Nucleic Acids Research, 2026, **54**, gkag185
https://doi.org/10.1093/nar/gkag185
RNA and RNA-protein complexes

Cryo-EM structures reveal a conserved architecture for raiA noncoding RNA

Yao He [1,2,3,*,†], **Janet Zhong** [1,†], **Yuan Yang** [1], **Robert P. Gunsalus** [2,4], **Z. Hong Zhou** [2,5], **Juli Feigon** [1,*]

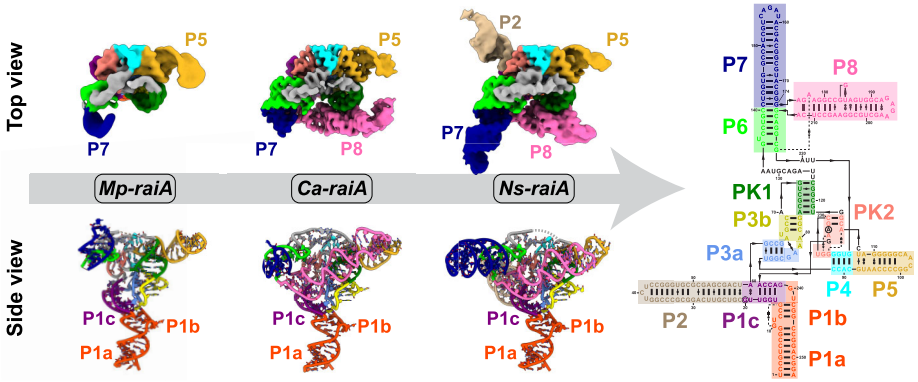
[1]Department of Chemistry and Biochemistry, University of California Los Angeles, Los Angeles, CA 90095-1569, United States
[2]Department of Microbiology, Immunology, and Molecular Genetics, University of California Los Angeles, Los Angeles, CA 90095, United States
[3]Present address: Department of Microbiology and Molecular Genetics, School of Medicine, University of California, Irvine, CA 92617, United States
[4]UCLA-DOE Institute, Los Angeles, CA 90095, United States
[5]The California NanoSystems Institute (CNSI), University of California Los Angeles, Los Angeles, CA 90095, United States

*To whom correspondence should be addressed. Email: feigon@mbi.ucla.edu
Correspondence may also be addressed to Yao He. Email: yaoh20@uci.edu
†The first two authors should be regarded as Joint First Authors.

Abstract

RaiA motif RNA is a family of bacterial noncoding RNAs (ncRNAs) found in over 2700 bacterial species. Although its cellular abundance is comparable to that of rRNAs and tRNAs in the human pathogen *Clostridioides difficile* and its knockout results in pronounced phenotypes, its function remains unknown. Sequence conservation analysis predicted a consensus secondary structure of *raiA* motif RNA with several major subtypes that differ in the number and composition of stems. Here, we present cryogenic electron microscopy (cryo-EM) structures of three *raiA* motif RNAs from three bacterial species, one from each subtype, at 3.0–3.5 Å resolution, as well as a minimal variant with 113 nucleotides at ∼8 Å resolution. Comparison of the structures reveals a conserved architecture, with a compact core comprising stems P3a–P3b bent by an asymmetric internal loop, P4, pseudoknot 1 (PK1), and PK2 with unusual tertiary interactions. While most of the peripheral stems vary, the length, structure, and tertiary interactions of the closing P1 are remarkably conserved, suggesting an essential role. Our study defines the conserved structural framework of *raiA* motif RNAs and provides a foundation for structure-based functional studies. This work also highlights the utility of cryo-EM for *de novo* structure determination of ncRNAs.

Graphical abstract



Introduction

Noncoding RNAs (ncRNAs) play essential roles in numerous cellular processes, including gene regulation, catalysis, and molecular sensing [1]. Despite extensive functional studies revealing their diverse biological activities, the structural characterization of most ncRNAs remains limited due to the lack of high-resolution structures. Cryogenic electron microscopy (cryo-EM) has recently been applied to the study of RNA-only structures [2–4]. Compared to other methods such as X-ray crystallography and nuclear magnetic resonance (NMR),

Received: September 20, 2025. **Revised:** February 5, 2026. **Accepted:** February 7, 2026

cryo-EM requires less sample, eliminates the need for crystallization, is particularly well suited for larger RNAs, and, most importantly, can capture multiple conformations under near-physiological conditions [5]. This ability to visualize structural heterogeneity offers valuable insights into RNA dynamics and functional mechanisms, thereby bridging the gap between structural and functional studies of ncRNAs [6–12].

RaiA motif RNA is a bacterial ncRNA first discovered in 2017 through bioinformatic analyses [13] and later identified in over 2700 bacterial species [14]. It is predominantly found in the phyla *Bacillota* and *Actinomycetota* and has an average length of ∼200 nucleotides [14]. Transcriptome analyses revealed that *raiA* motif RNA is highly abundant in cells despite its gene being present at only a single locus per genome; in *Clostridium acetobutylicum*, it is the fourth most abundant RNA, excluding rRNAs and tRNAs [14], and in the human pathogen *Clostridioides difficile*, its abundance is comparable to that of some rRNAs and tRNAs [15]. Knockout of *raiA* motif RNA in these bacteria leads to defects in sporulation and aggregation, as well as a delayed transition from exponential growth into stationary phase in *C. difficile* [14, 15]. These phenotypes can be rescued by restoring the expression of *raiA* motif RNA, indicating that it functions *in trans*, rather than as a *cis*-regulatory element like a riboswitch [14, 15]. In *C. difficile*, the sporulation defect caused by the loss of *raiA* motif RNA (also known as ModT) was linked to an increased cellular level of the second messenger cyclic di-GMP (c-di-GMP) [15]. Grad-seq analysis in *C. difficile* revealed that *raiA* motif RNAs are abundant in a form of ribonucleoprotein (RNP) complexes, and pulldown assays using *raiA* motif RNA as bait significantly enriched two non-specific RNA binding proteins (RBPs), KhpA and KhpB [16]. The widespread distribution and high cellular abundance of *raiA* motif RNA suggest it plays an essential role in fundamental physiological processes. However, its function remains unknown.

A consensus secondary structure of *raiA* motif RNA, consisting of five stems (paired region, P), P1, P3, P4, P5, and P6, and two pseudoknot (PK) stems, PK1 and PK2, was proposed based on sequence conservation analysis [14]. Subtypes include variants with an extra stem 2 (P2) or a multi-stem domain containing stem 7 (P7) and stem 8 (P8) or both [14]. Recently, cryo-EM structures of the *raiA* motif RNA from *Clostridium* were reported, determined either alone [17, 18] or with the assistance of a large RNA scaffold [19]. However, these structures all belong to a single subtype that has P7–P8 but not P2. Here, we present four cryo-EM structures of *raiA* motif RNAs, three at resolutions of 3.0–3.5 Å and one minimal variant at ∼8 Å resolution. These RNAs, derived from four different bacterial species, represent distinct structural subtypes. Comparison of these structures reveals a highly conserved core with an extruding P1 stem. This work not only establishes the conserved three-dimensional architecture of *raiA* motif RNAs, which is important to fulfill its function, but also provides a structural foundation for future structure-based function studies of this widely distributed and highly abundant bacterial ncRNA.

Materials and methods

RNA preparation

All *raiA* motif RNA constructs (Supplementary Table S1) were prepared by *in vitro* transcription using T7 RNA polymerase P266L mutant [20] or G47A + 884G mutant [21] with synthetic DNA templates (Integrated DNA Technologies), as described previously [22]. Two G nucleotides were added to the 5′ ends of each construct to increase transcription efficiency. Transcription mixtures (40 mM Tris–HCl, pH 8.0, 1 mM spermidine, 0.01% Triton X-100, 2.5 mM DTT, 20 mM $MgCl_2$, 4 mM each rATP, rUTP, rGTP, and rCTP, 0.5 μM DNA template, T7 RNAP) were incubated at 37°C for 6 h. Transcribed RNAs were purified by 15% denaturing polyacrylamide gel electrophoresis (PAGE). The band corresponding to the desired product was excised from the gel, crushed, and soaked five times in elution buffer [20 mM Tris–HCl, pH 7.4, 0.3 M sodium acetate, 1 mM ethylenediaminetetraacetic acid (EDTA)]. Collected RNA eluents were concentrated and buffer exchanged into nuclease-free water using an Amicon filtration system with a 3 kDa molecular weight cutoff membrane (Millipore), then into high salt buffer (10 mM sodium phosphate, pH 7.2, 1 mM EDTA, 1.5 M KCl), and finally buffer exchanged back into nuclease-free water. Diluted RNA (<100 μM) solutions were heated at 95°C for 5 min and snap-cooled on ice for 1 h. RNAs were then concentrated and stored at −20°C. RNA samples were diluted to ∼30 μM in EM buffer (20 mM HEPES–HCl, pH 7.5, 50 mM KCl, 10 mM $MgCl_2$, 0.05% Igepal CA-630) before preparing cryo-EM samples.

Cryo-EM specimen preparation and data collection

To prepare cryo-EM samples, 3 μl of the purified RNA samples were applied to glow-discharged Quantifoil grids (UltrAuFoil R0.6/1.0, 300 mesh), blotted with filter paper (595 filter paper, Ted Pella), and flash-frozen in liquid ethane using an FEI Vitrobot Mark IV at 10°C and 100% humidity. Cryo-EM grids were loaded into a ThermoFisher Titan Krios G1 electron microscope operated at 300 kV for automated data collection using SerialEM [23]. Movies of dose-fractionated frames were acquired with a Gatan K3 direct electron detector in super-resolution mode at a pixel size of 0.856 Å. A Gatan Imaging Filter (GIF) was inserted between the electron microscope and the K3 camera and operated in zero-loss mode with a slit width of 20 eV. The defocus was set between −0.5 μm and −2.5 μm. The total dose on the sample was set to ∼50 electrons/$Å^2$, which was fractionated into 40 frames.

