Cryo-EM structures of full-length *Tetrahymena* ribozyme at 3.1 Å resolution

https://doi.org/10.1038/s41586-021-03803-w

Received: 13 January 2021

Accepted: 6 July 2021

Published online: 11 August 2021

Check for updates

Zhaoming Su[1,2,7]✉, Kaiming Zhang[2,3,7], Kalli Kappel[4,7], Shanshan Li[2,3], Michael Z. Palo[4], Grigore D. Pintilie[2], Ramya Rangan[4], Bingnan Luo[1], Yuquan Wei[1], Rhiju Das[4,5]✉ & Wah Chiu[2,6]✉

Single-particle cryogenic electron microscopy (cryo-EM) has become a standard technique for determining protein structures at atomic resolution[1–3]. However, cryo-EM studies of protein-free RNA are in their early days. The *Tetrahymena thermophila* group I self-splicing intron was the first ribozyme to be discovered and has been a prominent model system for the study of RNA catalysis and structure–function relationships[4], but its full structure remains unknown. Here we report cryo-EM structures of the full-length *Tetrahymena* ribozyme in substrate-free and bound states at a resolution of 3.1 Å. Newly resolved peripheral regions form two coaxially stacked helices; these are interconnected by two kissing loop pseudoknots that wrap around the catalytic core and include two previously unforeseen (to our knowledge) tertiary interactions. The global architecture is nearly identical in both states; only the internal guide sequence and guanosine binding site undergo a large conformational change and a localized shift, respectively, upon binding of RNA substrates. These results provide a long-sought structural view of a paradigmatic RNA enzyme and signal a new era for the cryo-EM-based study of structure–function relationships in ribozymes.

RNAs can fold into complex tertiary structures and participate in important biological processes, such as catalysis and transcriptional and translational regulation, in the absence of proteins[5,6]. However, our understanding of RNA structure–function relationships remains limited owing to a lack of RNA structural information. This paucity arises from the challenges posed by the intrinsic heterogeneity of RNAs to conventional X-ray crystallography and nuclear magnetic resonance[7,8]. Single-particle cryo-EM is an alternative method for determining structures, but its application to RNA has been limited, like that of previous methods. At present, fewer than ten protein-free RNA cryo-EM maps at resolutions better than 5 Å exist, with the best resolved at 3.7 Å resolution[9]. Recently, we developed an accelerated pipeline—Ribosolve—to determine 11 protein-free RNA structures through sub-nanometre-resolution cryo-EM maps, secondary structure mapping using M2-seq (mutate-and-map read out through next-generation sequencing), and Rosetta computational modelling[10]. A full-length structure of the wild-type *Tetrahymena* ribozyme at 6.8 Å resolution was a highlight of this approach.

In 1982, the *Tetrahymena* group I self-splicing intron was discovered as the first example of a protein-free RNA catalyst, and the term 'ribozyme' was coined[4]. This ribozyme catalyses two successive transesterification reactions to cleave the 5′ splice site and then ligate the 5′ and 3′ exons. Extensive studies indicated that the highly conserved core forms a compact structure[11–15] (and references therein). Metal ions, especially Mg^{2+}, are essential for the stabilization of RNA structures and for their catalytic reactions[12–14,16–22]. The ribozyme's peripheral regions have been predicted to form long-range interactions to stabilize the core[23–27]. Deletions and mutations of these regions affect the ribozyme's folding pathways and stability and allosterically regulate catalysis[23,24,26,28].

Despite extensive efforts, including ground-breaking crystal structures of subdomains[12–15,29–31], the complete structure of the ribozyme remains unavailable. Here we determine structures of the full-length *Tetrahymena* ribozyme at 3.1 Å resolution in both apo and holo states, the latter bound to two RNA substrates to mimic the second step of splicing. Our structures reveal tertiary interactions that have not previously been described, to our knowledge, and conformational changes in the internal guide sequence (IGS) and catalytic site upon substrate binding, providing structural and mechanistic insights into this classic RNA enzyme.

Apo L-21 ScaI ribozyme overall structure

We obtained a 3.1 Å cryo-EM structure of the apo L-21 ScaI ribozyme, a linear form of the self-splicing intron without its first 21 nucleotides (nts 22–409), with transcription termination at a ScaI restriction endonuclease site in the DNA template (Fig. 1a, Extended Data Fig. 1, Extended Data Table 1). The catalytic core P3–P7 (in which P denotes a 'paired' double helix region) is resolved to 3.0 Å, whereas flexible regions such as P9.2 and P13 are resolved to 4.6 Å. Further 3D classifications reveal modest

[1]The State Key Laboratory of Biotherapy and Cancer Center, Department of Geriatrics and National Clinical Research Center for Geriatrics, West China Hospital, Sichuan University, Chengdu, China. [2]Department of Bioengineering and James H. Clark Center, Stanford University, Stanford, CA, USA. [3]MOE Key Laboratory for Membraneless Organelles and Cellular Dynamics, Hefei National Laboratory for Physical Sciences at the Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, China. [4]Biophysics Program, Stanford University, Stanford, CA, USA. [5]Department of Biochemistry and Department of Physics, Stanford University, Stanford, CA, USA. [6]Division of CryoEM and Bioimaging, SSRL, SLAC National Accelerator Laboratory, Menlo Park, CA, USA. [7]These authors contributed equally: Zhaoming Su, Kaiming Zhang, Kalli Kappel. ✉e-mail: zsu@scu.edu.cn; rhiju@stanford.edu; wahc@stanford.edu

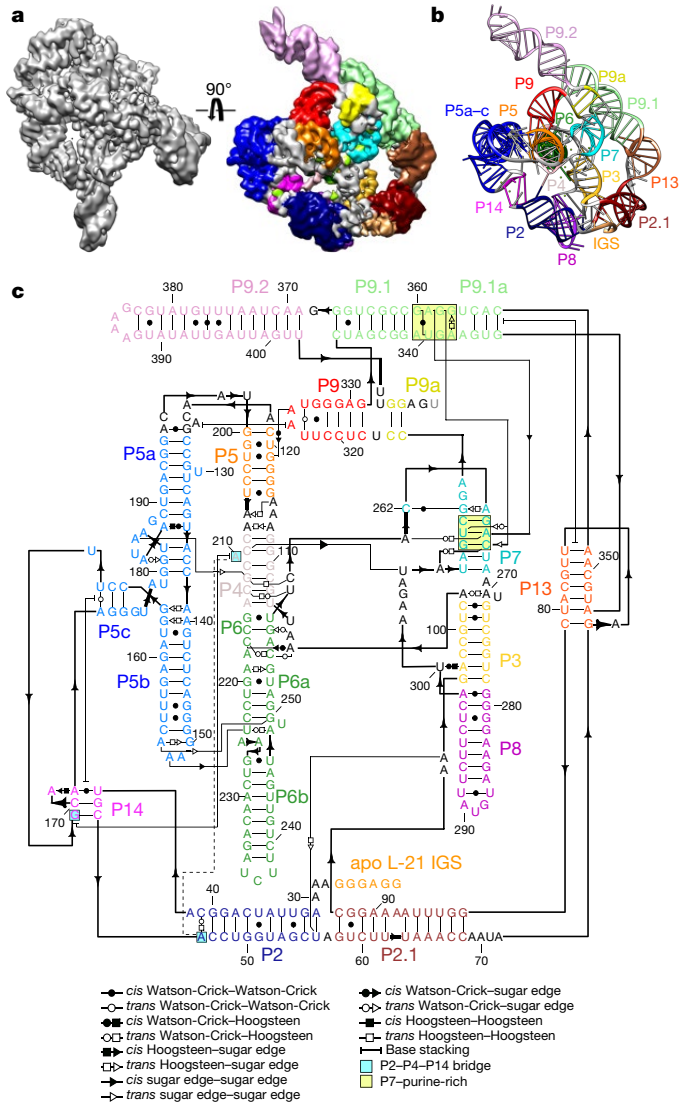


Fig. 1 | Cryo-EM reconstruction of apo L-21 ScaI ribozyme. a, Cryo-EM map at 3.1 Å (left) and the segmented map (right) coloured according to the secondary structure. **b**, Coloured cryo-EM model. **c**, Secondary structure based on the cryo-EM model. Thick lines with arrows denote base continuity; other lines with different symbols denote different base pairings and stackings. The dashed line indicates a hydrogen bond in the P2–P4–P14 bridge.

conformational heterogeneity (Extended Data Fig. 2). The resolvability of the cryo-EM density allowed us to refine the initial Ribosolve model[10] (Fig. 1b, c) to a structure that is well validated by *Q*-score analysis[32] (which measures the resolvability of individual atoms or clusters of atoms in cryo-EM maps) (Extended Data Fig. 3).

The apo L-21 ScaI ribozyme consists of the P3–P9 domain and newly resolved peripheral regions (Fig. 1c). The ribozyme strand starts from the 5′ IGS that binds the oligonucleotide substrate, followed by the 11-base-pair (bp) P2, in which the loop (L2) forms a 3-bp pseudoknot (PK) with L5c called P14. After P2, the 11-bp P2.1 presents L2.1, which forms a 7-bp PK P13 with L9.1, and connects to P2.1 via a 4-nt bulge. After P2.1, the ribozyme continues to residues 96–331, which includes the intricately pseudoknotted P3–P9 section that forms substrate binding sites. The strand continues to P9.1, which exhibits a novel tertiary interaction with the P7 catalytic site. P9.2 is a long stem adjacent to P9.1 that points away from the rest of the ribozyme. It is followed by the 2-bp P9a, which returns the ribozyme's 3′ end close to the catalytic site (Supplementary Video 1).

The cryo-EM structure of the catalytic core is largely consistent with the crystal structures of other group I ribozymes (Extended Data Figs. 4–6, Supplementary Table 1). The inclusion of the peripheral domains provides structural and functional insights, as described below.

Unpredicted noncanonical interactions

The functional relevance of the peripheral regions has been explored in previous investigations of the allosteric regulation of the catalytic process in the *Tetrahymena* ribozyme[24,33]. Our structure explains previous structural modelling and biochemical findings (Fig. 2). At a domain level, previous work predicted that the peripheral regions would stack coaxially and wrap around the catalytic core[27], which generally agrees with our structure, although there is an almost 90° bend involving L2.1 between two long domains (P5c–P14–P2–P2.1 and P13–P9.1–P9.2) (Fig. 2, centre).

At a nucleotide level, we resolve numerous noncanonical interactions that are likely to influence function, and we highlight here two base triples as examples. First, previous crystal structures of the P4–P6 domain revealed a noncanonical 'A-platform' motif in A171–A172 of L5c, which was predicted to potentially interact with peripheral regions[12,29]. Our structure shows that the A171–A172 platform is indeed involved in a U43–A171–A172 base triple that connects P5c and the peripheral helix P2 via P14 (Fig. 2a–c, Extended Data Fig. 6e). Second, we identified a base triple A31–U56–A95 that is stacked on both sides and positioned at the P2–P2.1 junction (Fig. 2d, e), which is the 'anchor point' for the ribozyme's IGS without substrates. Disruption of this junction would result in incorrect splice sites through changes in tertiary structure (Supplementary Video 1), consistent with mutational analysis[34].

New interactions with allosteric effects

The cryo-EM structure illuminates two tertiary contacts that have not been previously seen, to our knowledge, and both of which have potential importance for how peripheral domains affect catalysis. One contact, termed here 'P2–P4–P14 bridge', involves A210 flipping out of P4 and forming a hydrogen bond with A46 in P2 and a base stack with G169 in P14 (Fig. 2b, c). This provides an explanation for the existence of the bulged A210, which was found to be destabilizing in the context of an isolated P4–P6 domain but important for folding and function of the full-length *Tetrahymena* ribozyme[15]. Disruption of P14 formation led to a decrease in oligonucleotide substrate docking via an allosteric effect of P2 and the P2–P2.1 junction[24,33], which is now explained directly by the P2–P4–P14 bridge.

