncovering a broad landscape of sequence plasticity that expands the functional variant space, identifying activity-enhancing ribozyme variants. Structural modeling of the full-length ribozyme further reveals how peripheral elements modulate global folding and fine-tune catalytic efficiency. Collectively, these findings establish the first structural framework of a human lincRNA-encoded ribozyme and uncover a new RNA folding motif, revealing the structural principles underlying hovlinc ribozyme catalysis, demonstrating how subtle evolutionary variations tune RNA catalysis, and providing a framework for the discovery and engineering of functional RNA elements.

Key words: hovlinc ribozyme, RNA tertiary structure, catalytic mechanism, pseudoknot, noncoding RNA, Rainbow-type pseudoknot

Background

In addition to serving as carriers of genetic information, non-coding regions of the genome can be transcribed and processed into diverse forms of RNA molecules that participate in the regulation of cellular activities. Among them, long non-coding RNAs (lncRNAs) perform a wide range of cellular functions, and their expansion across species during evolution suggests a potential role in shaping organismal complexity (Chen and Kim, 2024; Mattick et al., 2023). Molecular and structural biology research indicates that RNAs can fold into complex three-dimensional architectures that enable precise molecular recognition, regulation, and catalysis. Ribozymes are catalytic RNA molecules whose discovery reshaped the long-standing view that biological catalysis is exclusively mediated by proteins and provided compelling support for the RNA world hypothesis (Gilbert, 1986).

Since their initial discovery, distinct classes of small self-cleaving ribozymes including the hammerhead, hairpin, hepatitis delta virus (HDV), glmS, pistol, twister, twister-sister, and hatchet ribozymes, have been identified in viral genomes, bacteria, or mobile genetic elements (Cervera and de la Pena, 2020; Jimenez et al., 2015; Roth et al., 2014; Weinberg et al., 2015). Despite performing the same transesterification reaction and generating a 2′,3′-cyclic phosphate and a 5′-hydroxyl product, different ribozymes, each defined by distinct conserved sequence and motifs, adopt strikingly diverse structures and catalytic strategies, highlighting the remarkable structural plasticity of RNA and demonstrating that catalysis can emerge from multiple independently evolved architectures.

The recent discovery of the first ribozyme, hovlinc, embedded within a long non-coding RNA of the human genome represents a significant conceptual advance, expanding the known repertoire of RNA catalysis into a new genomic and biological context and also suggesting that RNA-based catalysis, often viewed as a relic of the RNA world, may persist in modern genomes in subtle and previously unrecognized forms (Chen et al., 2021). Although the physiological function of hovlinc has not yet been established, its phylogenetic distribution is particularly intriguing. Comprehensive genomic analyses indicate that homologous sequences are present only in some placental mammals, with the highest sequence conservation observed in primates (Chen et al., 2021). This conservation raises intriguing questions about why catalytic RNA elements are retained within long non-coding transcripts under evolutionary selective pressure and how their activity is structurally encoded.

Here, we determine the tertiary structure of the hovlinc ribozyme and elucidate the folding principles that underlies its self-cleavage activity. Our structural analysis reveals that hovlinc adopts a compact catalytic architecture composed of two

extended helices, stabilized by extensive long-range interactions and a previously unrecognized “double-clamp” pseudoknot topology, which we termed it as Rainbow-type (R-type) pseudoknot. These structural features position the cleavage site at the core of the ribozyme and embed it within a rigid, continuous stacking framework optimized for catalysis. Guided by this architecture, we employed systematic truncation, sequence engineering, and SELEX-based selection to generate additional hovlinc variants in diverse forms that retain or even enhance catalytic activity. We also modeled the tertiary structure of the full-length 124-nucleotide wild-type ribozyme.

Together, our findings provide the first structural insight into a ribozyme encoded within a human long non-coding RNA and demonstrate that the human genome retains the capacity to encode catalytically competent RNAs. Moreover, the identification of previously uncharacterized RNA structural motifs provides a substantial conceptual advance in understanding the fundamental folding principles of RNA architecture and in guiding future structure analyses, prediction, and rational design. Our systematic engineering of the hovlinc ribozyme also establishes a blueprint for rationally editing RNA modules, demonstrating how peripheral and core elements can be tuned to precisely control catalytic activity. These results not only illuminate fundamental principles of RNA design but also pave the way for potential applications of hovlinc variants in biotechnology and synthetic biology.

Results

Tertiary structure determination of hovlinc ribozyme

Compared with other reported self-cleaving ribozymes, the hovlinc ribozyme exhibits relatively lower cleavage activity (Chen et al., 2021), which enables the preparation of predominantly pre-cleaved RNA samples via *in vitro* transcription. To maintain the ribozyme in a low activity state prior to crystallization while ensuring proper folding, RNA samples were annealed in a neutral buffer containing 5 mM $MgCl_2$. Based on the sequence conservation analysis and biochemical characterization, we designed a series of catalytically active hovlinc variants for structural studies. Among these constructs, an 81-nucleotide variant, hovlinc-81 (termed as Hov-81 in the following description) yielded well-diffracting crystals, allowing tertiary structure determination (Figure 1A). In Hov-81, the native L2 loop is replaced with a variable loop, and the link regions between stems P1 and P4 is truncated to facilitate crystallization (Figure 1A and Supplementary Figure 1A). Cleavage assays indicated that Hov-81 retained robust catalytic activity, with a cleavage rate of 0.0139±0.00010 min^{-1} (Figure 1B), comparable to that of the previously reported catalytically active truncated Hovlinc ribozyme variant (Chen et al., 2021), which we designate as Hov-109 in this study

(Supplementary Figure 1A-B).

The crystallographic phase problem was solved by molecular replacement using the stem P4 (containing 9 bp) tertiary motif generated from coot as an initial search model. The structure of the Hov-81 was refined at 2.94 Å resolution, with R_{work} and R_{free} values of 0.22 and 0.26. Each asymmetric unit contains one single Hov-81 molecule (Mol A), which interacts with a neighboring symmetry-related molecule (Mol B) to form a dimer (Figure 1C and Supplementary Figure 1C). The overall architecture of this dimer, together with a schematic representation, are shown in Figure 1C and Figure 1D respectively.

In each monomer, the two segments P1-L1-P2-L2 and P4-L4 splay out at the variable linker between stems P1 and P4 and engage in long-range base-pairing interaction (PK1 and PK2) with the other monomer within the dimer (Figure 1C and Supplementary Figure 1C). To probe the intrinsic properties of hovlinc ribozyme cleavage site, Hov-81 construct was split between stems P1 and P4 to generate the presumed functional unit containing separate enzyme and substrate strands (Supplementary Figure 1D). To facilitate *in vitro* transcription, the first two base pairs at the 5′-ends of P1 and P4 were replaced with G-C pairs (Supplementary Figure 1D), followed by individual purification. Cleavage assays demonstrated that the resulting bimolecular ribozyme remains catalytically active, exhibiting a cleavage rate approximately one quarter that of unimolecular Hov-81 under a 1:1 substrate-to-enzyme ratio (Figure 1B). Based on these findings, we defined two complementary bimolecular assemblies within the dimeric tertiary structure (Figure 1C-D). Specifically, the P1-L1-P2-L2 segment of molecule A together with the P4′-L4′ region of molecule B constitutes a novel bimolecular assembly, designated molecule A′ (Mol A′) (Figure 1D). Conversely, the P1′-L1′-P2′-L2′ segment of molecule B together with the P4-L4 region of molecule A form a second bimolecular assembly, designated molecule B′ (Mol B′) (Figure 1D). Each bimolecular assembly contains a single functional cleavage site (Figure 1D).

Long-range interactions in the Hov-81 ribozyme structure

As mentioned above, the crystal structure reveals that Hov-81 forms a dimer, within which Mol A′ and Mol B′ constitute the functional catalytic units (Figure 1C-D). Given that Mol A′ and Mol B′ adopt identical interaction modes and structural organizations, our subsequent structural analysis and description focus on Mol A′ (Figure 2A), with its corresponding schematic secondary structure shown in Figure 2B. The overall structure comprises two nearly parallel, extended helices (Figure 2A-B). PK2 stacks coaxially on P1, forming an elongated helical assembly, termed Helix P1-PK2, while

PK1 bridges P4 and L2 into a contiguous helix designated Helix P4'-PK1-P2-L2 (Supplementary Figure 2A-B). The cleavage site is positioned at the structural center of the ribozyme (highlighted in yellow in Figure 2A-B), with both catalytic residues participating in the continuous stacking interaction of the two major helical elements.

Within the L1 segment, residues A6 and A40 form a non-canonical base pair (Figure 2C), which facilitates continuous stacking between P1 and PK2 (Figure 2A-B and Supplementary Figure 2A). In addition, C7 and G72 form a canonical Watson-Crick base pair that initiates PK1 formation and promotes stacking continuity between P4 and PK1 (Figure 2D, 2E and Supplementary Figure 2B). Furthermore, A14 and U32 in L1 form a canonical base pair (Supplementary Figure 2C), thereby enabling continuous stacking between PK1 and P2 (Figure 2A-B). Given that the A6-A40 and C7-G72 pairs are directly involved in formation of the catalytic pocket, detailed analyses of cleavage activity for the corresponding mutants will be provided in the following section. Here, we focused on the A14-U32 base pair, which mediates coaxial stacking between PK1 and P2. Despite reaching a cleavage rate comparable to ~92% of that of the wild-type (0.0129±0.00063 min^{-1}), the A14U mutant displayed a pronounced reduction in cleavage efficiency at the end of the reaction, dropping from ~90% (wild-type) to only ~60% (Supplementary Figure 3A).

Compared to the predicted secondary structure of hovlinc ribozyme (Figure 1A), the tertiary structure of Hov-81 reveals substantial structural rearrangements within PK1 and PK2 (Supplementary Figure 1D). PK1 is stabilized by seven base pairs (Figure 2E), among which the first base pair, C7-G72, directly contributes to formation of the catalytic pocket (Figure 2A). Within the PK1 stem, one sheared non-canonical base triplet, C68·(C11-A67) (Supplementary Figure 2D), is stacked between four previously predicted canonical base pairs below and two additional G-C base pair above. This arrangement stabilizes PK1 stem and enables its alignment with stem P2 and P4', forming the long helix P4'-PK1-P2-L2 (Figure 2B & Supplementary Figure 2B).

To detect the contribution of the C68·(C11-A67) base triple to the cleavage activity, we performed mutation analyses. Because the electron density for the base of C68 was not observed, suggesting it adopts a flexible conformation in the tertiary structure, we first deleted C68. Deletion of C68 (C68Del) caused a significant reduction in cleavage activity, yielding a rate of 0.0050±0.00006 min^{-1}, approximately 36% of that measured for Hov-81 (Supplementary Figure 3B). Mutations C11U and A67G, which converted the C11•A67 pair into either a U-A or C-G base pair, impaired the cleavage activity to a similar extent (0.0103±0.00104 min^{-1} and 0.0088±0.00069 min^{-1}) (Supplementary Figure 3C-D). Further deletion of C68 in C11U and A67G mutations severely reduced (C11U/C68Del, 0.0066±0.00085 min^{-1}) or nearly abolished

(A67G/C68Del) cleavage activity (Supplementary Figure 3E-F). These results suggest that the formation of the C68·(C11-A67) base triple is essential to stabilize the scaffold of the P4'-PK1-P2-L2 helix, or for providing backbone flexibility required for the L4 loop to participate the formation of both PK1 and PK2 concurrently.

PK2 consists of six base pairs (Figure 2F), deviating from the prediction by one nucleotide. Specifically, U32, originally predicted to participate in pairing to A61, instead forms a canonical Watson–Crick base pair with A14 (Supplementary Figure 2C), while G38·G55 constitutes the first base pair of PK2 (Figure 2G). The G38·G55 base pair stacks above the A6·A40, contributing to shaping the catalytic pocket (Figure 3A). Together, the formation of both PK1 and PK2 is critical for stabilizing the overall tertiary fold and maintaining catalytic activity. Consistent with this structural observation, previous studies have shown that disruption of base pairing in either pseudoknot abolishes ribozyme activity (Chen et al., 2021).

Identification of novel RNA structural motifs in the hovlinc ribozyme

RNA pseudoknots are the typical tertiary structural elements formed between nucleotides in a loop pair with complementary sequences outside the stem–loop. Unlike simple hairpins, pseudoknots create interwoven helices that bring distant regions of an RNA molecule into close spatial proximity, often stabilizing compact folds and enabling catalytic or regulatory functions. Over time, structural analyses have revealed that pseudoknots can adopt diverse topologies. The hovlinc ribozyme adopts a distinct and more complex architectural organization (Figure 2H). Rather than forming a single pseudoknot interaction, hovlinc establishes two independent pseudoknot pairing stems between the terminal loop L4′ of stem P4′ and the internal loop L1 (Figure 1A and 2A-B). These long-range interactions cooperatively constrain the global fold, effectively locking the RNA into a compact tertiary conformation. This arrangement gives rise to a unique "double-clamp" topology, in which the two pseudoknots act in concert to stabilize the relative orientation of the major helical elements and position the catalytic core at the structural center of the ribozyme (Figure 2H-I). Following the established pseudoknot nomenclature (Bon et al., 2008), the hovlinc ribozyme scaffold can be represented as ABCADCBD, with each letter denoting a distinct helix along the RNA backbone from the 5′ to 3′ end. We designate this previously unrecognized topology as the Rainbow-type (R-type) pseudoknot (Figure 2I).

R-type pseudoknots are topologically distinct from previously characterized pseudoknot classes, including H-, K-, L-, and M-types, which can be represented as ABAB, ABACBC, ABCABC and ABCADBCD, respectively (Figure 2J and

Supplementary Figure 4) (Bon et al., 2008; Kucharik et al., 2016). Classical H-type and K-type pseudoknots represent the canonical form, in which a hairpin loop base-pairs with a single-stranded sequence (H-type) or a stem-loop (K-type) located downstream, forming a second stem that crosses the first (Figure 2J and Supplementary Figure 4). L-type pseudoknots involve more complex strand trajectories, where the loop engages in base pairing that does not simply form a second stem opposite the first but instead threads around or alongside existing helices (Figure 2J and Supplementary Figure 4). Compared with these classical pseudoknot architectures, R-type pseudoknot exhibits a unique connectivity, in which the helices are interleaved in a noncanonical order with PK1 and PK2 mediating critical tertiary contacts (Figure 2H-I).

M-type pseudoknots were defined very early in theoretical classifications (Bon et al., 2008; Kucharik et al., 2016) and are characterized by multiple crossing interactions and/or additional nested helices layered onto the basic pseudoknot scaffold (Figure 2J and Supplementary Figure 4). Their architectures can accommodate more than two helical elements, forming intricate tertiary cores that stabilize global RNA folding and support specialized functions (Figure 2J). Both M-type and R-type pseudoknots consist of two stem–loops that form two pseudoknot stems oriented in distinct directions (Figure 2H-J and Supplementary Figure 4). To further investigate the tertiary structural differences between these two motifs, we performed a systematic search across all experimentally determined RNA structures. M-type pseudoknots scaffolds are presumed to exist based on theoretical classifications of RNA topology. However, no M-type pseudoknot corresponding scaffolds have been observed in currently reported experimentally determined RNA structures. Similarly, the R-type pseudoknot corresponding architectures have not been identified in any previously determined structures. Notably, the hovlinc ribozyme represents the first experimentally observed example of an R-type pseudoknot, thereby establishing it as a novel RNA structural motif (Figure 2I).

This previously unrecognized R-type pseudoknot in hovlinc ribozyme is not only structurally unique but also essential to support the organization of the catalytic core (Figure 2A and 2H-I), thereby expanding the known repertoire of RNA architectures and demonstrating how RNA can exploit multiple long-range pairing interactions to achieve precise spatial organization and biological function. To date, RNA structural space remains incompletely explored, and our current understanding still captures only a limited fraction of its architectural and functional diversity. The discovery of the new pseudoknot type suggests that additional, previously unrecognized RNA folds likely exist, with important implications for RNA biology, as well as for future efforts in molecular design and drug discovery.

Organization of the Catalytic Core

The cleavage site of hovlinc ribozyme is located within the internal loop L1 between P1 and PK1 in the secondary structure and is embedded in the core of the tertiary fold, where it lies at the junction of multiple stems—particularly between the pseudoknot stems PK1 and PK2 and the supporting stems P1 and P4′ (Figure 2A and 3A). At the cleavage site, the A6-C7 step adopts a splayed-apart conformation (Figure 3A-B). Both A6 and C7 intercalate into the two extended helices, Helix P1-PK2 and Helix P4'-PK1-P2-L2, respectively and participate in long-range interactions that define the architecture of the catalytic core (Figure 2A-B and Figure 3A-B). A6 forms a non-canonical reverse base pair with A40, which is further sandwiched between the first non-canonical base pair of PK2, G38-G55 and the terminal base pair of P1, U5-A41 (Figure 3A-B). This arrangement promotes continuous stacking across adjacent helical elements (stems P1 and PK2) and helps stabilize the local fold. C7 forms a canonical Watson-Crick pair with G72, constituting the first base pair of PK1 (Figure 3B), which nucleates formation of the PK1 pseudoknot and anchors the cleavage site within the core domain of the ribozyme. The C7-G72 base pair stacks directly on the terminal base pair of stem P4, A73-U54 (Figure 3A), facilitating coaxial stacking across the pseudoknot junction.