Cryo-EM data processing

Cryo-EM data processing workflows are outlined in Supplementary Figs S1–S4 for the structure determination of *Ns-raiA* (*Nocardioides sp.* Iso805N), *Ca-raiA* (*Clostridium acetobutylicum*, ATCC 824), *Mp-raiA* (*Mogibacterium pumilum*, ATCC 700696), and *To-raiA* (*Thermosediminibacter oceani*, DSM 16646), respectively. Data were preprocessed on-the-fly using cryoSPARC Live (v4.5.1) during the data collection sessions and then further processed using cryoSPARC [24], including steps of 2D classification, *ab initio* reconstruction, heterogeneous refinement, homogeneous refinement, CTF refinement, reference-based motion correction, and non-uniform refinement [25]. Topaz [26] was used for particle picking within the framework of cryoSPARC. Local resolution was estimated using cryoSPARC. Data collection and processing statistics are summarized in Supplementary Table S2.

Model building and refinement

RNA models were built manually against cryo-EM density maps in Coot (version 0.9.8.7) [27]. The distal regions of stem–loops with local resolution lower than 4 Å were modeled with the assistance of AlphaFold3 [28]. All models were refined using PHENIX (version 1.20.1) [29] in real space with base pair restraints, manually adjusted using ISOLDE [30] to resolve clashes and geometry outliers, and refined again with PHENIX. The final models were validated with PHENIX and the wwPDB validation server, and visualized by ChimeraX [31]. Refinement statistics of the models are summarized in Supplementary Table S2. Q-score analyses shown in Supplementary Fig. S5 were performed using chimeraxqscore [32]. Dot-bracket notations shown in Supplementary Fig. S6A were generated from cryo-EM structures using RNApdbee [33]. Sequence alignments were visualized using Jalview [34].

Protein expression and purification

The cDNAs encoding *C. acetobutylicum* KhpA (UniProt ID: Q97I96) and KhpB (Q97CW1) were purchased from IDT (Integrated DNA Technologies), and cloned into pETDuet plasmids separately. KhpA has an N-terminal 6×His tag followed by a TEV protease cleavage site, and KhpB has an N-terminal 6×His–MBP tag followed by a TEV protease cleavage site. The plasmids were transformed into BL21-GOLD(DE3) competent cells for protein expression. Cultures were grown at 37°C to an optical density at 600 nm (OD_{600}) of 0.6, shifted to 18°C for 1 h, and induced overnight with 0.5 mM IPTG. Cells were harvested by centrifugation and resuspended in lysis buffer (50 mM Tris–HCl, pH 7.5, 300 mM NaCl, 30 mM imidazole, 1 mM TCEP, 5% glycerol) supplemented with protease inhibitor cocktail (Thermo Scientific Pierce, A32965); for KhpA, lysis buffer with 500 mM NaCl and 10% glycerol was used. Cells were lysed by sonication, and the lysates were clarified by centrifugation and filtration through a 0.45 μm syringe filter. The supernatant was loaded onto a pre-equilibrated 5 ml Ni-Sepharose affinity column (HisTrap HP, GE Healthcare). The column was washed with 100 mM imidazole and bound proteins eluted with 300 mM imidazole. Eluted fractions were dialyzed overnight against dialysis buffer (20 mM HEPES, pH 7.5, 200 mM NaCl, 1 mM TCEP, 10% glycerol) in the presence of 1 mg TEV protease to cleave the N-terminal tag. The dialyzed sample was reapplied to a 5 ml Ni-Sepharose column to separate cleaved proteins from tag and uncleaved material, followed by size-exclusion chromatography (SEC; HiLoad 16/600 Superdex 75, GE Healthcare). For KhpA, an additional purification step using a 5 ml heparin affinity column (HiTrap HP, GE Healthcare) was performed to remove nucleic acid contaminants prior to SEC. Fractions of KhpB collected from SEC were mixed with KhpA at 1:1.1 ratio and incubated at 4°C for 1 h before another round of SEC. The heterodimer was collected from peak fractions and analyzed by SDS polyacrylamide gel electrophoresis. Proteins were concentrated using Amicon centrifugal filters (Millipore), and final concentrations were determined by absorbance at 280 nm.

Electrophoretic mobility shift assay

RaiA motif RNAs and the KhpA–KhpB heterodimer, as well as KhpA and KhpB individually, were prepared separately in binding buffer (20 mM HEPES, pH 7.5, 200 mM NaCl, 1 mM TCEP, 5% glycerol). RNA and protein were mixed at various ratios in a total volume of 10 μl, with the RNA concentration at 0.2 μM for each mixture. The mixtures were incubated at room temperature for 15 min prior to gel electrophoresis on a 5.7% polyacrylamide gel (37.5:1 crosslinking ratio) with 0.5× TBE buffer (45 mM Tris–borate, 1 mM EDTA, pH 8.3). Gels were subsequently stained with SYBR Gold (Invitrogen) and imaged by a Pharos FX Plus scanner (Bio-Rad). The signal intensity of the free RNA was quantified with ImageJ software. The apparent dissociation constant (Kd) was determined by fitting the fraction bound values (θ) to the following equation:

$$\theta = P_f{}^n / \left(P_f{}^n + K_d^n\right),$$

$$P_f = P_t - \left(P_t + R_t + K_d - \sqrt{(P_t + R_t + K_d)^2 - 4 * P_t * R_t}\right)/2,$$

where P_f is the unbound protein concentration, P_t is the total protein concentration, R_t is the total RNA concentration, and n is the Hill coefficient.

Results

Cryo-EM structure of raiA motif RNA with the highest secondary structure complexity

The previously reported bioinformatics study of 2377 distinct *raiA* motif RNA sequences revealed a consensus secondary structure with five stems (P1, P3, P4, P5, and P6) and two pseudoknot stems (PK1 and PK2), while 42% of the *raiA* motif RNAs contain an additional stem 2 (P2), and 68% have a larger multi-stem domain consisting of stem 7 (P7) and stem 8 (P8) in addition to P6 [14]. To investigate the structure of a *raiA* motif RNA with all eight possible stems (P1 to P8), we selected the RNA from *Nocardioides sp. Iso805N* (hereafter referred to as *Ns-raiA*) for cryo-EM study, which is 253 nucleotides in length (82 kDa). UltrAuFoil holey grids were used to prepare the cryo-EM sample, and surfactant (0.05% Igepal CA-630) was added to optimize particle distribution on the grids (Supplementary Fig. S1A). After extensive 2D classification and heterogeneous refinement using cryoSPARC, 115 827 particles were selected for the final refinement, yielding a reconstruction with an overall resolution of 3.0 Å and local resolution of the rigid core better than 2.5 Å (Fig. 1A and Supplementary Fig. S1). The high-resolution features of the density map (Fig. 1B) enable *ab initio* model building of 195 nt (out of 253 nt) and allow accurate assignment of base geometries for both canonical Watson–Crick base pairs and non-canonical base pairs formed through the Hoogsteen or sugar edges of the bases (Supplementary Fig. S5). The distal regions of P1 and peripheral stem–loops P2, P5, P7, and P8 exhibited lower resolution (~4–5 Å, Supplementary Fig. S1F), likely due to increased local flexibility, as indicated by 3D variability analysis (Supplementary Movie S1).

The 3D model (Fig. 2A) and 2D schematics of *Ns-raiA* (Fig. 2B and C) reveal an unusually complex architecture, in which most nucleotides participate in base-pairing interactions. The eight stems of *Ns-raiA* are connected by one four-way junction (4WJ) and two three-way junctions (3WJ), J1, J2, and J3, respectively: J1 connects P1c, P2, P3a, and P4; J2 connects P4, P5, and P6; J3 connects P6, P7, and P8 (Fig. 2B). P5 and P6 are connected by a long linker (PK-L) that can be divided into three regions: the 5′ region interacts with the linker between P1c and P4, forming the PK2 stem; the middle region

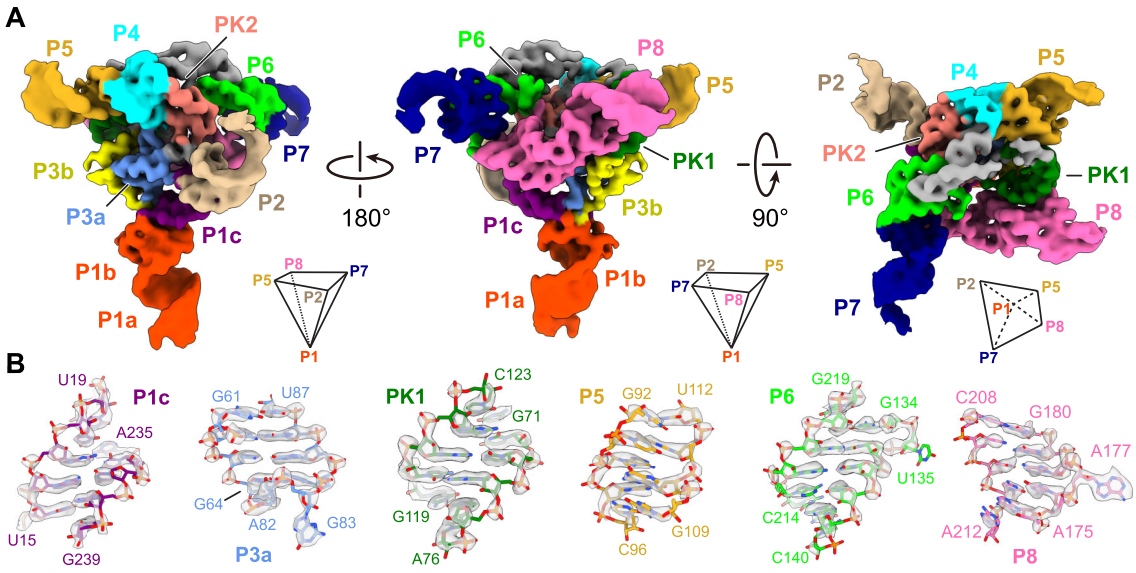


Figure 1. Cryo-EM structure of *raiA* motif RNA from *Nocardioides* sp. Iso805N (*Ns-raiA*) at 3.0 Å resolution. (**A**) Cryo-EM density map of *Ns-raiA* shown in three different views. Individual stem–loops are color-coded as indicated. Pyramid diagrams indicate the orientation of *Ns-raiA* structure. (**B**) Overlaid cryo-EM densities and models of P1c, P3a, PK1, P5, P6, and P8.