The second contact is termed here the 'P7–purine-rich interaction'. Previous observations suggested that P9.1–P9.2 affected the folding process and stability of the catalytic site P3–P7[35]. Our structure reveals that a purine-rich internal loop in P9.1 (nts 339–342 and 357–360) forms a noncanonical G341–G357 pair that resembles an internal loop E structure[36] (Extended Data Fig. 6f), while G358, A359 and G360 form noncanonical base pairs with the minor groove of P7 (Fig. 2f–i). P7 contains the guanosine binding site and forms base triples A265–U310–A261 and C266–G309–A306 at the catalytic site[14] (Fig. 2f, h, i). The P7–purine-rich interaction represents a tertiary contact that is closer (17 Å) to the catalytic core than the previously proposed five tertiary contacts (23–48 Å), which might explain why mutations of P9.1–P9.2 lead to the largest effect on the stability of catalytic core folding[24,33,35].

Rearrangement of extended IGS

A classic crosslinking study indicated that the IGS of the ribozyme moves by 37 Å upon substrate binding, providing the earliest evidence that RNA can undergo large-scale conformational changes and leading to the important hypothesis that such RNA flexibility is conserved in ancient machines such as the ribosome and spliceosome[37].

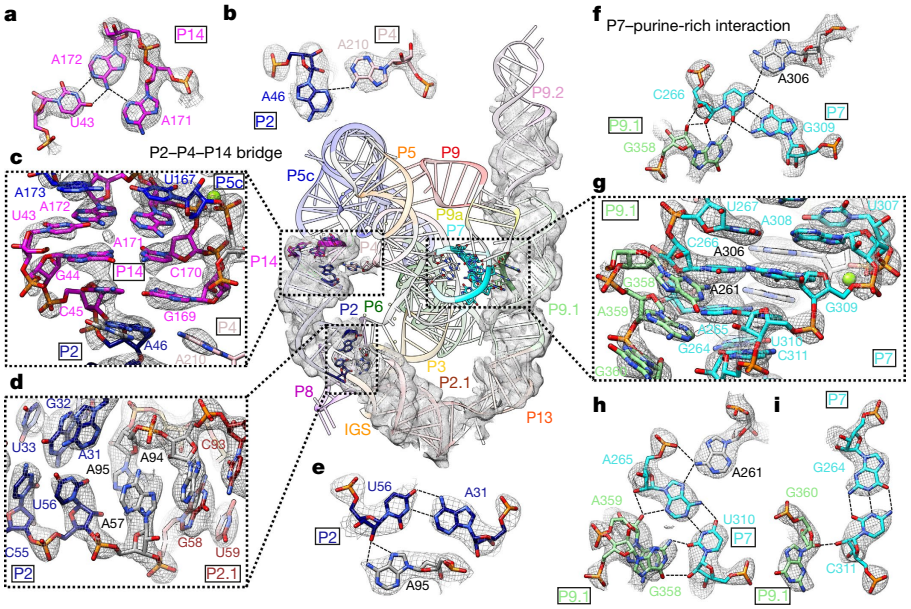


Fig. 2 | Structural insights in the peripheral regions. a, Base triple consisting of U43–A171–A172 connects P5c with P14. **b**, A46 and A210 form a P2–P4–P14 bridge. **c**, P14 coaxially stacks with P5c and P2; region includes base triple in **a** and P2–P4–P14 bridge in **b**. **d**, P2 forms a A31–U56–A95 base triple to connect with P2.1. **e**, The A31–U56–A95 base triple, part of **d**. **f**, G358 interacts with C266–G309–A306 to form a base quartet. **g**, The P7–purine-rich interaction, which includes the interaction in **f**. **h**, G358 and A359 interact with A265–U310–A261 to form base quartets, part of **g**. **i**, G360 interacts with G264–C311 to form a base triple, also part of **g**. The peripheral region is presented as a cryo-EM map visualized at 3σ threshold. Local regions for all subpanels are visualized at 1σ threshold except for **c** (2σ) and **d**, **g** (1.5σ). Dashed lines, hydrogen bonds.

To understand this substantial conformational change of the ribozyme, we obtained a 3.1 Å cryo-EM map of the L-16 ScaI ribozyme (nts 17–409) bound to two RNA substrates. Substrate S1 contains the terminal G414 (ωG) of the intron and the 3′ splice site, whereas substrate S2 comprises the cleaved 5′ exon. We introduced a phosphorothioate substitution at the scissile phosphate in S2 to isolate the docked complex, thereby mimicking the start of the second step of splicing[38–40] (Fig. 3a–c). The holo ribozyme adopts an identical overall architecture with similar flexible regions compared to the apo ribozyme (Fig. 3d, Extended Data Fig. 7), except that the extended IGS (nts 17–27) undergoes a substantial conformational change to form a 4-bp P10 and a 6-bp P1 with substrates S1 (8 nt) and S2 (6 nt, defined as P in previous studies[40,41]) (Fig. 3c, e–g, Supplementary Video 2). The 5′ end of S1 (nts 412–414) mimics the *Tetrahymena* intron's 3′ terminus, forming a 2-bp P7 extension to place ωG in the guanosine binding site (Fig. 3e)—the same site that binds an exogenous guanosine that is the nucleophile for the first step of splicing[13,14,17,18,20]. The 3′ end of S1 mimics the 3′ exon (Fig. 3a, b) and forms P10 with the 5-nt IGS extension (nts 17–21) (Fig. 3e). S2 mimics the 5′ exon of the *Tetrahymena* RNA (Fig. 3a, b) and forms P1 with the 6-nt IGS (nts 22–27) (Fig. 3f). Comparison between the apo and holo structures shows an approximately 60° conformational change of the IGS, moving its 5′ end from A87, A88 and A89—previously observed sites of crosslinking—into the catalytic core (Fig. 3g, Supplementary Video 2).

Local shifts in guanosine binding site

Previous studies based on pre-steady-state enzymology[39,41] have shown that the binding of guanosine substrate is slow and may require conformational changes of the guanosine binding site. Comparison of the apo structure with the holo structure and other group I introns shows that the guanosine binding site is largely preformed (Fig. 4a, b, Extended Data Fig. 8). However, the apo structure shows weak density for the C262 base (Fig. 4a), in contrast to a well-resolved C262 in the holo structure (Fig. 4b). In addition, a peripheral metal ion (named M_2 below) coordinates to the C262 phosphate in both structures. M_2 and C262 shift between the apo and holo structures, providing a structural picture of proposed conformational changes upon guanosine binding (Extended Data Fig. 8d).

Metal ions in the *Tetrahymena* ribozyme

The *Tetrahymena* ribozyme is a metalloenzyme: metal ions are essential for structure stabilization and catalytic reactions[13–23]. A total of 31 metal ions (M_1 to M_{31}) are identified in the holo structure (Methods), four of which are absent in the apo structure (M_{27}, M_{29}–M_{31}; Supplementary Table 2). Most metal ions are in close proximity to oxygen atoms with *Q*-scores mostly greater than 0.8 (Extended Data Fig. 9a–c), and more than half are consistent with previous crystal structures (Extended Data Fig. 9d–g).

Some metal ions appear to be relevant for catalysis. Metal ion rescue experiments and ribozyme structures have identified five metal ions (M_A–M_E, in which M_A, M_C and M_E correspond to M_{26}, M_{27} and M_{28} in the cryo-EM structures[16,22]) as important for the enzymatic reactions of the ribozyme[16,19,22,42–47]. In the holo structure, we found that M_2, M_A and M_C coordinate with multiple non-bridging phosphate oxygens of C208, C262, A304 and A306 (Fig. 4a, b). Most importantly, M_A coordinates with the nucleophilic 3′-OH oxygen of the attacking uracil (conventionally termed u(−1)), whereas M_C coordinates with the 2′-OH oxygen of ωG to establish linear nucleophilic attack (Fig. 4c). M_E coordinates with U307 and A308. M_A and M_E overlap with those identified in the *Azoarcus* ribozyme (0.74 and 1.06 Å deviations)[17,20], whereas M_C is shifted by 2.67 Å (Extended Data Fig. 8e). This may be due to differences in sequence or chemical modifications between ribozyme studies. Analogues of M_B and M_D are not observed in any structures[17,18,20].

Discussion

The *Tetrahymena* ribozyme is a very well-characterized catalytic RNA, but its complete structure remains unresolved. Here, we have determined 3.1 Å structures of full-length apo L-21 and holo L-16 ScaI ribozymes, with most nucleotides and metal ions unambiguously resolved. Novel tertiary interactions suggest explanations for allosteric

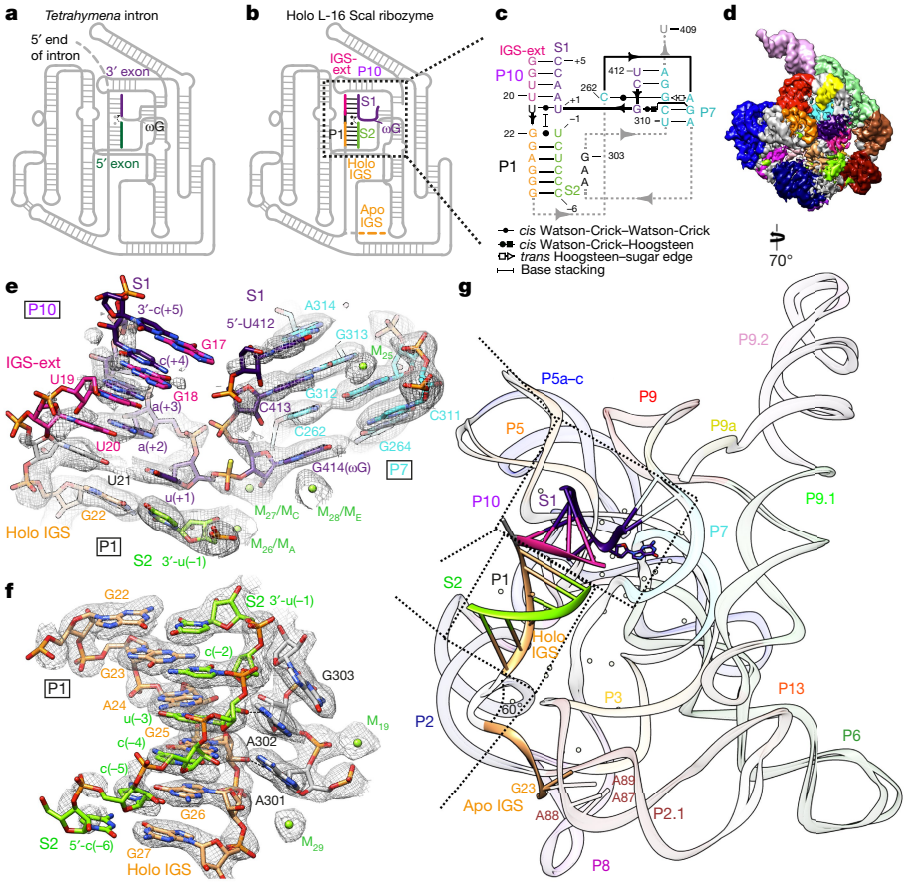


Fig. 3 | Cryo-EM structure of holo L-16 ScaI ribozyme reveals docked P1–P10 and conformational change of IGS. a, **b**, Schematics of *Tetrahymena* intron's second step of splicing (**a**) and holo L-16 ScaI ribozyme mimicking **a** (**b**). IGS-ext, IGS extension. **c**, Secondary structure of docked P1–P10 in the catalytic site. **d**, Coloured cryo-EM map. **e**, Substrate S1 (dark purple) forms P10 with IGS extension (deep pink), P7 extension with G313 and A314 (cyan), and occupies the guanosine binding site; 0.9σ threshold. **f**, Substrate S2 (chartreuse) and IGS (sandy brown) form P1; 1σ threshold. **g**, Superposition of apo and holo structures reveals conformational change of IGS from a substrate-free site close to A87, A88 and A89. Dashed arrows in **a**, **b** indicate nucleophilic attack direction of the second step of splicing; dashed lines indicate hydrogen bonds.

effects connecting the catalytic core and peripheral regions that were noted in previous studies (Fig. 2).