To identify residues involved in the formation and stabilization of the cleavage site, we performed structure-based residue-specific mutagenesis. Given A40 and A6 form a trans base pair at the cleavage site using their Watson-Crick edges (Figure 2C), we first mutated A40 and A6 to uridine to convert this interaction into the canonical trans A-U Watson-Crick base pair. However, the A6U and A40U mutants resulted in a n almost or complete loss of cleavage activity (Supplementary Figure 5A-B). We further generated A6C and A40C mutants to disrupt the pairing interaction, respectively. Cleavage assays show that both A6C and A40C mutants displayed approximately four- or three-fold reductions in cleavage rates (Supplementary Figure 5C-D). As shown in Figure 3A-B, C7 from L1 and G72 from L4 form a canonical Watson–Crick base pair at the cleavage site. Disruption of this canonical pairing by generation the C7G and G72C mutants resulted in a complete or near-complete loss of cleavage activity Supplementary Figure 5E-F). Further substitutions of C7 or G72 to A or U, which alter the pairing interactions, showed that C7A nearly abolished cleavage activity (Supplementary Figure 5G), whereas G72A and G72U retained about 30% of wild-type activity (Supplementary Figure 5 I-J). C7U, which can form a wobble base pair with G72, retained about 50% of wild-type activity (Supplementary Figure 5H).

Above the A6-A40 pair, G38 and G55 also form a trans base pair using their Watson-Crick edges and engage in hydrogen-bonding interaction (Figure 2G). To

further investigate the role of G38 in catalysis, we generated G38A, G38U, and G38C mutants. Substitution of G38 with either A or U abolished cleavage activity (Supplementary Figure 6A-B), which is consistent with previous studies that mutation of G38 to A, corresponding to a naturally occurring homolog of the hovlinc ribozyme in *gorillas* eliminates the cleavage activity (Chen et al., 2021). Although G38 is highly conserved in the sequence, our finding that the G38C mutant retains approximately 80% of cleavage activity (0.0113±0.00055 min) while with a substantial decrease in cleavage efficiency at the end of the reaction like A14U (Supplementary Figure 6C), suggesting that the ribozyme may tolerate a broader range of sequence variation than previously recognized. We noticed that the formation of the A6-A40 and G38-G55 base pairs extrudes C39 from the structure (Figure 3A). Although C39 does not participate directly in base interactions, deletion of C39 (C39Del) results in an almost complete loss of cleavage activity (Supplementary Figure 6D). We therefore propose that the presence of C39 provides additional spatial accommodation, which may be required to support the conformational geometry associated with catalysis. To examine this possibility, we measured the distances between the C1′ atoms of the paired nucleotides. The C1'–C1' distances for the G55–G38 and A6–A40 pairs are 13.8 Å and 14.2 Å, respectively (Supplementary Figure 7A-B, 7E), which are substantially longer than those of the observed canonical base pairs, such as G56–C37 (10.6 Å) and U5–A41 (10.4 Å) (Supplementary Figure 7C-D, 7E). Although the precise structural basis for these effects remains unclear, these mutational analyses suggest that the hovlinc ribozyme can tolerate substantial variation at the cleavage-site residues, distinguishing it from previously characterized ribozymes.

To identify the optimal sequence surrounding the cleavage site, we employed SELEX (Systematic Evolution of Ligands by EXponential enrichment) method to screen a sequence library centered on this region (Figure 3C-D). Sequencing of enriched clones revealed a consensus motif, with the top four sequences ranked as A40G, A40G/A6G, WT (A40/A6), and A6C, among which A40G emerged as the most frequently selected variant (Figure 3E). Based on this, we generated A40G mutant, which displayed markedly increased cleavage rates (0.0251±0.00099 min^{-1}) (Figure 3F and 3H), approximately 1.8-fold higher than that of Hov-81. The double mutant A6G-A40G restored cleavage activity to levels comparable to Hov-81 (Figure 3G-H), consistent with the SELEX results in which this variant ranked as the second-highest hit (Figure 3E), highlighting the importance of coordinated base-pairing interactions at the cleavage site. Considering the symmetric interaction between A6 and A40, we also tested the A6G single mutant, which exhibited minimal activity (Figure 3I), likely due to

a mismatch with C39 that disrupts proper catalytic pocket formation. We noticed that A6C was also enriched in SELEX as a top hit (Figure 3E), which retained catalytic activity around 25% in our previous assays (Supplementary Figure 5C).

Structure-guided transformation and redesign of the hovlinc ribozyme

As described above, Hov-81 forms dimer in the tertiary structure (Figure 1C-D and Supplementary Figure 1C). Based on our structural and cleavage analysis, we defined Mol A′ and Mol B′ as the functional catalytic units and confirmed that separating Hov-81 into substrate and enzyme strands preserves cleavage activity (Figure 1). However, the spatial separation between 3′ end of stem P1 to the 5′ end of stem P4 in the functional catalytic unit (Figure 2A) precludes the direct covalent linkage, as shown in Hov-81 sequence (Figure 1B). To further assess its solution behavior of Hov-81, we perform size-exclusion chromatography, which indicated that Hov-81 predominantly exists as dimer (Supplementary Figure 8A). This dimerization is likely attributable to the complete removal of the sequence linking stems P1 and P4. We therefore proposed that retaining R-type pseudoknot scaffold while introducing an additional linker between stem P1 and P4 could promote monomer formation. To test this, we perform gel-filtration analysis of Hov-109, a catalytically active variant, containing an additional stem P3 between stem P1 and P4 (Supplementary Figure 1A and 8B). In addition, we further test the full-length hovlinc ribozyme variant Hov-124, which contains the same additional stem P3 as in Hov-109 as well as an extra stem P5 (Figure 5F). Gel-filtration analysis indicated that both constructs predominately exist as monomers in solution while preserving cleavage activity (Figure 4A, 4D-F and Supplementary Figure 8B). Together, these results demonstrate that linker length could modulate the oligomeric state of the hovlinc ribozyme without substantially affecting catalytic function.

In the tertiary structure of Hov-81, two Hov-81 molecules undergo domain swapping at the linker between stems P1 and P4, resulting in intermolecular pseudoknot formation between loops L1 and L4. This arrangement brings the 5′ and 3′ ends of each Hov-81 molecule into close proximity, with stem P2 positioned at the terminus of the overall structure (Figure 1C-D and Supplementary Figure 1C). Guided by this tertiary organization, we connected the 5′ end of stem P1 to the 3′ end of stem P4 and opened the L2 stem–loop, thereby generating the Hov-81o construct (Figure 4A-B and Supplementary Figure 8C). Cleavage assays showed that Hov-81o retains catalytic activity comparable to that of Hov-81 (Figure 4C, 4F). Based on these observations, we introduced a GAAA tetraloop into Hov-124 and left the L2 stem-loop open as Hov-81o, generating the Hov-124o construct (Figure 4A and Supplementary

Figure 8D). Cleavage assays showed that Hov-124o retains catalytic activity comparable to that of Hov-124 (Figure 4C-F), while gel-filtration analysis indicated that it predominantly exists as a monomer in solution, similar to Hov-124 (Supplementary Figure 8B).

It is noted that in the tertiary structure, stem P2 stacks above PK1, with loop L2 positioned at the terminus of the molecule. To evaluate the impact of stem P2 on the activity of hovlinc ribozyme, we systematically truncated and mutated P2 and L2. In one construct, L2 was replaced with a GAAA tetraloop, which formed a dimer in solution while retaining ~80% of the cleavage activity (Hov-74, 0.0114±0.00015 min^{-1}) (Supplementary Figure 9A-C, 9G). Complete removal of P2 and retain GAAA tetraloop retained about half of activity (Hov-68, 0.0068±0.00019 min^{-1}) (Supplementary Figure 9D, 9G). We further replaced the GAAA tetraloop with two nucleotides (A and U) or removed both P2 and L2, generating two constructs designated Hov-66 and Hov-64, respectively (Supplementary Figure 9A). Cleavage assay showed that Hov-66 partially restored activity (0.0046±0.00059 min^{-1}), whereas Hov-64 retained only ~10% of the Hov-81 activity (0.0012±0.00014 min^{-1}) (Supplementary Figure 9E-G). These results suggest that a short linker within L1, separating the two segments that participate in PK1 and PK2 formation, is required to prevent steric interference during tertiary R-type pseudoknot formation. Proper spatial constraints and structural support are therefore essential for maintaining the catalytic activity of the hovlinc ribozyme.

Together, our structural and biochemical analyses not only support our structural observations but also demonstrate that the hovlinc ribozyme can be engineered into diverse architectures while preserving catalytic activity. Guided by its tertiary structure, we show that peripheral elements can be truncated, rearranged, or replaced to generate variants with distinct oligomeric states and structural properties. Importantly, such structural plasticity provides a framework for tailoring hovlinc into different formats—such as monomeric or modular constructs—suited for specific biochemical and biotechnological applications, highlighting its potential as a versatile platform for RNA engineering.

Structure prediction and modeling of the full-length hovlinc ribozyme

Biomolecular structure prediction has progressed rapidly in recent years, driven by advances in artificial intelligence and deep learning. The accuracy of protein structure prediction has been dramatically improved and are now being extended to RNA and other macromolecules (Abramson et al., 2024; Shen et al., 2024). As Hov-81 represents a newly determined RNA structure, we used it as a test case to evaluate the current performance of AlphaFold3 in RNA structure prediction (Figure 5A-E).

We performed tertiary structure prediction of Hov-81 using the full 81-nt sequence (Figure 1A) as input. As shown in Figure 5B and Supplementary Figure 10, the helical stems and pseudoknot interactions were successfully predicted; however, the overall backbone folding of the predicted models deviates significantly from the experimentally determined crystal structure. In the crystal structure, Hov-81 forms a dimer, in which each Hov-81 molecule adopts an extended conformation and exchanges their loops L1 and L4 with the opposing molecule to form the PK1 and PK2 stems, while L1 and L4′ adopt a “front-to-back” spatial arrangement (Figure 5A). In contrast, the AlphaFold3 models represent a monomeric conformation, with loops L4 and L2 intercrossing and nesting within each other (Figure 5B). Moreover, in the predicted structures, loops L1 and L4 exhibit severe steric clashes, and the conformation around the cleavage site is distinct from that observed in the crystal structure. For example, in the predicted model 1, although the A6-C7 step adopts an extended conformation, A6 interacts with C39, whereas by contrast A6 interact with A40 and C39 flips outward in the crystal structure (Figure 3A and Supplementary Figure 10A). In addition, U54, which forms the first base pair of stem P4 in the crystal structure, is instead predicted to pair with A40 (Supplementary Figure 10A). In the predicted model 2, A6 and C7 stack directly with each other; although C7 still forms a canonical Watson–Crick pair with G72, A6 interacts with A73, resulting in substantial steric hindrance (Supplementary Figure 10B).

Because structure prediction generally becomes more challenging for longer RNA sequences, and based on our observation that the split Hov-81 enzyme/substrate (E/S) system retains cleavage activity (Figure1B), we further used AlphaFold3 to predict the tertiary structure of Hov-81-E/S by providing the “enzyme” and “substrate” strands as separate inputs (Supplementary Figure 1D). The “substrate” strand was predicted to adopt a long helical conformation, while stem P4 and part of loop L4 in the “enzyme” strand also formed helical structures (Figure 5C-D). When the “enzyme” and “substrate” strands were predicted together, the overall shape of the model roughly resembled the crystal structure (Figure 5E); however, substantial discrepancies and unrealistic features remained. For example, in predicted models 1 and 2, the steric clash between loops L1 and L4 persisted (Supplementary Figure 11). Around the cleavage site, model 1 shared features with predicted model 2 of Hov-81, whereas in model 2, A6 and U54 exhibited clear steric conflict (Supplementary Figure 11B). Collectively, these results highlight a considerable gap between the AlphaFold3-predicted models and the experimentally determined crystal structure, indicating that the reliability of current predictions for complex RNA architectures still requires further improvement.

Although the prediction of Hov-81 encountered significant difficulties, we still tried to predict the structure of the full-length ribozyme, Hov-124 (Figure 5F and Supplementary Figure 1A), in order to gain additional structural insights. However, among the five representative output models, none successfully formed the PK1 and PK2 pseudoknots (Supplementary Figure 12). Apart from stem P4, the remaining helical elements—including P1, P2, and P0-were not properly predicted, resulting in overall structural models that appear clearly inconsistent with the experimentally determined architecture.

To obtain a tertiary structure model of Hov-124, we performed structural modeling based on the crystal structure of Hov-81 followed by energy minimization. As shown in Figure 5G, the calculated model places stem-loop P3-L3 stacked beneath the long P1-PK2 helix, while P5 stacks below the P4-PK1-P2-L2 module, yielding an overall fold highly similar to that observed in the crystal structure. Near the cleavage site, A6 and C7 adopt an extended conformation, with C7 forming a canonical Watson–Crick base pair with G72 that mediates continuous helical stacking between PK1 and P4 (Figure 5F & Supplementary Figure 13). A6 still interacts with A40; however, the interaction between G38 and G55 above the A6·A40 pair is weakened, retaining only a single hydrogen bond (Supplementary Figure 13B). In this configuration, the base of G55 is oriented nearly perpendicular, rather than parallel, to the base of A6 (Figure 3A and Supplementary Figure 13B). Moreover, the conformation of A6 differs markedly from that observed in the crystal structure, with its 2′-OH positioned closer to G55 than to the phosphate group of U5 (Figure 3B and Supplementary Figure 13B).

Discussion

Comparative sequence analysis revealed that homologs of hovlinc ribozyme are restricted to a subset of placental mammals (75 Xenarthra and Boreoeutheria species), with 22 species sharing over 90% sequence identity with the human ribozyme (Chen et al., 2021). Given that the linker between stems P1 and P4, as well as P2 and L2, is dispensable for catalysis, we focused on sequence variations within P1, L1, P4, and L4 elements (Supplementary Figure 14 and 15). Structure-guided analysis allowed us to classify the differences between inactive homologs and human hovlinc into four categories: (1) mutations in P1, which compromise stem stability and the formation of the catalytic pocket; (2) alterations in PK1 and PK2, which disrupt pseudoknot formation and destabilize the global fold; (3) changes in P4, which perturb stem integrity and overall tertiary structure; and (4) mutations at the cleavage site, which interfere with the precise geometry of the catalytic pocket, abolishing activity (Supplementary Figure 16 and Supplementary Table 3).

Previous studies have shown that among these 22 homologs, only the sequences from *bonobo* and *chimpanzee* exhibit cleavage activity (Chen et al., 2021). These two sequences are identical to that of human hovlinc ribozyme in the catalytic core region, with only minor variations in peripheral regions flanking P1 and P4 (Supplementary Figure 15). By contrast, the *gorilla* homolog carries a natural mutation corresponding to A38G in Hov-81 (Supplementary Figure 15), which abolishes activity entirely. Among the remaining 19 inactive homologs, we identified a conserved mutation in P1 at the terminal base pair, equivalent to A41C in Hov-81, highlighted by a yellow star in Supplementary Figure 14. As shown in the structure of Hov-81, the terminal base pair U5-A41 stacks beneath A6·A40 at the cleavage site and contributes to formation of catalytic pocket. Consistently, in vitro cleavage assays showed that the A41C mutant nearly abolishes activity (Supplementary Figure 6E). In addition to A41C, these inactive homologs carry diverse combinations of mutations in P1, PK1, PK2, and P4, collectively undermining catalytic function (Supplementary Figure 14-15 and Supplementary Table 3). Additionally, the hovlinc variant containing SNP rs72720496 (Supplementary Figure 15), which corresponds to mutation type 1 of PK1 (Supplementary Figure 16), has been shown to exhibit weaker activity than the wild type sequence.

Using SELEX, we identified a Hov-81 nucleotide variant, A40G, that enhances the cleavage activity of the hovlinc ribozyme by approximately two-fold (Figure 3F and 3H). Notably, variants including A40G, the double mutant A40G-A6G, and A40C reveal positions within the catalytic site that are highly tunable, providing valuable insights into the sequence plasticity of hovlinc and suggesting potential evolutionary trajectories in which subtle base changes could modulate catalytic efficiency without disrupting the overall fold. While the precise physiological function of hovlinc remains unknown, the discovery of activity-enhancing variants raises intriguing questions about how natural sequence variations may influence ribozyme function across species or within human populations. Moreover, these findings could guide the rational design of synthetic ribozymes for potential biotechnological applications, such as RNA-based sensors, gene regulation tools, or modular catalytic systems. Overall, the combination of SELEX-driven optimization and structural understanding not only illuminates the evolutionary potential of the hovlinc ribozyme, but also provides a blueprint for exploring and engineering functional RNA elements more broadly.

Hovlinc ribozyme exhibits remarkable intrinsic structural flexibility of RNA molecules. Our structure-guided mutation assays and SELEX experiments reveal that the cleavage site tolerates multiple mutations. In addition, our transformation and redesign engineering experiments indicate that the ribozyme's oligomeric state can be

modulated, highlighting the ribozyme's structural adaptability. Although the linker is not essential for catalytic activity of hovlinc ribozyme, it plays an important role in regulating the oligomeric state of the ribozyme, highlighting the intrinsic structural flexibility of RNA molecules. A shortened linker between stems P1 and P4 may restrict intramolecular pseudoknot formation between loops L1 and L4, whereas a longer linker provides the spatial flexibility required for proper folding of the monomeric hovlinc ribozyme. By modulating the linker length, the oligomeric state can be tuned, shifting between monomeric and dimeric assemblies. This observation underscores how subtle architectural elements, such as short connecting sequences, can influence higher-order RNA organization without abolishing catalytic activity. Such tunability may provide a useful strategy for engineering RNA molecules with controllable assembly states for future biochemical and biotechnological applications.