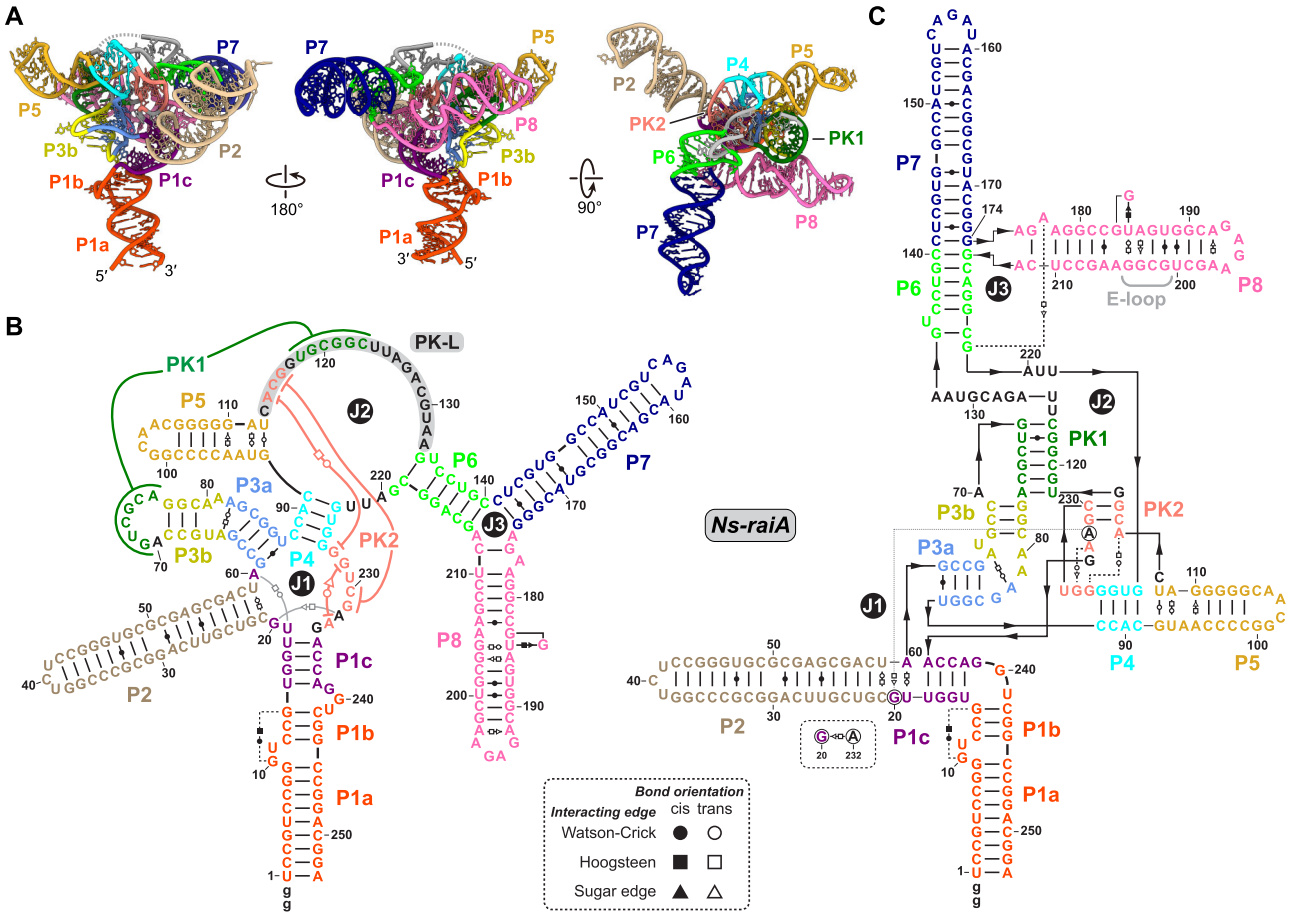


Figure 2. Tertiary and secondary structures of *Ns-raiA*. (**A**) Atomic model of *Ns-raiA* shown in three different views. Individual stem–loops are color-coded as indicated. Nucleotides in the conserved UUAGACGUAA linker connecting PK1 and P6 not resolved in the cryo-EM map are shown as a dotted line. Secondary structures of *Ns-raiA* are shown in the canonical layout (**B**) as proposed in [14] or in a layout that more closely reflects the tertiary structure (**C**). The nucleotides are colored as in the structure in panel (A). Black arrowheads indicate the backbone direction. Non-Watson–Crick base pairs are indicated with Leontis–Westhof nomenclature symbols (inset) [55].

interacts with the distal loop of P3b, forming the PK1 stem; the 3′ region is the longest single-stranded region (10 nt) in the whole structure and is solvent exposed (Fig. 2B). Notably, the 3′ region exhibits elevated flexibility, especially for the central four nucleotides (A126-C129, shown as dashed lines in Fig. 2A), as evidenced by their weaker densities compared to neighboring nucleotides.

The overall structure of *Ns-raiA* adopts a trapezoidal-based pyramidal architecture, with the central core comprising P3a–P3b, P4, PK1 stem, and PK2 stem (Figs 1A and 2A). The remaining stems, P1, P2, P5, P6–P7, and P8, form five extended arms that either protrude from or attach to the core. P1, which is formed by pairing of the 5′ and 3′ ends of the RNA, is divided into three segments by two bulges: P1a and P1b stack together as a continuous helix, and P1c stacks with P2 (Fig. 2C). These two helices, P1a–P1b and P1c–P2, are anchored to the core via P1c, while most of P1a, P1b, and P2 make no contact with the core. On the opposite side of the core, P8 extends perpendicularly from the junction of the stacked P6–P7 (Fig. 2A). P4 and P5, which stack on each other to form a continuous helix, and together with P8 sandwich the core, with P5 and the end of P8 extending out like a crab claw (Fig. 2A). Taken together, we obtained a complete model of *Ns-raiA* that encompasses all eight predicted stems.

Overall, the cryo-EM structure of *Ns-raiA* largely agrees with the predicted consensus secondary structure of *raiA* motif RNAs [14], including the Watson–Crick base pairs in the two pseudoknot stems PK1 and PK2 (Supplementary Fig. S6A). However, nucleotides within the predicted internal loops engage in extensive non-Watson–Crick base-pairing and long-range stacking interactions that help stabilize the central core and maintain the pyramidal architecture, as described in the following sections.

Structural comparison of three sub-types of raiA motif RNA reveals a conserved architecture

The above cryo-EM structure of *Ns-raiA* represents the subtype of *raiA* motif RNA that contains both P2 and P8 (Figs 2 and 3A). To investigate whether other subtypes of *raiA* motif RNA share the same overall architecture as *Ns-raiA* and to identify the minimal structural elements required to maintain this architecture, we selected two additional *raiA* motif RNAs for cryo-EM analysis: one from *Clostridium acetobutylicum* (*Ca-raiA*, 204 nt), representing the subtype without P2, and another from *Mogibacterium pumilum* (*Mp-raiA*, 163 nt), representing the subtype lacking both P2 and P8. The RNA samples were synthesized and prepared for cryo-EM following the protocols described earlier and in the "Materials and methods" section, and their cryo-EM structures were determined to an overall resolution of 3.0 Å for *Ca-raiA* (66 kDa) and 3.5 Å for *Mp-raiA* (53 kDa) (Fig. 3B–E and Supplementary Figs S2, S3, and S5). Notably, despite using a comparable number of particles for the final reconstructions, the structure of *Mp-raiA* was resolved at a lower resolution than *Ca-raiA* (Supplementary Figs S2 and S3), likely due to its smaller size, which poses greater challenges for single-particle cryo-EM data processing.

As expected, *Ca-raiA*, which lacks P2, presents only four arms (P1a–P1b, P4–P5, P6–P7, and P8) surrounding the core (Fig. 3B and D), consistent with recently reported cryo-EM structures of the same RNA at 2.9 and 3.0 Å resolution [18, 19] and the closely related *raiA* motif RNA from *Clostridium* sp. CAG 138 at 3.3 Å resolution [17]. Comparison of these structures reveals small differences at the distal ends of P5 and P8 (Supplementary Fig. S6B), likely due to lower local resolution in these regions across all structures, as well as inconsistencies in the conformations of some bases between 9G7C and the other three structures (Supplementary Fig. S6C and D), perhaps reflecting its origin from a different organism.