The apo structure shows an overall largely preorganized structure. A holo structure reveals docked P1 and P10, and four metal ions, M_2, M_A, M_C and M_E near the catalytic site. Conformational changes of IGS, C262, and M_2 occur upon substrate binding (Figs. 3g, 4a, b). These results provide possible explanations of changes in cross-linking patterns during oligonucleotide substrate binding and slow guanosine binding that were found to be important for function and specificity during *Tetrahymena* intron self-splicing[37,39,41].

Our results demonstrate the capability and potential of cryo-EM for expanding our knowledge of RNA structure–function relationships. It now appears feasible to study this ribozyme with experimental structures that illuminate all of its aspects—substrate binding, metal ions and folding pathways—and thereby to complete a foundational RNA enzymology research program that began 40 years ago.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03803-w.

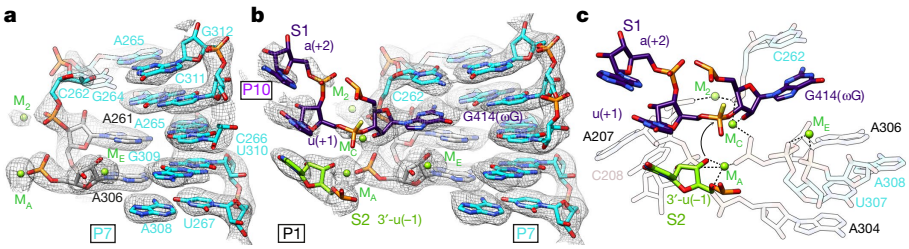


Fig. 4 | Localized shifts and mechanistic insights. a, **b**, Cryo-EM maps and models of apo L-21 ScaI ribozyme with preorganized catalytic site and weak density of C262 base (**a**), and of holo L-16 ScaI ribozyme catalytic site with docked substrate S1, which contains ωG and the 3′ splice site (**b**), both at 1σ threshold. **c**, Mechanistic insights into the second catalytic reaction. The model illustrates one stereoisomer of phosphorothioate (yellow). Black dashed lines indicate inner-sphere coordination. Black arrow indicates the direction of nucleophilic attack.

1. Nakane, T. et al. Single-particle cryo-EM at atomic resolution. *Nature* **587**, 152–156 (2020).
2. Yip, K. M., Fischer, N., Paknia, E., Chari, A. & Stark, H. Atomic-resolution protein structure determination by cryo-EM. *Nature* **587**, 157–161 (2020).
3. Zhang, K., Pintilie, G. D., Li, S., Schmid, M. F. & Chiu, W. Resolving individual atoms of protein complex by cryo-electron microscopy. *Cell Res.* **30**, 1136–1139 (2020).
4. Kruger, K. et al. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* **31**, 147–157 (1982).
5. Cech, T. R. The RNA worlds in context. *Cold Spring Harb. Perspect. Biol.* **4**, a006742 (2012).
6. Cech, T. R. & Steitz, J. A. The noncoding RNA revolution—trashing old rules to forge new ones. *Cell* **157**, 77–94 (2014).
7. Zhang, H. & Keane, S. C. Advances that facilitate the study of large RNA structure and dynamics by nuclear magnetic resonance spectroscopy. *Wiley Interdiscip. Rev. RNA* **10**, e1541 (2019).
8. Zhang, J. & Ferré-D'Amaré, A. R. New molecular engineering approaches for crystallographic studies of large RNAs. *Curr. Opin. Struct. Biol.* **26**, 9–15 (2014).
9. Zhang, K. et al. Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 Å resolution. *Nat. Commun.* **10**, 5511 (2019).
10. Kappel, K. et al. Accelerated cryo-EM-guided determination of three-dimensional RNA-only structures. *Nat. Methods* **17**, 699–707 (2020).
11. Cech, T. R. Ribozymes, the first 20 years. *Biochem. Soc. Trans.* **30**, 1162–1166 (2002).
12. Cate, J. H. et al. Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* **273**, 1678–1685 (1996).
13. Golden, B. L., Gooding, A. R., Podell, E. R. & Cech, T. R. A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme. *Science* **282**, 259–264 (1998).
14. Guo, F., Gooding, A. R. & Cech, T. R. Structure of the *Tetrahymena* ribozyme: base triple sandwich and metal ion at the active site. *Mol. Cell* **16**, 351–362 (2004).
15. Juneau, K., Podell, E., Harrington, D. J. & Cech, T. R. Structural basis of the enhanced stability of a mutant ribozyme domain and a detailed view of RNA–solvent interactions. *Structure* **9**, 221–231 (2001).
16. Piccirilli, J. A., Vyle, J. S., Caruthers, M. H. & Cech, T. R. Metal ion catalysis in the *Tetrahymena* ribozyme reaction. *Nature* **361**, 85–88 (1993).
17. Adams, P. L., Stahley, M. R., Kosek, A. B., Wang, J. & Strobel, S. A. Crystal structure of a self-splicing group I intron with both exons. *Nature* **430**, 45–50 (2004).
18. Golden, B. L., Kim, H. & Chase, E. Crystal structure of a phage Twort group I ribozyme-product complex. *Nat. Struct. Mol. Biol.* **12**, 82–89 (2005).
19. Shan, S., Yoshida, A., Sun, S., Piccirilli, J. A. & Herschlag, D. Three metal ions at the active site of the *Tetrahymena* group I ribozyme. *Proc. Natl Acad. Sci. USA* **96**, 12299–12304 (1999).
20. Stahley, M. R. & Strobel, S. A. Structural evidence for a two-metal-ion mechanism of group I intron splicing. *Science* **309**, 1587–1590 (2005).
21. Steitz, T. A. & Steitz, J. A. A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl Acad. Sci. USA* **90**, 6498–6502 (1993).
22. Forconi, M., Piccirilli, J. A. & Herschlag, D. Modulation of individual steps in group I intron catalysis by a peripheral metal ion. *RNA* **13**, 1656–1667 (2007).
23. Beaudry, A. A. & Joyce, G. F. Minimum secondary structure requirements for catalytic activity of a self-splicing group I intron. *Biochemistry* **29**, 6534–6539 (1990).
24. Benz-Moy, T. L. & Herschlag, D. Structure-function analysis from the outside in: long-range tertiary contacts in RNA exhibit distinct catalytic roles. *Biochemistry* **50**, 8733–8755 (2011).
25. Doherty, E. A. & Doudna, J. A. Ribozyme structures and mechanisms. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 457–475 (2001).
26. Joyce, G. F., van der Horst, G. & Inoue, T. Catalytic activity is retained in the *Tetrahymena* group I intron despite removal of the large extension of element P5. *Nucleic Acids Res.* **17**, 7879–7889 (1989).
27. Lehnert, V., Jaeger, L., Michel, F. & Westhof, E. New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the *Tetrahymena thermophila* ribozyme. *Chem. Biol.* **3**, 993–1009 (1996).
28. Russell, R. et al. The paradoxical behavior of a highly structured misfolded intermediate in RNA folding. *J. Mol. Biol.* **363**, 531–544 (2006).
29. Cate, J. H. et al. RNA tertiary structure mediation by adenosine platforms. *Science* **273**, 1696–1699 (1996).
30. Cate, J. H., Hanna, R. L. & Doudna, J. A. A magnesium ion core at the heart of a ribozyme domain. *Nat. Struct. Biol.* **4**, 553–558 (1997).
31. Shoffner, G. M., Wang, R., Podell, E., Cech, T. R. & Guo, F. In crystallo selection to establish new RNA crystal contacts. *Structure* **26**, 1275–1283.e3 (2018).
32. Pintilie, G. et al. Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nat. Methods* **17**, 328–334 (2020).
33. Shi, X. et al. Roles of long-range tertiary interactions in limiting dynamics of the *Tetrahymena* group I ribozyme. *J. Am. Chem. Soc.* **136**, 6643–6648 (2014).
34. Downs, W. D. & Cech, T. R. A tertiary interaction in the *Tetrahymena* intron contributes to selection of the 5′ splice site. *Genes Dev.* **8**, 1198–1211 (1994).
35. Zarrinkar, P. P. & Williamson, J. R. The P9.1-P9.2 peripheral extension helps guide folding of the *Tetrahymena* ribozyme. *Nucleic Acids Res.* **24**, 854–858 (1996).
36. Correll, C. C., Freeborn, B., Moore, P. B. & Steitz, T. A. Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. *Cell* **91**, 705–712 (1997).
37. Wang, J. F., Downs, W. D. & Cech, T. R. Movement of the guide sequence during RNA catalysis by a group I ribozyme. *Science* **260**, 504–508 (1993).
38. Lipchock, S. V. & Strobel, S. A. A relaxed active site after exon ligation by the group I intron. *Proc. Natl Acad. Sci. USA* **105**, 5699–5704 (2008).
39. Karbstein, K. & Herschlag, D. Extraordinarily slow binding of guanosine to the *Tetrahymena* group I ribozyme: implications for RNA preorganization and function. *Proc. Natl Acad. Sci. USA* **100**, 2300–2305 (2003).
40. Karbstein, K., Lee, J. & Herschlag, D. Probing the role of a secondary structure element at the 5′- and 3′-splice sites in group I intron self-splicing: the *Tetrahymena* L-16 ScaI ribozyme reveals a new role of the G.U pair in self-splicing. *Biochemistry* **46**, 4861–4875 (2007).
41. Shi, X., Mollova, E. T., Pljevaljcić, G., Millar, D. P. & Herschlag, D. Probing the dynamics of the P1 helix within the *Tetrahymena* group I intron. *J. Am. Chem. Soc.* **131**, 9571–9578 (2009).
42. Shan, S., Kravchuk, A. V., Piccirilli, J. A. & Herschlag, D. Defining the catalytic metal ion interactions in the *Tetrahymena* ribozyme reaction. *Biochemistry* **40**, 5161–5171 (2001).
43. Shan, S. O. & Herschlag, D. An unconventional origin of metal-ion rescue and inhibition in the *Tetrahymena* group I ribozyme reaction. *RNA* **6**, 795–813 (2000).
44. Sjögren, A. S., Pettersson, E., Sjöberg, B. M. & Strömberg, R. Metal ion interaction with cosubstrate in self-splicing of group I introns. *Nucleic Acids Res.* **25**, 648–653 (1997).
45. Weinstein, L. B., Jones, B. C., Cosstick, R. & Cech, T. R. A second catalytic metal ion in group I ribozyme. *Nature* **388**, 805–808 (1997).
46. Forconi, M., Lee, J., Lee, J. K., Piccirilli, J. A. & Herschlag, D. Functional identification of ligands for a catalytic metal ion in group I introns. *Biochemistry* **47**, 6883–6894 (2008).
47. Hougland, J. L., Kravchuk, A. V., Herschlag, D. & Piccirilli, J. A. Functional identification of catalytic metal ion binding sites within RNA. *PLoS Biol.* **3**, e277 (2005).