Despite rapid advances in AI-driven structure prediction, accurately modeling RNA structures remains a major challenge. One key limitation is the scarcity of high-resolution RNA structures, which constrains the training and validation of predictive algorithms. Limited structural information has historically constrained our understanding of RNA folding and functional motifs. The identification of the R-type pseudoknot in the hovlinc ribozyme represents the first example of a complex "double-clamp" tertiary architecture, in which two independent pseudoknots cooperatively stabilize the catalytic core. This discovery not only expands the known repertoire of RNA structural motifs but also provides a valuable framework for future RNA structural analysis, computational prediction, and de novo RNA design. In addition to the limited structural information, RNA molecules' distinctive features further complicate the tertiary structure prediction. Long-range tertiary interactions play critical roles in RNA folding, yet these interactions cannot be reliably inferred from base-pairing alone. Moreover, given RNA molecules are negatively charged, metal ions are always involved in their folding and structural stabilization, and often contribute directly to their functional activity. Although RNA is composed of only four nucleotides, its residues can participate in diverse hydrogen-bonding networks, generating intricate and highly context-dependent interactions that are difficult to capture computationally. For example, even in the best-performing predictions of the hovlinc ribozyme, AlphaFold3 approximately predicted the global backbone fold but failed to capture the detailed side-chain interactions critical for catalytic function. These observations underscore the urgent need for additional experimentally determined RNA structures, discovery of new RNA motifs, and rationally designed variants to guide and benchmark the next generation of RNA structure prediction and engineering.

Methods

RNA purification and Crystallization

All relevant sequences of the hovlinc ribozyme were cloned into the pUT7 vector downstream of the T7 promoter. To ensure homogeneity of the 5' end of the transcribed RNA, a hammerhead ribozyme or an HDV ribozyme sequence was inserted immediately after the hovlinc ribozyme sequence. All sequences used in this study are detailed in Supplementary Table 4. Following in vitro transcription, the RNA products were separated using Urea-PAGE. The target RNA bands were excised and purified by successive precipitation with isopropanol and ethanol to obtain high-purity RNA samples. Prior to crystal screening, the RNA samples at a concentration of 0.4 mM were annealed in a buffer containing 50 mM HEPES pH 7.0, 50 mM NaCl, and 5 mM $MgCl_2$, followed by cooling on ice for 30 min. Crystals appeared approximately one week after initial screening and were obtained under the condition including 0.1 M MES pH 6.0, 3.2 M $(NH_4)_2SO_4$.

X-ray data collection and refinement

To solve the phase problem, we performed co-crystallization of the Hov-81 RNA sample with $Ir(NH_3)_6^{3+}$ during optimization of the crystal growth conditions. Although no anomalous scattering signal from Ir was detected, we ultimately resolved the phase problem by molecular replacement (MR). Native X-ray diffraction data of Hov-81 were collected at 100 K using beamline BL-19U1 at the Shanghai Synchrotron Radiation Facility (SSRF) and processed with autoPX. The space group was $C222_1$. Using the P4 duplex generated in Coot as the initial search model for molecular replacement, along with auto-build and refinement in PHENIX and adjusted manually in Coot, we finally obtained the tertiary structure model of Hov-81. The X-ray statistics of the Hov-81 crystal is listed in Supplementary Table 1.

Cleavage assays and cleavage kinetics

To determine the observed first-order rate constant (k_{obs}) for the hovlinc ribozyme and its variants, the in vitro cleavage assay was performed. 5 μM RNA sample was refolded in a buffer containing 30 mM Tris-HCl pH 8.0 and 150 mM KCl. After incubation on ice for 15 minutes, the cleavage reaction was initiated by adding $MgCl_2$ to a final concentration of 10 mM. The reaction mixture was incubated at 37 °C for 0, 0.25, 0.5, 1, 2, 3, and 5 hours in a total volume of 10 μL. At each time point, the reaction was quenched by adding a three-fold volume of RNA loading buffer containing 7 M urea, 50 mM EDTA, and 0.1% xylene cyanol. An EDTA control was included, in which $MgCl_2$

was replaced with 10 mM EDTA (pH 8.0) and incubated for 5 hours.

The treated RNA samples were separated and detected on 15% or 20% Urea-PAGE. The gel was stained with 2× GelRed for 15 minutes at room temperature, imaged using Tanon Chemi Dog Ultra, and quantified with its Gel Image system. The k_{obs} values were obtained by nonlinear regression analysis in Prism using the equation $F_t = F_0 + F_\infty(1 - e^{-kt})$, where F_t represents the cleaved fraction at time t, F_0 and F_∞ denote the cleaved fractions at time zero and the reaction endpoint, and k corresponds to k_{obs}. All experiments were performed in duplicate. The k_{obs} of all variants are summarized in Supplementary Table 2.

Size Exclusion Chromatography

Prior to size exclusion analysis, the RNA sample was pretreated following the same protocol as that used for crystallization trials. The size exclusion chromatography was performed using a Superose™ 6 Increase 10/300 GL column with a buffer containing 50 mM HEPES pH 7.0, 50 mM NaCl, and 5 mM $MgCl_2$.

One-round SELEX on Hov-81 active center

We performed one round of SELEX targeting the active center of Hov-81. To facilitate the SELEX, we linked the 5' and 3' ends of the original Hov-81 sequence and introduced an opening terminal in the loop region of P2. First, we synthesized a random mutagenesis library containing N substitutions at positions G38, C39, A40, and G55. Through PCR, we introduced the T7 promoter and constructed four random DNA libraries (DNA-lib-A6/G6/C6/T6-4N) under A6, G6, C6, and T6 conditions, respectively. Subsequently, four initial RNA libraries (RNA-lib-A6/G6/C6/U6-4N) were obtained via in vitro transcription. All RNA libraries were annealed in a buffer containing 30 mM Tris-HCl pH 8.0, 150 mM KCl, and 10 mM $MgCl_2$ at a final concentration of 50 μM in a volume of 400 μL. The RNA samples were then subjected to a cleavage reaction at 37 °C for 3 hours, after which the reaction was stopped by adding RNA loading buffer. The above mixtures were separated on 15% Urea-PAGE. The bands corresponding to the cleaved 5' products were excised and eluted using 0.5×TAE buffer. The RNA samples were purified by successive precipitation with isopropanol and ethanol, lyophilized, and dissolved in DEPC-treated water. The cleaved products from the four RNA libraries were individually reverse transcribed, and the resulting cDNA was amplified by PCR to obtain DNA libraries of active sequences. Following purification, the four DNA libraries were pooled and sent to Shanghai Sangon Biotech for high-throughput sequencing.

Structure modeling

The full-length 3D structure of the 124-nt hovlinc ribozyme was modeled using the trRosettaRNA algorithm(Wang et al., 2023; Wang et al., 2026a; Wang et al., 2026b). Specifically, in addition to the nucleotide sequence, the secondary structure and the crystal structure of the core region (provided as an inter-nucleotide C1' distance map) were input into trRosettaRNA as geometric restraints. The predicted structures were then refined using a PyRosetta(Chaudhury et al., 2010)-based fast relaxation protocol to reduce structural violations, such as non-ideal bond lengths and steric clashes.

Data availability

The atomic coordinates and structure factors have been deposited to the Protein Data bank (www.rcsb.org) under the following accession codes 28CP (Hov-81 structure). All study data are included in the article and Supplementary Materials. Source data are provided with the paper.

References:

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J.*, et al.* (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature *630*, 493-500.

Bon, M., Vernizzi, G., Orland, H., and Zee, A. (2008). Topological classification of RNA structures. Journal of molecular biology *379*, 900-911.

Cervera, A., and de la Pena, M. (2020). Small circRNAs with self-cleaving ribozymes are highly expressed in diverse metazoan transcriptomes. Nucleic Acids Res *48*, 5054-5064.

Chaudhury, S., Lyskov, S., and Gray, J.J. (2010). PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. Bioinformatics *26*, 689-691.

Chen, L.L., and Kim, V.N. (2024). Small and long non-coding RNAs: Past, present, and future. Cell *187*, 6451-6485.

Chen, Y., Qi, F., Gao, F., Cao, H., Xu, D., Salehi-Ashtiani, K., and Kapranov, P. (2021). Hovlinc is a recently evolved class of ribozyme found in human lncRNA. Nature chemical biology *17*, 601-607.

Gilbert, W. (1986). Origin of life: The RNA world. Nature *319*, 618-618.

Jimenez, R.M., Polanco, J.A., and Luptak, A. (2015). Chemistry and Biology of Self-Cleaving Ribozymes. Trends Biochem Sci *40*, 648-661.

Kucharik, M., Hofacker, I.L., Stadler, P.F., and Qin, J. (2016). Pseudoknots in RNA folding landscapes. Bioinformatics *32*, 187-194.

Mattick, J.S., Amaral, P.P., Carninci, P., Carpenter, S., Chang, H.Y., Chen, L.L., Chen, R., Dean, C., Dinger, M.E., Fitzgerald, K.A.*, et al.* (2023). Long non-coding RNAs: definitions, functions, challenges and recommendations. Nat Rev Mol Cell Biol *24*, 430-447.

Roth, A., Weinberg, Z., Chen, A.G., Kim, P.B., Ames, T.D., and Breaker, R.R. (2014). A widespread self-cleaving ribozyme class is revealed by bioinformatics. Nat Chem Biol *10*, 56-60.

Shen, T., Hu, Z., Sun, S., Liu, D., Wong, F., Wang, J., Chen, J., Wang, Y., Hong, L., Xiao, J.*, et al.* (2024). Accurate RNA 3D structure prediction using a language model-based deep learning approach. Nat Methods *21*, 2287-2298.

Wang, W., Feng, C., Han, R., Wang, Z., Ye, L., Du, Z., Wei, H., Zhang, F., Peng, Z., and Yang, J. (2023). trRosettaRNA: automated prediction of RNA 3D structure with transformer network. Nature Communications *14*, 7266.

Wang, W., Liu, X., Peng, Z., and Yang, J. (2026a). The trRosettaRNA server for RNA structure prediction. Nature Protocols.

Wang, W., Peng, Z., and Yang, J. (2026b). Predicting RNA 3D structure and conformers using a pre-trained secondary structure model and structure-aware attention. Nature Machine Intelligence.

Weinberg, Z., Kim, P.B., Chen, T.H., Li, S., Harris, K.A., Lunse, C.E., and Breaker, R.R. (2015). New classes of self-cleaving ribozymes revealed by comparative genomics analysis. Nature chemical biology *11*, 606-610.

Acknowledgments

We thank the staff members of the Large-scale Protein Preparation System, BL-17B, BL-17U1, BL18U1, BL-19U1 and BL02U1 beamlines at the National Facility for Protein Science in Shanghai (NFPS), Zhangjiang Lab, China for providing technical support and assistance in data collection and analysis. We thank the staff of the BL02U1 beamline at the National Center for Protein Sciences Shanghai (NCPSS) at SSRF for their assistance in X-ray data collection. We thank the technical assistance from the core facility of the Life Sciences Institute (LSI), Zhejiang University. We thank the technical assistance from the core facility of the College of Life Sciences, Anhui Normal University.

Funding

This work was supported by the National Natural Science Foundation of China [32325029 to A.R., 32201057 to L.Z.], the National Key Research and Development Project of China [2023YFC2604300 to A.R.], the major program of Zhejiang Province [LD25C050002 to A.R.], the Anhui Provincial Natural Science Foundation [2508085Y014 to L.Z.], and the Austrian Science Fund FWF [F8011−B to R.M., 10.55776/F80].

Author contributions

L.Z. undertook all crystallographic experiments and collected the diffraction data with the assistance of W.L., Z.H., X.B. and H.Y. under the supervision of A.Ren. The structure was solved and analyzed by L.Z. and A.Ren, and L.Z. and A. Rázková performed SELEX and cleavage assays with the help of W.L, G.Z. and Y.T. under the supervision of A.Ren and R.M. The structure modeling and searching was performed by W. W., H. Z., Z. Q., and A. S. under the supervision of J.Y., P. X., D.P. and A.Ren. The paper was written jointly by L.Z. and A.R. with input from the remaining authors.

Competing interests

The authors declare no competing interests.

Figure Caption

Figure 1 I Self-cleavage activity and tertiary structure of the Hov-81 ribozyme.

A. Sequence and reported secondary structure of Hov-81. Each nucleotide is color-coded according to its structural segment in the tertiary structure shown in Figure 1C-D. The cleavage site is highlighted in yellow, and a red triangle indicates the self-cleavage site.

B. Time-course cleavage assays of Hov-81 and Hov-81-E/S analyzed by denaturing urea–PAGE, together with the corresponding kinetic profiles. Hov-81 is the construct used for crystallization, whereas Hov-81-E/S is a structure-guided bimolecular ribozyme designed for functional and structural analyses. In Hov-81-E/S, the substrate strand (Hov-81-S) corresponds to nucleotides 1-45 of Hov-81, while the enzyme strand (Hov-81-E) corresponds to nucleotides 46-81 of Hov-81.

C. Hov-81 ribozyme adopts a dimeric tertiary fold, composed of one molecule (Mol A) and a symmetry-related molecule (Mol B). Mol B is shown in a lighter shade than Mol A. The red dashed line indicates the division of the dimer into two functional monomers, termed Mol A′ and Mol B′, as illustrated in Figure 1D.

D. Schematic representation of the Hov-81 ribozyme dimer. Two Hov-81 molecules assemble into a dimer through reciprocal interactions between the L1 region of one molecule and the L4 region of the other, forming two pseudoknots that constitute a double-pseudoknot architecture. The cleavage site is highlighted in yellow. For clarity, the dimer is divided by a red dashed line into two functional monomeric units, termed Mol A′ and Mol B′, which exhibit identical tertiary interactions and catalytic architectures.

Figure 2 I Overall structure of the two-stranded Hov-81-E/S and long-range interaction between L1 and L4'.

A. Tertiary fold of Mol A′ of the Hov-81 ribozyme. The color coding is consistent with that in Figure 1C-D. The cleavage site is labeled in yellow. L1 in Mol A forms PK1 and PK2 with L4′ from Mol B.

B. Schematic representation of Mol A′ of the Hov-81 ribozyme. The interactions within PK1 and PK2 are highlighted in red.

C. Residues A6 at the cleavage site forms a non-canonical trans base pair with A40 in L1 through their Watson–Crick edges above stem P1.

D. Residue C7 from L1 form a canonical cis Watson–Crick pair with G72 from L4′ at the cleavage site above the P4 stem.

E. Residues C7–C13 from L1 form a seven-base-pair pseudoknot (PK1) with residues 65–72 in L4′. In the previously reported secondary structure, PK1 consisted of only

four base pairs (C7–A10 and U69–U70). The crystal structure reveals that PK1 is extended by three additional base pairs.

F. Residues C33–G38 in L1 form a six-base-pair PK2 with residues G55–G60 in L4′. In the previously reported secondary structure, PK2 comprised four base pairs (C33–C37 paired with G56–G60). The crystal structure reveals that PK2 is extended by one additional base pair.

G. G38 in L1 and G55 in L4′ form a non-canonical base pair through hydrogen-bonding interactions involving their Watson–Crick edges.

H. Schematic representation of the stem architecture of Hov-81 and formation of the two pseudoknots (PK1 and PK2).

I. Topology and secondary structure of the Rainbow-type (R-type) pseudoknot identified in the Hovlinc ribozyme. Each colored arc represents a distinct helix. The R-type topology is represented as ABCADCBD, where each letter (A, B, C, etc.) denotes a distinct helix encountered along the RNA backbone from the 5′ to 3′ end (Bon et al., 2008; Kucharik et al., 2016). The connectivity pattern of R-type pseudoknot is distinct from previously characterized pseudoknot architectures.

J. Topologies and secondary structures of H-, K-, L-, and M-type pseudoknots. These pseudoknot architectures can be uniquely represented as ABAB, ABACBC, ABCABC, and ABCADBCD, respectively (each letter corresponding to a helix as shown in Panel **I**).

Figure 3 I Organization of Hov-81 ribozyme cleavage site and cleavage assays on Hov-81 mutants.

A. Overall view of the catalytic center of the Hov-81 ribozyme.

B. Detailed view of the cleavage site. A6 and C7 adopt an extended conformation, with A6 forming a hydrogen bond with A40 and C7 interacting with G72 via Watson–Crick edges. The 2'-OH of A6 is positioned near the phosphate oxygen of U5.

C. Design of the random library for SELEX targeting the catalytic center of Hov-81. Core residues involved in non-canonical base pairs and the outwardly flipped C39 were randomized (G38N, C39N, A40N, A55N) during DNA library synthesis. A6N was introduced through four separate DNA libraries (DNA-lib-A6/C6/G6/T6-4N) generated by PCR.

D. Simplified flowchart of the SELEX workflow. The DNA library containing randomized residues was chemically synthesized, followed by addition of a T7 promoter and A6N sequence via PCR. The RNA library was generated by in vitro transcription, and cleaved 5′ products containing the randomized positions were separated from uncleaved RNA by 15% urea-PAGE. Cleaved RNA was reverse-transcribed into cDNA,

and double-stranded DNA libraries were prepared by PCR for sequencing.

E. Top four enriched sequences from high-throughput sequencing. The third-ranked sequence corresponds to the wild-type Hov-81 ribozyme.

F-I. Time-course cleavage assays of A40G, the A6G/A40G double mutant, and A6G on denaturing urea-PAGE gels. The A40G mutant (Top1 sequence) exhibits significantly enhanced cleavage activity. The A6G/A40G double mutant (Top2 sequence) restores activity to wild-type levels. The A6G mutant displays minimal cleavage activity.