In *Ca-raiA*, since there is no P2, J1 is a 3WJ instead of the 4WJ in *Ns-raiA*. The linker between P1c and P3b folds into a tetraloop-like structure, formed by G20, A21, A22, and A183, which is donated from PK2 region to form a base pair with G20 via 3D folding and closed by U19-A23 reverse Hoogsteen base pair (Fig. 3G). There is a sharp turn between G20 and A21, while A21, A22, and A183 stack sequentially, forming a fold whose overall shape is almost identical to the well-known GAAA tetraloop [35, 36]. We refer to this as a GAA(A) tetraloop. The long-range interaction from A183 to complete the GAA(A) tetraloop by pairing with G20 has precedent in the $tRNA^{Asp}$ structure [17, 37] and in the ATP aptamer structure [38], where an AMP ligand inserts to pair with G, although with a different base-pairing geometry. In contrast, in *Ns-raiA*, P1c and P2 form a coaxially stacked helix, with C21-U59, G20-A232 (equivalent to G20-A183 in *Ca-raiA*), and U19-A60 stacking continuously at the junction (Figs 2C and 3F). Remarkably, the long-range tertiary interaction between G20 in P1c and a bulge A (A183 in *Ca-raiA*) in PK2 is the same, and the connection to P3a is at the same position of P1c for *Ns-raiA* and *Ca-raiA* (Fig. 3F and G). In addition to the absence of P2, *Ca-raiA* has a much shorter P7—comprising only four base pairs—compared to the 14 base pairs found in *Ns-raiA* (Fig. 3A and B). Aside from these differences, the rest of the *Ca-raiA* structure (including P1, P3–P6, P8, and the two PK stems) is almost identical to that of *Ns-raiA* (Supplementary Fig. S7), with a root-mean-square deviation (RMSD) of 2.7 Å, indicating that the absence of P2 has minimal impact on the overall architecture of *raiA* motif RNA.

For *Mp-raiA*, which lacks both P2 and P8, only three protruding arms (P1a–P1b, P4–P5, and P6–P7) are present in the structure (Fig. 3C and E). The connection between P1c and P3 is also mediated by a GAA(A) tetraloop [G20-A21-A22-(A142)], just as in the *Ca-raiA*, and the same long-range interaction between G20 in the P1c loop and a bulge A (A142) in PK2 is present (Fig. 3G and H). In both *Ns-raiA* and *Ca-raiA*, P8 emerges from the coaxially stacked P6–P7 helix via a 3WJ (Fig. 3A and B) and docks onto the core through ribose-zipper interactions of its conserved E-loop motif [39, 40] and PK1 (Fig. 3I and J). In *Mp-raiA*, the absence of P8 (Fig. 3K) results in a 5-nt bulge that disrupts the stacking of P6 and P7 (Fig. 3C). The remainder of *Mp-raiA* (including P1, P3–P6, and the two PK stems) is almost identical to that of *Ns-raiA* and *Ca-raiA* (Supplementary Fig. S7), with RMSDs of 1.8 Å and 1.7 Å, respectively. Taken together, the structural comparisons among *Ns-raiA*, *Ca-raiA*, and *Mp-raiA* reveal a conserved overall architecture across the three subtypes of *raiA* motif RNAs.

Triangular configuration of the core

The most prominent feature among the conserved structural elements of *raiA* motif RNA is its central core (Fig. 4A and B and Supplementary Fig. S8). In all three high-resolution structures we determined, the backbone of P3 adopts a sharply bent configuration (Fig. 4C and Supplementary Fig. S8),

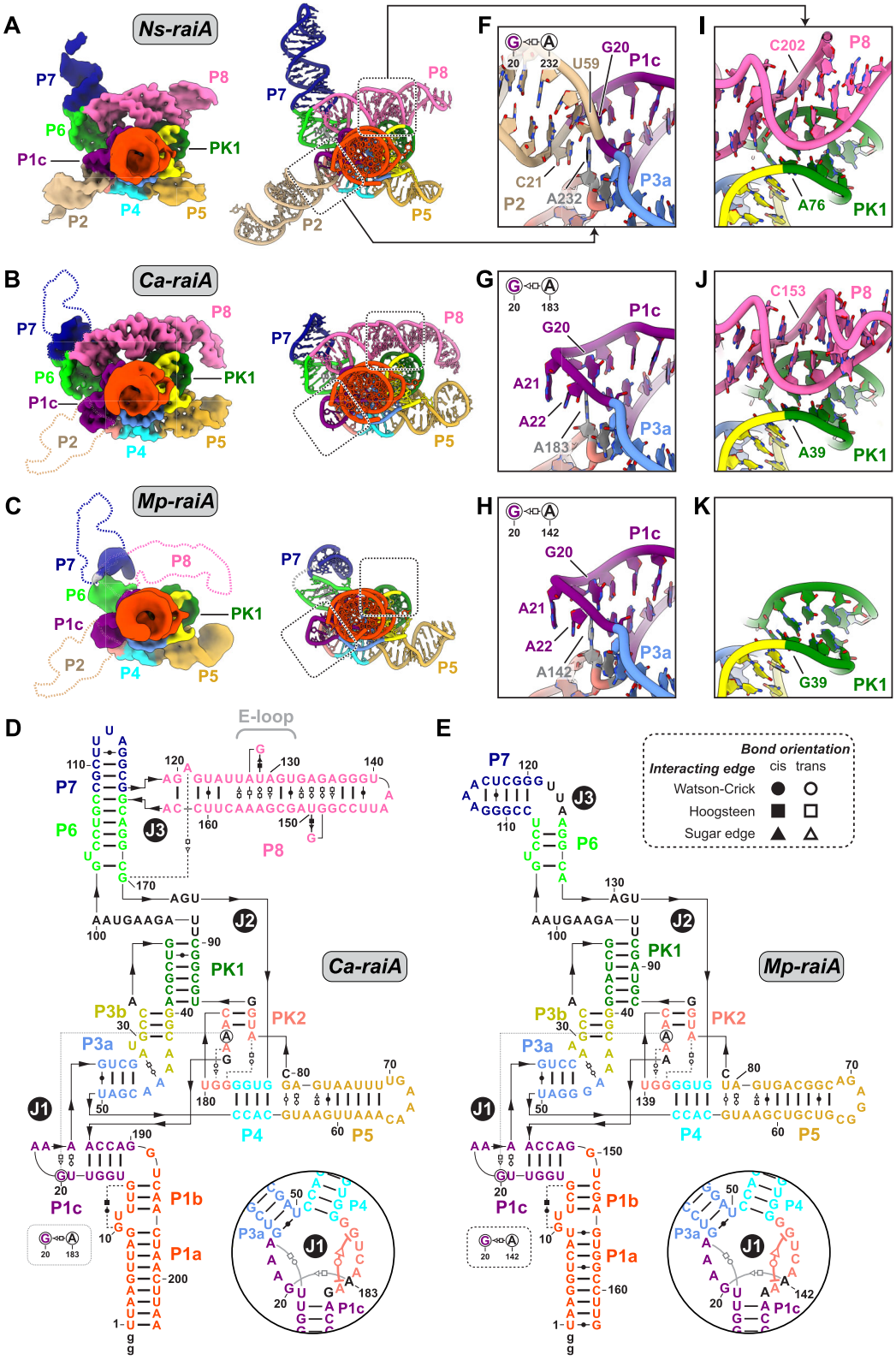


Figure 3. Structural comparison of *raiA* motif RNAs from *Nocardioides* sp. Iso805N (*Ns-raiA*), *Clostridium acetobutylicum* (*Ca-raiA*), and *Mogibacterium pumilum* (*Mp-raiA*). Cryo-EM density map (left) and atomic model (right) of *Ns-raiA* (**A**), *Ca-raiA* (**B**), and *Mp-raiA* (**C**). The absence of P2 and distal P7 in *Ca-raiA* and *Mp-raiA*, and the absence of P8 in *Mp-raiA* are indicated by dashed lines. Sequence and secondary structure of *Ca-raiA* (**D**), and *Mp-raiA* (**E**). Black arrowheads indicate the backbone direction. Non-Watson–Crick base pairs are labeled as indicated. Insets show the schematics of J1 regions. Zoom-in views of the P1c-P2-P3a junction in *Ns-raiA* (**F**), and the P1c-P3a junctions in *Ca-raiA* (**G**) and *Mp-raiA* (**H**), highlighting the GAA(A) tetraloop fold, shown in the same orientation. Zoom-in views of the interface of P8 and PK1 stems in *Ns-raiA* (**I**) and *Ca-raiA* (**J**), and the PK1 stem in *Mp-raiA* (**K**), shown in the same orientation.