Methods

RNA preparation

L-21 ScaI ribozyme was prepared as previously described[10]. In brief, the DNA template was amplified from the pT7L-21 plasmid[48], then RNA was prepared through in vitro transcription in a reaction containing 0.2 μM DNA template, 40 mM Tris·HCl, pH 8.1, 25 mM $MgCl_2$, 3.5 mM spermidine, 0.01% TritonX-100, 40 mM DTT, 4% PEG 8000, 3 mM NTPs, and 7.5 U/μl T7 RNA polymerase (New England Biolabs). The transcription reaction was incubated at 37 °C for 1 h. The RNA was then isolated by ethanol precipitation, then purified on an 8% 29:1 acrylamide:bis, 7 M urea polyacrylamide gel. The gel was allowed to set overnight, then the precipitated RNA was mixed with loading buffer containing 95% formamide, 10 mM EDTA, 0.1% xylene cyanol, and 0.1% bromophenol blue, and loaded on the gel. The gel was run at 25 W for 2 h, then visualized briefly with a 254-nm UV lamp, held far from the gel to minimize RNA damage[49]. RNA was eluted from the gel overnight in RNase-free water at 4 °C, then purified with Zymo RNA Clean and Concentrator columns (Zymo Research).

The DNA template of L-16 ScaI ribozyme was amplified from the pT7L-21 plasmid using forward primer 5′-TTCTAATACGACTCACTATA GGTTTGGAGGGAAAAGTTATCAGGCATGCACCTGGTAGC-3′ and reverse primer 5′- ACTCCAAAACTAATCAATATACTTTCGCATACAAATTAGTT CCCAGCGGCTCC-3′. RNA was prepared using the TranscriptAid T7 High Yield Transcription Kit (Thermo Scientific) according to the manufacturer's protocol, then purified with RNA Clean and Concentrator-25 columns (Zymo Research) and PAGE purified as described above for L-21 ScaI ribozyme. RNA was eluted from the gel using the ZR small-RNA PAGE Recovery Kit (Zymo Research) and then ethanol precipitated. The RNA oligonucleotide substrates S1 (5′-UCG*UAACC) and S2 (5′-CCCUCU), in which * indicates a phosphorothioate bond, were acquired from Integrated DNA Technologies. A phosphorothioate-substituted substrate was selected to capture the complex mimicking the second step of splicing, based on previous work showing that a similar ligation reaction catalysed by *Tetrahymena* ribozyme was significantly inhibited by substitution of phosphorothioate in the R_p isomeric form at the scissile phosphate (less than 5% product after 3 h at 10 μM ribozyme concentration)[50]. Similarly, phosphorothioate substitution in the equivalent substrate for catalysis of the second step of splicing by the *Azoarcus* ribozyme reduced the reaction rate by a factor of $>10^5$ and ~14 for the R_p and S_p isomeric forms, respectively[38].

Cryo-EM sample preparation

To prepare L-21 and L-16 ScaI ribozyme samples for cryo-EM analysis, RNAs (20 μM or 15 μM final concentration, respectively) were denatured at 90 °C for 3 min in 50 mM Na-HEPES, pH 8 and cooled to room temperature for 10 min. $MgCl_2$ was added to a final concentration of 10 mM and the samples were incubated at 50 °C for 30 min. Ribozyme samples were again cooled to room temperature for 10 min. At this point, the L-21 ScaI ribozyme was kept on ice, while substrates S1 and S2 (75 μM final concentration each) were added to L-16 ScaI ribozyme and the sample was incubated at room temperature for 20 min to form the holoenzyme complex before being placed on ice. A total of 3 μl of the *Tetrahymena* ribozyme sample was applied onto glow-discharged (30 s) 200-mesh R2/1 Quantifoil Cu grids. The grids were blotted for 3 s in 100% humidity with no blotting offset and rapidly frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher).

Cryo-EM single-particle data acquisition and data processing

The frozen grids of apo L-21 ScaI ribozyme were loaded into a Titan Krios (Thermo Fisher) operated at 300 kV, condenser lens aperture 50 μm, spot size 7, parallel beam with illuminated area of 0.85 μm in diameter. Microscope magnification was at 215,000× (corresponding to a calibrated sampling of 0.65 Å per physical pixel). Movie stacks were collected automatically using EPU software on a K2 direct electron camera equipped with a Bioquantum energy filter with an energy slit of 20 eV (Gatan), operating in counting mode at a recording rate of 5 raw frames per second and a total exposure time of 5 s, yielding 25 frames per stack, and a total dose of 75 $e^-/Å^2$. A total of 7,577 movie stacks were collected with defocus values ranging between −0.3 and −1.5 μm. These movie stacks were motion-corrected using Motioncor2[51]. After CTF correction by CTFFIND4[52], 7,469 micrographs were subjected to EMAN2.2 for neural network particle picking[53]. A total of 1,658,961 particles were extracted in Relion3[54] with a box size of 320 pixels. After two rounds of 2D classifications, the best classes by visual examination were subjected to EMAN2.2 to build the initial model, and a total of 1,559,933 particles were subjected to 3D classification in Relion3. The major class showing RNA features, including 415,918 particles, was subjected to autorefinement. The initial autorefinement result was subjected to Bayesian polishing followed by another round of autorefinement[55]. A sharpening *B* factor of −14 $Å^2$ was applied to the resulting cryo-EM map to yield the final sharpened map at 3.1 Å global resolution estimated by the 0.143 criterion of the Fourier shell correlation (FSC) curve.

Frozen grids of the holo L-16 ScaI ribozyme were loaded into a Titan Krios (Thermo Fisher) operated at 300 kV, condenser lens aperture 70 μm, spot size 5, parallel beam with illuminated area of 1.1 μm in diameter. Microscope magnification was at 105,000× (corresponding to a calibrated sampling of 0.86 Å per physical pixel). Movie stacks were collected automatically using EPU software on a K3 direct electron camera equipped with a Bioquantum energy filter with an energy slit of 15 eV (Gatan), operating in counting mode and a total exposure time of 2.5 s, yielding 30 frames per stack with a total dose of 50 $e^-/Å^2$. A total of 5,559 movie stacks were collected with defocus values ranging between −0.8 and −2.0 μm. The data were processed as described above, and a total of 230,386 particles were subjected to autorefinement, Bayesian polishing and postprocessing with a sharpening *B* factor of −30 $Å^2$ to yield the final sharpened map at 3.1 Å global resolution estimated by the 0.143 criterion of the FSC curve.

Both local resolution maps were determined in Relion3 and the final map was lowpass filtered accordingly and displayed in UCSF Chimera[56].

Cryo-EM model building and refinement

The initial *Tetrahymena* ribozyme models of both apo and holo cryo-EM structures were built with DRRAFTER[57], then manually adjusted and rebuilt with Coot as needed[58]. The models were refined with Phenix. real_space_refine[59], yielding an averaged model–map correlation coefficient (CCmask) of 0.79 and 0.82, respectively. The final model was validated by MolProbity[60] and the map and model correlation was confirmed by *Q*-score analysis[32]. Secondary structure diagrams were prepared with RiboDraw aided by manual adjustment (https://github.com/ribokit/RiboDraw).

Metal ion identification and validation in cryo-EM models

Additional densities in our cryo-EM maps after fitting the RNA models were modelled as metal ions. These densities were also observed in both half maps of the reconstruction, which is another means of validating these metal ions. Several metal ions in the cryo-EM structures validate previous biochemical results: M_4 and M_5 have been previously identified to adopt inner-sphere coordination by metal-ion rescue experiments[61]; residues that make inner-sphere contacts with M_2, M_{15}, M_{16}, M_{18} and M_{19} have been suggested as metal binding sites in a phosphorothioate interference assay[62,63].

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The cryo-EM maps and associated atomic coordinate models of the apo L-21 and holo L-16 ScaI ribozymes have been deposited in the wwPDB

OneDep System under EMD accession codes EMD-31385 and EMD-31386 and PDB ID codes 7EZ0 and 7EZ2, respectively.

48. Zaug, A. J., Grosshans, C. A. & Cech, T. R. Sequence-specific endoribonuclease activity of the *Tetrahymena* ribozyme: enhanced cleavage of certain oligonucleotide substrates that form mismatched ribozyme-substrate complexes. *Biochemistry* **27**, 8924–8931 (1988).
49. Kladwang, W., Hum, J. & Das, R. Ultraviolet shadowing of RNA can cause significant chemical damage in seconds. *Sci. Rep.* **2**, 517 (2012).
50. Rajagopal, J., Doudna, J. A. & Szostak, J. W. Stereochemical course of catalysis by the *Tetrahymena* ribozyme. *Science* **244**, 692–694 (1989).
51. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
52. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
53. Chen, M. et al. Convolutional neural networks for automated annotation of cellular cryo-electron tomograms. *Nat. Methods* **14**, 983–985 (2017).
54. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
55. Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* **6**, 5–17 (2019).
56. Pettersen, E. F. et al. UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
57. Kappel, K. et al. De novo computational RNA modeling into cryo-EM maps of large ribonucleoprotein complexes. *Nat. Methods* **15**, 947–954 (2018).
58. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
59. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D* **74**, 531–544 (2018).
60. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
61. Frederiksen, J. K., Li, N. S., Das, R., Herschlag, D. & Piccirilli, J. A. Metal-ion rescue revisited: biochemical detection of site-bound metal ions important for RNA folding. *RNA* **18**, 1123–1141 (2012).
62. Strauss-Soukup, J. K. & Strobel, S. A. A chemical phylogeny of group I introns based upon interference mapping of a bacterial ribozyme. *J. Mol. Biol.* **302**, 339–358 (2000).
63. Ortoleva-Donnelly, L., Szewczak, A. A., Gutell, R. R. & Strobel, S. A. The chemical basis of adenosine conservation throughout the *Tetrahymena* ribozyme. *RNA* **4**, 498–519 (1998).
64. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
65. Green, R. & Szostak, J. W. In vitro genetic analysis of the hinge region between helical elements P5-P4-P6 and P7-P3-P8 in the sunY group I self-splicing intron. *J. Mol. Biol.* **235**, 140–155 (1994).
66. Karbstein, K., Tang, K. H. & Herschlag, D. A base triple in the *Tetrahymena* group I core affects the reaction equilibrium via a threshold effect. *RNA* **10**, 1730–1739 (2004).
67. Tanner, M. A., Anderson, E. M., Gutell, R. R. & Cech, T. R. Mutagenesis and comparative sequence analysis of a base triple joining the two domains of group I ribozymes. *RNA* **3**, 1037–1051 (1997).
68. Tanner, M. A. & Cech, T. R. Joining the two domains of a group I ribozyme to form the catalytic core. *Science* **275**, 847–849 (1997).
69. Ikawa, Y., Yoshimura, T., Hara, H., Shiraishi, H. & Inoue, T. Two conserved structural components, A-rich bulge and P4 XJ6/7 base-triples, in activating the group I ribozymes. *Genes Cells* **7**, 1205–1215 (2002).
70. Ikawa, Y., Naito, D., Shiraishi, H. & Inoue, T. Structure–function relationships of two closely related group IC3 intron ribozymes from *Azoarcus* and *Synechococcus* pre-tRNA. *Nucleic Acids Res.* **28**, 3269–3277 (2000).
71. Michel, F., Ellington, A. D., Couture, S. & Szostak, J. W. Phylogenetic and genetic evidence for base-triples in the catalytic domain of group I introns. *Nature* **347**, 578–580 (1990).
72. Szewczak, A. A. et al. An important base triple anchors the substrate helix recognition surface within the *Tetrahymena* ribozyme active site. *Proc. Natl Acad. Sci. USA* **96**, 11183–11188 (1999).
73. Rangan, P., Masquida, B., Westhof, E. & Woodson, S. A. Assembly of core helices and rapid tertiary folding of a small bacterial group I ribozyme. *Proc. Natl Acad. Sci. USA* **100**, 1574–1579 (2003).
74. Green, R., Ellington, A. D. & Szostak, J. W. In vitro genetic analysis of the *Tetrahymena* self-splicing intron. *Nature* **347**, 406–408 (1990).
75. Mitchell, D., III, Jarmoskaite, I., Seval, N., Seifert, S. & Russell, R. The long-range P3 helix of the *Tetrahymena* ribozyme is disrupted during folding between the native and misfolded conformations. *J. Mol. Biol.* **425**, 2670–2686 (2013).