Figure 4 l Structure-guided rational redesign of peripheral region the Hov-81 ribozyme.

A. Schematic illustration of structure-guided remodeling of peripheral elements based on the Hov-81 scaffold. Three design strategies were explored: linker modification, terminal extension, and sequence opening or downsizing at P2-L2. Linker modifications alter the distance between stems P1 and P4, ranging from no linker to incorporation of the P3 region. Terminal extensions progressively add 5′ and 3′ flanking sequences, including the P5 stem. Sequence reopening redesigns the connectivity of the ribozyme by opening the P2/L2 region and joining the original 5′ and 3′ termini. Hov-81 contains no linker between P1 and P4. Hov-124 incorporates the P3 region between P1 and P4, and extends the terminal by adding the P5 stem. Hov-81o is generated by joining the original 5′ and 3′ termini of Hov-81 while opening the P2/L2 region. Hov-124o is generated by connecting the termini of Hov-124 through a GAAA tetraloop within P5 while simultaneously opening the P2/L2 region.

B. The schematic of Hov-81o variant.

C–E. Time-course cleavage assays of Hov-81o, Hov-124, and Hov-124o analyzed by denaturing urea-PAGE.

F. Kinetic profiles of Hov-81, Hov-81o, Hov-124, and Hov-124o.

Figure 5 l Prediction and modeling of the hovlinc ribozyme.

A. Backbone trajectory observed in the Hov-81 crystal structure. During pseudoknot (PK) formation, the Hov-81-E (enzyme) and Hov-81-S (substrate) strands adopt a front-to-back arrangement.

B. Alignment of five AlphaFold3-predicted Hov-81 structures. Residues 1–45 correspond to the Hov-81-S-strand, whereas residues 46–81 correspond to the Hov-81-E-strand. Formation of the double pseudoknot results in strand crossing between the S and E strands, generating a topologically knotted architecture.

C–D. Alignments of the Hov-81-S-strand (C) and Hov-81-E-strand (D) predicted by

AlphaFold3, highlighting conformational variability.

E. Alignment of five AlphaFold3-predicted Hov-81-E/S structures using Hov-81-E-strand and Hov-81-S-strand as separate inputs.

F-G. Schematic and modeled tertiary structure of Hov-124. The Hov-124 structure was modeled based on the experimentally determined Hov-81 tertiary fold (Figure 2A). Compared with Hov-81, Hov-124 includes an additional P3 region and a P5 segment while retaining all core secondary structure elements of the wild-type Hov-168.

H. Structural alignment of modeled Hov-124 with the Hov-81 crystal structure, illustrating overall conservation of the catalytic core.

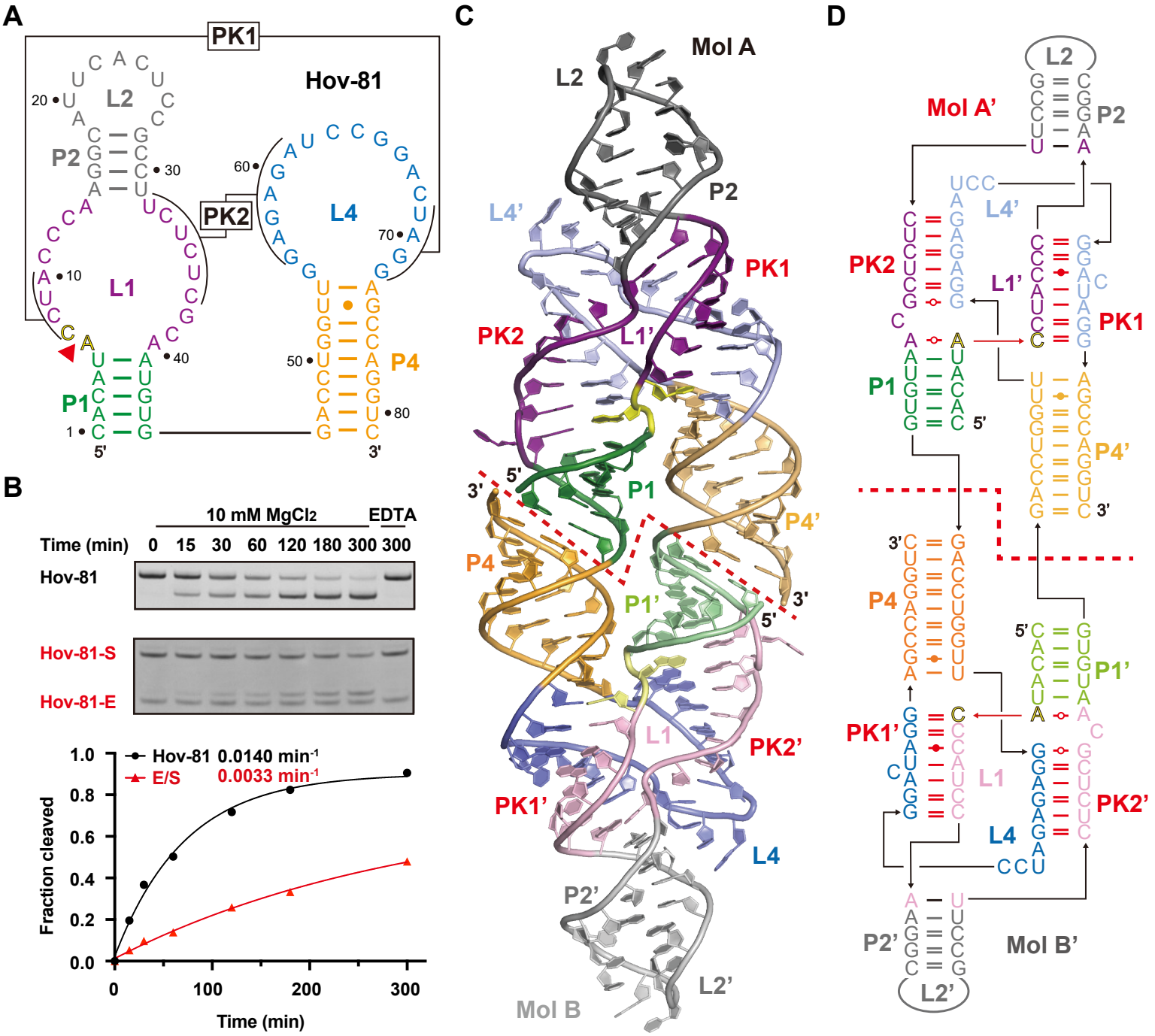

A
PK1
Hov-81
L2
P2
PK2
L4
L1
P1
P4
5'
3'
B
10 mM MgCl2
EDTA
Time (min) 0 15 30 60 120 180 300 300
Hov-81
Hov-81-S
Hov-81-E
Hov-81 0.0140 min-1
E/S 0.0033 min-1
Fraction cleaved
Time (min)
C
Mol A
Mol B
L2
P2
L4'
PK1
PK2
L1'
P1
P4'
P4
P1'
L1
PK2'
PK1'
L4
P2'
L2'
D
Mol A'
Mol B'
L2
P2
L4'
PK2
L1'
PK1
P1
P4'
P4
P1'
PK1'
L1
PK2'
L4
P2'
L2'

Figure 1

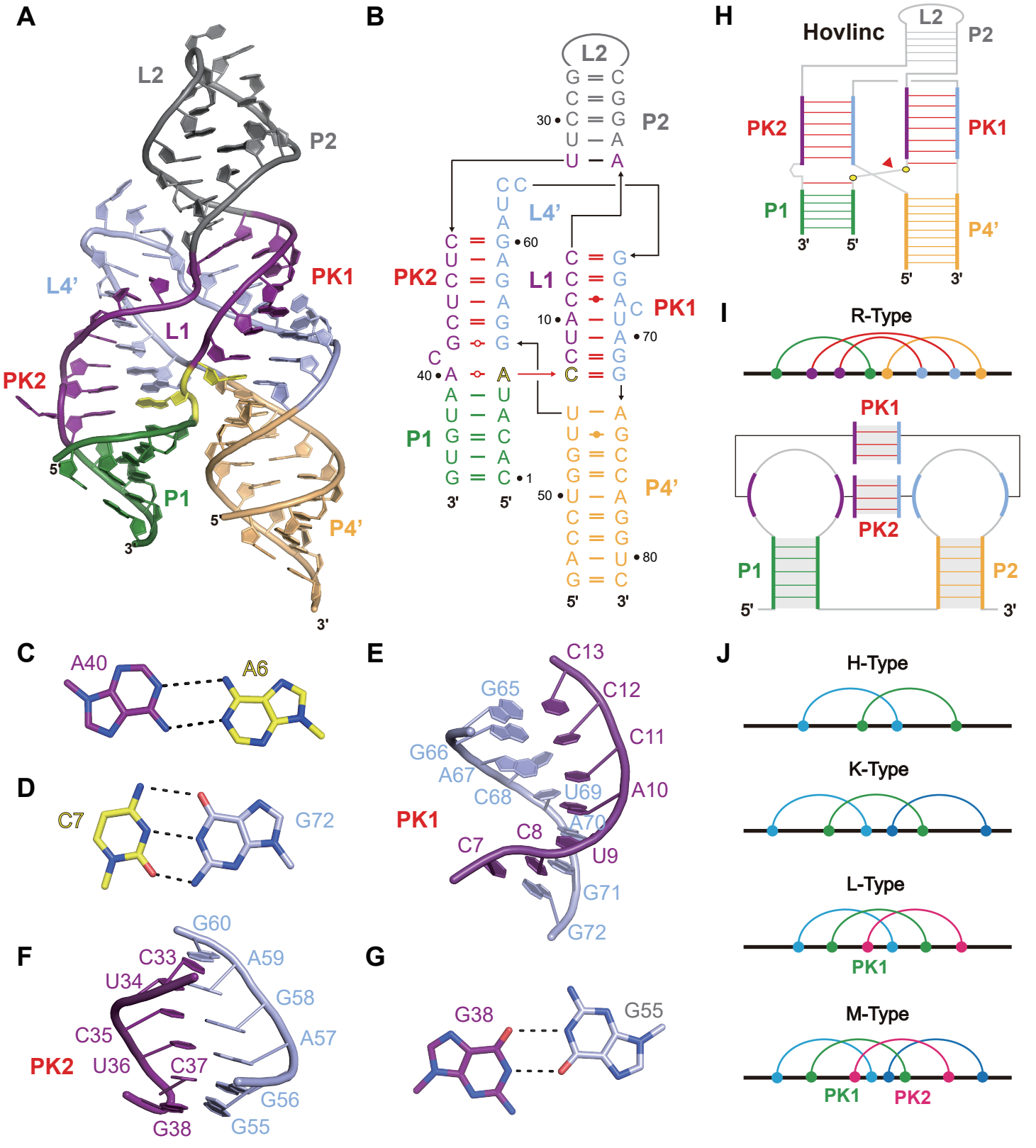

A
L2
P2
L4'
PK1
L1
PK2
5'
P1
3'
5'
P4'
3'
B
L2
P2
PK2
L4'
L1
PK1
P1
P4'
H
Hovlinc
L2
P2
PK2
PK1
P1
P4'
3'
5'
5'
3'
I
R-Type
PK1
PK2
P1
P2
5'
3'
C
A40
A6
D
C7
G72
E
C13
G65
C12
C11
G66
A67
C68
U69
A10
PK1
A70
C7
C8
U9
G71
G72
J
H-Type
K-Type
L-Type
PK1
M-Type
PK1
PK2
F
G60
C33
A59
U34
G58
C35
A57
PK2
U36
C37
G56
G38
G55
G
G38
G55

Figure 2

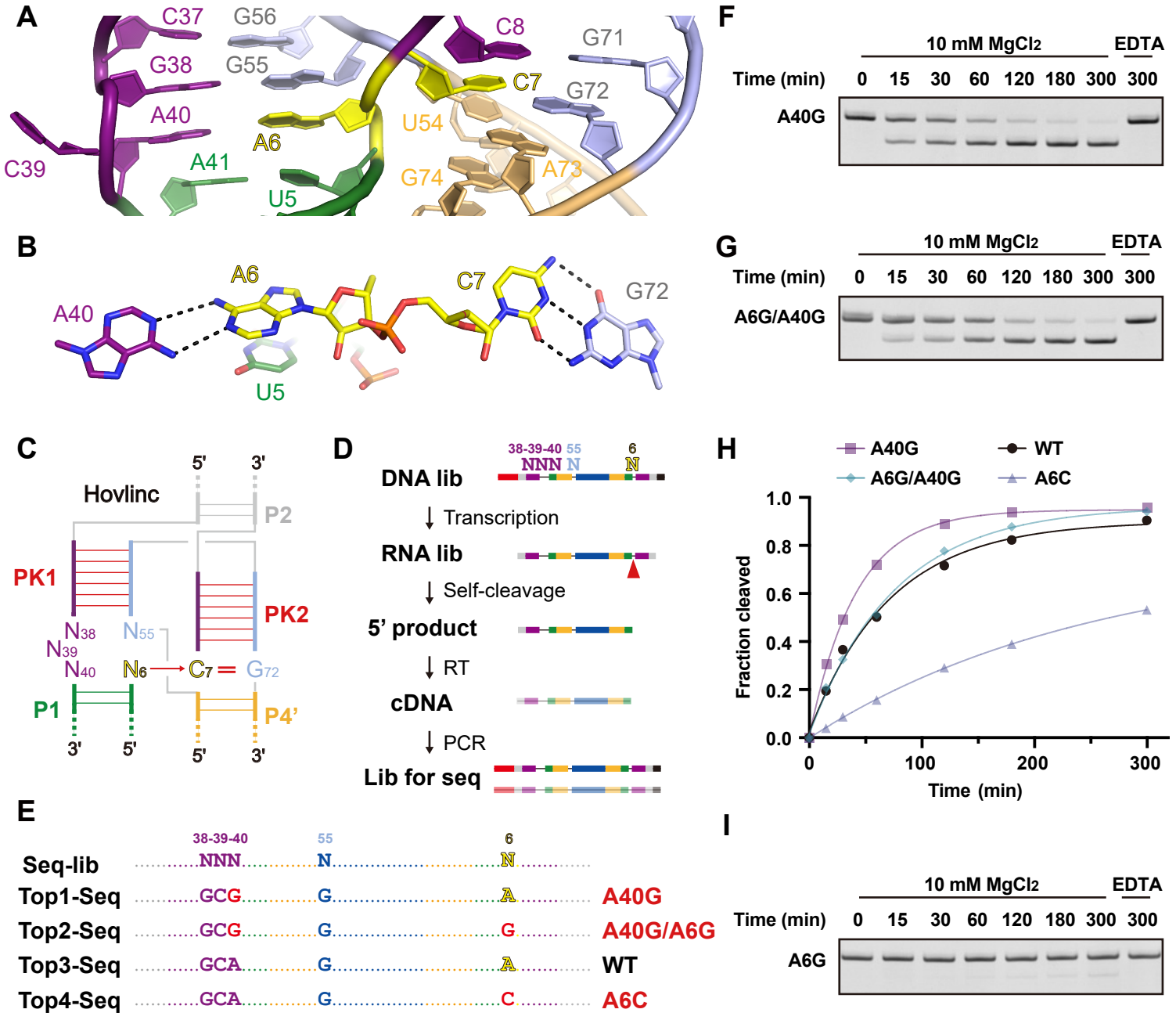

A
C37
G56
C8
G71
G38
G55
C7
G72
A40
A6
U54
C39
A41
G74
A73
U5
B
A6
C7
G72
A40
U5
C
5'
3'
Hovlinc
P2
PK1
PK2
N38
N55
N39
N40
N6
C7
G72
P1
P4'
3'
5'
5'
3'
D
38-39-40 55 6
NNN N N
DNA lib
Transcription
RNA lib
Self-cleavage
5' product
RT
cDNA
PCR
Lib for seq
E
38-39-40 55 6
Seq-lib NNN N N
Top1-Seq GCG G A A40G
Top2-Seq GCG G G A40G/A6G
Top3-Seq GCA G A WT
Top4-Seq GCA G C A6C
F
10 mM MgCl2
EDTA
Time (min) 0 15 30 60 120 180 300 300
A40G
G
10 mM MgCl2
EDTA
Time (min) 0 15 30 60 120 180 300 300
A6G/A40G
H
A40G
WT
A6G/A40G
A6C
Fraction cleaved
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
Time (min)
I
10 mM MgCl2
EDTA
Time (min) 0 15 30 60 120 180 300 300
A6G