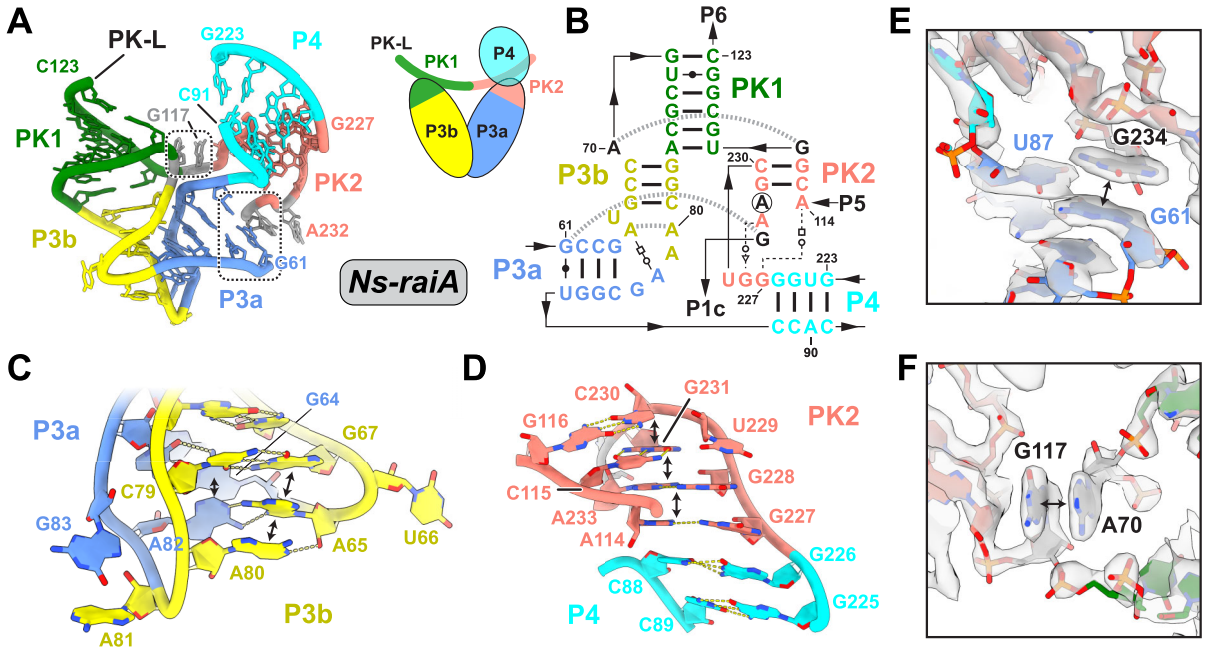


Figure 4. Structural details of the core of *Ns-raiA*. (**A**) Structure of the triangular-shaped core of *Ns-raiA* formed by P3, P4, PK1, and PK2 stems. Inset shows schematic of the structure, including single-strand loop linking PK1 and PK2 (PK-L). (**B**) Secondary structure of the core. Long-range stacking interactions are indicated by gray dashed lines. (**C**) Zoom-in view of the bend in the internal loop between P3a and P3b. (**D**) Zoom-in view of the junction between P4 and PK2 stem. (**E**, **F**) Zoom-in views of the long-range stacking interactions in the dashed box regions in panel (A). Double-headed arrows in C–F indicate base stacking.

resembling the kink-turn motifs originally identified in ribosomes [17, 41]. Its two stems, P3a and P3b, are connected by an asymmetric, adenosine-rich internal loop formed by A65-U66 on one strand and A80-A81-A82-G83 on the opposite strand (residue numbering follows *Ns-raiA* hereafter unless otherwise indicated). In the middle of the loop, A65 and A82 form a propeller-twisted base pair, with A65 stacking on G67 and A82 stacking on G64 (Fig. 4C). This A65-A82 base pair, together with the hydrogen-bond interactions between the P3a and P3b backbones, stabilizes the sharply bent architecture.

P4 is connected directly to P3a on its 5′ strand, while its 3′ strand is linked to P1c through an 8-nt linker (G227-G234) of notable structural complexity (Figs 2B and 4A and B). G228 and A233 form an internal base pair, orienting the backbone so that G227, G231, and C230 can pair with A114, C115, and G116 in the PK-L loop (Fig. 2B) between P5 and P6, respectively (Fig. 4D). Notably, only the C115-G231 and G116-C230 pairs were annotated as PK2 in the previously reported consensus secondary structure [14]. However, as these four base pairs stack continuously and are conserved across all three structures we determined, we include the two non-canonical G-A base pairs as part of the PK2 region (Fig. 4D and Supplementary Fig. S8).

Separated by a single nucleotide (G117), the subsequent nucleotides on PK-L (U118–C123) pair with the apical loop of P3b (G71–A76) to form PK1 stem—a six-base-pair helix that stacks directly on P3b (Figs 2B and 4A). Additional long-range stacking interactions further stabilize the core, including those between G234 and G61 at the end of P3a (Fig. 4E), A70 and G117 at the PK1–PK2 junction (Fig. 4F), and A65 and A80 at the P3a–P3b junction (Fig. 4C). Together, the triangular arrangement of P3a, P3b, and the two PK stems linked by PK-L, along with conserved tertiary interactions, locks the structure into a compact and stable core, which is preserved even in *Mp-raiA* that lacks the peripheral stems P2 and P8.

P1 is remarkably conserved

In addition to the triangular core, another conserved structural element across the three structures is P1, which is the closing stem formed by the 5′ and 3′ ends of *raiA* motif RNAs (Fig. 3A–C and Supplementary Fig. S9A–C). P1 consists of three sub-stems—P1a and P1b, which stack as a continuous helix, and P1c, which bends 90 degrees from P1b (Fig. 5A and B). These stems are connected by two highly conserved bulges: a GU bulge between P1a and P1b, and a GGU bulge between P1b and P1c (Fig. 5C and Supplementary Fig. S9D and F). In *Ns-raiA*, a long-range stacking interaction between G10 and G240 brings the two bulges together, while G10 also forms a major-groove base triple with G14 and C242 in P1b (Fig. 5B and C). Additionally, A81 and G83 on the internal loop between P3a and P3b insert into the junction of P1b and P1c, stacking with G240 on the P1b side and G239 on the P1c side, respectively (Fig. 5B and C). The G-G-C base triple and stacking interactions at the interface are also observed in the structures of *Ca-raiA* and *Mp-raiA* (Supplementary Fig. S9D and F) and are highly conserved (>97% identity) across all identified raiA motif RNAs [14], suggesting that they are important in maintaining the stable association of P1c with the core, as well as orienting P1a and P1b as a continuous long arm protruding from the core.

To investigate whether these structural features of P1 are conserved across *raiA* motif RNAs, we further analyzed the