Acknowledgements We thank R. N. Sengupta, D. Herschlag and I. Zheludev for discussions. Cryo-EM data were collected at SLAC and Stanford, and processed at SLAC, Stanford and Duyu High Performance Computing Center in Sichuan University. This work was supported by the National Institutes of Health (P41GM103832, R01GM079429, P01AI120943, and S10OD021600 to W.C.; R35GM112579 and R21AI145647 to R.D.); National Science Foundation (DGE-114747 to K.K. and DGE-1656518 to R.R.); Gabilan Stanford Graduate Fellowship to K.K.; Sichuan University start-up funding 20822041D4057 and Natural Science Foundation of China (NSFC82041016 and 32070049 to Z.S). We thank N. Lawless for manuscript editing.

Author contributions Z.S., K.Z., K.K., R.D., W.C. conceived the project; K.K. and M.Z.P. prepared RNA samples; Z.S. and K.Z. collected cryo-EM data; Z.S., K.Z., S.L., and Y.W. processed cryo-EM data; Z.S., K.K., S.L., B.L. and G.D.P. built and refined atomic models; G.D.P, B.L., S.L. and R.R. validated the models and map–model correlation. All authors contributed to the preparation of the manuscript.

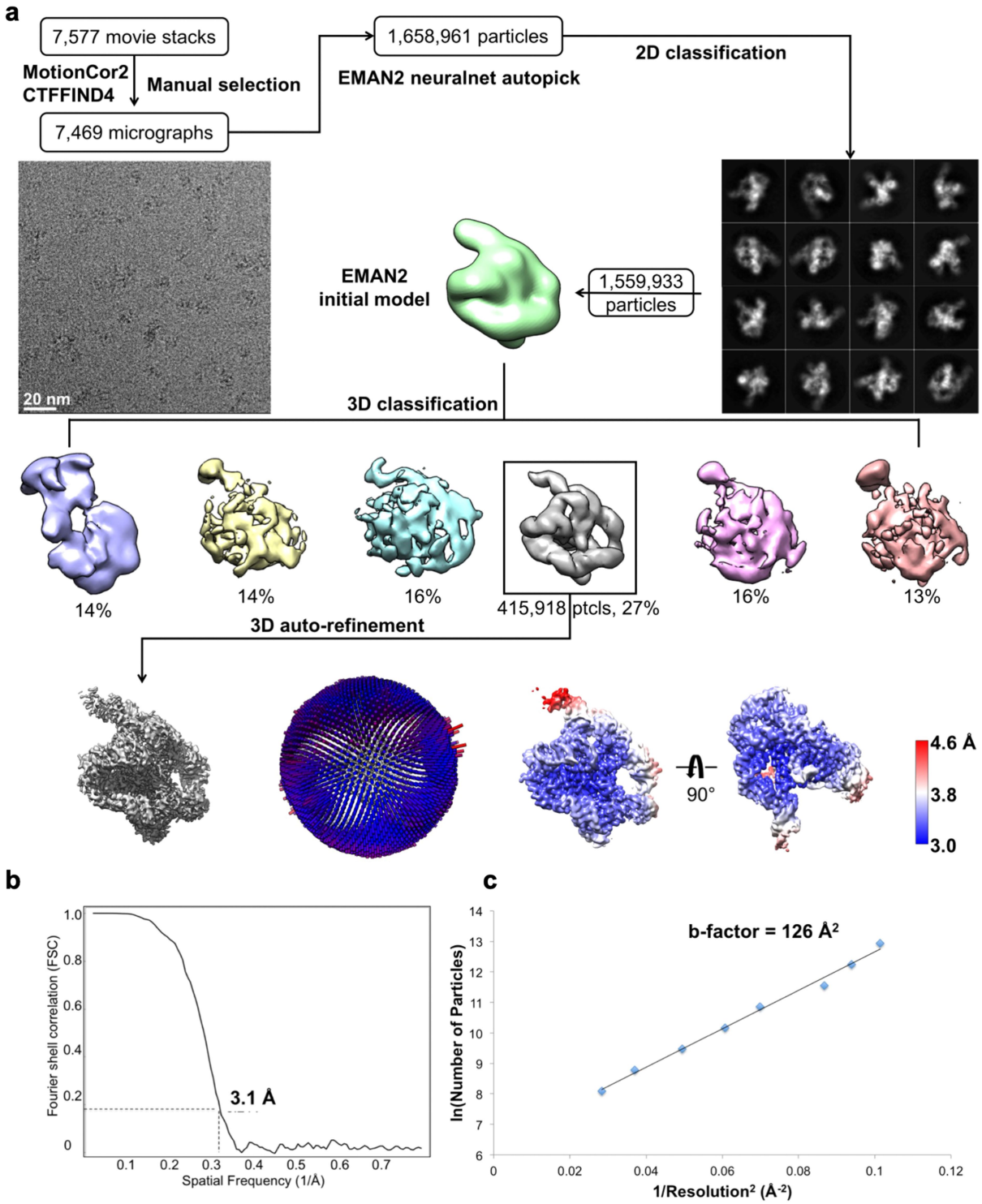
Competing interests The authors declare no competing interests.

Additional information

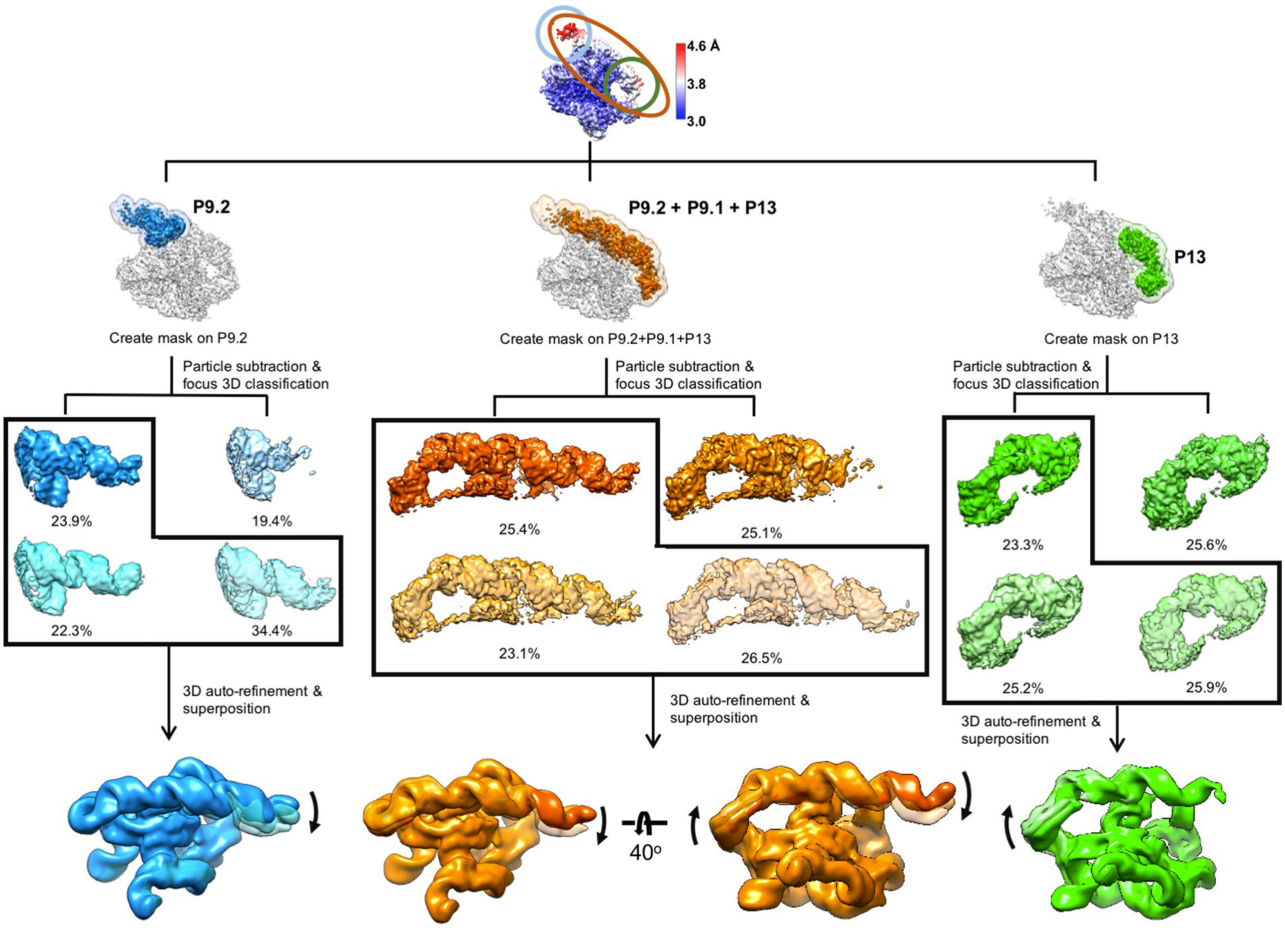
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-03803-w.

Correspondence and requests for materials should be addressed to Z.S., R.D. or W.C.

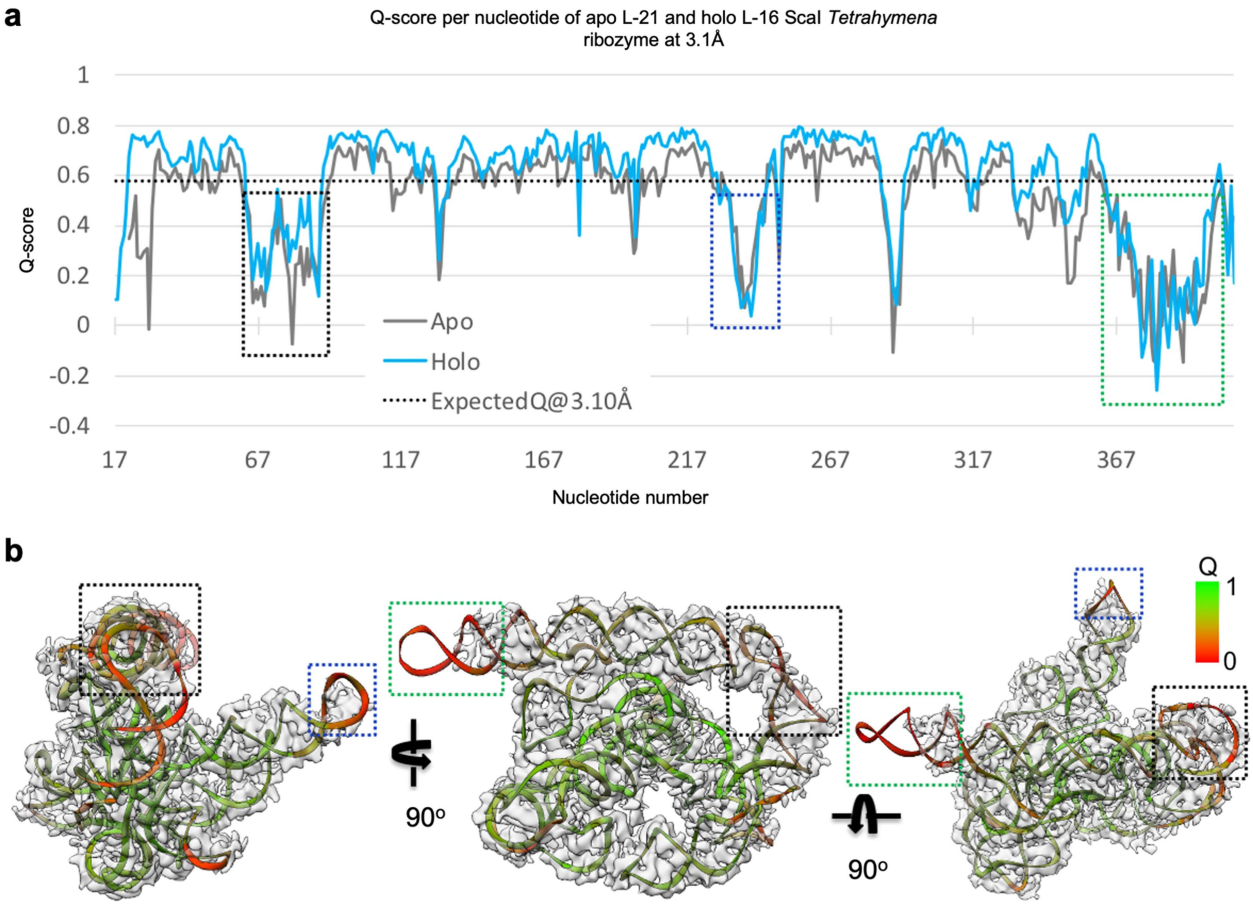
Peer review information *Nature* thanks Thomas Cech, Amedee des Georges and Andrej Lupták for their contribution to the peer review of this work. Peer reviewer reports are available.