Figure 3

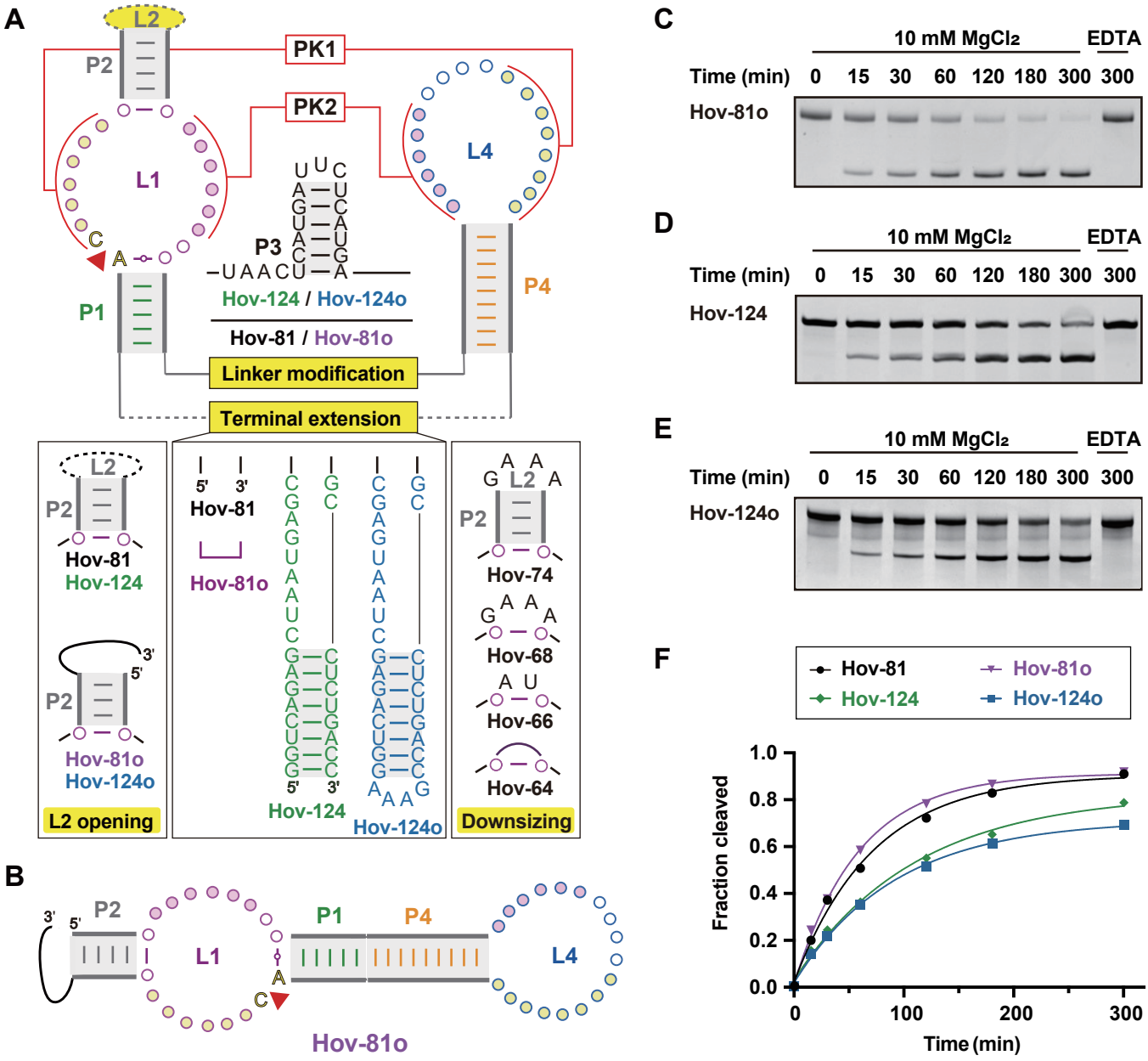

A
L2
P2
PK1
PK2
L1
L4
P3
P1
P4
Hov-124 / Hov-124o
Hov-81 / Hov-81o
Linker modification
Terminal extension
Hov-81
Hov-124
Hov-124
Hov-81o
Hov-124o
L2 opening
Hov-74
Hov-68
Hov-66
Hov-64
Downsizing
B
P2
P1
P4
L1
L4
Hov-81o
C
10 mM MgCl₂
EDTA
Time (min)
0 15 30 60 120 180 300 300
Hov-81o
D
Hov-124
E
Hov-124o
F
Hov-81
Hov-81o
Hov-124
Hov-124o
Fraction cleaved
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
Time (min)

Figure 4

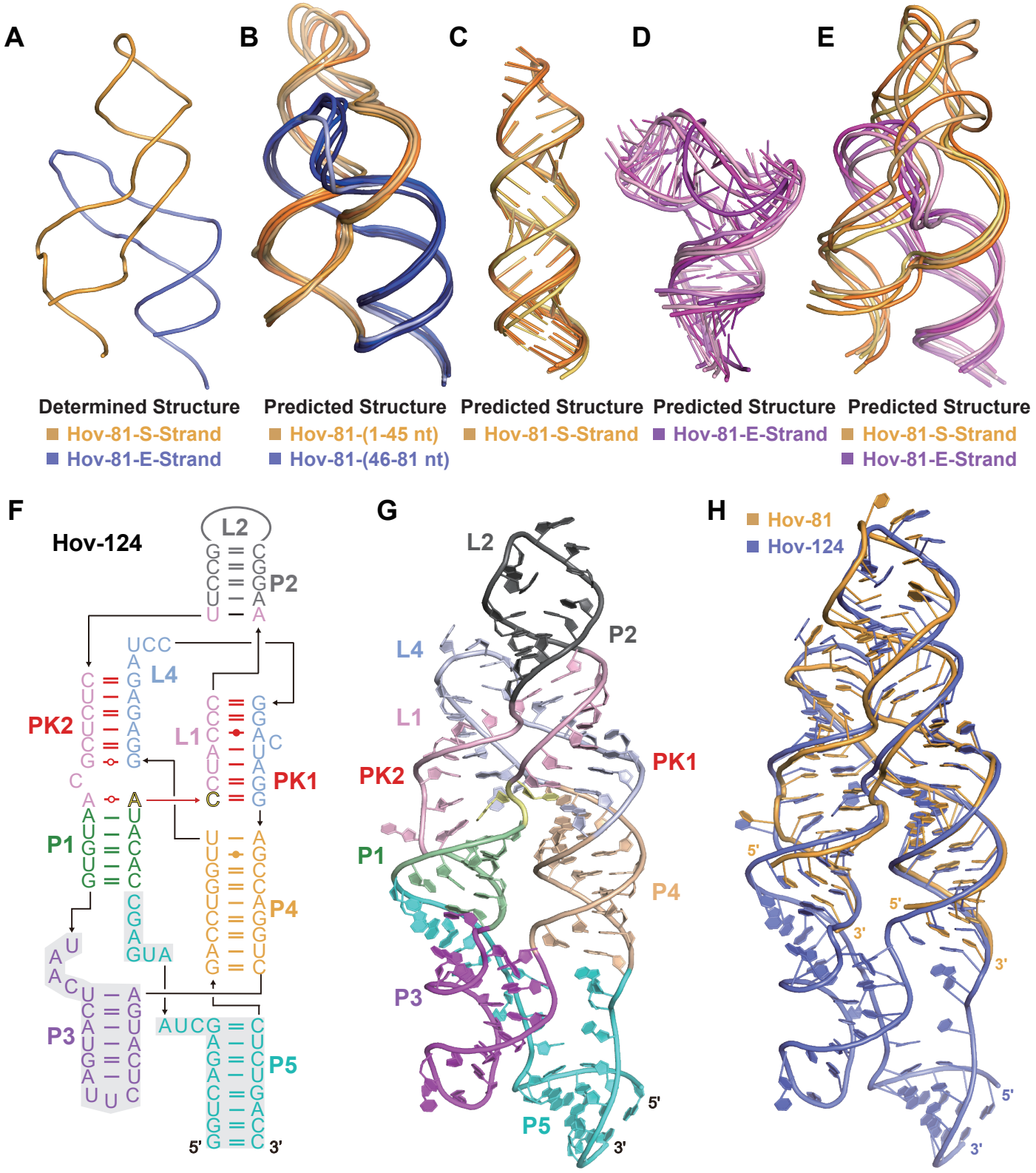

A
B
C
D
E
Determined Structure
Hov-81-S-Strand
Hov-81-E-Strand
Predicted Structure
Hov-81-(1-45 nt)
Hov-81-(46-81 nt)
Predicted Structure
Hov-81-S-Strand
Predicted Structure
Hov-81-E-Strand
Predicted Structure
Hov-81-S-Strand
Hov-81-E-Strand
F
Hov-124
L2
P2
L4
PK2
L1
PK1
P1
P4
P3
P5
5'
3'
G
L2
P2
L4
L1
PK2
PK1
P1
P4
P3
P5
5'
3'
H
Hov-81
Hov-124
5'
3'
5'
3'
5'
3'

Figure 5

Supplementary Information

A distinct pseudoknot fold defines the hovlinc self-cleaving ribozyme in human lncRNA

Luqian Zheng[1,2,*], Wenfei Li[1], Wenkai Wang[3], Heqin Zhu[4], Zejia Hu[1], Guiqi Zhi[2], Yuting Tang[2], Xinyue Bao[1], Huiqin You[1], Anna Rázková[5], Abhishek Suman[6], Zhenghao Qiao[7], Ronald Micura[5], Dinshaw Patel[6], Jianyi Yang[3], Peng Xiong[4] and Aiming Ren[1,*]

[1]Department of Cardiology of The Second Affiliated Hospital and Life Sciences Institute and School of Medicine and Liangzhu Laboratory, Zhejiang University, Hangzhou 310058, China

[2]College of Life Sciences, Anhui Provincial Key Laboratory of Molecular Enzymology and Mechanism of Major Metabolic Diseases, Anhui Normal University, Wuhu 241000, Anhui, China

[3]MOE Frontiers Science Center for Nonlinear Expectations, Shandong Province Key Laboratory of Financial Risk, Research Center for Mathematics and Interdisciplinary Sciences, Shandong University, Qingdao, 266237, China.

[4]Suzhou Institute for Advanced Research, University of Science and Technology of China, Suzhou, Jiangsu 215123, China

[5]Institute of Organic Chemistry, Center for Molecular Biosciences Innsbruck, Leopold Franzens University, Innsbruck A6020, Austria

[6]Structural Biology Program, Memorial Sloan-Kettering Center, New York, New York 10065, USA.

[7]Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou 510623, Guangdong, China

*To whom correspondence should be addressed. Tel: 86-571-88981228; Fax: 86-571-88981336;

Email: aimingren@zju.edu.cn (AR) and luqianzheng@ahnu.edu.cn (LZ)

Contents

Supplementary Tables

Supplementary Figures

Supplementary Tables

Supplementary Table 1 | Crystallographic statistics of hovlinc

Crystal	Hovlinc-81
Data collection	
Space group	*C* 222_1
a, b, c (Å)	49.08, 50.03, 209.26
α, β, ϒ (°)	90, 90, 90
Wavelength(Å)	1.099
Resolution (Å)	34.89-2.93(3.09-2.93)
R_{pim}	0.03(0.22)
I/dI	18.9(3.5)
Completeness (%)	99.9(99.3)
Redundancy	12.3(13.1)
$CC_{1/2}$	0.999(0.977)
Refinement	
Resolution (Å)	34.88-2.93(3.05-2.93)
No.reflections	5799
R_{work}/R_{free} (%)	21.9/25.7
No. of atoms	
RNA	1716
Mg^{2+}	1
Water	4
***B*-factors(Å^2)**	
RNA	98.8
Mg^{2+}	66.3
Water	57.3
R.m.s deviations	
Bond lengths (Å)	0.006
Bond angles (°)	1.27

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

Supplementary Table 2 | Summary of activity of hovlinc

Sequence	k_{obs}
Hov-81	0.0139±0.00010 min^{-1}
Hov-81-E/S	0.0032±0.00008 min^{-1}
Hov-124	0.0089±0.00019 min^{-1}
Hov-109	0.0093±0.00027 min^{-1}
Hov-85	0.0032±0.00021 min^{-1}
Hov-74	0.0114±0.00015 min^{-1}
Hov-68	0.0068±0.00019 min^{-1}
Hov-66	0.0046±0.00059 min^{-1}
Hov-64	0.0012±0.00014 min^{-1}
Hov-81o	0.0165±0.00067 min^{-1}
Hov-124o	0.0106±0.00033 min^{-1}
A6G	Almost loss of activity
A6C	0.0036±0.0004 min^{-1}
A6U	Loss of activity
C7G	Loss of activity
C7A	Almost loss of activity
C7U	0.0066±0.00019 min^{-1}
A40G	0.0251±0.00099 min^{-1}
A40U	Loss of activity
A40C	0.0046±0.00069 min^{-1}
G72A	0.0041±0.00027 min^{-1}
G72U	0.0040±0.00054 min^{-1}
G72C	Almost loss of activity
C39Del	Almost loss of activity
G38A	Loss of activity
G38C	0.0113±0.00055 min^{-1}
G38U	Loss of activity
C11U	0.0103±0.00104 min^{-1}
C68Del	0.0050±0.00006 min^{-1}
A67G	0.0088±0.00069 min^{-1}
C11U/C68Del	0.0066±0.00085 min^{-1}
A67G/C68Del	Almost loss of activity
A14U	0.0129±0.00063 min^{-1}
A41C	Almost loss of activity
A6GA40G	0.0135±0.00024 min^{-1}

Supplementary Table 3 | Summary of variations in human hovlinc homologous sequences

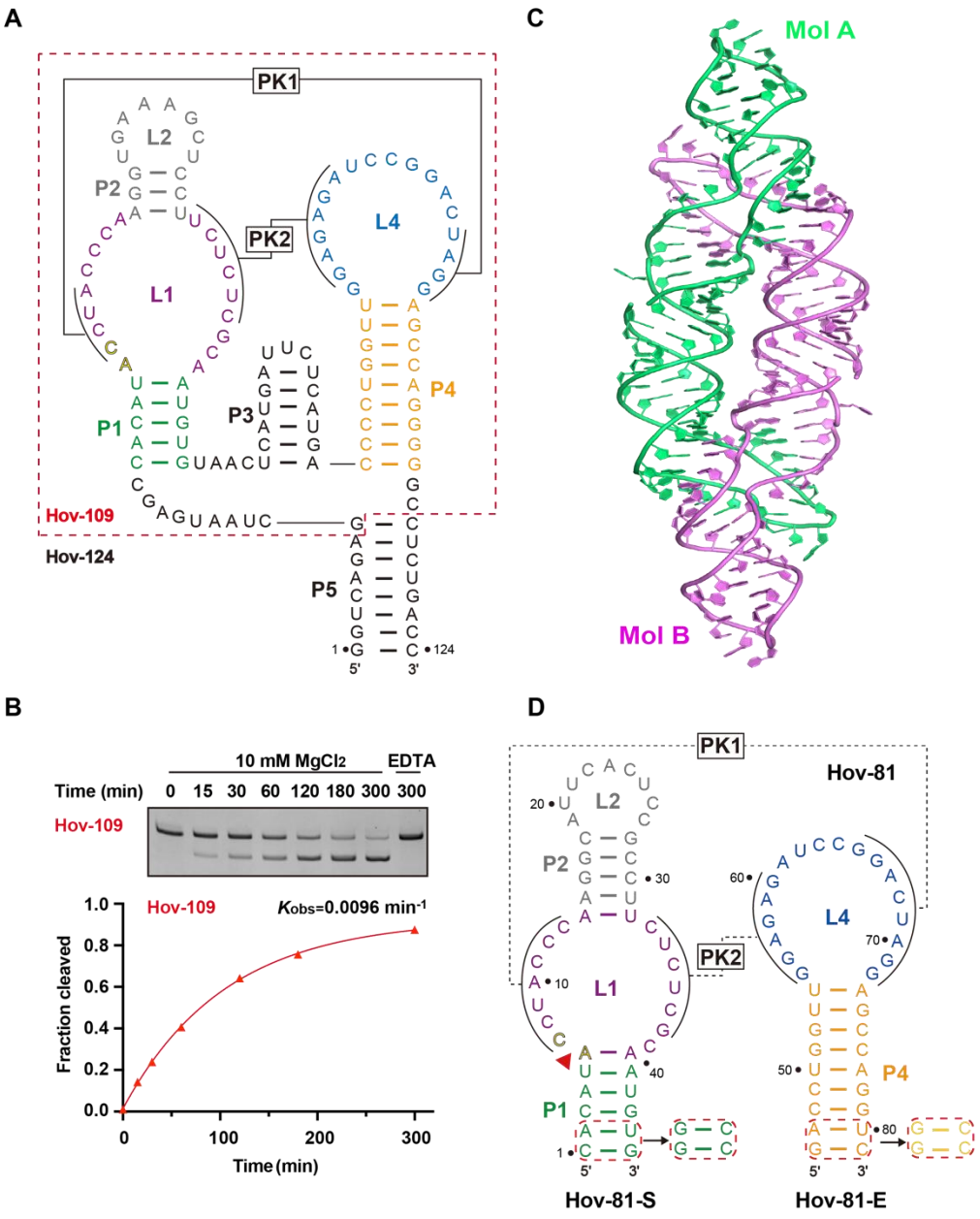
Species	P1	PK1	PK2	P4	Cleavage site
Human	/	/	/	/	/
Human SNP rs72720496	/	PK1-M1	/	/	/
Bonobo	/	/	/	/	/
Chimpanzee	/	/	/	/	/
Gorilla	/	/	PK2-M1	/	/
Gibbon	P1-M1	/	/	P4-M1	/
Orangutan	P1-M2	PK1-M2	/	/	/
Bolivian_squirrel_monkey	P1-M1	/	PK2-M2	P4-M2	M1
Capuchin	P1-M1	/	PK2-M3	P4-M3	M1
Ma's_night_monkey	P1-M1	/	PK2-M4	P4-M4	M1
Marmoset	P1-M3	PK1-M3	PK2-M5	/	M1
Ugandan_red_Colobus	P1-M3	PK1-M4	PK2-M6	P4-M5	/
Vervet-AGM	P1-M3	PK1-M4	PK2-M6	P4-M6	/
Angola_colobus	P1-M1	PK1-M4	PK2-M6	P4-M7	/
Gelada	P1-M3	PK1-M4	PK2-M6	P4-M8	/
Golden_snub-nosed_monkey	P1-M1	PK1-M4	PK2-M6	P4-M7	/
Hamadryas_Baboon	P1-M4	PK1-M4	PK2-M6	P4-M8	/
Olive_baboon	P1-M4	PK1-M4	PK2-M6	P4-M8	/
Sooty_mangabey	P1-M3	PK1-M4	PK2-M6	P4-M8	/
Black_snub-nosed_monkey	P1-M1	PK1-M4	PK2-M6	P4-M7	/
Crab-eating_macaque	P1-M3	PK1-M5	PK2-M6	P4-M8	/
Drill	P1-M3	PK1-M4	PK2-M6	P4-M8	/
Pig-tailed_macaque	P1-M3	PK1-M5	PK2-M6	P4-M8	/
Macaque	P1-M3	PK1-M5	PK2-M6	P4-M8	/