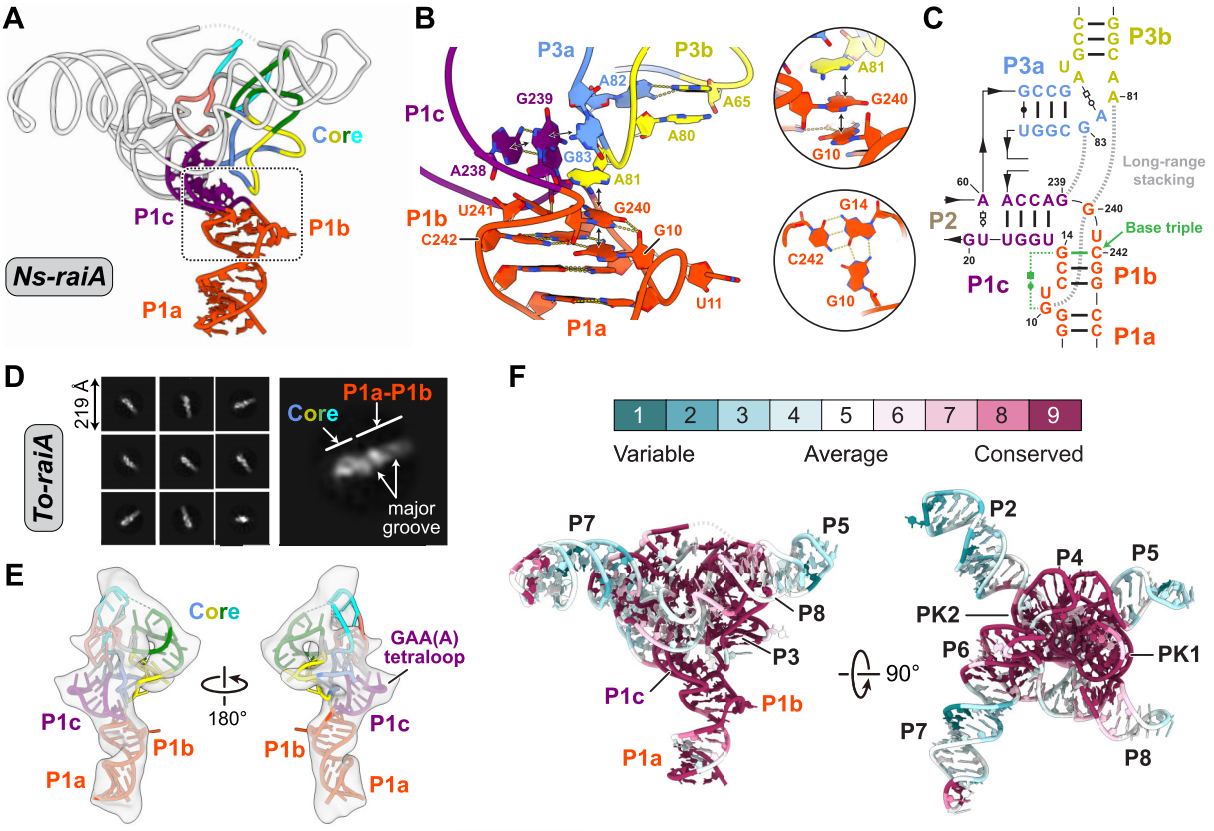


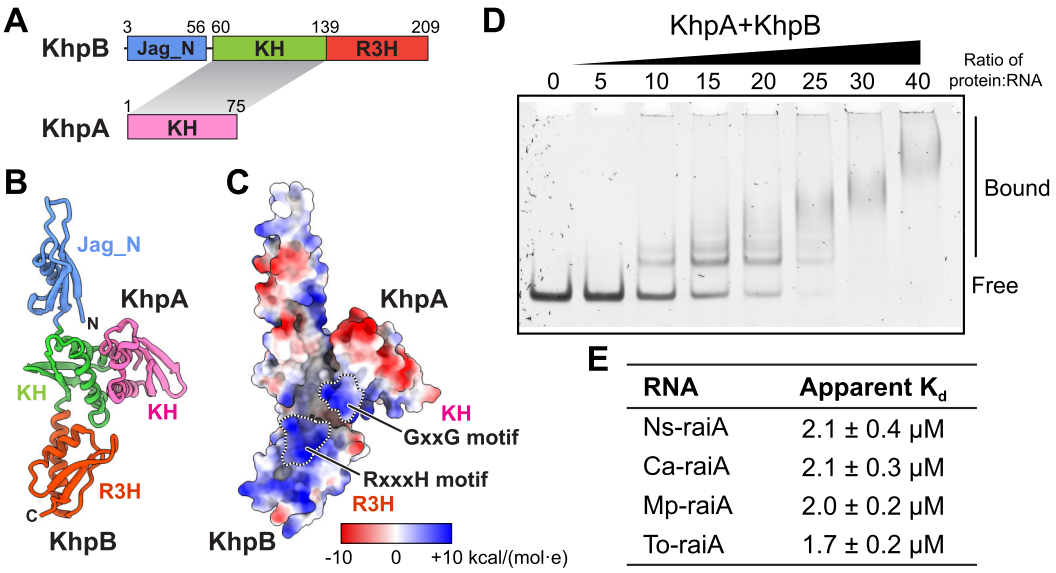
Figure 5. Structural details of P1 and its interactions with the core. (**A**) Overall view of P1 (shown as colored ribbon for backbone and filled bases and sugars) and its position relative to the core (colored ribbon) in the structure of *Ns-raiA*. Other stems are shown as white ribbons. (**B**) Close-up view of the interface between P1 and the core, in dashed box region in panel (A). The two insert panels highlight the long-range A81-G240-G10 stacking and the G10-G14-C242 base triple, respectively. (**C**) Secondary structure representation of the region shown in panel (B). Long-range stacking interactions are indicated by gray dashed lines, while base triple interactions are marked with green lines. (**D**) Representative 2D class averages of *To-raiA* (left) and enlargement with structure features labeled (right). (**E**) Cryo-EM density map and ribbon model of *To-raiA*. (**F**) Sequence conservation of *raiA* motif RNA mapped onto the *Ns-raiA* structure. View on left highlights conservation of P1 and on right conservation of the core. Conservation scores were calculated using the ConSurf server [56].

RNA from *Thermosediminibacter oceani* (*To-raiA*)—a "minimal" subtype of only 113 nucleotides—by cryo-EM (Fig. 5D and E and Supplementary Fig. S4). Based on the predicted secondary structure, *To-raiA* retains P3a, P3b, P4, and potentially PK1 and PK2 stems that collectively form the core, but lacks all peripheral stems except P1. 2D class averages of *To-raiA* revealed an elongated duplex with clear features for its major and minor grooves, corresponding to the stacked helix of P1a and P1b (Fig. 5D). Although the subsequent 3D refinement yielded a reconstruction at only ~8 Å resolution, likely limited by its small size of 37 kDa, all predicted structural elements of *To-raiA* can be assigned in the density map (Fig. 5E). Several other conserved features are discernible at this resolution, including the bend from P1b to P1c and the kink of the GAA(A) tetraloop in J1 (Fig. 5E and Supplementary Fig. S4G and H). Notably, the sequences of P1b and P1c are highly conserved, comparable to P3, P4, PK1, and PK2 in the core region, in contrast to the low sequence conservation observed in all the other peripheral stems, including P2, P5, P7, and P8 (Fig. 5F). Based on sequence conservation, the G-G-C base triple in P1 and the long-range A-G-G and G-G stacking interactions from P3 to P1 described above (Fig. 5C) also appear to be present in the minimal *To-raiA*.

KhpA and KhpB bind raiA motif RNAs as a heterodimer

Grad-seq analysis of ncRNAs and RBPs in *C. difficile*, a major Gram-positive human pathogen, revealed that *raiA* motif RNAs are highly abundant and form stable RNP complexes [16]. Subsequent pulldown assays using *raiA* motif RNA as bait significantly enriched two KH domain-containing proteins, KhpA and KhpB [16]. Previous studies have shown that these two proteins can function cooperatively, in the form of a heterodimer, to bind a broad range of RNAs in *S. pneumoniae* [42, 43] and *F. nucleatum* [44], and they are also capable of each forming homodimers *in vitro* [43, 45].

To investigate the interaction of KhpA and KhpB with *raiA* motif RNA and to explore whether they are part of a functional "*raiA* RNP," we purified the full-length *C. acetobutylicum* KhpA and KhpB recombinantly expressed in *Escherichia coli*, reconstituted the KhpA–KhpB heterodimer *in vitro*, and tested its interaction with *Ca-raiA* using electrophoretic mobility shift assays (EMSAs). KhpA is a small protein with a single KH domain, while KhpB contains an N-terminal Jag_N domain, a central KH domain, and a C-terminal R3H domain (Fig. 6A and B). Size exclusion chromatography results indicate that KhpA and KhpB can form a stable complex with a 1:1 stoichiometry



RNA	Apparent K_d
Ns-raiA	2.1 ± 0.4 μM
Ca-raiA	2.1 ± 0.3 μM
Mp-raiA	2.0 ± 0.2 μM
To-raiA	1.7 ± 0.2 μM

Figure 6. KhpA–KhpB heterodimer binds with *raiA* motif RNAs. (**A**) Domain diagrams of KhpA and KhpB from *C. acetobutylicum*. The predicted interface between the two proteins is indicated with gray shading. (**B**) AlphaFold3 model of KhpA–KhpB heterodimer shown as ribbon diagram. (**C**) Electrostatic surface of KhpA–KhpB heterodimer shown in panel (B). (**D**) Representative EMSA of *Ca-raiA* with KhpA–KhpB heterodimer. The RNA concentration was 0.2 μM for each sample, and the protein:RNA ratios are labeled for each lane. Gels and binding curves for *Ca-raiA* with KhpA only and KhpB only are shown in Supplementary Fig. S10D. (**E**) Summary of apparent dissociation constants of *C. acetobutylicum* KhpA–KhpB with all four *raiA* motif RNAs used in this paper. K_d values were determined as described in the "Materials and methods" section. Results were based on three biological replicates. Gels and binding curves for *Ns-raiA*, *Mp-raiA*, and *To-raiA* are shown in Supplementary Fig. S10E.

(Supplementary Fig. S10A and B). The KhpA–KhpB heterodimer structure predicted by AlphaFold3 [28] shows that KhpA and KhpB interact with each other through their KH domains (Fig. 6B), consistent with the interface previously identified in their homologs in *S. pneumoniae* [43]. In the predicted model, the two KH domains adopt a pseudo-C2 arrangement, with two α-helices from each protein contributing to the interface and jointly forming a four-helix bundle (Fig. 6B). Similar interactions between two KH domains were observed in the crystal structures of KhpB homodimer from *Clostridium symbiosum* (PDB ID 3GKU) and *Thermus thermophilus* [45] (Supplementary Fig. S11A and B). The predicted structure of KhpA–KhpB heterodimer also revealed two prominent positively charged surfaces: one on the KH domain of KhpA, encompassing the conserved GxxG loop motif, and the other on the R3H domain of KhpB, containing the signature RxxxH motif (Fig. 6C and Supplementary Fig. S11D and E). These positively charged surfaces may collectively engage with RNA substrates, providing a potential mechanism for cooperative recognition, resembling the RNA-binding surface of KhpB homodimer [45] (Supplementary Fig. S11A).

We tested the binding of KhpA–KhpB heterodimer with *Ca-raiA* using EMSA. We incubated *Ca-raiA* with increasing concentrations of KhpA–KhpB and observed a clear mobility shift of the RNA, with an apparent dissociation constant (Kd) of 2.1 ± 0.3 μM (Fig. 6D and E). At protein:RNA ratio of 10:1, a shifted complex was observed, whereas at higher protein:RNA ratios, the complexes migrated as a broad, higher molecular weight smear, indicating the formation of larger, heterogeneous assemblies (Fig. 6D). For KhpB only, no detectable gel shift was observed within the tested protein:RNA ratios, although there was some decrease in intensity of the RNA band, precluding reliable Kd fitting (Supplementary Fig. 10D). This is in agreement with a previous finding that the KhpA–KhpB complex binds small noncoding RNAs more strongly than KhpB alone in *F. nucleatum* [44]. For KhpA only, a small decrease in binding affinity compared to KhpA–KhpB was observed (Kd = 2.9 ± 0.6 μM). Previous studies suggested that KhpA alone lacks detectable RNA-binding activity [44] and is difficult to purify in isolation [45]. Our results indicate that KhpA alone can engage RNA, although its binding mode may differ from that within the KhpA–KhpB heterodimer, e.g. through homodimerization. We also tested the binding of KhpA–KhpB heterodimer with *Ns-raiA*, *Mp-raiA*, and *To-raiA* using EMSA (Supplementary Fig. S10E). They all exhibited clear mobility shifts with Kds of 2.1 ± 0.4 μM, 2.0 ± 0.2 μM, and 1.7 ± 0.2 μM, respectively, comparable to that of *Ca-raiA* (Fig. 6E). Similar to *Ca-raiA*, higher KhpA–KhpB concentrations drove the formation of heterogenous complexes (Supplementary Fig. S10E). Taken together, these results establish that KhpA and KhpB form a stable heterodimer *in vitro* and engage *raiA* motif RNAs with low-micromolar affinity.