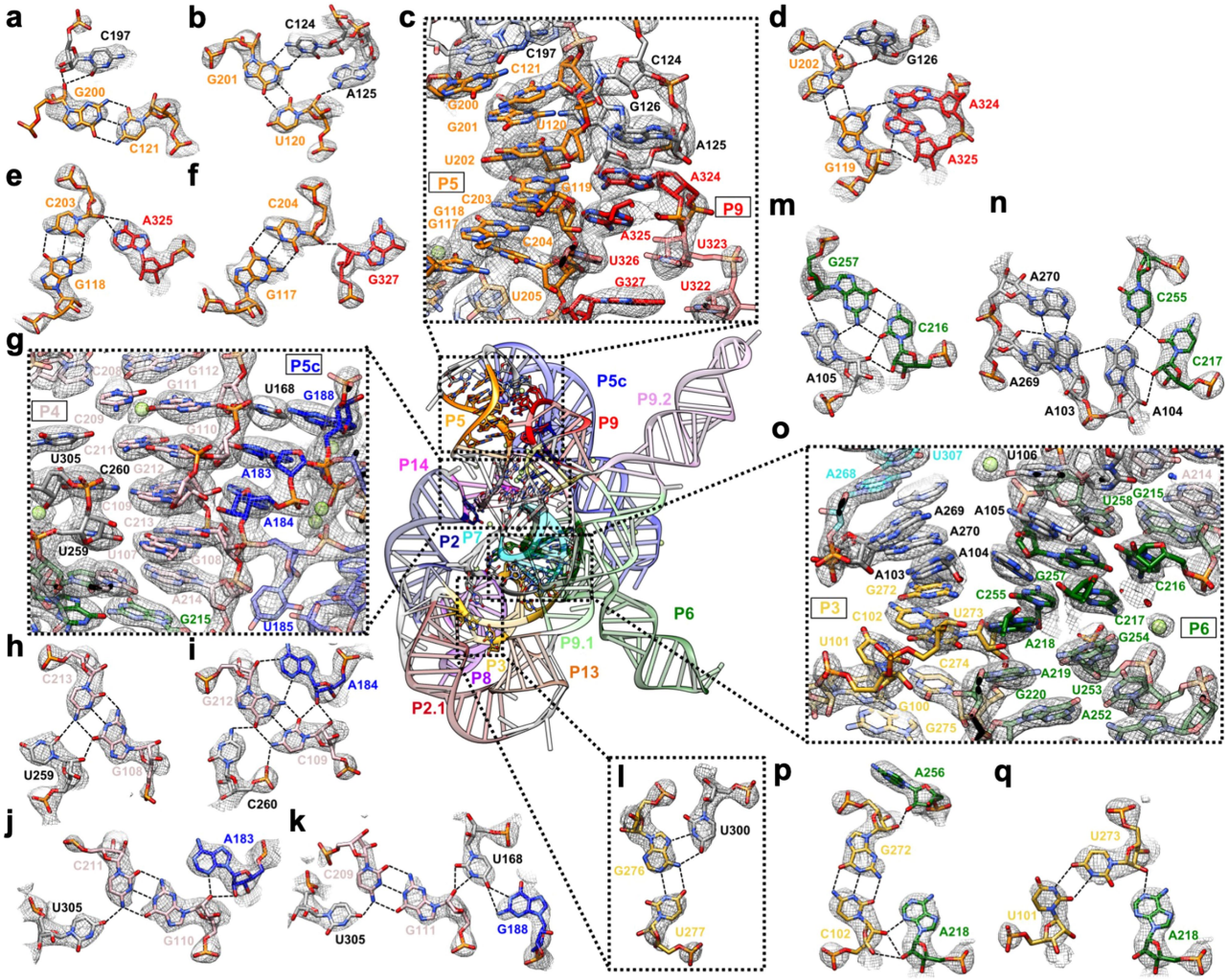
Extended Data Fig. 1 | Cryo-EM single-particle reconstruction of the apo L-21 ScaI ribozyme. Related to Fig. 1. **a**, Single-particle pipeline yields the final cryo-EM reconstruction with the corresponding angular distribution and local resolution map. The local resolution map shows more flexibility and lower resolution in the 4-nt bulge that connects stem P2.1 and P13, and towards the end of stem P6 and P9.2. **b**, FSC curve shows 3.1 Å resolution according to the 0.143 cutoff. **c**, Cryo-EM *B* factor[64], which relates the number of particles to the map resolution attributed to cumulative experimental and computational factors that affect the final reconstruction.



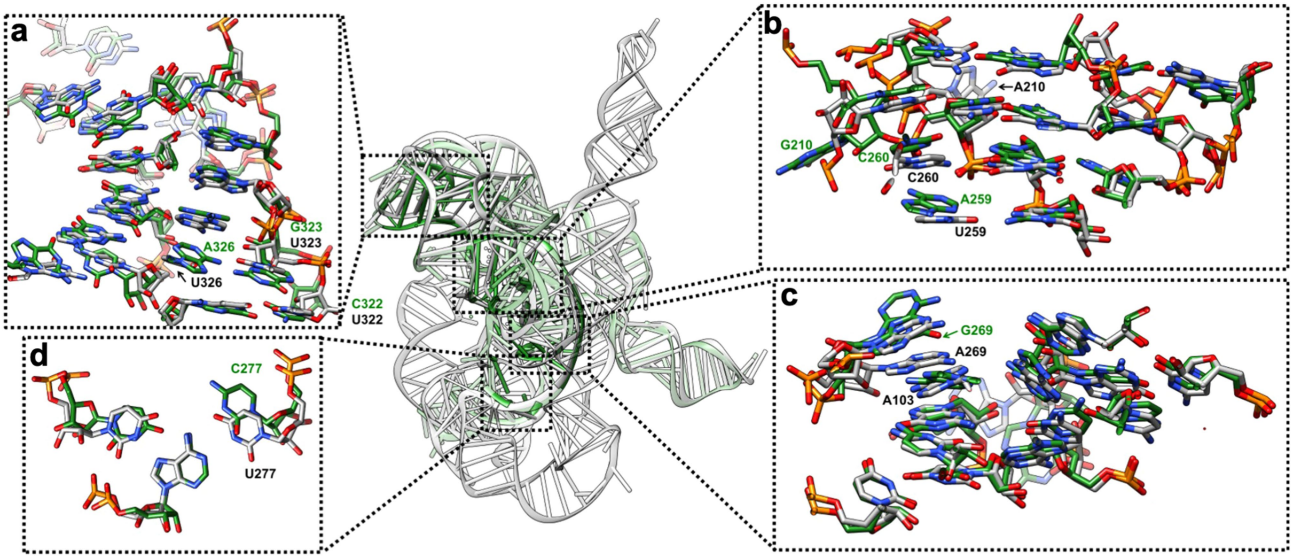
Extended Data Fig. 2 | Focused 3D classification of apo L-21 ScaI ribozyme reveals local conformational dynamics. The regions of low local resolution, P9.2 (blue), P9.2–P9.1–P13 (orange) and P13 (green), were extracted. Focused 3D classification was performed and different classes were superimposed to show rotational and translational motions on P9.2 (left), P9.2–P9.1–P13 (middle) and P13 (right).



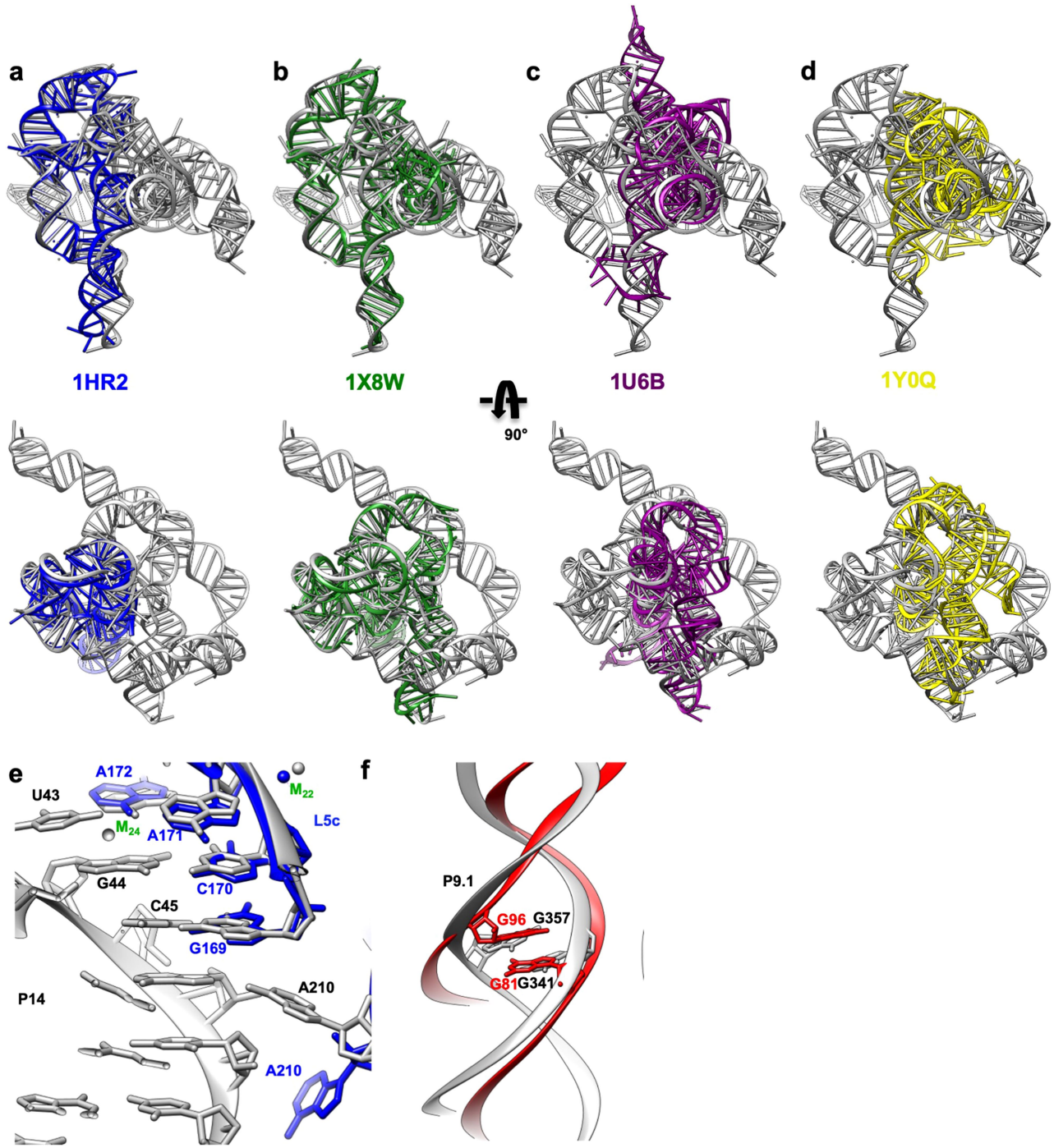
Extended Data Fig. 3 | *Q*-score analyses of cryo-EM maps and models of both the apo L-21 and holo L-16 ScaI ribozymes. Related to Fig. 1. **a**, *Q*-score analyses per residue of the apo L-21 (grey) and holo L-16 (blue) ScaI ribozyme cryo-EM models and maps. Black dashed line indicates average *Q*-score from nucleic acid cryo-EM models and maps at 3.1 Å resolution in the PDB. **b**, Cryo-EM model of apo L-21 ScaI ribozyme coloured according to *Q*-score per residue. Dashed boxes (black, blue and green) correspond to regions in the cryo-EM model with low *Q*-scores in **a**.