Supplementary Table 4 | Summary of sequences of hovlinc variants

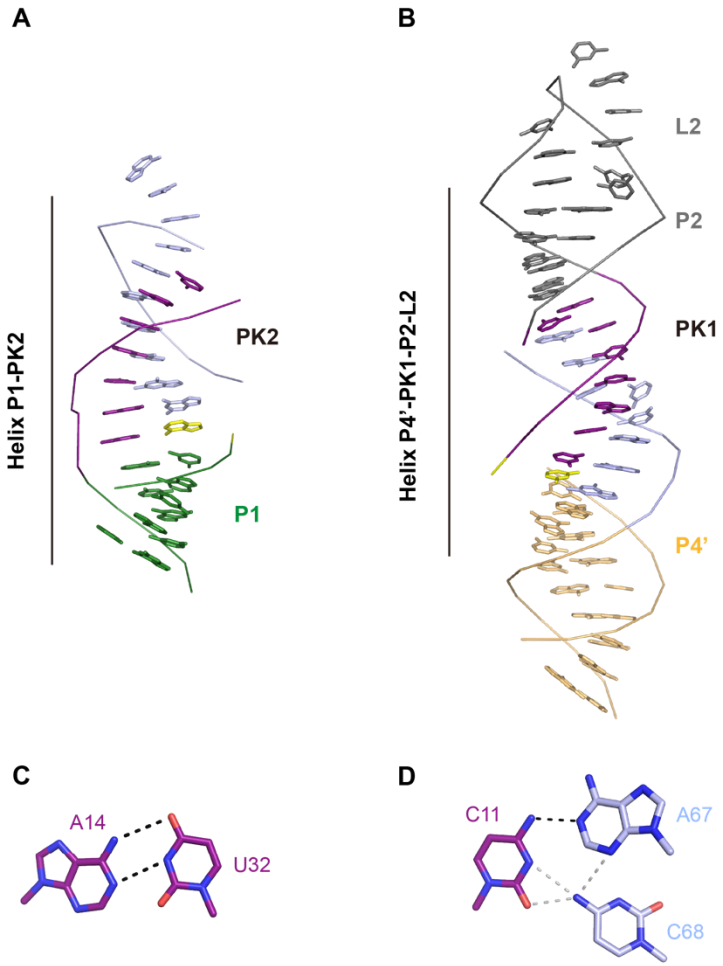
Name	sequences (5' to 3')
Hov-81	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
Hov-81-S	GGCAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGCC
Hov-81-E	GGCCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGCC
Hov-124	GGUCAGAGCUAAUGAGCCACAUACCUACCCAAGGUGAAAGCUCCUUC UCUCGCAAUGUGUAACUCAUGAUUCUCAUGACCCCUGGUUGGAGAGA UCCGGACUAGGAGCCAGGGGGCCUCUGACC
Hov-109	GCUAAUGAGCCACAUACCUACCCAAGGUGAAAGCUCCUUCUCUCGCA AUGUGUAACUCAUGAUUCUCAUGACCCCUGGUUGGAGAGAUCCGGAC UAGGAGCCAGGGGGC
Hov-85	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGUA ACGACCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
Hov-81o	GGGCCUUCUCUCGCAAUGUGGACCUGGUUGGAGAGAUCCGGACUAG GAGCCAGGUCCACAUACCUACCCAAGGCCCAAGCUU
Hov-124o	GGACCUUCUCUCGCAAUGUGUAACUCAUGAUUCUCAUGACCCCUGGU UGGAGAGAUCCGGACUAGGAGCCAGGGGGCCUCUGACCGAAAGGUC AGAGCUAAUGAGCCACAUACCUACCCAAGGUCCAAGCUUAUCG
Hov-74	CACAUACCUACCCAAGGGAAACCUUCUCUCGCAAUGUGGACCUGGUU GGAGAGAUCCGGACUAGGAGCCAGGUC
Hov-68	CACAUACCUACCCAGAAAUCUCUCGCAAUGUGGACCUGGUUGGAGAG AUCCGGACUAGGAGCCAGGUC
Hov-66	CACAUACCUACCCAAUUCUCUCGCAAUGUGGACCUGGUUGGAGAGAU CCGGACUAGGAGCCAGGUC
Hov-64	CACAUACCUACCCAUCUCUCGCAAUGUGGACCUGGUUGGAGAGAUCC GGACUAGGAGCCAGGUC
A6G	CACAUGCCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
A6C	CACAUCCCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
A6U	CACAUUCCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
C7G	CACAUAGCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
C7A	CACAUAACUACCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
C7U	CACAUAUCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
A40G	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCGAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
A40U	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCUAUGUGGA

	CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
A40C	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCCAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
G72A	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGAAGCCAGGUC
G72U	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGUAGCCAGGUC
G72C	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGCAGCCAGGUC
C39Del	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCGAAUGUGGAC CUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
G38A	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCACAAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
G38C	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCCCAAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
G38U	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCUCAAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
C11U	CACAUACCUAUCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
A67G	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGGCUAGGAGCCAGGUC
C11U/C68Del	CACAUACCUAUCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGAUAGGAGCCAGGUC
A67G/C68Del	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGGUAGGAGCCAGGUC
C68Del	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGAUAGGAGCCAGGUC
A14U	CACAUACCUACCCUAGGCAUUCACUCCGCCUUCUCUCGCAAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
A41C	CACAUACCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCACUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC
A6GA40G	CACAUGCCUACCCAAGGCAUUCACUCCGCCUUCUCUCGCGAUGUGGA CCUGGUUGGAGAGAUCCGGACUAGGAGCCAGGUC

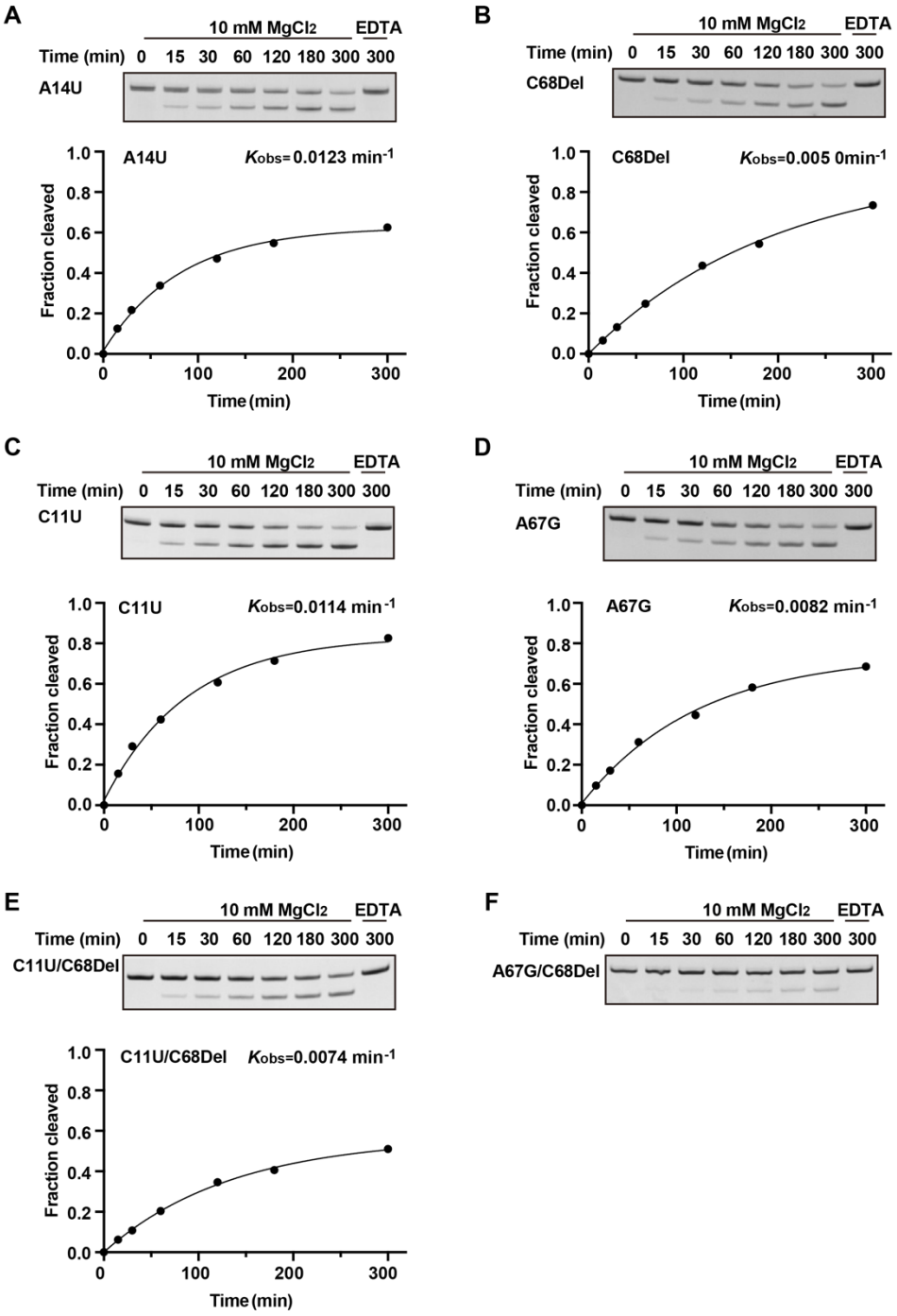
Supplementary Figures



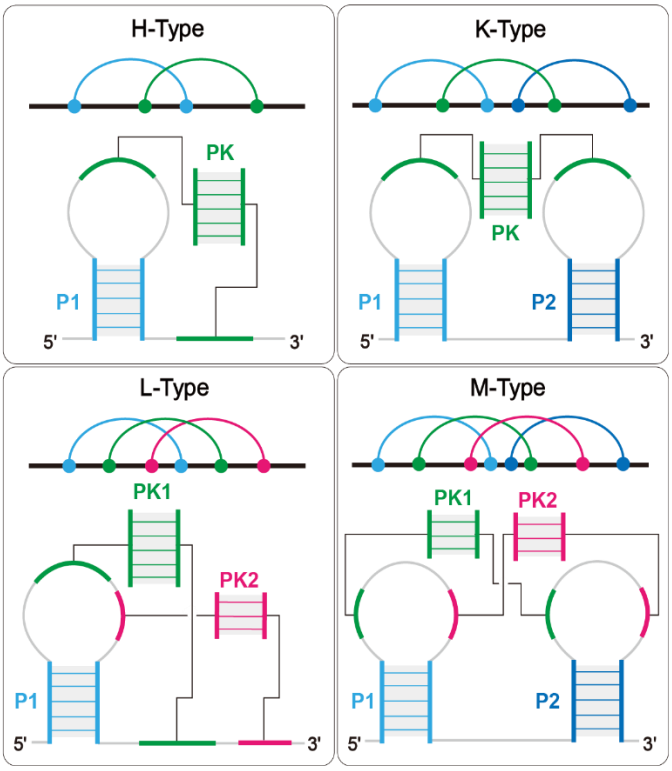
Supplementary Figure 1 | Secondary structure and cleavage assays of hovlinc ribozyme variants and dimer state in crystal for Hov-81. A. Secondary structure of Hov-109 and Hov-124, with Hov-109 enclosed in the red dashed box. B. Time-course cleavage analysis of Hov-109 on a denaturing Urea-PAGE gel, together with the corresponding kinetic profile. C. In the crystal structure of Hov-81, Mol A forms a dimer with Mol B from a neighboring asymmetric unit. D.Secondary structure schematic of Hov-81-E/S based on the tertiary structure. L1 forms additional base pairs above P1 and below P2, and both PK1 and PK2 are extended compared to the original secondary structure model. For the preparation of the two RNA samples of Hov-81-E/S, the bottom two base pairs of P1 and P4 were each replaced with two G-C pairs to facilitate transcription.



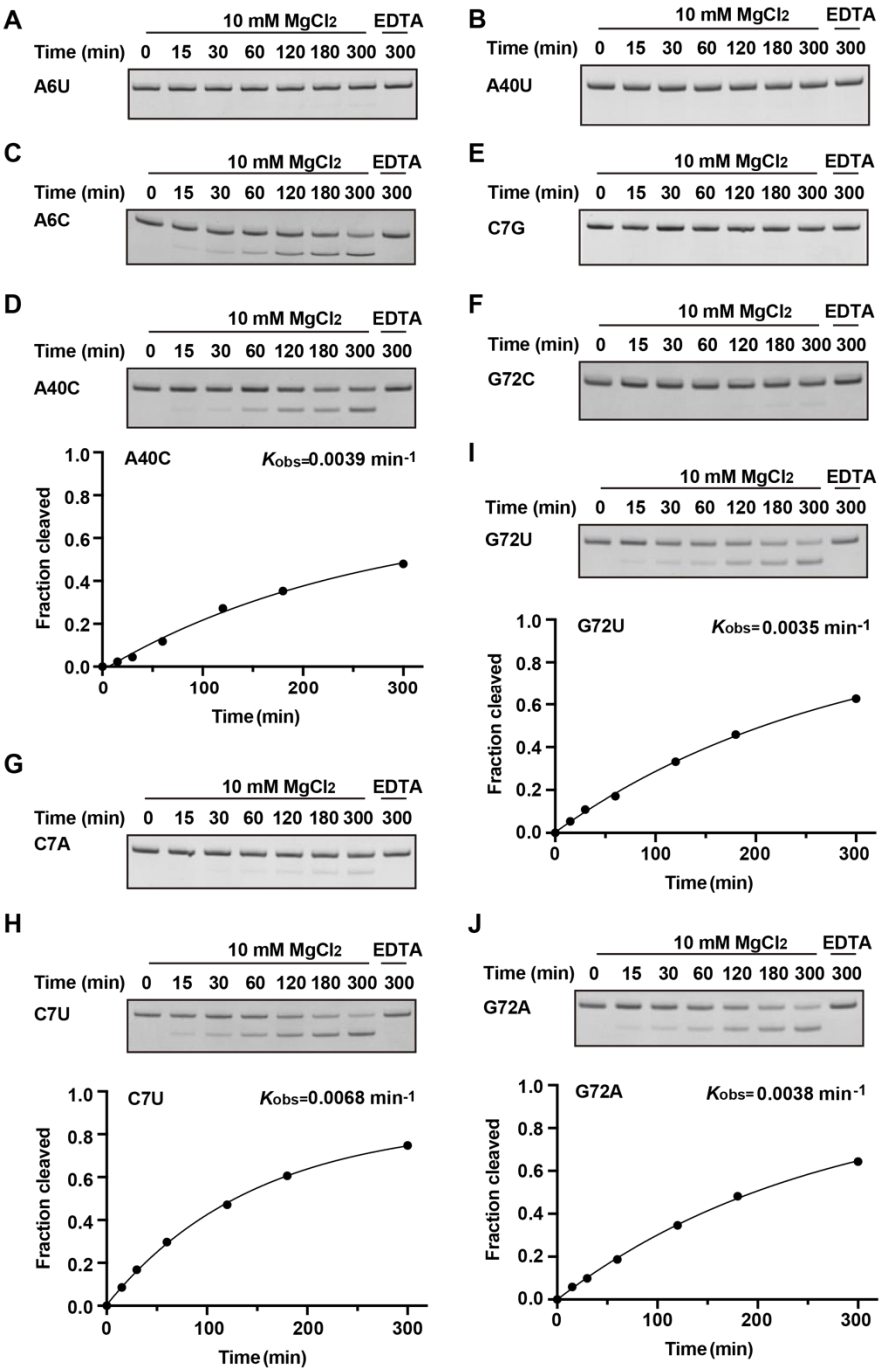
Supplementary Figure 2 | Long helices in Hov-81 tertiary structure. A-B. In the structure, P1 and PK2 coaxially stack to form a continuous long helix, designated Helix P1-PK2. Likewise, P4', PK1, and P2-L2 coaxially stack to form another continuous long helix, designated Helix P4'-PK1-P2-L2. C. A14 and U32 in L1 form a canonical cis Watson–Crick pair between stem P1 and PK1. D. In the middle of PK1, C11 from L1 forms a base triple with A67–C68 from L4′. The 4-NH_2 of C11 forms a hydrogen bond with N1 of A67. The map destiny of C68 base is not observed in the crystal structure; it may form hydrogen bonds with C11 and A67 via its 4-NH_2 group.