Discussion

Structured ncRNAs often adopt intricate folds that enable diverse functions, as seen in self-cleaving ribozymes with conserved catalytic cores, riboswitch aptamers with specific ligand-binding pockets, and the recently discovered large RNA nanocages assembled from the ROOL and GOLLD ncRNAs [18, 46–48]. Here, with the four cryo-EM structures of *raiA* motif RNAs representing different subtypes, we define the consensus architecture of *raiA* motif RNA and identify conserved structural elements and interactions critical for fold stabilization. Previously reported DMS-MaP experiments of *raiA* motif RNA in *C. difficile* showed that the fold is preserved both *in vitro* and *in vivo* [15]. Our structures reveal that the intrinsic stability of *raiA* motif RNA originates from

the compact core formed by P3a–P3b, P4, PK1, and PK2, together with the interactions between the core and the extruding P1 stem, specifically P1b and P1c. By contrast, contributions from the other peripheral stems are limited, as the overall fold is maintained in *Ca-raiA* (lacking P2), *Mp-raiA* (lacking P2 and P8), and *To-raiA* (lacking P2 and P5–P8).

Although the sequence of P1a is less conserved, its length is highly conserved [14]. This is unusual, as P1a protrudes entirely from the core without direct contact and would typically be expected to contribute only to bringing together the 5′ and 3′ ends for structural stabilization, where strict length conservation is generally not required. Based on these observations, we speculate that P1 may play a functional role, potentially by mediating interactions or undergoing conformational changes that are important for the yet-unclear function of *raiA* motif RNA. Notably, unlike the single form of *Ca-raiA* detected in *C. acetobutylicum*, two isoforms of *raiA* motif RNA were detected in some bacteria, including *C. difficile* [15, 16]. The longer isoform, which is likely generated by read-through transcription, has an additional stem loop attached to P1 at the 3′ end [15]. DMS probing of the two isoforms revealed a consensus structure of the shared region [15], similar to that of *Ca-raiA*, suggesting that this additional terminal stem would not alter the structure and the potential function of P1.

RaiA motif RNA has been shown to function in *trans*, instead of as a *cis* element like a riboswitch, as the phenotypes caused by its deletion can be rescued by restoring its expression [14, 15]. In *C. difficile*, *raiA* motif RNA regulates sporulation and the transition from exponential growth into stationary phase by modulating the cellular level of the second messenger c-di-GMP [15]. The consensus structure of *raiA* motif RNA resolved in this work does not resemble that of known c-di-GMP riboswitches [49–51], suggesting that it is unlikely to function as a reservoir for the second messenger. Nevertheless, we cannot rule out the possibility that it binds c-di-GMP in a non-canonical manner or that c-di-GMP binding induces a substantial conformational change. These possibilities remain to be investigated.

In the three high-resolution structures of *raiA* motif RNAs reported here, the 10-nt single-stranded linker between PK1 and P6 is solvent-exposed and less well-resolved than its neighboring elements, indicating substantial flexibility (Fig. 2A and B). Our data show that KhpA–KhpB heterodimer binds *raiA* motif RNAs with low-micromolar affinity (Fig. 6E). We propose that this exposed linker may be the binding site of KhpA–KhpB, based on the following evidence: (i) both KH and R3H domains preferentially engage single-stranded RNA [52, 53]; (ii) *Thermus thermophilus* KhpB binds single-stranded RNAs longer than 7 nucleotides, with binding affinity saturating at 9 nucleotides [45]; (iii) this linker is the only solvent-exposed single-stranded region longer than 7 nucleotides in the structure (Fig. 2B and C). Supporting this idea, docking of the AlphaFold3-predicted KhpA–KhpB heterodimer structure onto this region, based on the recently reported KhpB homodimer–RNA complex structure [45], revealed no steric clashes. To further probe this interaction, we generated 179 models of KhpA–KhpB bound to *Ca-raiA* using HADDOCK [54], which were clustered into 11 groups. These models revealed two major binding modes of KhpA–KhpB heterodimer on the interface that includes the single-strand region, both of which exhibit buried surface areas over 1000 $Å^2$ at the interface (Supplementary Fig. S11C). The conserved residues of KhpA–KhpB at the RNA binding interface (Supplementary Fig. S11D and E) would explain why *C. acetobutalicum* KhpA–KhpB can bind *raiA* motif RNAs from different bacteria with comparable affinity (Fig. 6E), suggesting a shared recognition mechanism that likely involves the compositionally conserved, solvent-exposed single-stranded linker between PK1 and P6. Notably, the sequence of this linker, which is UUAGA(C/A)GUAA in the three structures, is highly conserved in all *raiA* motif RNAs [14]. This conserved sequence does not appear to be critical for maintaining the overall fold but may serve as a recognition site for sequence-specific protein or RNA binding, which could be important for the yet-to-be-discovered function of *raiA* motif RNA.

Together, the high-resolution structures of *raiA* motif RNAs presented in this study reveal conserved three-dimensional structural elements and interactions essential for maintaining the overall fold, providing the structural basis for investigating the function of this abundant bacterial ncRNA. The architecture of *raiA* motif RNA is novel and intrinsically stable, which offers a promising framework for rational functional RNA design.

Acknowledgements

We thank M. Summers for the gift of the plasmid encoding the T7 RNA polymerase G47A + 884G mutant.

Author contributions: Y.H.: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. J.Z.: Methodology, Formal analysis, Investigation, Data curation, Writing – review & editing. Y.Y.: Methodology, Investigation, Validation, Writing – review & editing. R.P.G.: Conceptualization, Resources, Writing – review & editing. Z.H.Z.: Resources, Writing – review & editing, Funding acquisition. J.F.: Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing, Supervision, Project Administration, Funding acquisition.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

This work was supported by National Institutes of Health (NIH) grants R35GM131901 to J.F. and R01GM071940 to Z.H.Z, and by the Department of Energy (DOE) BER award DE-FC02-02ER63421 to the UCLA DOE Institute. The Electron Imaging Center for Nanomachines was supported in part by instrumentation grants from NIH (S10RR23057) and NSF (DBI-1338135 and DMR-1548924). Funding to pay the Open Access publication charges for this article was provided by National Institutes of Health (NIH) grants R35GM131901 to J.F.

Data availability

Cryo-EM density maps of *raiA* motif RNAs have been deposited in the Electron Microscopy Data Bank under accession codes EMD-47657 (*Ns-raiA*), EMD-47655

(*Ca-raiA*), EMD-47658 (*Mp-raiA*), and EMD-72663 (*To-raiA*), respectively. The atomic models have been deposited in the Protein Data Bank under accession codes 9E74 (*Ns-raiA*), 9E73 (*Ca-raiA*), and 9E75 (*Mp-raiA*).