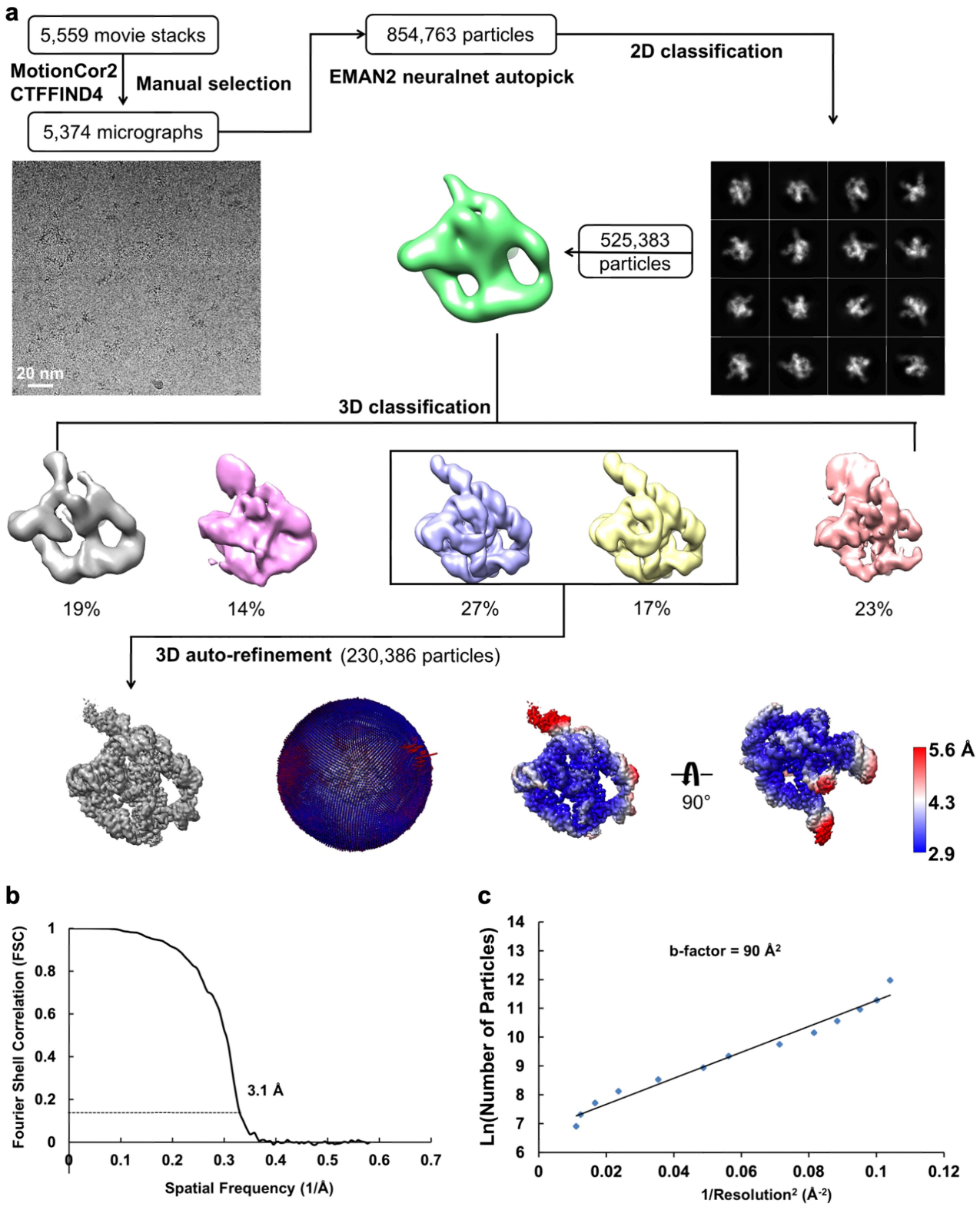
Extended Data Fig. 4 | Detailed tertiary interactions in the core region of the apo L-21 ScaI ribozyme. a–**f**, The P5–J5/5a–L9 region has a highly structured J5–5a junction in previous structures[12,15]. The cryo-EM structure shows tertiary interactions of C197 (**a**), C124 and A125 (**b**), G126, A324 and A325 (**d**, **e**) and G327 (**f**) in the minor groove of P5. **g**–**k**, Previous studies have shown that A183 and A184 in the A-rich bulge of the metal core, U259 and C260 from J6/J7 and U305 from J8/J7 are conserved and essential for catalytic site formation and splicing reactions[65–71]. The cryo-EM structure shows that U259 (**h**), C260 (**i**) and U305 (**j**, **k**) stack continuously and interact with P4 base triples. **k**, U168 from P5c stacks on the A-rich bulge and interacts with P4 in the minor groove, while pairing with G188 in P5a. **l**, The Hoogsteen base triple U277–A97–U300 is essential for substrate helix recognition[72]. **m**–**q**, In the P3–J3/4–P6–J7/3 region, the A-rich J3/4 and J7/3 were previously found to interact with P6 in the *Azoarcus* ribozyme[17,73]. Base triples formed by J3/4 were critical for catalysis[65,70,71,74], and alterations in these regions result in RNA misfolding[28,75]. In the cryo-EM structure, A104 and A105 form A-minor interactions with P6, whereas A103 and A104 join A269 and A270 from J7/3 to form an adenosine cluster (**m**, **n**). The proposed A103–U271 reverse-Hoogsteen pair is not found; instead we observed a noncanonical A103–A270 pair[74]. **p**, **q**, The previously observed A-platform of A218–A219 is disrupted in the cryo-EM structure with P3 present[12,29]. A218 forms two A-minor interactions with C102–G272 and U273–U101 from P3, which also supports the conservation of this C–G pair in group Ib introns[74]. Black dashed lines indicate hydrogen bonds. The cryo-EM maps of all subpanels are visualized at 1σ threshold except for **c**, **g**, **o** (1.5σ).



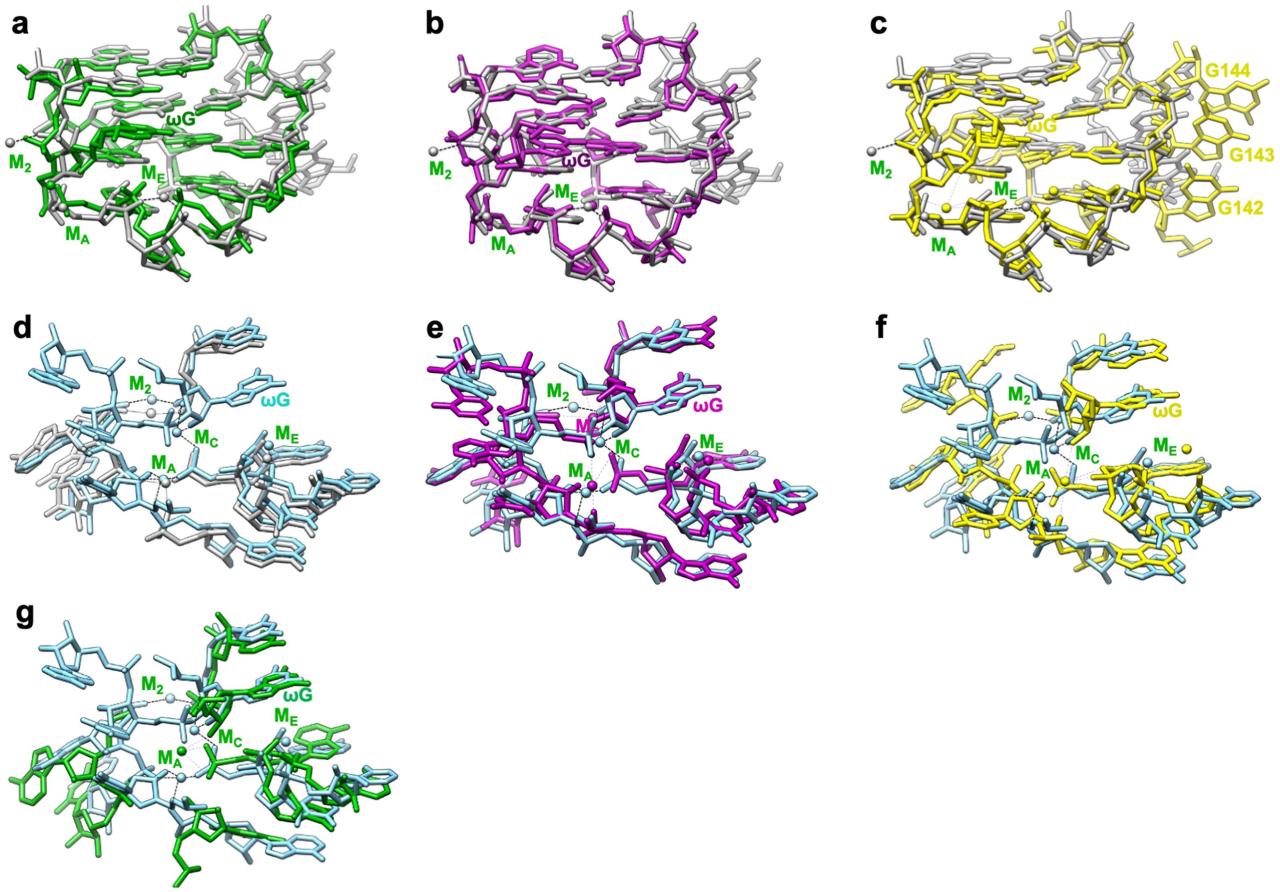
Extended Data Fig. 5 | Comparison between the previous 3.8 Å crystal structure of the mutated *Tetrahymena* ribozyme catalytic core (green) and the cryo-EM structure of the wild-type apo L-21 ScaI ribozyme construct (grey) shows minor differences. Related to Extended Data Fig. 4. The overall r.m.s.d. for the catalytic core region (stem P3–P9) is 6.6 Å. **a**, The same view of the P5–J5/5a–P9 region as in Extended Data Fig. 4c. The nucleotide conformations generally agree well between two models; three mutations (U322C, U323G and U326A) are not involved in tertiary interactions. The r.m.s.d. for this region is 4.9 Å. **b**, The same view of the P4–P5a–J6/7–J8/7 region as in Extended Data Fig. 4g. In the crystal structure, U259A is slightly moved away from the G108–C213 base pair and disrupts this base triple interaction. A210G is moved far away from the wild-type position of A210, because there is no A46 in stem P2 to interact with in the crystal structure. The very top base quartet is much more compact in the cryo-EM structure compared to the crystal structure, probably owing to the presence of the peripheral domain that wraps around the catalytic core to make it more compact. The r.m.s.d. for this region is 5.7 Å. **c**, The same view of the P3–J3/4–P6–J7/3 region as in Extended Data Fig. 4o. The overall nucleotide conformations agree very well between the two models, except that A269G and A270 in the crystal structure are completely moved away and disrupt their interactions with A103, which is observed in the cryo-EM structure. The r.m.s.d. for this region is 1.7 Å. **d**, The same view as in Extended Data Fig. 4l. U277C disrupts the U277–A97–U300 base triple. The r.m.s.d. for this base triple is 2.7 Å. See also Supplementary Table 1.