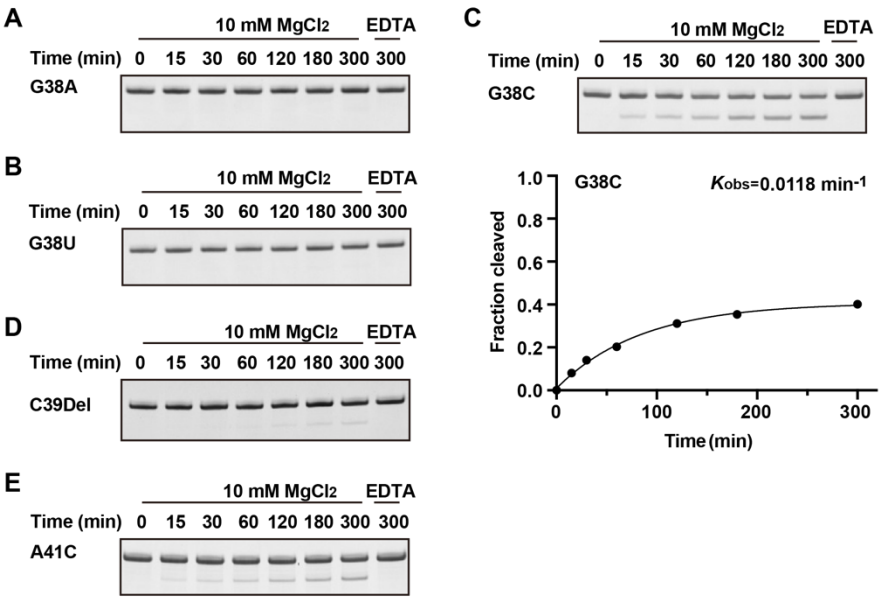
Supplementary Figure 3 | Cleavage assays of Hov-81 mutations within PK1 and L1. A-F. Time-course cleavage analysis of A11U, C68Del, C11U, A67G, C11U/C68Del and A67G/C68Del on denaturing Urea-PAGE gels, together with their kinetic profiles.



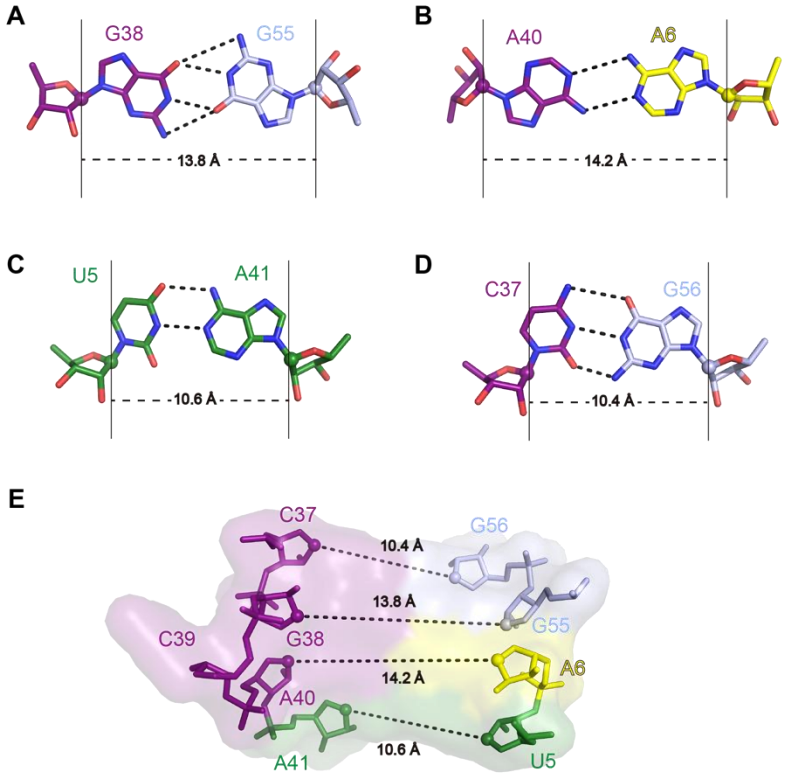
Supplementary Figure 4 | Classic H-, K-, L- and M-Type PK. Topologies and secondary structures of H-, K-, L-, and M-type pseudoknots. These pseudoknot architectures can be uniquely represented as ABAB, ABACBC, ABCABC, and ABCADBCD, respectively, where each letter (A, B, C, etc.) denotes a distinct helix encountered along the RNA backbone from the 5′ to 3′ end (Bon et al., 2008; Kucharik et al., 2016).



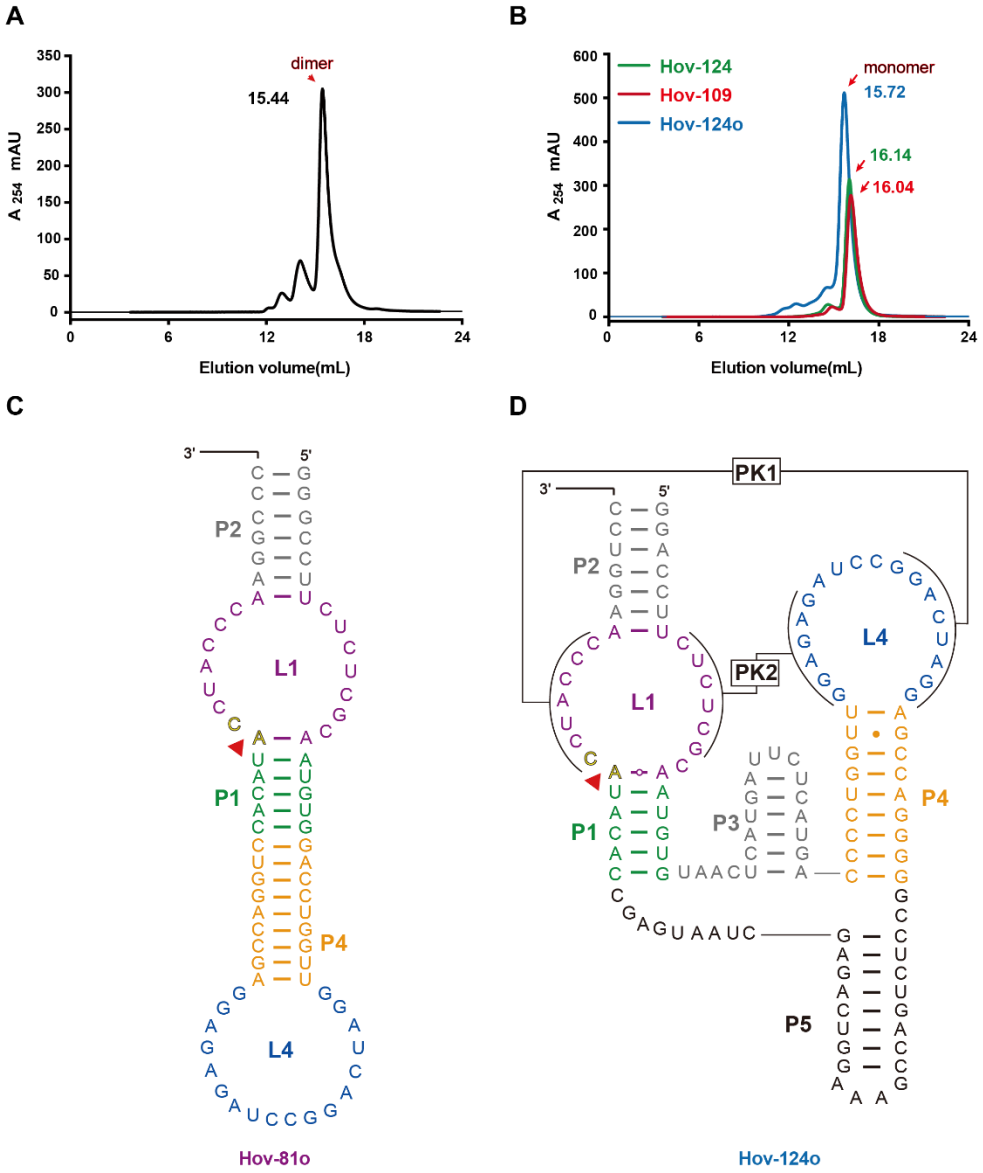
Supplementary Figure 5 | Cleavage assays of Hov-81 mutations at cleavage site. A-H. Time-course cleavage analysis of A6U, A40U, A6C, A40C, C7G, C7A, G72C, G72U, C7U, and G72A on a denaturing Urea-PAGE gel, together with their kinetic profiles.



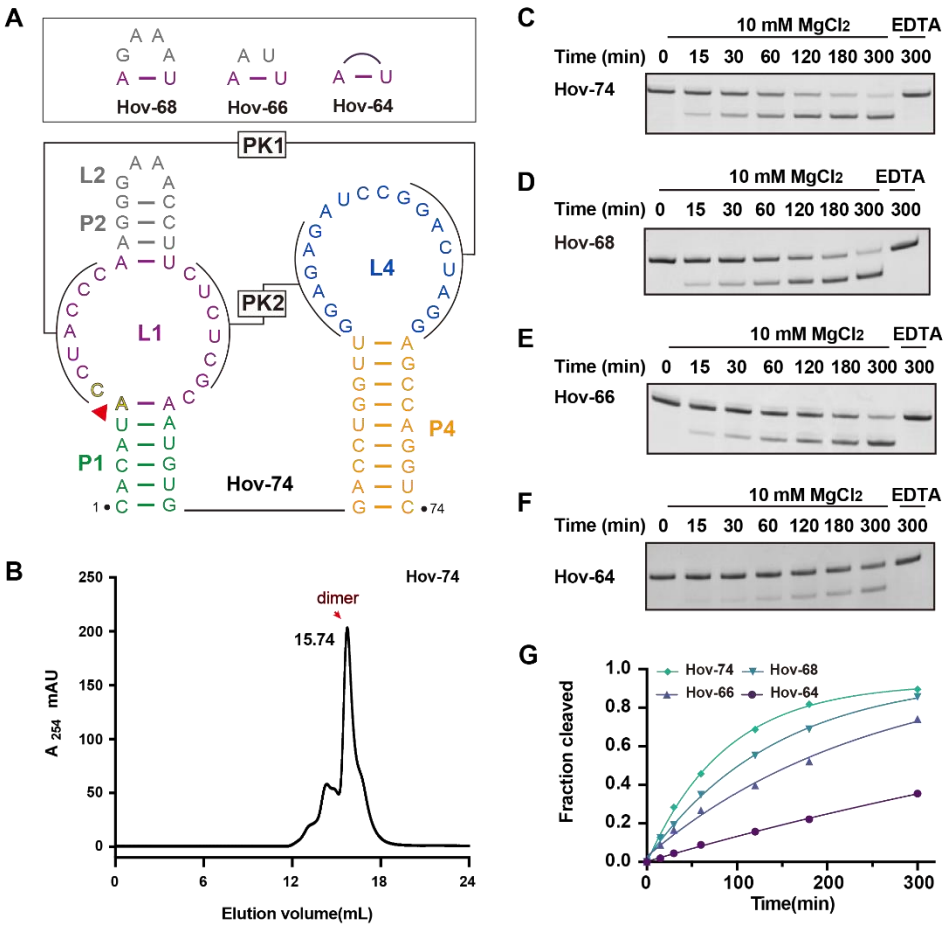
Supplementary Figure 6 | Cleavage assays of Hov-81 mutations involved in catalytic core formation. A-F. Time-course cleavage analysis of G38A, G38U, G38C, C39Del and A41C on a denaturing Urea-PAGE gel.



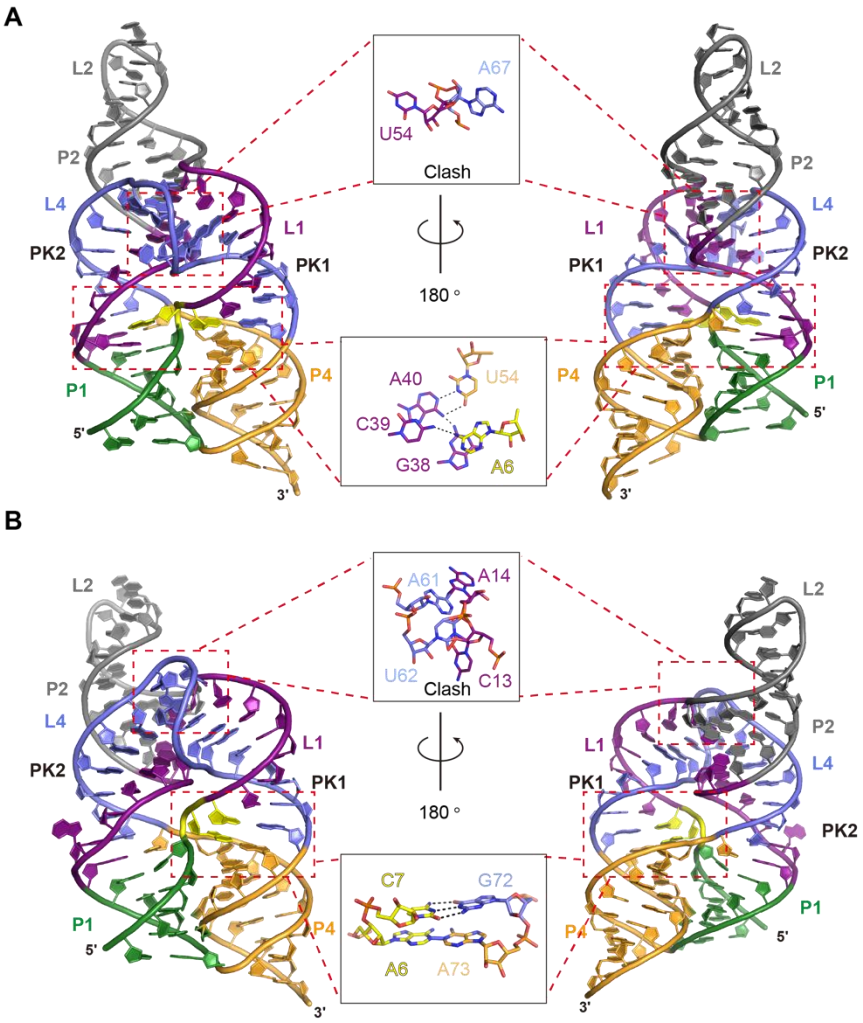
Supplementary Figure 7 | Conformational geometry of backbone at cleavage site. A. The C1'-C1' distances for the G55-G38 pairs is13.8 Å. **B**. The C1'-C1' distance for the A6-A40 pair is 14.2 Å. **C-D**. The C1'-C1' distances for the canonical base pairs G56-C37 and U5-A41 are 10.4 Å and 10.6 Å, respectively. **E**. Local conformational geometry of the backbone around cleavage site A6.



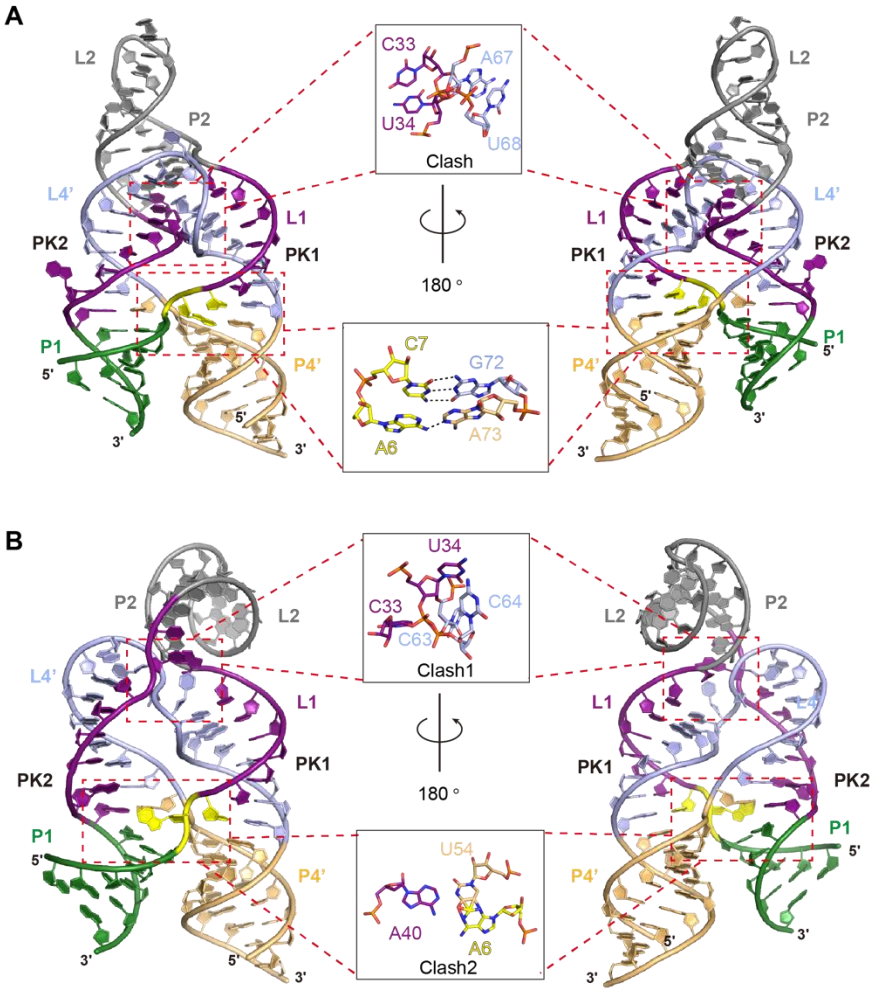
Supplementary Figure 8 | Cleavage assays and secondary structure of Hov-81 variants with peripheral structure changes. A-B. Solution states of Hov-81, Hov-109, Hov-124 and Hov-124o. Size-exclusion chromatography results show that Hov-81 exists predominantly as a dimer in solution. Hov-109, Hov-124, and Hov-124o all exist primarily as monomers in solution. The size-exclusion experiments were performed using a Superose™ 6 Increase 10/300 GL column with a running buffer containing 50 mM HEPES pH 7.0, 50 mM NaCl, and 5 mM $MgCl_2$, which is identical to crystal screening. **C.** Secondary structure of Hov-81o, connecting the 5' and 3'ends of Hov-81 while opening L2. **D.** Secondary structure of Hov-124o, connecting the 5' and 3' ends of P5 via a GAAA loop from Hov-124 and simultaneously opening L2.



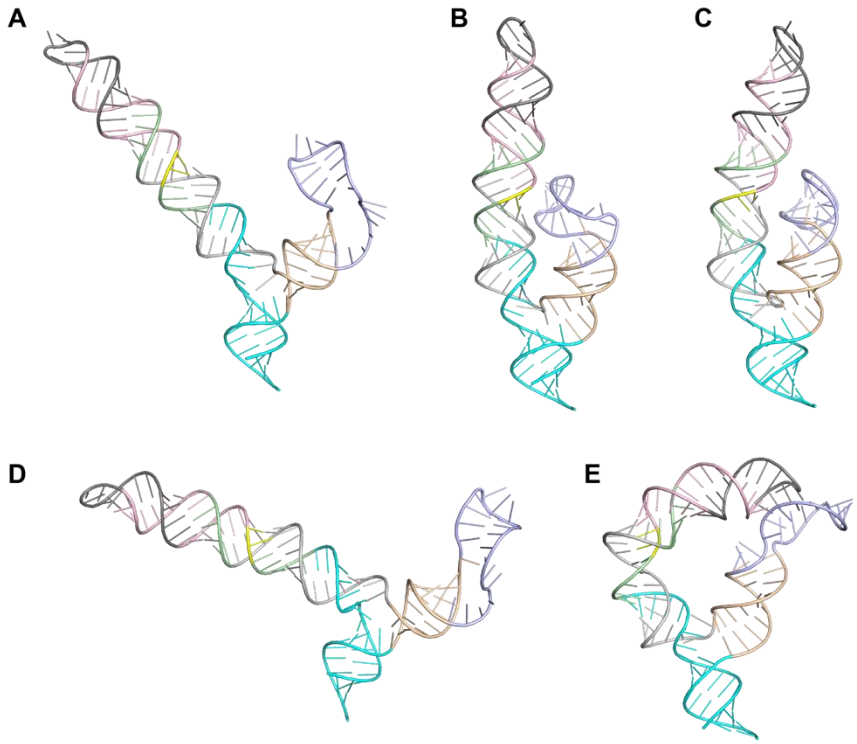
Supplementary Figure 9 | Cleavage assays and secondary structure of Hov-81variants with P2-L2 truncations. A. Secondary structure of Hov-74, with one base pair deleted from P2 and the L2 loop replaced by a GAAA loop, compared to Hov-81. **B.** Size-exclusion chromatography results show that Hov-74 exists predominantly as a dimer in solution. C-G Time-course cleavage analysis of Hov-74, Hov-68, Hov-66 and Hov-64 on denaturing Urea-PAGE gels, together with their kinetic profiles.



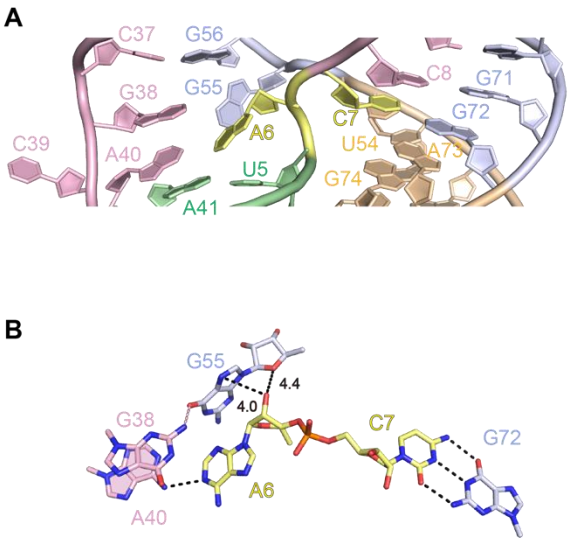
Supplementary Figure 10 | Global architecture and steric clashes of representative structures of Hov-81 predicted by AlphaFold3. A. Overall structure of representative structure 1 of Hov-81 predicted by AlphaFold3. In the upper region, L1 and L4 exhibit a clear clash. At the cleavage site, A6 interacts with C39, whereas A40, which originally interacts with A6 in the crystal structure, is predicted to pair with U54 in the stem region of P4. **B.** Overall structure of representative structure 2 of Hov-81 predicted by AlphaFold3. Similar to structure 1, L1 and L4 show a clash in the upper part. At the cleavage site, A6 and C7 no longer adopt the extended conformation observed in the crystal structure. C7 is still predicted to pair with G72, but A7 and A73 exhibit a steric clash.



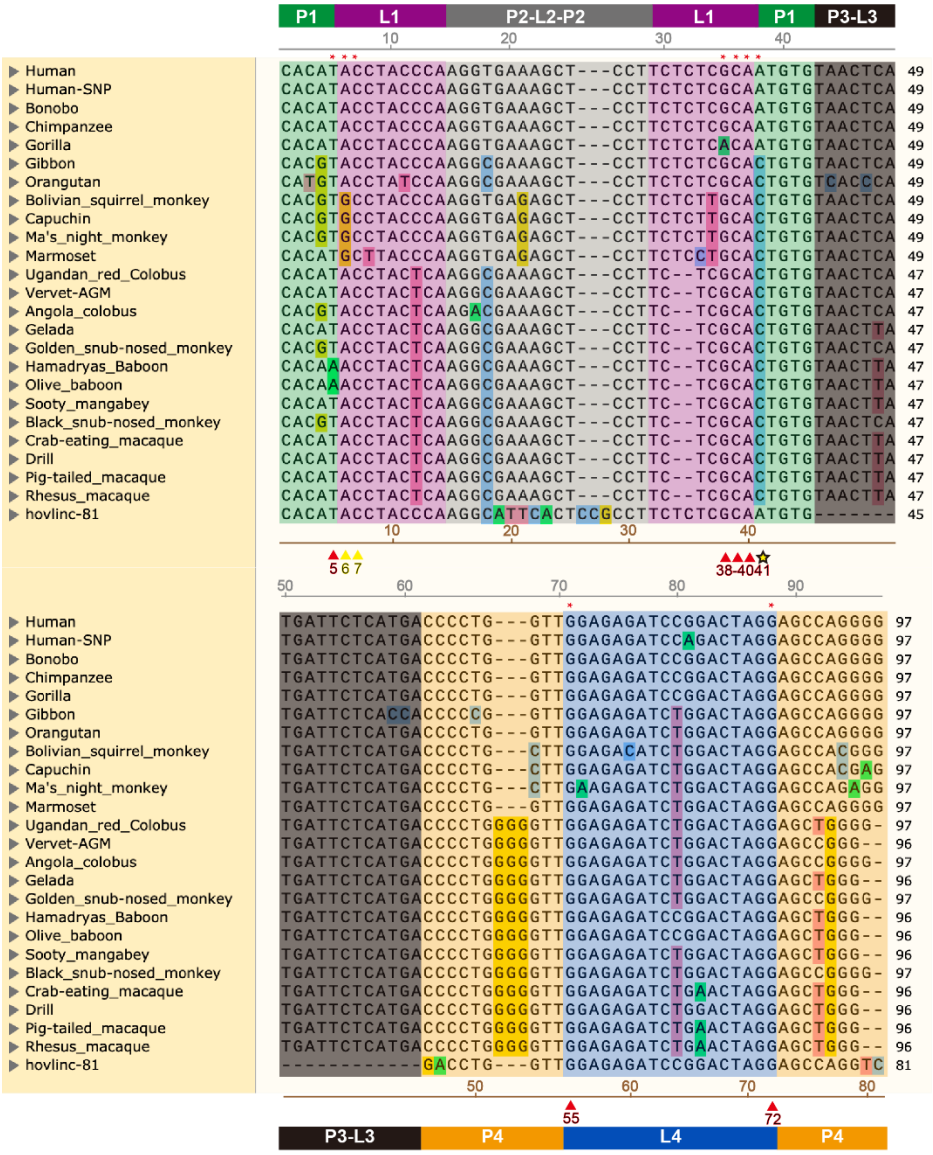
Supplementary Figure 11 | Global architecture and steric clashes of representative structures of Hov-81-E/S predicted by AlphaFold3. A. Overall structure of representative structure 1 of Hov-8-E/S predicted by AlphaFold3. In the upper region, L1 and L4 exhibit a clear clash. At the cleavage site, A6 and C7 no longer adopt the extended conformation observed in the crystal structure. C7 is still predicted to pair with G72, but A7 is predicted to interacts with A73 in the stem region of P4. **B.** Overall structure of representative structure 2 of Hov-81 predicted by AlphaFold3. Clashes were observed between L1 and L4, as well as at the cleavage site.



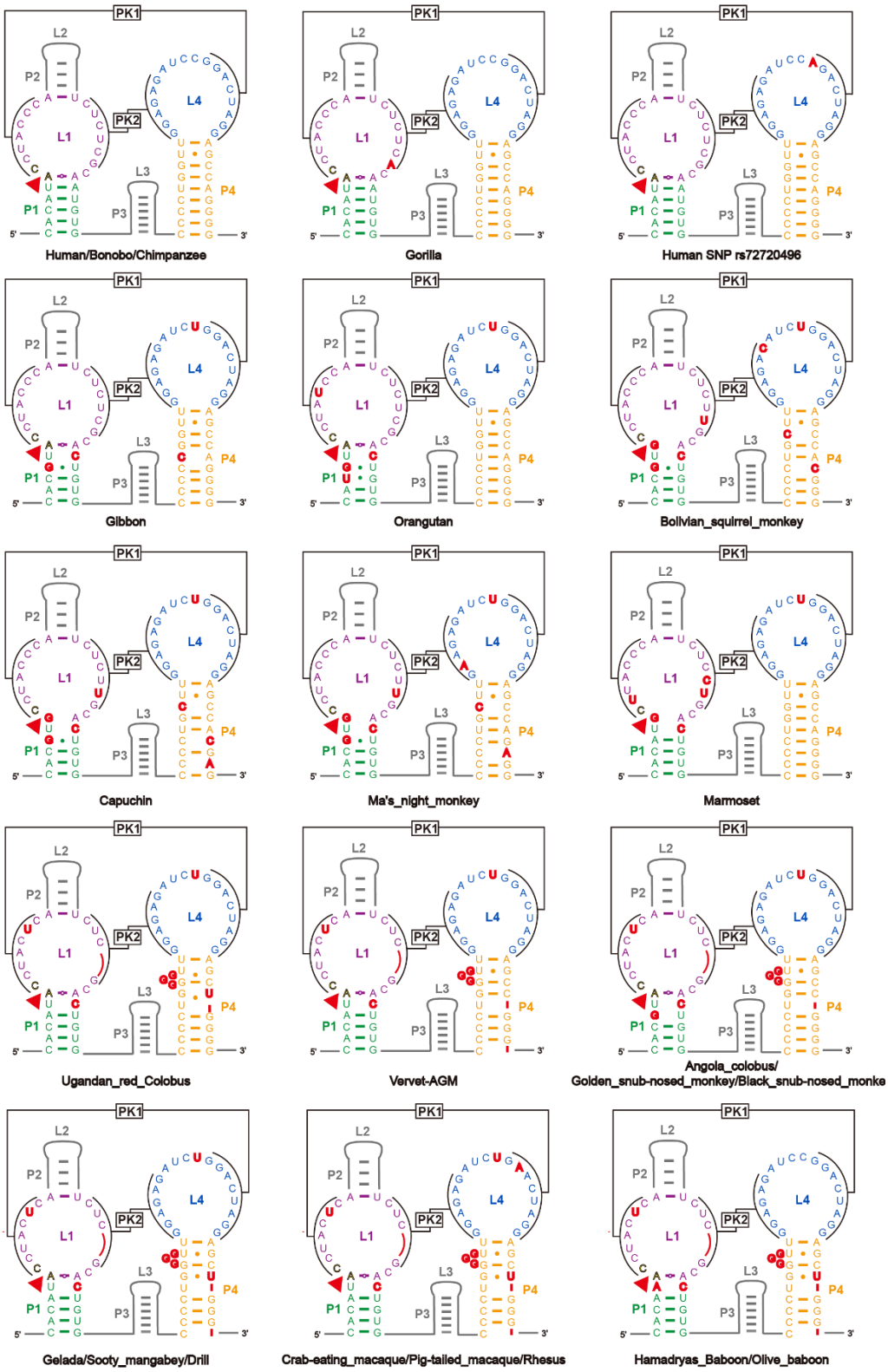
Supplementary Figure 12 | Structure prediction of Hov-124 by AlphaFold3. A-E. Overall structures of 5 representative structures. The colored elements are corresponding to that presented in Figure 5F. None of the structures were able to predict the formation of PK1 and PK2.



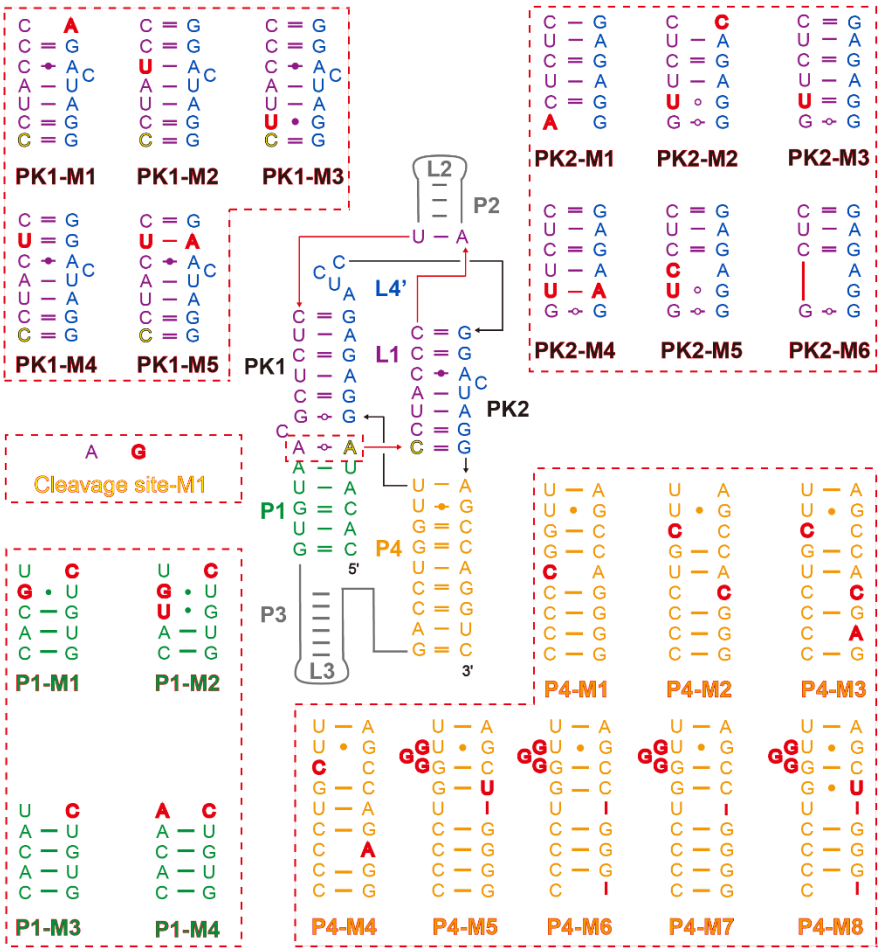
Supplementary Figure 13 | Organization of the catalytic center of modeled Hov-124. A. Overall structure and organization of of the catalytic center of modeled Hov-124. **B.** Structure details at cleavage sites of modeled Hov-124. A6 and C7 adopt the extended conformation observed in the crystal structure. C7-G72 forms base pair consistent with the crystal structure. A6 still interacts with A40, but the conformations of A6 and G55 are significantly altered, with the 2'-OH of A6 moving closer to G55.



Supplementary Figure 14 | Multiple sequence alignment of human hovlinc ribozyme homologs. Analysis of homologous sequences from different species that share over 90% sequence identity with the human hovlinc ribozyme, along with the corresponding crystal structure sequence and a sequence contain human SNP rs72720496. Comparison of the core structural modules (P1, L1, P4, and L4) revealed that, aside from the previously validated active homologs from bonobo and chimpanzee, all other homologs exhibited various degrees of sequence variation within these core modules. Within the core module region, a common mutation was identified across the homologous sequences from 19 species. This mutation corresponds to the A41C substitution in the crystal structure. Residue A41 pairs with U5, forming the terminal base pair of the P1 helix, which stacks directly beneath the A6-A40 at cleavage site and contributes to the architecture of the catalytic core. Key nucleotides critical for the cleavage activity of the Hov-81 ribozyme are marked with triangles in this figure, while those in the human sequence are indicated by red asterisks. The universally observed A41C mutation is highlighted with a yellow star.



Supplementary Figure 15 | Secondary structures and variation sites of homologous sequences of the human hovlinc ribozyme. The secondary structures of the humanSNP rs72720496 and the homologous sequences from 22 species are presented, along with the sequence information of the core modules P1, L1, P4, and L4. Changes relative to the human hovlinc ribozyme are marked with red nucleotides or line segments.



Supplementary Figure 16 | Mutations found in homologous sequences of the human hovlinc ribozyme. For the human SNP rs72720496 and the 22 homologous sequences analyzed, we observed four types of mutations in PK1, six in PK2, four in P1, eight in P4, and a single mutation type at cleavage site.

References

Bon, M., Vernizzi, G., Orland, H., and Zee, A. (2008). Topological classification of RNA structures. Journal of molecular biology *379*, 900-911.

Kucharik, M., Hofacker, I.L., Stadler, P.F., and Qin, J. (2016). Pseudoknots in RNA folding landscapes. Bioinformatics *32*, 187-194.