References

1. Cech TR, Steitz JA. The noncoding RNA revolution-trashing old rules to forge new ones. *Cell* 2014;157:77–94. https://doi.org/10.1016/j.cell.2014.03.008
2. Su Z, Zhang K, Kappel K *et al.* Cryo-EM structures of full-length Tetrahymena ribozyme at 3.1 A resolution. *Nature* 2021;596:603–7. https://doi.org/10.1038/s41586-021-03803-w
3. Bonilla SL, Sherlock ME, MacFadden A *et al.* A viral RNA hijacks host machinery using dynamic conformational changes of a tRNA-like structure. *Science* 2021;374:955–60. https://doi.org/10.1126/science.abe8526
4. Zhang K, Li S, Kappel K *et al.* Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 A resolution. *Nat Commun* 2019;10:5511. https://doi.org/10.1038/s41467-019-13494-7
5. Bonilla SL, Kieft JS. The promise of cryo-EM to explore RNA structural dynamics. *J Mol Biol* 2022;434:167802. https://doi.org/10.1016/j.jmb.2022.167802
6. Luo BN, Zhang C, Ling XB *et al.* Cryo-EM reveals dynamics of *Tetrahymena* group I intron self-splicing. *Nat Catal* 2023;6:298–309. https://doi.org/10.1038/s41929-023-00934-3
7. Li S, Palo MZ, Zhang X *et al.* Snapshots of the second-step self-splicing of *Tetrahymena* ribozyme revealed by cryo-EM. *Nat Commun* 2023;14:1294. https://doi.org/10.1038/s41467-023-36724-5
8. Yang W, Yi R, Yao J *et al.* Structural insights into dynamics of the BMV TLS aminoacylation. *Nat Commun* 2025;16:1276. https://doi.org/10.1038/s41467-025-56612-4
9. Ling X, Yao Y, Ma J. Structures of a natural circularly permuted group II intron reveal mechanisms of branching and backsplicing. *Nat Struct Mol Biol* 2025;32:1101–10. https://doi.org/10.1038/s41594-025-01489-6
10. Zhang X, Li S, Pintilie G *et al.* Snapshots of the first-step self-splicing of *Tetrahymena* ribozyme revealed by cryo-EM. *Nucleic Acids Res* 2023;51:1317–25. https://doi.org/10.1093/nar/gkac1268
11. Ding J, Deme JC, Stagno JR *et al.* Capturing heterogeneous conformers of cobalamin riboswitch by cryo-EM. *Nucleic Acids Res* 2023;51:9952–60. https://doi.org/10.1093/nar/gkad651
12. Bonilla SL, Vicens Q, Kieft JS. Cryo-EM reveals an entangled kinetic trap in the folding of a catalytic RNA. *Sci Adv* 2022;8:eabq4144. https://doi.org/10.1126/sciadv.abq4144
13. Weinberg Z, Lunse CE, Corbino KA *et al.* Detection of 224 candidate structured RNAs by comparative analysis of specific subsets of intergenic regions. *Nucleic Acids Res* 2017;45:10811–23. https://doi.org/10.1093/nar/gkx699
14. Soares LW, King CG, Fernando CM *et al.* Genetic disruption of the bacterial raiA motif noncoding RNA causes defects in sporulation and aggregation. *Proc Natl Acad Sci USA* 2024;121:e2318008121. https://doi.org/10.1073/pnas.2318008121
15. Lence T, Sulzer J, Andress K *et al.* The conserved noncoding RNA ModT coordinates growth and virulence in *Clostridioides difficile*. *PLoS Biol* 2024;22:e3002948. https://doi.org/10.1371/journal.pbio.3002948
16. Lamm-Schmidt V, Fuchs M, Sulzer J *et al.* Grad-seq identifies KhpB as a global RNA-binding protein in *Clostridioides difficile* that regulates toxin production. *Microlife* 2021;2:uqab004. https://doi.org/10.1093/femsml/uqab004
17. Badepally NG, de Moura TR, Purta E *et al.* Cryo-EM structure of raiA ncRNA from *Clostridium* reveals a new RNA 3D fold. *J Mol Biol* 2024;436:168833. https://doi.org/10.1016/j.jmb.2024.168833
18. Kretsch RC, Wu Y, Shabalina SA *et al.* Naturally ornate RNA-only complexes revealed by cryo-EM. *Nature* 2025;643:1135–42. https://doi.org/10.1038/s41586-025-09073-0
19. Haack DB, Rudolfs B, Jin S *et al.* Scaffold-enabled high-resolution cryo-EM structure determination of RNA. *Nat Commun* 2025;16:880. https://doi.org/10.1038/s41467-024-55699-5
20. Guillerez J, Lopez PJ, Proux F *et al.* A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc Natl Acad Sci USA* 2005;102:5958–63. https://doi.org/10.1073/pnas.0407141102
21. Dousis A, Ravichandran K, Hobert EM *et al.* An engineered T7 RNA polymerase that produces mRNA free of immunostimulatory byproducts. *Nat Biotechnol* 2023;41:560–8. https://doi.org/10.1038/s41587-022-01525-6
22. Wang Y, He Y, Wang Y *et al.* Structure of LARP7 protein p65–telomerase RNA complex in telomerase revealed by Cryo-EM and NMR. *J Mol Biol* 2023;435:168044. https://doi.org/10.1016/j.jmb.2023.168044
23. Mastronarde DN. Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* 2005;152:36–51. https://doi.org/10.1016/j.jsb.2005.07.007
24. Punjani A, Rubinstein JL, Fleet DJ *et al.* cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* 2017;14:290–6. https://doi.org/10.1038/nmeth.4169
25. Punjani A, Zhang H, Fleet DJ. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat Methods* 2020;17:1214–21. https://doi.org/10.1038/s41592-020-00990-8
26. Bepler T, Morin A, Rapp M *et al.* Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat Methods* 2019;16:1153–60. https://doi.org/10.1038/s41592-019-0575-8
27. Emsley P, Lohkamp B, Scott WG *et al.* Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 2010;66:486–501. https://doi.org/10.1107/S0907444910007493
28. Abramson J, Adler J, Dunger J *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;630:493–500. https://doi.org/10.1038/s41586-024-07487-w
29. Adams PD, Afonine PV, Bunkoczi G *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 2010;66:213–21. https://doi.org/10.1107/S0907444909052925
30. Croll TI. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* 2018;74:519–30. https://doi.org/10.1107/S2059798318002425
31. Goddard TD, Huang CC, Meng EC *et al.* UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci* 2018;27:14–25. https://doi.org/10.1002/pro.3235
32. Pintilie G, Zhang K, Su Z *et al.* Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nat Methods* 2020;17:328–34. https://doi.org/10.1038/s41592-020-0731-1
33. Antczak M, Zok T, Popenda M *et al.* RNApdbee—a webserver to derive secondary structures from pdb files of knotted and unknotted RNAs. *Nucleic Acids Res* 2014;42:W368–72. https://doi.org/10.1093/nar/gku330
34. Waterhouse AM, Procter JB, Martin DM *et al.* Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 2009;25:1189–91. https://doi.org/10.1093/bioinformatics/btp033
35. Jucker FM, Heus HA, Yip PF *et al.* A network of heterogeneous hydrogen bonds in GNRA tetraloops. *J Mol Biol* 1996;264:968–80. https://doi.org/10.1006/jmbi.1996.0690
36. Cate JH, Gooding AR, Podell E *et al.* Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* 1996;273:1678–85. https://doi.org/10.1126/science.273.5282.1678

37. Chan CW, Badong D, Rajan R *et al.* Crystal structures of an unmodified bacterial tRNA reveal intrinsic structural flexibility and plasticity as general properties of unbound tRNAs. *RNA* 2020;26:278–89. https://doi.org/10.1261/rna.073478.119
38. Dieckmann T, Suzuki E, Nakamura GK *et al.* Solution structure of an ATP-binding RNA aptamer reveals a novel fold. *RNA* 1996;2:628–40.
39. Leontis NB, Westhof E. A common motif organizes the structure of multi-helix loops in 16 S and 23 S ribosomal RNAs. *J Mol Biol* 1998;283:571–83. https://doi.org/10.1006/jmbi.1998.2106
40. Leontis NB, Westhof E. The 5S rRNA loop E: chemical probing and phylogenetic data versus crystal structure. *RNA* 1998;4:1134–53. https://doi.org/10.1017/S1355838298980566
41. Klein DJ, Schmeing TM, Moore PB *et al.* The kink-turn: a new RNA secondary structure motif. *EMBO J* 2001;20:4214–21. https://doi.org/10.1093/emboj/20.15.4214
42. Zheng JJ, Perez AJ, Tsui HT *et al.* Absence of the KhpA and KhpB (JAG/EloR) RNA-binding proteins suppresses the requirement for PBP2b by overproduction of FtsA in *Streptococcus pneumoniae* D39. *Mol Microbiol* 2017;106:793–814. https://doi.org/10.1111/mmi.13847
43. Winther AR, Kjos M, Stamsas GA *et al.* Prevention of EloR/KhpA heterodimerization by introduction of site-specific amino acid substitutions renders the essential elongasome protein PBP2b redundant in *Streptococcus pneumoniae*. *Sci Rep* 2019;9:3681. https://doi.org/10.1038/s41598-018-38386-6
44. Zhu Y, Ponath F, Cosi V *et al.* A global survey of small RNA interactors identifies KhpA and KhpB as major RNA-binding proteins in *Fusobacterium nucleatum*. *Nucleic Acids Res* 2024;52:3950–70. https://doi.org/10.1093/nar/gkae010
45. Fukui K, Murakawa T, Baba S *et al.* KH-R3H domain cooperation in RNA recognition by the global RNA-binding protein KhpB. *Nat Commun* 2025;16:8028. https://doi.org/10.1038/s41467-025-62302-y
46. Wang L, Xie J, Gong T *et al.* Cryo-EM reveals mechanisms of natural RNA multivalency. *Science* 2025;388:545–50. https://doi.org/10.1126/science.adv3451
47. Ling X, Golovenko D, Gan J *et al.* Cryo-EM structure of a natural RNA nanocage. *Nature* 2025;644:1107–15. https://doi.org/10.1038/s41586-025-09262-x
48. Zhang S, Yi R, An L *et al.* Structural insights into higher-order natural RNA-only multimers. *Nat Struct Mol Biol* 2025;32:2012–21. https://doi.org/10.1038/s41594-025-01650-1
49. Kulshina N, Baird NJ, Ferre-D'Amare AR. Recognition of the bacterial second messenger cyclic diguanylate by its cognate riboswitch. *Nat Struct Mol Biol* 2009;16:1212–7. https://doi.org/10.1038/nsmb.1701
50. Smith KD, Lipchock SV, Ames TD *et al.* Structural basis of ligand binding by a c-di-GMP riboswitch. *Nat Struct Mol Biol* 2009;16:1218–23. https://doi.org/10.1038/nsmb.1702
51. Smith KD, Shanahan CA, Moore EL *et al.* Structural basis of differential ligand recognition by two classes of bis-(3′–5′)-cyclic dimeric guanosine monophosphate-binding riboswitches. *Proc Natl Acad Sci USA* 2011;108:7757–62. https://doi.org/10.1073/pnas.1018857108
52. Nicastro G, Taylor IA, Ramos A. KH–RNA interactions: back in the groove. *Curr Opin Struct Biol* 2015;30:63–70. https://doi.org/10.1016/j.sbi.2015.01.002
53. Grishin NV. The R3H motif: a domain that binds single-stranded nucleic acids. *Trends Biochem Sci* 1998;23:329–30. https://doi.org/10.1016/S0968-0004(98)01258-4
54. Honorato RV, Trellet ME, Jimenez-Garcia B *et al.* The HADDOCK2.4 web server for integrative modeling of biomolecular complexes. *Nat Protoc* 2024;19:3219–41. https://doi.org/10.1038/s41596-024-01011-0
55. Leontis NB, Westhof E. Geometric nomenclature and classification of RNA base pairs. *RNA* 2001;7:499–512. https://doi.org/10.1017/S1355838201002515
56. Ashkenazy H, Abadi S, Martz E *et al.* ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* 2016;44:W344–50. https://doi.org/10.1093/nar/gkw408

Received: September 20, 2025. **Revised:** February 5, 2026. **Accepted:** February 7, 2026