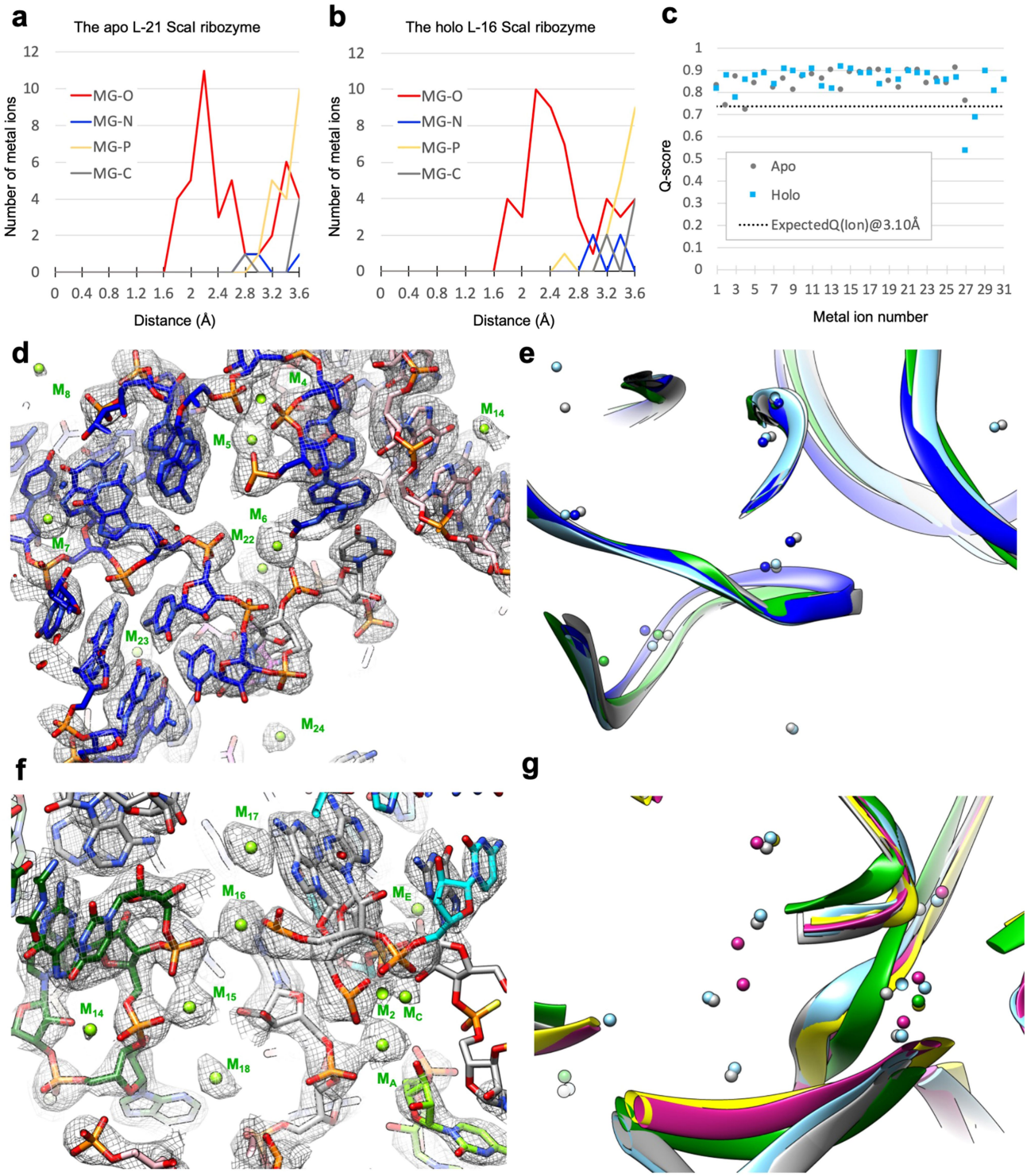
Extended Data Fig. 6 | Superposition of the apo L-21 ScaI ribozyme cryo-EM structure (grey) with previous crystal structures of the truncated and/or mutated *Tetrahymena* ribozyme, other group I introns and 5S rRNA loop E show global and local structural similarities. a–**d**, Overlays of the cryo-EM structure (grey) with the *Tetrahymena* ribozyme P4–P6 Delta C209 (**a**; blue, PDB 1HR2); the mutated *Tetrahymena* ribozyme P3–P9 (**b**; green, PDB 1X8W); the *Azoarcus* ribozyme (**c**; violet, PDB 1U6B) and the phage Twort ribozyme (**d**; yellow, PDB 1Y0Q). **e**, P5c region of the wild type P4–P6 crystal structure (blue, PDB 1GID). **f**, 5S rRNA loop E crystal structure (red, PDB 354D).



Extended Data Fig. 7 | Cryo-EM single-particle reconstruction of the holo L-16 ScaI ribozyme. Related to Fig. 3. **a**, Single-particle pipeline yields the final cryo-EM reconstruction with the corresponding angular distribution and local resolution map. The local resolution map shows more flexibility and lower resolution in the 4-nt bulge that connects stem P2.1 and P13, and towards the end of stem P6 and P9.2. **b**, FSC curve shows 3.1 Å resolution according to the 0.143 cutoff. **c**, Cryo-EM *B* factor.



Extended Data Fig. 8 | Comparisons of apo L-21 and holo L-16 ScaI ribozyme cryo-EM models with previous crystal structures show structural conservation and metal ion shifts in the guanosine binding site among group I introns. Related to Fig. 4. **a**–**c**, The apo L-21 ScaI ribozyme adopts a preorganized guanosine binding site (grey) that superimposes with previous crystal structures of mutated P3–P9 of the *Tetrahymena* ribozyme (**a**; green, PDB 1X8W), the *Azoarcus* ribozyme (**b**; violet, PDB 1U6B), and the phage Twort ribozyme (**c**; yellow, PDB 1y0q). **d**–**g**, The holo L-16 ScaI ribozyme (sky blue) superimposes with apo L-21 ScaI ribozyme (**d**; grey), the *Azoarcus* ribozyme (**e**; violet), the phage Twort ribozyme (**f**; yellow), and mutated P3–P9 of the *Tetrahymena* ribozyme (**g**; green). M_C is absent in the apo L-21 ScaI ribozyme, whereas M_C in the *Azoarcus* ribozyme is shifted compared to the holo L-16 ScaI ribozyme. Dash line indicates metal ion coordination with surrounding atoms.



Extended Data Fig. 9 | Metal ion validations by distance and *Q*-score analysis, and illustrations of the apo L-21 and holo L-16 ScaI ribozyme cryo-EM structures compared with previous crystal structures. **a**, Distances between metal ions and other atoms in the apo L-21 ScaI ribozyme model. **b**, Distances between metal ions and other atoms in the holo L-16 ScaI ribozyme model. **c**, *Q*-score analysis per metal ion of the apo L-21 and holo L-16 ScaI ribozyme cryo-EM models and maps. **d**, Metal core region of the holo L-16 ScaI ribozyme, visualized at 1.1σ threshold. **e**, Comparisons of the apo L-21 (grey) and holo L-16 (sky blue) ScaI ribozyme cryo-EM models with P4–P6 Delta C209 (blue, PDB 1HR2) and mutated P3–P9 of the *Tetrahymena* ribozyme (green, PDB 1X8W) in the same view as **d**. **f**, Catalytic site of the holo L-16 ScaI ribozyme, visualized at 1.4σ threshold. **g**, Comparisons of the apo L-21 (grey) and holo L-16 (sky blue) ScaI ribozyme cryo-EM models with the *Azoarcus* ribozyme (violet, PDB 1U6B), mutated P3–P9 of the *Tetrahymena* ribozyme (green, PDB 1X8W), and the phage Twort ribozyme (yellow, PDB 1Y0Q) in the same view as **f**. See also Supplementary Table 2.

Extended Data Table 1 | Cryo-EM data collection, processing, and model refinement statistics of the apo L-21 and holo L-16 ScaI ribozymes

Cryo-EM data collection and processing	The apo L-21 ScaI ribozyme	The holo L-16 ScaI ribozyme
Microscope	Titan Krios	Titan Krios
Voltage (kV)	300	300
GIF Quantum energy filter slit width (eV)	20	15
Detector	Gatan K2	Gatan K3
Nominal Magnification	215,000 ×	105,000 ×
Pixel size (Å)	0.65	0.86
Symmetry imposed	C1	C1
Defocus range (μm)	-0.3 – -1.5	-0.8 – -2.0
Electron exposure ($e^-/Å^2$)	75	50
Micrographs (acquired/used)	7,577/7,469	5,559/5,374
Number of extracted particles	1,658,961	854,763
Number of particles after 2D classifications	1,559,933	525,373
Number of particles going to 3D refinement	415,918	230,386
Map resolution at 0.143 FSC criterion (Å)	3.1	3.06
Local resolution range (Å)	3.0-4.7	2.9-5.6
Sharpening B-factor ($Å^2$)	-14	-30
Model refinement		
Atoms	8,306	8,698
Residues	387	406
CC_{mask}	0.79	0.82
$Resolution_{FSC\ map\ vs.\ model\ @\ 0.5}$ (Å)	3.2	3.0
r.m.s. deviations		
Bond lengths (Å)	0.003	0.002
Bond angles (°)	0.547	0.523
Clash score	2.65	1.91
MolProbity score	2.20	2.10

nature research

Corresponding author(s): Zhaoming Su, Rhiju Das and Wah Chiu

Last updated by author(s): Jun 2, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☒	☐ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	EPU2.5 (Thermofisher)
Data analysis	All software are publicly accessible. Data processing: MotionCor2 1.1.0, CtfFind4.1.13, EMAN2.2, Relion3.0. Visualization: Chimera 1.14, Biomovie 0.9.2. Model refinement and validation: DRRAFTER from Rosetta version 3.10, Coot0.8.7, Phenix1.15.2-3472, Molprobity4.5.1, Q-score v1.6.4, mapQ 1.6.4. Secondary structure: Ribodraw

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The cryo-EM maps and associated atomic coordinate models of the apo L-21 and holo L-16 ScaI Tetrahymena ribozyme have been deposited in the wwPDB OneDep System under EMD accession code 31385, 31386 and PDB ID code 7EZ0, 7EZ2.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	3D reconstructions were calculated from 7,577 images (415,918 particles) for the apo L-21 ScaI Tetrahymena ribozyme, 5,559 images (230,386 particles) for the holo L-16 ScaI Tetrahymena ribozyme.
Data exclusions	For cryo-EM analysis, particles that do not belong to the class of interest or have poor qualities based on well established cryo-EM principle were excluded after rounds of 2D and 3D classification.
Replication	Cryo-EM 3D reconstruction methods are approved to reconstruct high resolution structures. Our results can be reproduced if the procedure (see Methods section for details) was exactly followed.
Randomization	No grouping required for our studies.
Blinding	Blinding was not necessary for structure determination by single-particle cryo-EM method.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging