





Crystal structures and insights into precursor tRNA 5'-end processing by prokaryotic minimal protein-only RNase P

Yangyang Li¹, Shichen Su², Yanqing Gao ¹, Guoliang Lu², Hehua Liu¹, Xi Chen¹, Zhiwei Shao¹, Yixi Zhang¹, Qiyuan Shao¹, Xin Zhao¹, Jie Yang¹, Chulei Cao¹, Jinzhong Lin ², Jinbiao Ma ² & Jianhua Gan ¹✉

Besides the canonical RNA-based RNase P, pre-tRNA 5'-end processing can also be catalyzed by protein-only RNase P (PRORP). To date, various PRORPs have been discovered, but the basis underlying substrate binding and cleavage by HARPs (homolog of Aquifex RNase P) remains elusive. Here, we report structural and biochemical studies of HARPs. Comparison of the apo- and pre-tRNA-complexed structures showed that HARP is able to undergo large conformational changes that facilitate pre-tRNA binding and catalytic site formation. *Planctomycetes bacterium* HARP exists as dimer in vitro, but gel filtration and electron microscopy analysis confirmed that HARPs from *Thermococcus celer*, *Thermocrinis minervae* and *Thermocrinis ruber* can assemble into larger oligomers. Structural analysis, mutagenesis and in vitro biochemical studies all supported one cooperative pre-tRNA processing mode, in which one HARP dimer binds pre-tRNA at the elbow region whereas 5'-end removal is catalyzed by the partner dimer. Our studies significantly advance our understanding on pre-tRNA processing by PRORPs.

¹Shanghai Public Health Clinical Center, State Key Laboratory of Genetic Engineering, Collaborative Innovation Center of Genetics and Development, Department of Biochemistry and Biophysics, School of Life Sciences, Fudan University, Shanghai 200438, China. ²State Key Laboratory of Genetic Engineering, Collaborative Innovation Center of Genetics and Development, Department of Biochemistry, School of Life Sciences, Fudan University, Shanghai 200438, China. ✉email: ganjhh@fudan.edu.cn

Transfer RNA (tRNA) decodes mRNA and transfers the genetic information from mRNA to protein. The precursor of tRNA (pre-tRNA) undergoes a series of post-transcriptional processes, including nucleobase modification^{1–4}, splicing^{5–7}, 5'-end cleavage^{8–10} and 3'-end cleavage and conserved CCA addition^{11–14}. Correct processing is essential for the function of tRNA. In addition to pre-tRNA coding genes, mutations in many pre-tRNA processing enzymes have also been linked with serious diseases^{15–21}. RNase P catalyzes the 5'-end cleavage of pre-tRNA. The canonical RNase P was first discovered in Bacteria^{22,23}, it was later found in Archaea and Eukarya²⁴. The canonical RNase P occurs as ribonucleoprotein (RNP) composed of one RNA and various number of proteins^{24,25}, but pre-tRNA 5'-end is all cleaved by the RNA molecule.

Interestingly, recent studies showed that pre-tRNA 5'-end can also be cleaved by protein-only RNase Ps (PRORP). To date, PRORP has been found in the mitochondria of human, *Caenorhabditis elegans* (*C. elegans*), *Drosophila* and many other animals^{26,27}. These mitochondrial PRORPs are composed of three proteins, MRPP1–3. Pre-tRNA 5'-end cleavage is catalyzed by MRPP3 (Fig. 1A), but it strictly requires the assistance of MRPP1 and MRPP2. PRORP is also present in plants and protists²⁸, such as *Arabidopsis thaliana* (*A. thaliana*) and *Trypanosoma brucei* (*T. brucei*)²⁹. Although AtPRORPs and TbPRORPs are homologous to human MRPP3 (Fig. 1A), they are active on their own.

Bacteria and Archaea do not express eukaryotic-like PRORP proteins, but in vitro pre-tRNA 5'-end cleavage activity has been confirmed for protein Aq_880 (Fig. 1A) encoded by the hyperthermophilic bacterium *Aquifex aeolicus* (*A. aeolicus*)³⁰. Besides *A. aeolicus*, bioinformatic analysis showed that homologs of Aquifex RNase P (termed HARP hereafter) are also present in some other Bacteria and are even more wide-spread in Archaea^{30,31}. Like the eukaryotic PRORPs, HARPs are members of the PIN domain superfamily, but belong to a different subgroup (PIN_5 cluster)³². Eukaryotic PRORP contains three domains: the pentatricopeptide repeat (PPR) domain, the central domain and the catalytic metallonuclease domain. Compared to eukaryotic PRORP, HARP is much shorter and it lacks both PPR and the central domains. The catalytic metallonuclease domain is conserved, but it only shares very limited sequence similarity (<25%) between HARP and eukaryotic PRORP.

In addition to Aq_880, the in vitro pre-tRNA 5'-end cleavage activity has also been confirmed for HARPs from *Haloferax volcanii* (*H. volcanii*) and *Matheanosarcina mazei* (*M. mazei*), which are members of the euryarchaeotes³³. Aq_880 shares conserved residues with HvHARP, MmHARP and HARPs from many other species (Supplementary Fig. 1A). Knockout of the HARP gene showed no growth defects in *H. volcanii* and *M. mazei*, indicating that the RNA-based RNase P is the main and essential RNase P in these archaea³³. Different from euryarchaeotes, HARP is the only RNase P expressed in *A. aeolicus* and many related *Aquificaceae*, such as *Hydrogenobacter thermophilus* and *Thermocrinis albus* DSM. Aq_880 is able to replace the RNase P complex in budding yeast *Saccharomyces cerevisiae* and rescue the growth of *E. coli* strain BW, in which the chromosomal expression of RNase P RNA is switched off by replacing arabinose with glucose in the medium³⁰. To improve our understanding on pre-tRNA processing by HARP, we performed extensive structural and biochemical studies. Here, we report the crystal structures of *Planctomycetes bacterium* HARP (*PbHARP*) and *Thermococcus celer* HARP (*TcHARP*). Comparison of the apo- and pre-tRNA-complexed *PbHARP* structures showed that HARP is able to undergo large conformational changes that facilitate pre-tRNA binding and catalytic site formation. Electron microscopy analysis of *Thermocrinis ruber* HARP (*TrHARP*) showed that the dimer of HARP can assemble into oligomers.

Structural analysis, mutagenesis, in vitro cleavage and cross-linking assays all suggested a HARP pre-tRNA binding and cleavage mode involving the cooperation of two dimers.

Results

Pre-tRNA 5'-end cleavage activity is conserved in HARPs. To investigate whether the pre-tRNA 5'-end cleavage activity is conserved in HARPs, we constructed and purified four HARPs, *PbHARP*, *TcHARP*, *TmHARP* and *TrHARP*, which are coded by *Planctomycetes bacterium* GWF2_40_8 (*P. bacterium*), *Thermococcus celer* (*T. celer*), *Thermocrinis minervae* (*T. minervae*) and *Thermocrinis ruber* (*T. ruber*), respectively. *T. celer* is a hyperthermophilic archaeon, *T. minervae* and *T. ruber* are all members of *Aquificaceae*. *PbHARP*, *TcHARP*, *TmHARP* and *TrHARP* share 60–80% sequence similarities with Aq_880, and the sequence identity is about 30% among the five proteins (Supplementary Fig. 1B).

All proteins were purified to homogeneity. During purification, we noticed that the oligomerization state of HARP varies (Supplementary Fig. 2). *PbHARP* exists as dimer in buffer containing 300–500 mM salt; at lower concentration of salt, it precipitates out. The oligomerization state of *TcHARP* is changeable. At higher concentration of salt (200–500 mM), it exists as dimer; however, it assembles into larger oligomers in buffer containing 100 mM salt. Using 1.0 μ M *Thermus thermophilus* pre-tRNA^{Gly} (Supplementary Fig. 3A) purified in the laboratory, we performed in vitro cleavage assay (Fig. 1B) for *PbHARP*, *TcHARP*, *TmHARP*, *TrHARP* and Aq_880, which served as a positive control. The HARP concentrations are all fixed at 0.05 μ M. Like Aq_880, other four HARPs can also cleave the 5'-end off from the pre-tRNA substrate. At a reaction time of 20 min, about 74%, 79% and 48% pre-tRNA^{Gly} was cleaved by *TcHARP*, *TmHARP* and *TrHARP*, respectively. Although the cleavage activity of *PbHARP* is relatively weak, it can cleaves about 6% pre-tRNA^{Gly} at a reaction time of 20 min.

Overall folding and dimerization of apo- *PbHARP* and *TcHARP*.

Puzzled by their different oligomerization states and catalytic efficiencies, we performed crystallographic studies for all the HARPs. No crystal grew for Aq_880, *TmHARP* or *TrHARP*, but we successfully solved the apo-form structures of *PbHARP* and *TcHARP* at atomic resolution (Supplementary Table 1). *PbHARP* crystal belongs to P3₁ space group, per asymmetric unit contains two *PbHARP* dimers. *TcHARP* crystal belongs to P2₁2₁2₁ space group and there is only one *TcHARP* dimer in the asymmetric unit. Both *PbHARP* and *TcHARP* are of α/β fold in nature (Fig. 1C–F, Supplementary Fig. 4); they share six β -strands (β 1– β 6) and ten α -helices (α 1– α 10). Compared to *TcHARP*, *PbHARP* is 14 amino acid longer at its N-terminus, which forms one short extra helix, α 0. β 1– β 3 and β 5– β 6 are parallel to each other, forming one flat β -sheet at the center of the metallonuclease domain. Helices α 1– α 6 and α 10 reside on one side of the β -sheet, whereas α 9 and the C-terminus of α 8 locate on the opposite site.

β 4 and α 7 are not involved in the metallonuclease domain formation. β 4 is linked to the metallonuclease domain through one short linker that is perpendicular to the C-terminus of α 8 (Fig. 1C). α 7 is separated from α 8 by a short linker. Indicated by the weak electron density and different conformations (Supplementary Fig. 4C), the α 7– α 8 connecting linker is flexible. Both α 7 (amino acids 97–122) and α 8 (amino acids 131–165) are very long, forming a four-helix bundle (HB) with α 7 and α 8 from the partner molecule. Due to the presence of Lys or Arg residues, the surface of HB is highly positive in charge (Fig. 1D). The

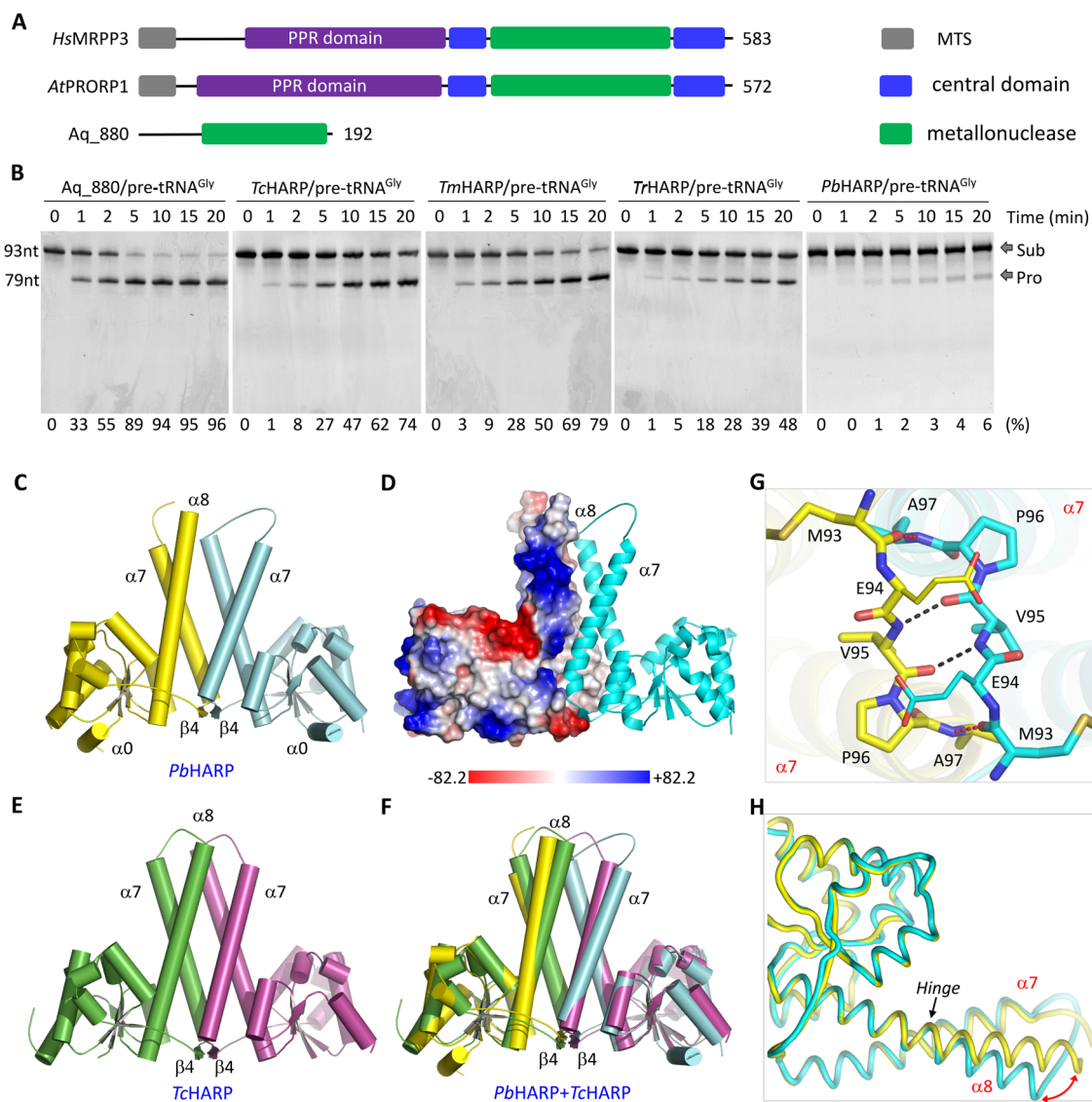


Fig. 1 pre-tRNA 5'-end processing and folding of HARPs. **A** Domain architectures of PRORPs. **B** In vitro pre-tRNA^{Gly} cleavage assays catalyzed by different HARPs. The substrate cleavage percentage (%) is shown at the bottom of the gels. Experiments were repeated independently twice with similar results. Source data are provided in Source Data file. **C, D** Cartoon and surface presentation showing the folding of PbHARP AB dimer observed in the apo-form structure. **E** Cartoon presentation of the apo-form TcHARP structure. **F** Structural superposition of PbHARP and TcHARP. The PbHARP AB dimer is colored as in panel (C), the TcHARP dimer is colored as in panel (E). **G** H-bond interactions involved in PbHARP dimerization in the apo-form structure. **H** Structural superposition showing the conformational difference between the HB domains of PbHARP A and B monomers. In panels (G), (H), the C-atoms are colored in yellow and cyan in the A and B monomers, respectively.

metallonuclease domain is also positive on the surface distal from HB, whereas is highly negative at the region facing HB.

$\beta 4$ plays an important role in PbHARP dimerization. As depicted in Fig. 1G, the $\beta 4$ strands of the two monomers form four hydrogen (H) binding interactions: two between the backbone carbonyl and amino functions of the two V95 residues and the other two mediated by M93 and A97 residues. These H-bond interactions are very stable, supported by their short distances (2.7–2.85 Å). The dimerization manner of TcHARP is similar to that of PbHARP. However, structural analysis revealed some conformational differences between the PbHARP and TcHARP dimers, especially in the HB regions (Fig. 1F). Further analysis showed that the overall folds of the two TcHARP monomers are similar to that of PbHARP monomer A (or C) (Supplementary Fig. 4D), the conformational differences are mainly caused by the bending of $\alpha 8$ of PbHARP monomer B (or D), hinged at residue E147 (Fig. 1H). The root-mean-square deviation (RMSD)

value between TcHARP and PbHARP A or C monomers is 1.0 Å, whereas is 1.4 Å between TcHARP and PbHARP B or D monomers.

Pre-tRNA recognition by PbHARP. Besides apo- proteins, we also performed co-crystallization trials using pre-tRNA^{Gly}, but no crystal was obtained. We then transcribed and purified another three pre-tRNAs, including *E. coli* pre-tRNA^{His}, mitochondrial pre-tRNA^{Cys} and chloroplastic pre-tRNA^{Phe} (Supplementary Fig. 3A). After extensive trials, we solved the complex structure of PbHARP/pre-tRNA^{His} at 2.8-Å resolution (Supplementary Table 1). The crystal belongs to P4₁2₁2 space group and it contains two PbHARP dimers (AB and CD) and one pre-tRNA^{His} molecule in the asymmetric unit (Fig. 2A). As indicated by the low RMSD value (0.6 Å), the overall folds of PbHARP AB and CD dimers are very similar (Supplementary Fig. 5A). Out of the 85 nucleotides of pre-tRNA^{His}, G-10 to U-2, U32 to A36 and C73/

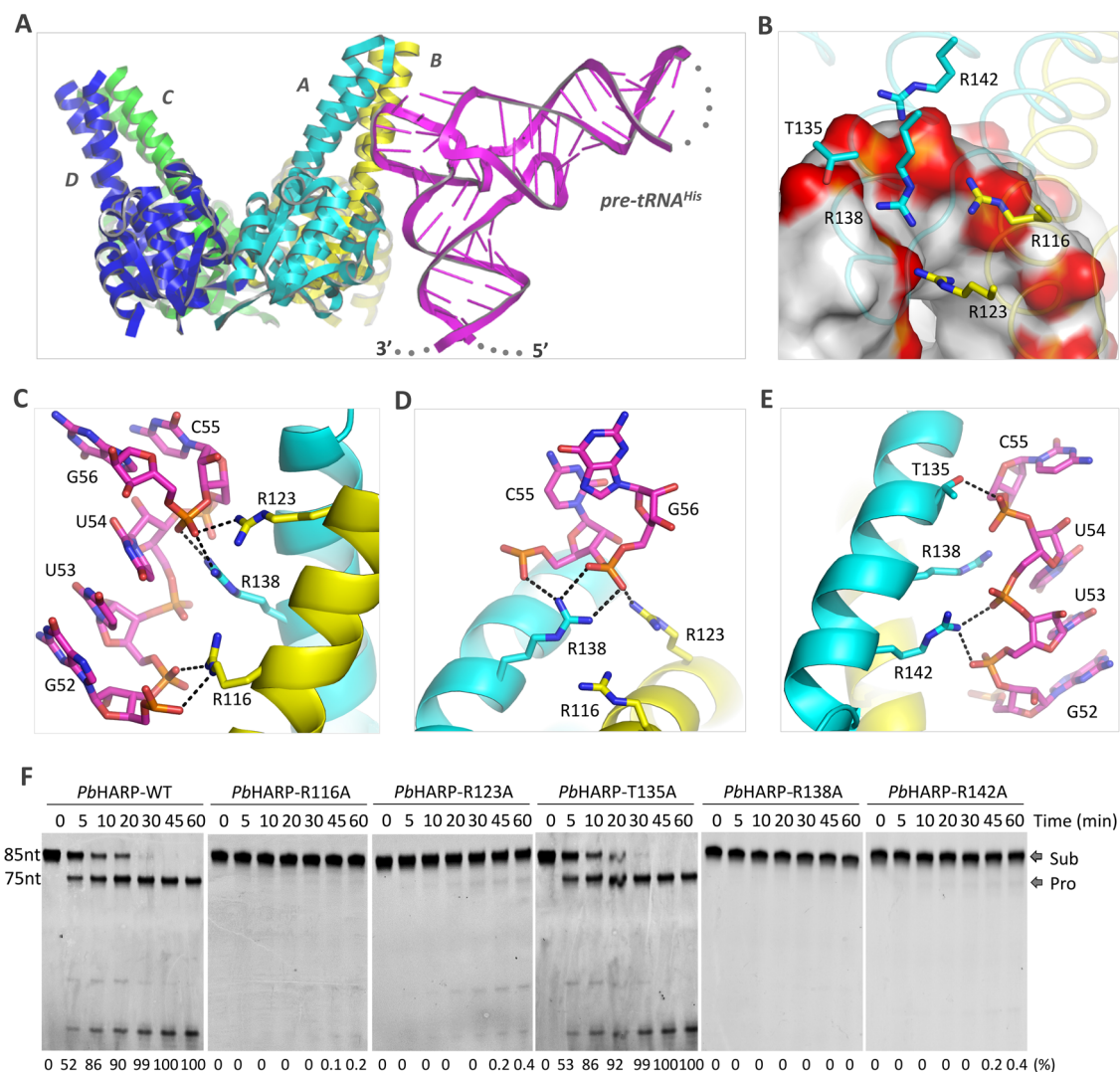


Fig. 2 The complex structure of *PbHARP/pre-tRNA^{His}*. **A** Cartoon presentation showing the overall fold and assembly of *PbHARP/pre-tRNA^{His}* complex. **B** The relative orientation between *pre-tRNA^{His}* and the conserved Arg residues. *Pre-tRNA^{His}* is shown as surface, the phosphate backbones are colored in red. **C–E** The detailed interactions between *pre-tRNA^{His}* TψC-loop and *PbHARP* AB dimer. *Pre-tRNA^{His}* and all the interacting Arg and Thr residues are shown as sticks. The C-atoms of *pre-tRNA^{His}*, the A and B monomers of *PbHARP* are colored in magenta, yellow, and cyan, respectively. **F** In vitro *pre-tRNA^{His}* cleavage assays catalyzed by wild-type and mutated *PbHARP*. The protein and substrate concentrations are all fixed at 1 μM. The substrate cleavage percentage (%) is shown at the bottom of the gels. Experiments were repeated independently three times with similar results. Source data are provided in Source Data file.

A74 located at the 5'-end, anticodon loop and the 3'-end are disordered, whereas other nucleotides are all well-defined (Supplementary Fig. 5B).

Pre-tRNA^{His} is recognized by the AB dimer of *PbHARP*. Although the overall folds of the metallonuclease domains and the HB domains are conserved, their relative orientations are different within the two monomers (Supplementary Fig. 5C). The asymmetric dimerization orients the side chains of four HB domain Arg residues, including R116 and R123 of monomer B and R138 and R142 of monomer A, toward the *pre-tRNA^{His}* TψC-loop where they form extensive H-bond interactions with the phosphate backbones of G52–G56 region (Fig. 2B). In details, R116 forms two H-bonds: one with the OP1 atom of G52 and the other with the OP2 atom of U53 (Fig. 2C). Both R123 and R138 interact with G56. Unlike G52–C55 region, the phosphate backbone of G56 is severely bent and packs against the nucleobase of U54 (Fig. 2C). R123 interacts with the OP1 atom of G56, whereas R138 interacts with both OP1 and OP2 atoms. Besides G56, R138 also interacts with the OP2 atom of C55

(Fig. 2D). R142 forms two H-bonds: one with the OP1 atom of U53 and the other with the OP2 atom of U54 (Fig. 2E).

In the complex structure, the acceptor stem and the predicted 5'-end cleavage site of *pre-tRNA^{His}* are far away from *PbHARP* (Fig. 2A), which prompted us to ask whether *PbHARP* has *pre-tRNA^{His}* cleavage activity. To this end, we carried out in vitro cleavage assay using 0.05 μM *PbHARP* and 1.0 μM *pre-tRNA^{His}* (Supplementary Fig. 3B). Although it is weak, *PbHARP* can remove the 5'-end of *pre-tRNA^{His}*. Besides *pre-tRNA^{His}*, we showed that *PbHARP* is also able to process *pre-tRNA^{Cys}* and *pre-tRNA^{Phe}* under identical reaction conditions (Supplementary Fig. 3B). We then asked whether the *pre-tRNA* cleavage by *PbHARP* is specific. To address this question, we performed *pre-tRNA^{Phe}* (5.0 μM) cleavage assay using 0.25 μM *PbHARP*, Aq_880 or *E. coli* RNase P. The reaction mixtures were separated on 20% denaturing UREA-PAGE gels (Supplementary Fig. 3C). No cleavage product was observed in the absence of RNase P. Two 5'-leader cleavage products were observed, one being identical in size to that generated by Aq_880 and *E. coli* RNase

P (Supplementary Fig. 3C, outer right lane). This shows that *PbHARP* cleaves pre-tRNA at the canonical cleavage site in addition to some apparent miscleavage at the neighboring upstream site.

To verify whether the interactions observed in the complex are functionally relevant, we constructed five *PbHARP* mutants, including R116A, R123A, T135A, R138A and R142A. T135 interacts with pre-tRNA^{His} in the complex (Fig. 2E), but it is not conserved (Supplementary Fig. 1). Surprisingly, in vitro binding assays showed that Ala substitution of the five pre-tRNA recognizing residue has no obvious impacts on pre-tRNA binding by *PbHARP* (Supplementary Fig. 6A). Unlike many other homology proteins, *PbHARP* is longer at its N-terminus. The extra *PbHARP* terminus is highly positive in charge; out of the total 13 residues, 7 are Arg or Lys (Supplementary Fig. 1B). To test whether the extra terminus contributes to the unexpected pre-tRNA binding results, we constructed one truncated protein, *PbHARP*_{ΔN} (aa 14–203). Compared to the full-length protein, the pre-tRNA binding ability of *PbHARP*_{ΔN} is weaker (Supplementary Fig. 6B). Started from *PbHARP*_{ΔN}, we then constructed several *PbHARP* mutants with pre-tRNA recognizing residue mutation. Although it has no strong impact on the overall pre-tRNA binding ability, Ala substitution of either R116 (for the *PbHARP*_{ΔN}_R116A mutant) or R123 (for the *PbHARP*_{ΔN}_R123A mutant) altered the pre-tRNA binding mode of *PbHARP*, indicated by the smeared bands on the gel. The pre-tRNA binding ability of *PbHARP*_{ΔN}_R138A is weaker than that of *PbHARP*_{ΔN}. No obvious difference was observed for the *PbHARP*_{ΔN}_R142A mutant, but the *PbHARP*_{ΔN}_R138A/R142A mutant showed weaker pre-tRNA binding ability, compared to the *PbHARP*_{ΔN}_R138A mutant.

The above mutation and binding assay results indicated that R116, R123, R138 and R142 affect pre-tRNA binding by *PbHARP*. To further confirm the functional importance of these residues, we performed in vitro cleavage assay using equimolar (1.0 μM) pre-tRNA^{His} and proteins (Fig. 2F). Substitution of T135 by Ala has no obvious impacts on pre-tRNA^{His} cleavage, compared to the wild-type (WT) *PbHARP*. In contrast to T135, R116, R123, R138 and R142 are all highly conserved (Supplementary Fig. 1). Ala substitution of either R116, R123 or R142 caused a dramatic reduction of the protein's pre-tRNA^{His} cleavage activity. No detectable pre-tRNA^{His} cleavage activity at all was observed for the R138A mutant. Altogether, these observations suggested that R116, R123, R138 and R142 are critical for assembly of the catalytically active *PbHARP*/pre-tRNA complex.

Domain rearrangement and active site assembly of *PbHARP*.

HARP contains one metallonuclease domain, whose activity is normally dependent on cations³⁴. Previous sequence analysis and mutagenesis studies³⁰ suggested that the catalytic center of Aq_880 is composed of three aspartate residues, corresponding to D151, D155 and D173 in *PbHARP* (Supplementary Fig. 1B). The functional importance of these catalytic residues could be confirmed by the in vitro cleavage assays using equimolar (1 μM) protein and pre-tRNA^{His} (Supplementary Fig. 7). *PbHARP* has no pre-tRNA^{His} cleavage activity in the absence of divalent cation. In the presence of Mg²⁺ (the common cofactor of metallonucleases), *PbHARP* can efficiently remove pre-tRNA^{His} 5'-end. At a reaction time of 5 min, about 60% of the substrate were cleaved; only trace amounts of intact substrate remained at a reaction time of 20 min. In contrast to WT *PbHARP*, no pre-tRNA^{His} cleavage activity was observed for the D151A, D155A and D173A mutants (Supplementary Fig. 7).

In the apo-form *PbHARP* structure, both D151 and D155 reside at the central region of α8 and D173 locates at the N-terminus of α9 (Fig. 3A, Supplementary Fig. 8A). The side chains of D155 and D173 are close to each other, forming water-mediated H-bond interactions. In contrast, the side chain of D151 points away from D155, the averaged distance between their carboxylate groups exceeding 7 Å. R110 locates at the middle of α7; its side chain points toward D155, forming two direct H-bond interactions between their guanidine and carboxylate groups. Such conformations are incompatible with Mg²⁺ coordination with D151, D155 and D173, suggesting that the apo-form *PbHARP* structure is actually in an inactive state.

As demonstrated for many nucleases, such as RNase III^{35,36}, RNase D³⁷ and RNase H³⁸, Ca²⁺ is insufficient in supporting the cleavage activity but can mimic Mg²⁺ in coordination. The crystal of *PbHARP*/pre-tRNA^{His} was grown in the presence of 10 mM CaCl₂. In the structure, several well-defined Ca²⁺ ions were captured (Supplementary Fig. 8B). As demonstrated by the A monomer of *PbHARP* (Fig. 3B), Ca²⁺ directly coordinates with the side chain OD1 atom of D151. In addition, it also coordinates with the main chain O atoms of R146 and I149. The side chains of D151, D155 and D173 are close to each other. Although not observed in the structure, this arrangement may allow *PbHARP* to coordinate a second divalent cation, thereby utilizing the common two metal ion-assisted mechanism of catalysis^{35,37,38}.

Divalent metal ion coordination, substrate binding and mutagenesis studies all indicated that *PbHARP* in the complex is in the active state. To reveal the structural basis underlying the inactive and active state switching, we compared the apo- and the pre-tRNA^{His}-complexed *PbHARP* structures. As depicted in Fig. 3C, the relative orientations of the metallonuclease domains are preserved, but the orientations of the HB domains are significantly different in the two structures. Compared to the apo-structure, all the HB domain helices are anticlockwisely rotated in the complex. Further structural analysis (Supplementary Fig. 8C–E) revealed that the inactive and active state switching of *PbHARP* is mainly mediated by the rearrangement of the HB domain, including rotation of α7 and shifting of the N-terminus of α8. Due to the structural rearrangement, α8 is split into two helices (termed α8a and α8b) in the complex. α8a and α8b are linked by one short loop, starting at the hinge residue E147 and ending at the catalytic residue D151 (Fig. 3D).

Proper dimerization is essential for the function of *PbHARP*.

PbHARP exists as dimer in both the apo- and pre-tRNA^{His}-complexed structures. The dimerization interactions mediated by β4 (Fig. 1F, G) are maintained in both structures; however, due to the rearrangement of the HB domain, the dimerization interactions mediated by α7 and α8 varied. In the apo-form structure, L103, I104, V107 and I111 of α7 and L145 of α8 all locate around the central axis of HB, forming extensive hydrophobic interactions (Fig. 4A). V107 of the two partner molecules directly interact with each other, whereas L103 and I111 interact with I104 and L145 of the partner molecule, respectively. In the tRNA-complexed structure (Fig. 4B), L103 of the two *PbHARP* monomers interact with each other. Stable hydrophobic interactions are also observed between the two V107 residues. Compared to the apo-form structure, the averaged distance between I111 of the two monomers is increased by 2 Å, altering the interactions between I111 and L145 in the complex.

In the apo-form structure, the N-termini of α8 helices are close to each other (Fig. 1C). However, the C-termini of α7 helices gathered together in the complex structure, allowing cooperative pre-tRNA binding by R116, R123, R138 and R142 of the two monomers (Fig. 2B–E, Supplementary Fig. 5A, C). Like the four

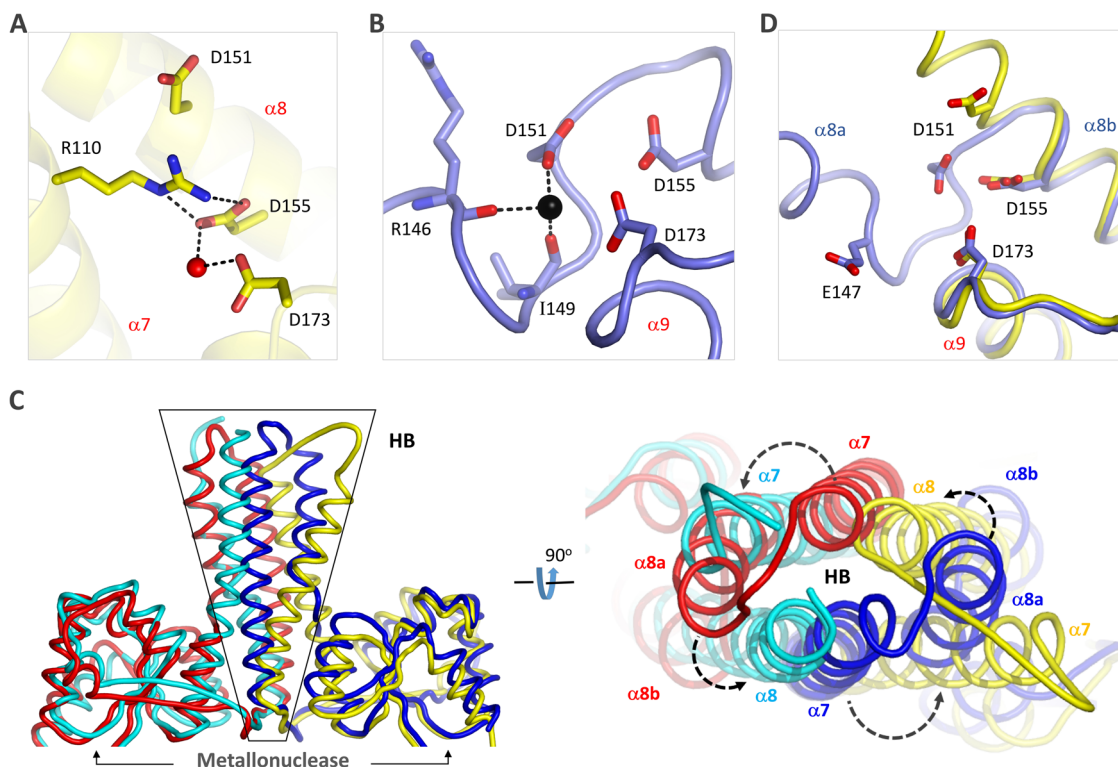


Fig. 3 Conformational change of *PbHARP*. **A** Conformation of the Asp residues observed in the apo-form *PbHARP* structure. **B** Conformation and cation coordination of the Asp residues observed in the pre-tRNA-complexed structure. Ca^{2+} ion is shown as black sphere, the Ca^{2+} -coordinating residues are shown as sticks. **C**, **D** Structural superposition showing the conformational changes of *PbHARP*. The apo-form structure is colored in yellow and cyan, whereas the pre-tRNA-complexed structure is colored in blue and red.

Arg residues, A118 is highly conserved (Supplementary Fig. 1). In the complex structure, the A118 residues center around the HB domain central axis, the distance between their side chains is only 3.6 Å (Fig. 4C). Except A118, many other hydrophobic residues, including A121 and Y141, are also arranged around the central axis (Fig. 4C).

To investigate the functional importance of HARP dimerization, we constructed and purified one *PbHARP* mutant, L103E/V107E, in which the hydrophobic L103 and V107 are simultaneously substituted by negatively charged Glu residues. Different from WT *PbHARP*, the size exclusion chromatographic study indicated that the L103E/V107E mutant mainly exists as monomer with an elution volume of 15.6 mL (Fig. 4D). Compared to WT *PbHARP*, the L103E/V107E mutant is less stable, indicated by its lower melting temperature (51.1 °C) measured by the Nano-DSF method (Supplementary Fig. 9A, B). In addition, we also performed CD spectra analysis for the two *PbHARP* proteins (Supplementary Fig. 10). In consistent with the crystal structure (Supplementary Fig. 4A, B), the CD results (Supplementary Fig. 10A, B) showed that WT *PbHARP* possesses ~50% α -helices and 10% β -strands (Parallel + Antiparallel) at 20 °C; the calculated T_m value is ~70.0 °C. Different from the WT protein, the L103E/V107E mutant (Supplementary Fig. 10C, D) possesses ~30% α -helices and 20% β -strands (Parallel + Antiparallel) at 20 °C; the T_m value of the mutant is about 53.0 °C. These results suggested that the L103E/V107E mutation affected the proper folding of *PbHARP*, which may play important role in the lower stability of the L103E/V107E mutant.

In addition to L103E/V107E, we also constructed the A118E/A121E mutant of *PbHARP*, in which A118 and A121 are substituted by Glu. The elution volume (14.4 mL) of A118E/A121E is similar to that of WT *PbHARP*, suggesting that it also exists as dimer (Fig. 4D). However, different from WT *PbHARP*,

the pre-tRNA binding affinity of A118E/A121E is weak; no pre-tRNA binding at all was observed for the L103E/V107E mutant (Supplementary Fig. 11A). WT *PbHARP* can cleave pre-tRNA, but detectable pre-tRNA^{His} cleavage activity was neither observed for the L103E/V107E nor for the A118E/A121E mutant under identical reaction conditions (Fig. 4E), probably due to their low stability and/or difficulty to form a functional dimer.

Length of the acceptor stem affects pre-tRNA cleavage by HARPs. Compared to pre-tRNA^{Gly}, pre-tRNA^{Cys} and pre-tRNA^{Phe}, the pre-tRNA^{His} cleavage activity of *PbHARP* is significantly weaker (Fig. 1B, Supplementary Fig. 3B). In the structure of the complex, *PbHARP* mainly recognizes the shape and backbone of the T Ψ C-loop (Fig. 2B–E), a structural element conserved in all tRNAs. However, compared to other pre-tRNAs (Supplementary Fig. 3A), the acceptor stem of pre-tRNA^{His} is one base pair longer, due to the pairing of G-1 and C72 which is not visible in Supplementary Fig. 5B. During pre-tRNA 5'-end maturation, the phosphodiester bond between nucleotides -1 and +1 is usually cleaved. To investigate whether the extra base pair contributes to the weaker pre-tRNA^{His} cleavage activity of *PbHARP*, we transcribed three pre-tRNA^{His} variants (Fig. 5A), His-7bp, His-6bp and His-9bp, which have seven, six and nine base pairs in the acceptor stem region, respectively.

Using 0.05 μM *PbHARP* and 1 μM native or mutated pre-tRNA^{His}, we performed in vitro cleavage assays. As depicted in Fig. 5B, no His-9bp cleavage activity was observed for *PbHARP*. Although it can be cleaved, the cleavage efficiency of native pre-tRNA^{His} was much weaker than those of His-7bp and His-6bp. In fact, the cleavage efficiency of His-7bp was comparable to those of pre-tRNA^{Cys} and pre-tRNA^{Phe} under identical reaction conditions (Supplementary Fig. 3B). Besides *PbHARP*, we also performed in vitro cleavage assay using 0.05 μM *TrHARP*

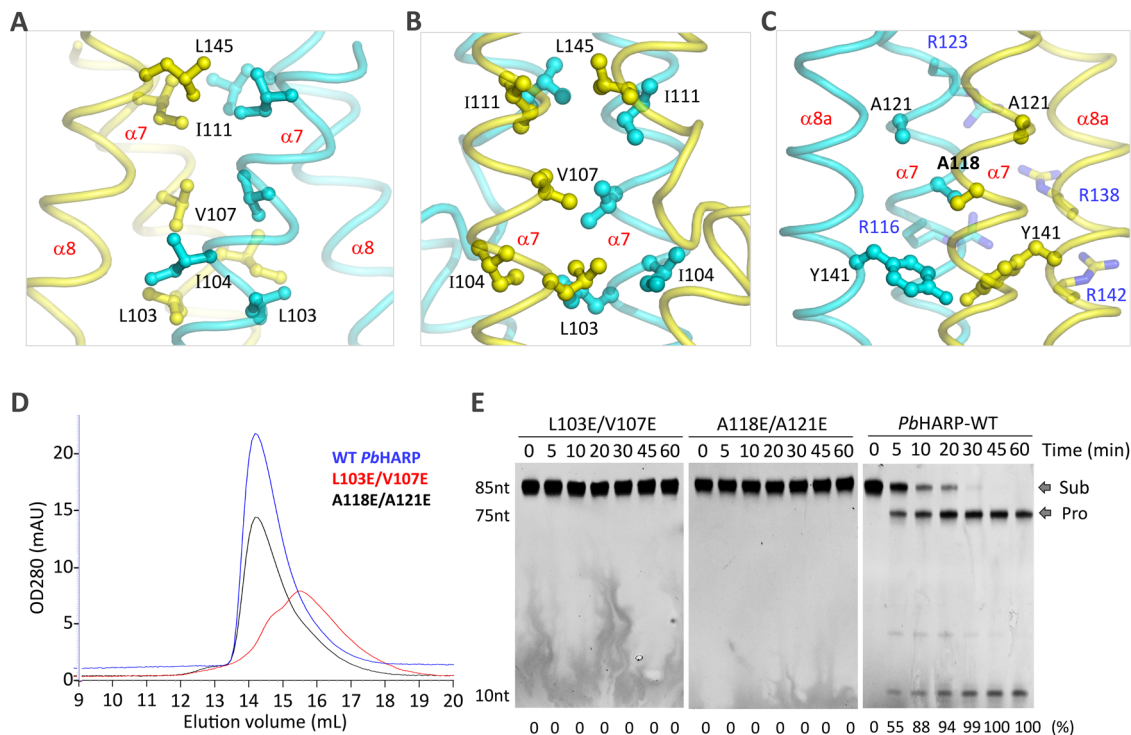


Fig. 4 Verification of the functional importance of *PbHARP* dimerization. **A** Interactions involved in *PbHARP* dimerization in the apo-form structure. **B, C** Interactions involved in *PbHARP* dimerization in the pre-tRNA-complexed structure. **D** Size-exclusion chromatographic analysis of wild-type, L103E/V107E and A118E/A121E mutants of *PbHARP*. **E** In vitro pre-tRNA^{His} cleavage assays catalyzed by WT *PbHARP*, L103E/V107E and A118E/A121E mutants. The substrate cleavage percentage (%) is shown at the bottom of the gels. Experiments were repeated independently three times with similar results. Source data are provided in Source Data file.

(Fig. 5C). Similar to *PbHARP*, *TrHARP* failed to cleave His-9bp and also showed a clear preference for His-7bp.

Oligomerization and assembly of *TrHARP*. Both *TrHARP* and *PbHARP* preferably cleaved the His-7bp variant (Fig. 5B, C), but their oligomerization states are different. *TrHARP* eluted from the size exclusion column at ~11.4 mL in all buffers containing 100, 200 or 500 mM salt (Fig. 6A). A previous study suggested the possibility that Aq_880 may form either six trimers or three hexamers³⁰ (inferred from size exclusion chromatography profiles), but both *PbHARP* and *TcHARP* formed dimers in the crystal structures (Fig. 1C–F). To gain more insights into the assembly of HARPs, *TrHARP* collected from the peak of the size-exclusion columns was subjected to electron microscopy. As depicted in Fig. 6B and S12A, *TrHARP* produced many well-defined particles on the negative stain CCD images, which can be divided into 20 different classes (Fig. 6C).

The 2D class average showed an open ring-like conformation of *TrHARP*. The shape and size of the ring matched well with an assembly of six *PbHARP* dimers that arrange in a left-handed orientation (Fig. 6D). Oligomerization of *TrHARP* is mainly mediated by the metallo-nuclease domains, which form the main body of the ring; the HB domains of *TrHARP* all point toward the outside and are roughly vertical to the central rotation axis of the ring. Dimer_1 and Dimer_6 reside near the cleft of the ring, the rotation angle between them is ~80° (Supplementary Fig. 12B). The rotation angles are ~60°–70° between Dimer_4 and the two neighboring dimers (Dimer_3 and Dimer_5). The rotation angles are ~50° between Dimer_1 and Dimer_2, Dimer_2 and Dimer_3, and Dimer_5 and Dimer_6. On the average, the rotation angle between all the interacting dimers is ~55°.

HARP binds and cleaves pre-tRNA in a cooperative mode. In addition to the AB dimer that interacts with pre-tRNA^{His}, there is one more *PbHARP* dimer (the CD dimer) within the asymmetric unit of the *PbHARP*/pre-tRNA^{His} complex. Possibly due to crystal packing, no pre-tRNA^{His} is found near the CD dimer. Similar to *TrHARP*, the AB and CD dimers of *PbHARP* interact with each other via their metallo-nuclease domains; the rotation angle between the two dimers is ~55° (Supplementary Fig. 12C), which well matched with the averaged angle observed in the *TrHARP* model (Supplementary Fig. 12B). *PbHARP* appeared as dimer during the size-exclusion chromatographic analysis, but the complex structure and its similarity with *TrHARP* suggested that *PbHARP* has the potential to form larger oligomers under proper conditions. The incorporation of additional *PbHARP* dimers can be easily modeled by rotating the complex structure along the 2-fold axis of either the AB dimer (Fig. 7A) or the CD dimer.

Based on pre-tRNA^{His}, the original AB dimer and the predicted EF dimer, one plausible pre-tRNA binding and cleavage model can be produced (Fig. 7B). In the model, the AB dimer is responsible for pre-tRNA binding, whereas the processing is performed by the EF dimer. The model can well explain the substrate preference of *PbHARP* (Fig. 5B). For His-7bp, the acceptor stem is composed of 7 base pairs. When the elbow region of His-7bp is recognized by *PbHARP* AB dimer, the first base pair of the acceptor stem, G1:C71, reaches the catalytic site of molecule E of the EF dimer. The phosphate backbone of G1 can readily coordinate with the catalytic cations, allowing the cleavage reaction to occur. G-1 pairs with C72 in native pre-tRNA^{His}, and such pairing may interfere with the conformational change and metal ion coordination at the scissile phosphodiester bond between G-1 and G1. The acceptor stem of His-9bp is 2 bp longer than that of His-7bp; the longer acceptor stem likely

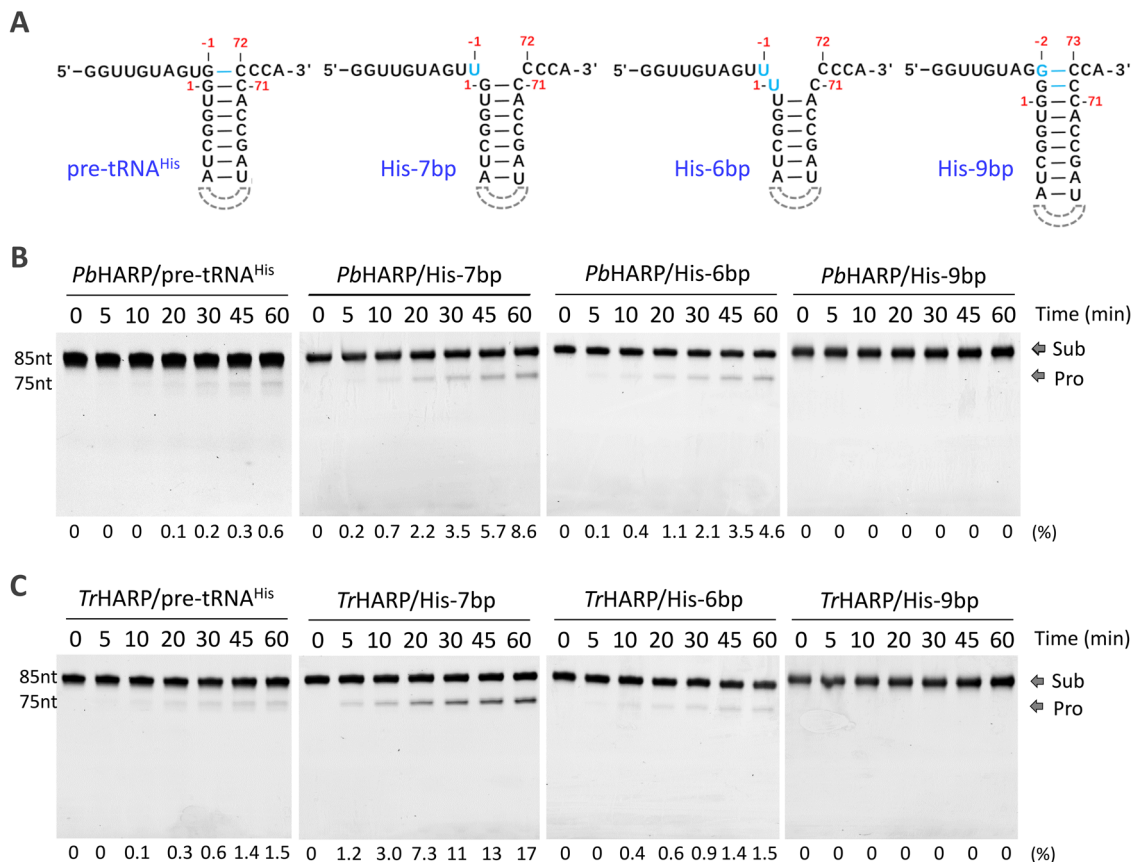


Fig. 5 Pre-tRNA^{His} mutation and cleavage by HARP. **A** Sequence and secondary structure of 5'-end, acceptor stem and 3'-end of the native and mutated pre-tRNA^{His}. **B** In vitro cleavage assays using 1 μ M pre-tRNA and 0.05 μ M *Pb*HARP. **C** In vitro cleavage assays using 1 μ M pre-tRNA and 0.05 μ M *Tr*HARP. The substrate cleavage percentage (%) is shown at the bottom of the gels. Experiments were repeated independently three times (**B**, **C**) with similar results. Source data are provided in Source Data file.

prevents His-9bp from simultaneously interacting with the AB and EF dimers.

Encouraged by the consistence between the in vitro cleavage assay results (Fig. 5B, C) and the pre-tRNA binding and cleavage model (Fig. 7B), we did further sequence and structural analysis. Besides the four Arg residues involved in pre-tRNA elbow recognition (Fig. 2B–E), two other Arg residues are also highly conserved in HARPs, corresponding to R108 and R146 in *Pb*HARP (Supplementary Fig. 1B). In the proposed pre-tRNA binding and cleavage model, the side chains of R146 of molecule E and R108 of molecule F are very close to the phosphate backbones of G-1 and G1 (Fig. 7C). To investigate the potential function of these Arg residues, we constructed two *Pb*HARP mutants, R108A and R146A. Although Ala substitution of R108 and R146 had no strong impact on pre-tRNA binding by *Pb*HARP (Supplementary Fig. 11B), no detectable pre-tRNA^{His} cleavage activity was observed for the R108A or the R146A mutant (Fig. 7D).

Pre-tRNA^{His} is recognized by the *Pb*HARP AB dimer in the complex, but neither R108 nor R146 of the AB dimer is involved in the interaction. These structural observations combined with the inactivity of these mutants further suggested that *Pb*HARP binds and cleaves pre-tRNA in a cooperative mode. To better demonstrate the cooperation between *Pb*HARP dimers, we performed in vitro crosslinking assays using WT *Pb*HARP and His-7bp (Supplementary Fig. 13). In addition to the monomer, *Pb*HARP dimers were also observed on the SDS-PAGE gel upon the addition of suberic acid bis (3-sulfo-N-hydroxysuccinimide ester) sodium salt (BS3), a chemical reagent that can crosslink the

side chains of two lysine residues. Although not clearly detectable in the absence of His-7bp, bands tentatively assigned to the *Pb*HARP tetramer (complex of two *Pb*HARP dimers) were clearly enhanced in the presence of His-7bp.

Discussion

Here we present results of structural and biochemical analysis of HARPs from four organisms, *P. bacterium*, *T. celer*, *T. minervae* and *T. ruber*. Like Aq_880 from *A. aeolicus*³⁰, in vitro cleavage assays showed that *Pb*HARP, *Tc*HARP, *Tm*HARP and *Tr*HARP can all remove the 5'-end of pre-tRNA, suggesting that pre-tRNA 5'-end processing is one conserved activity of HARP. HARP and the eukaryotic PRORP are all members of the PIN domain superfamily³². However, unlike PRORP enzymes constituting the PRORP subfamily, HARPs belong to the PIN_5 cluster subgroup. The metallo-nuclease domain is shared by all PIN superfamily members. In the reported *At*PRORP1³⁴ and *Hs*MRPP3³⁹ structures, the metallo-nuclease domains adopt one well conserved conformation (Supplementary Fig. 14A, B). Similar to all PIN superfamily members, the metallo-nuclease domain of HARP is of α/β fold in nature, but only two central β strands (β 1 and β 4) and one α helix (α 8b) can be superimposed with the corresponding regions in the eukaryotic PRORP structure (Supplementary Fig. 14C, D).

The activities of HARP and eukaryotic PRORP proteins are all cation-dependent. The apo-form *Hs*MRPP3 structure is in an inactive form, the active center is blocked by one of its own Arg residues³⁹. Activation of *Hs*MRPP3 requires the help of the two

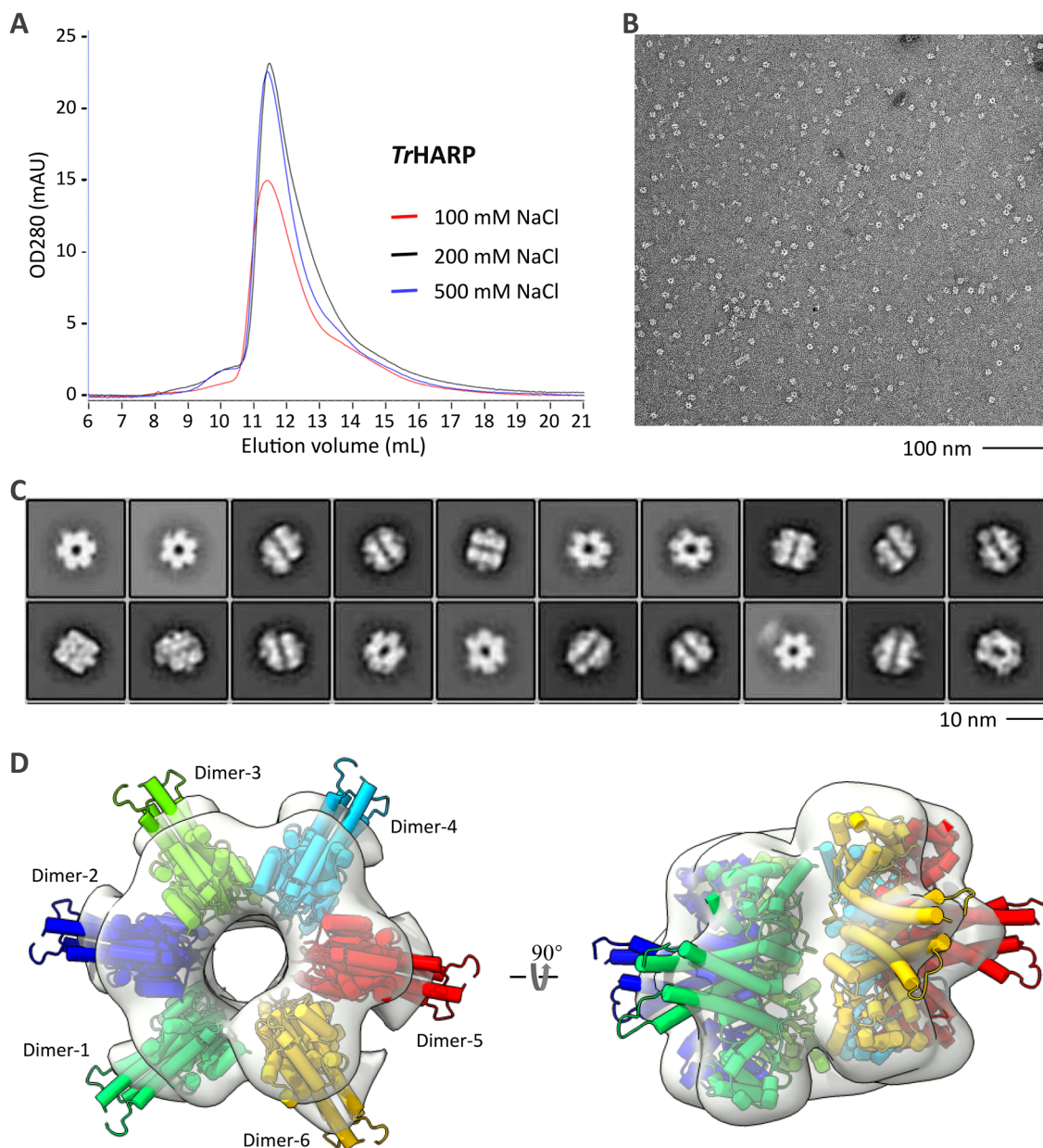


Fig. 6 Electron microscopy analysis of *TrHARP*. **A** Size-exclusion chromatographic analysis of *TrHARP*. Source data are provided in Source Data file. **B** A typical negative stain CCD image of *TrHARP*. Experiments were repeated independently twice with similar results. **C** The reference-free two-dimensional class averages of *TrHARP*. **D** The electron microscopy density maps of *TrHARP*, fitted with *PbHARP* dimer.

partner proteins, *HsMRPP1* and *HsMRPP2*. The apo-form structures of *PbHARP* and *TcHARP* are also in an inactive form. However, *in vitro* cleavage assays confirmed that all HARPs are active on their own (Fig. 1B). Structural comparison of the apo- and the pre-tRNA-complexed *PbHARP* structure suggested that HARP undergo self-activation via a large conformational change to the catalytically active form (Fig. 3B–D). When the catalytic *PbHARP* and *AtPRORP1* structures are superimposed based on their central β strands and α helix, D155 and D173 of *PbHARP* can align well with D475 and D493 of *AtPRORP1*, respectively. However, the third catalytic Asp residue (D151) of *PbHARP* is positioned in mirror image to D399 of *AtPRORP1* (Supplementary Fig. 14E). This mirror-image position of D151/D399 and the self-activation ability distinguish HARP from eukaryotic PRORPs.

As observed in their crystal structures, both *AtPRORP1* and *HsMRPP3* contain an N-terminal PPR domain. *AtPRORP1* PPR

is composed of five intandem PPR motifs; it is one of the typical proteins with multiple PPR motifs. In the complex structure of *PbHARP*, the pre-tRNA^{His} is recognized by the HB domain. Similar to the typical PPR motif, the HB domain is mainly composed of α -helices. Dimerization of HARP brings two HB domains together, structurally mimicking two PPR motifs arranged in tandem. Comparison with the recently reported *AtPRORP1* PPR/tRNA^{Phe} complex⁴⁰ showed that the HB domain of HARP and *AtPRORP1* PPR both recognize the T Ψ C-loop of pre-tRNA (Supplementary Fig. 15A, B), but the orientations of the HB domain and PPR domain are roughly vertical to each other (Supplementary Fig. 15C). Sequence and structure-based alignment indicated that the helices of the HARP HB and the *AtPRORP1* PPR domain are arranged in a reversed order (Supplementary Fig. 15D, E). However, one Arg residue is positioned at identical position in the two protein/tRNA complex structures (Supplementary Fig. 15F).

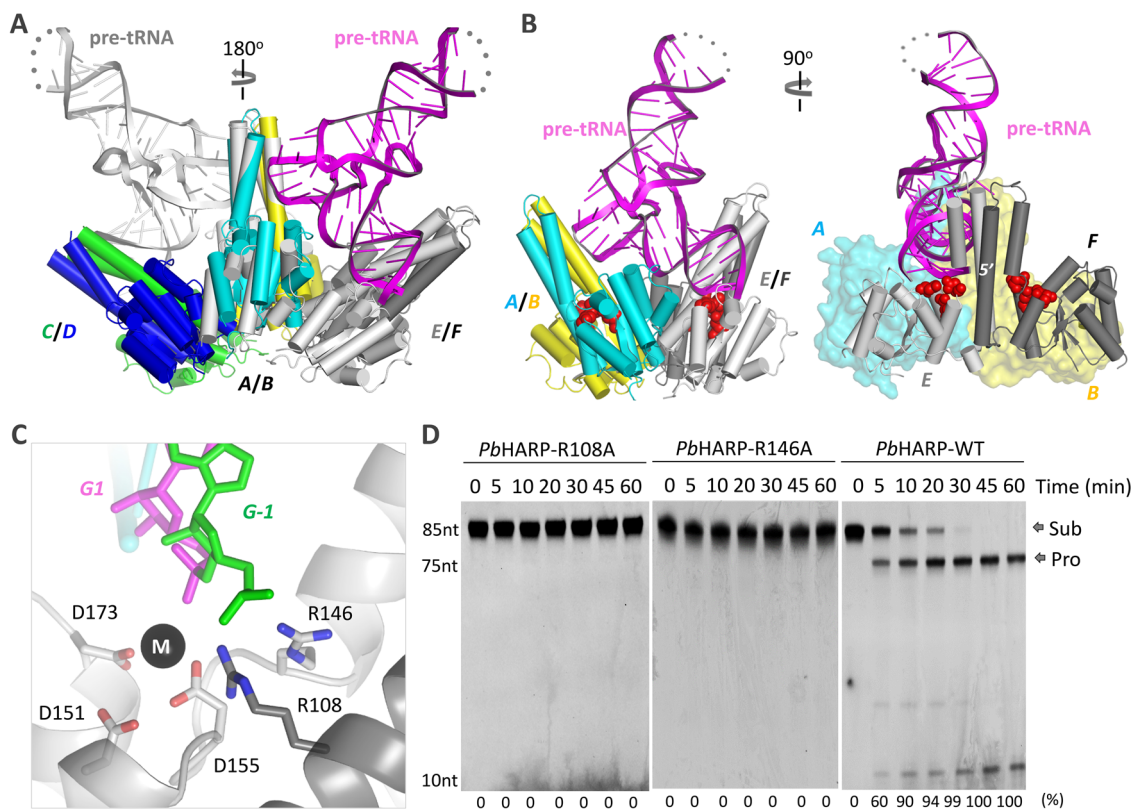


Fig. 7 Model for pre-tRNA 5'-end processing by HARPs. **A** Cartoon presentation of the original *PbHARP*/pre-tRNA^{His} complex and the one generated through rotation along the 2-fold axis of *PbHARP* AB dimer. **B** Proposed pre-tRNA 5'-end processing model. Pre-tRNA and the AB dimer of *PbHARP* are directly adopted from the complex, the EF dimer of *PbHARP* is generated through rotation of the structure. **C** Conformation of cation, cleavage site nucleotides, catalytic Asp residues, R108 and R146 in the proposed model. **D** In vitro pre-tRNA^{His} cleavage assays catalyzed by R108A and R146A mutants of *PbHARP*. The protein and pre-tRNA concentration are all fixed at 1.0 μM. The substrate cleavage percentage (%) is shown at the bottom of the gels. Experiments were repeated independently three times with similar results. Source data are provided in Source Data file.

The apo-form *AtPRORP1* structure adopts a V-shaped conformation. The two arms are formed by the metallonuclease domain and the PPR domain, linked by the central domain at the bottom. Although it has not been experimentally verified, it is likely that the V-shaped conformation is important for the function of *AtPRORP1*. With such arrangement, once pre-tRNA is recognized by the PPR domain, the acceptor stem and 5'-end of pre-tRNA will be placed near the active site of the metallonuclease domain. Neither a single monomer nor dimer of HARP is sufficient for binding and cleavage of pre-tRNA substrate, but the structure, mutagenesis and crosslinking assay results suggest that two dimers should be sufficient for the function of HARP. Interestingly, although not identical to *AtPRORP1*, the two HARP dimers also seem to adopt a V-shaped conformation (Fig. 7B).

The total size of HARP is only about 1/3 of *AtPRORP1* and *HsMRPP3*, the latter even requiring two additional protein cofactors (*HsMRPP1* and *HsMRPP2*) for in vivo activity. However, HARP performs the same reaction as *AtPRORP1* and the *HsMRPP* complex. A similar phenomenon has also been observed for some other RNA processing enzymes, such as RNase III and Trax. Bacterial RNase III is around 200 amino acids in length and functions as a homodimer. The human RNase III homology protein, Dicer, is about 2000 amino acids in length. Although different in size and domain architecture, bacterial RNase III and human Dicer catalyze the identical dsRNA cleavage reaction³⁵. In the RNAi pathway in human and *Drosophila*, the nicked RNA passenger strand is degraded by C3PO, the complex of Trax and Translin^{41,42}. Via oligomerization, Trax alone is

capable of RNA degradation in *Nanoarchaeum equitans*⁴³. Very recently, the cryo-EM structures of Aq_880 and HARP from *Halorhodospira halophila SL1* (Hh) were reported^{44,45}. Similar to TrHARP, Aq_880 and HhHARP assembled into dodecameric structures (Fig. 6D and Supplementary Fig. 16A, B). Structural comparison revealed that the overall folding of the individual HhHARP dimer is similar to that of *PbHARP* in the apo-form structure (Supplementary Fig. 16C), whereas the folding of Aq_880 dimer is more close to *PbHARP* in the complex structure (Supplementary Fig. 16D). Superposition of Aq_880 and the *PbHARP*/pre-tRNA^{His} complex structures further confirmed the cooperative pre-tRNA binding and cleavage mode of HARP (Supplementary Fig. 16E). Both our and the reported structures of HARP support the idea that oligomerization and division of function between mono- or dimers may be a more common way (than previously thought) to provide small enzymes with substrate binding and catalytic abilities.

RNase P exists in two distinct forms (PRORP and the RNA-based RNase P) and has been considered the most excellent example for the evolutionary transition theory³¹. Owing to their functional importance, the structure and catalytic mechanism of the RNA-based RNase P have been extensively studied^{46–50}. The existence of PRORP was proposed over 30 years ago^{51,52}, but the identification and characterization of PRORP were mainly reported in recent years^{28,29,53}. HARP is not essential for Archaea and Bacteria that also encode an RNA-based RNase P, whereas it is the only RNase P expressed in *A. aeolicus* and many related *Aquificaceae*, indicating their important role in tRNA maturation. Compared to PRORPs from animal mitochondria, plant

chloroplast, and trypanosomal mitochondria, very little is known about HARP from Archaeal and Bacteria. The *PbHARP* and *TcHARP* structures we report here are the only available crystal structures of HARP. Since no metallo-nuclease domain is present in the *AtPRORP1* PPR/tRNA^{Phe} complex⁴⁰, the *PbHARP*/pre-tRNA^{His} structure represents the only tRNA-complexed structure of a full-length protein-only RNase P. Instead of being dodecamers, *PbHARP* and *TcHARP* exist as dimer in the structures (Fig. 1C, E). In vitro size-exclusion chromatography showed that the oligomerization state of *TcHARP* is changeable (Supplementary Fig. 2B). We believe that a tetramer is the minimal catalytic unit of HARP in vivo, whereas dodecamerization may allow HARP to bind and cut multiple pre-tRNAs simultaneously, as proposed recently^{44,45}. Our study reveals the molecular basis for substrate binding and cleavage by HARP proteins.

Methods

Construction of recombinant plasmids. All genes that contain codon-optimized DNA sequences of Aq_880, *PbHARP*, *TcHARP*, *TmHARP*, *TrHARP* and *E. coli* RnpA (Supplementary Table 2) were purchased from Shanghai Genaray Biotech Co., Ltd, China. The target fragments were amplified by PCR, cleaved by BamHI (New England Biolabs, R3136L) and XhoI (New England Biolabs, R0146L), and cloned into pET28a-Sumo vectors. All recombinant proteins contain one 6 × His-Sumo tag at their N-termini. The recombinant plasmids of *PbHARP* mutants were constructed by overlap PCR using the wild-type (WT) *PbHARP* plasmid as template and primers listed in Supplementary Table 3. For transcription of *E. coli* RNase P RNA, the *mpB* gene with T7 promoter and hammerhead ribozyme (Supplementary Table 2) was inserted into a pUC18 vector at the HindIII and EcoRI restriction sites. Sequences of all plasmids were confirmed by DNA sequencing.

Expression and purification of recombinant proteins. All HARP proteins were expressed and purified using similar procedures. The recombinant plasmids were transformed into *Escherichia coli* BL21(DE3) competent cells and cultured in Lysogeny broth (LB) medium supplemented with 50 µg/mL kanamycin at 37 °C. When the OD₆₀₀ reached 0.6, isopropyl β-D-1-thiogalactopyranoside (IPTG, final concentration of 0.2 mM, Sangon Biotech, A600168-0100) was added to induce target protein expression. For accumulation of recombinant proteins, the cultures were incubated at 18 °C for an additional 18 h. For expression of Selenomethionine (Se-Met) substituted *TcHARP* protein, cells were cultured in M9 medium supplemented with 30 mg/L L-selenomethionine (J&K Scientific, S0442).

Cells were collected by centrifugation, resuspended in Buffer A (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 25 mM Imidazole), and lysed by ultrahigh-pressure homogenizer. Cellular debris was removed by centrifugation and the supernatant was loaded onto a HisTrap HP column (GE Healthcare). After washed by high salt buffer (20 mM Tris-HCl pH 8.0, 2 M NaCl), the bound protein was eluted with a linear gradient of buffer B (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 500 mM imidazole). The proteins were treated with Ulp1 and dialyzed against buffer C (20 mM Tris-HCl pH 8.0, 500 mM NaCl). The samples were loaded onto the HisTrap HP column again, the flow-through was collected, concentrated, and applied to a Hiloal 16/600 Superdex 200 column (GE Healthcare). For *E. coli* RnpA purification, the target protein eluted from HisTrap HP column was treated with Ulp1, diluted with 20 mM Tris-HCl pH 8.0 buffer and loaded onto a HiTrap Heparin column (GE Healthcare). The bound protein was eluted with a linear gradient of 20 mM Tris pH 8.0, 1 M NaCl buffer. The eluted protein was concentrated and applied to a Hiloal 16/600 Superdex 75 column (GE Healthcare) equilibrated with 20 mM Tris pH 8.0, 200 mM NaCl buffer. Peak fractions containing the target proteins were pooled and concentrated. The purity of protein was analyzed using 15% SDS-PAGE gel, the concentration was measured using a UV-spectrophotometer at 280 nm.

In vitro RNA transcription. All pre-tRNAs and *E. coli* RNase P RNA were synthesized by in vitro transcription using the same method. The template that contains a T7 promoter sequence upstream of the pre-tRNA or RNase P RNA coding region was obtained by overlap PCR using the primers listed in Supplementary Table 4. In vitro transcription reaction was carried out by mixing T7 RNA polymerase (prepared in laboratory), NTPs and template together and incubated at 37 °C overnight. Next day, the DNA template was digested by RNase-free DNase I (TaKaRa, 2270 A) at 37 °C. To remove the hammerhead ribozyme and increase RNase P RNA yield, the DNase I-digested transcription sample was incubated overnight at 4 °C. All RNA samples were mixed with equal volume of formamide, heated at 95 °C for 10 min, and chilled on ice. The in vitro transcripts were then separated by 10% (for pre-tRNA) or 6% (for RNase P RNA) denaturing UREA-PAGE gels (60 W for 6 h) in 1 × TBE buffer. The target pre-tRNA or RNase P RNA band was excised from the gel. The gel slice was eluted using elution buffer (300 mM sodium acetate pH 5.2, 1 mM EDTA pH 8.0) at 8 °C overnight. The RNA

was precipitated by mixing with 1/10 volume sodium acetate pH 5.2 and 1 volume of isopropanol and freezing at −80 °C. The sample was centrifuged at 13000 × g for 30 min at 4 °C, the supernatant was discarded and the pellet was washed by 75% ethanol and air-dried. Finally, the RNA was dissolved in RNase-free water, quantified by absorbance, and stored at −80 °C. All purified pre-tRNAs were analyzed by 10% denatured UREA-PAGE gels and 6% native PAGE gels.

In vitro RNA binding assays. Electrophoretic mobility shift assays (EMSA) were used to analyze the pre-tRNA binding ability of wild type full-length, truncated and/or mutated *PbHARP* proteins. Assays with a final volume of 10 µL were set up. The reaction mixtures that included 1 µM pre-tRNA and protein serially diluted from 1 µM to 10 µM (or 20 µM) were incubated on ice for 1 h in 20 mM Tris-HCl pH 8.0, 200 mM NaCl, 5% (v/v) glycerol buffer. After incubation, 2 µL 10 × DNA loading buffer (Vazyme) was added. Samples were loaded onto a precooled 1% agarose gel and then separated by electrophoresis (100 V for 20 min at 4 °C) in 1 × TAE buffer. The gel was stained by Gel Red and imaged using gel scanner. Intensities of the substrate bands were quantified by ImageJ. The percentage of binding, for each protein concentration, was calculated. Data were then fitted to the equation $Y = B_{\max} * X^h / (K_d^h + X^h)$ using nonlinear regression (curve fit) in GraphPad Prism. The dissociation constants (K_d) were determined from the regression curve.

In vitro pre-tRNA cleavage assays. *PbHARP* and mutants were serially diluted to 10 µM (other HARPs were 0.5 µM) with processing buffer [20 mM Tris-HCl pH 9.0, 100 mM NaCl, 5% (v/v) glycerol] and pre-incubated at 37 °C for 5 min. The mixture of 33.75 µL processing buffer, 2.25 µL 100 mM MgCl₂, and 4.5 µL 10 µM pre-tRNA was also pre-incubated at 37 °C for 5 min. The reactions were initiated through addition of 4 µL diluted protein. The final concentrations of HARP, pre-tRNA, MgCl₂ are 1 µM (or 0.05 µM), 1 µM, 5 mM, respectively. At specific time points, 5 µL aliquots of the reaction were quenched with 5 µL 2 × Formamide loading buffer [90% (v/v) formamide deionized, 20 mM EDTA, 1% SDS, 0.02% bromophenol blue, 0.02% xylene cyanol] and heated at 95 °C for 5 min. Samples were centrifuged and separated by 10% denaturing UREA-PAGE gels (10 W for 2 h) in 1 × TBE buffer. The bands were stained by Gel Red and visualized by Typhoon FLA 9000 Imaging Scanner.

For cleavage site specificity assay, *E. coli* RNase P holoenzyme was prepared by mixing RnpA and RNase P RNA with a 1:1 molar ratio in processing buffer. The mixture was incubated at 4 °C for 10 min. RNase P (*E. coli* RNase P holoenzyme, *PbHARP* or Aq_880) and the reaction mixtures (15 µL processing buffer, 1 µL 100 mM MgCl₂ and 2 µL 50 µM pre-tRNA^{Phe}) were incubated at 37 °C for 5 min. The reactions were initiated through addition of 2 µL diluted RNase P. The final concentrations of RNase P, pre-tRNA^{Phe} and MgCl₂ were 0.25 µM, 5 µM and 5 mM, respectively. The reactions were terminated after 15 min by heating to 95 °C for 5 min. Samples were centrifuged and separated by 20% denaturing UREA-PAGE gels (10 W for 2 h 25 min) in 1 × TBE buffer.

In vitro chemical crosslinking assays. The buffer of *PbHARP* was exchanged with 20 mM HEPES pH 7.9, 500 mM NaCl through protein concentration in an ultrafiltration tube. For crosslinking of complex, *PbHARP* (the final concentration is 1 mg/ml) and His-7bp were mixed with 10 mM CaCl₂ in gel filtration buffer at a molar ratio of 1:1.2. For cross-linking of apo-*PbHARP*, *PbHARP* was mixed with corresponding volume of water in gel filtration buffer. The mixtures were incubated at 4 °C for 40 min and divided into 10 µL per tube. Chemical crosslinking experiments were started by addition of BS3 (ThermoFisher Scientific, 21580) reagent at a final concentration of 0 mM, 0.01 mM, 0.03 mM, respectively. Samples were incubated at 25 °C for 1 h. In the crosslinking system, the final concentration of *PbHARP* and RNA is 37.6 µM and 45.1 µM, respectively. To terminate the crosslinking reactions, 20 mM Glycine was added and then incubated at 25 °C for 15 min. Finally, 3 µL of 5×SDS PAGE loading buffer [250 mM Tris-HCl, pH 6.8, 50% (v/v) glycerol, 10% (w/v) SDS, 0.05% bromophenol blue, 5% (v/v) β-mercaptoethanol] was added into each tube and 5 µL samples were loaded to a 12% SDS-PAGE (300 V for 30 min) for analysis.

Size-exclusion chromatography and electron micrographic analysis. To analyze the oligomerization state of HARPs, 500 µL HARP proteins (300 µg) were applied to Superdex 200 Increase 10/300 GL column (GE Healthcare) equilibrated with buffers containing 20 mM Tris-HCl pH 8.0 and 100–500 mM NaCl. The flow rate was set at 0.4 mL/min. Besides HARP proteins, we also performed size-exclusion chromatography analysis for 100 µL standard marker proteins using the same column equilibrated with 20 mM Tris-HCl pH 7.0 and 150 mM NaCl buffer. To further determine the oligomerization state of *TrHARP*, 400 µg protein was applied to Superdex 200 Increase 10/300 GL column equilibrated with 20 mM Tris-HCl pH 8.0 and 200 mM NaCl. The peak fraction was diluted to 40 µg/mL using the same buffer. 5 µL sample was absorbed onto the glow-discharged carbon-coated copper grids (200 mesh) purchased from Beijing Zhongjingkeyi Technology Co., Ltd, China and incubated for 1 min, the excess solution was blotted off. After washing, the sample on the grid was negatively stained with 0.75% uranyl formate for 3.5 min and then air-dried. Finally, the grid was viewed with the Talos L120C transmission electron microscope (ThermoFisher Scientific) equipped with a

4 K × 4 K CETA CCD camera (FEI) at an accelerating voltage of 120 kV. Data were collected and images were recorded at a nominal magnification of 92 000×, corresponding to a pixel size of 1.55 Å. The contrast-transfer function (CTF) parameters of each micrograph were determined using CTFIND4. Single particles were picked and processed using RELION3.0⁵⁴. 29647 particles were picked by using LoG-based autopicking from RELION3.0. Reference-free 2D classification was generated using RELION. 2D class-averages were used to generate de novo 3D model and 3D classification by RELION.

Nano-differential scanning fluorimetry (nano-DSF) analysis. Nano-DSF experiments were performed using a Prometheus NT.48 instrument from Nano-Temper Technologies (Munich, Germany). 20 µL of 1.0 mg/mL WT *PbHARP* and the L103E/V107E mutant samples were prepared. Each sample was dissolved in 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 5% (v/v) glycerol buffer. For each analysis, 10 µL samples were loaded into PR NT.48 standard capillaries from Nano-Temper Technologies. Samples were subjected to a temperature gradient of 1 °C/min from 20 °C to 90 °C and the fluorescence was constantly monitored. Protein unfolding was detected by following the change in tryptophan fluorescence at emission wavelengths of 330 and 350 nm. The ratio between the emission intensities at 350 nm and 330 nm was used to track the structural changes with increasing temperature. T_m value was determined from the first derivative nano-DSF curve.

Circular dichroism (CD) spectra analysis. Both WT *PbHARP* and the L103E/V107E mutant were dissolved in 5 mM Tris pH 8.0, 500 mM NaF buffer with a final concentration 0.2 mg/mL. 300 µL sample or buffer was loaded into a quartz cuvette of 1 mm path length (Hellma, Germany). The CD spectra was recorded on a nitrogen-flushed Chirascan CD spectrometer (Applied Photophysics, UK). The wavelength range of 190–260 nm was recorded. The step and band width were set at 1 nm. The scanning time per point was 0.5 s. For temperature control, the heating rate was set at 1 °C/min in a range of 20 °C–80 °C (with a temperature interval measurement of 2 °C). Secondary structure content (%) was calculated by CDNN software. T_m value was determined using nonlinear regression (curve fit) in GraphPad Prism software.

Crystallization and data collection. For crystallization of apo-form HARP, 10 mg/ml protein samples were used. For crystallization of HARP/pre-tRNA complex, HARP was mixed with pre-tRNA and 10 mM CaCl₂ in gel filtration buffer, the molar ratio between HARP and pre-tRNA is 1:1.2. The mixture was incubated at 4 °C for 1 h. The initial crystallization conditions were identified at 16 °C using the Gryphon crystallization robot system and commercial crystallization kits. The sitting-drop vapor diffusion method with the 3-drop Intelliplate plates were utilized during both initial screening and optimization process. The crystallization condition for *TcHARP* and Se-*TcHARP* was 30% (w/v) PEG 3000, 100 mM Tris-HCl pH 7.0, 200 mM NaCl. The crystals of apo-form *PbHARP* grew in the buffer composed of 2000 mM Ammonium sulfate, 100 mM Sodium citrate/Citric acid pH 5.5, whereas the crystallization conditions are composed of 8% (v/v) Tacsimate pH 8.0, 20% (w/v) PEG 3350 for the *PbHARP*/pre-tRNA^{His} crystals.

All crystals were cryoprotected in reservoir solution supplemented with 25% (v/v) glycerol and snap-frozen in liquid nitrogen. The X-ray diffraction data were collected on beamlines BL17U and BL19U at the Shanghai Synchrotron Radiation Facility (SSRF). HKL3000 program⁵⁵ was used to process the data. The data collection and processing statistics are summarized in Supplementary Table 1.

Structure determination and refinement. The apo-form Se-Met substituted *TcHARP* structure was solved by the single-wavelength anomalous diffraction (SAD) method with the Autosol program embedded in the Phenix suit⁵⁶. The initial model was built using the Autobuilt program and then refined against the diffraction data using the Refmac5 program of the CCP4 suite⁵⁷. The 2F_o - F_c and F_o - F_c electron density maps were regularly calculated and used as guide for the building of the missing amino acids using COOT⁵⁸. The apo-form *PbHARP* structure was solved by molecular replacement using the apo-*TcHARP* structure as the search model with the phaser program of the CCP4 suite. The pre-tRNA^{His}-complexed *PbHARP* structures was solved by molecular replacement using the apo-form *PbHARP* structure as the search model. Nucleic acids, ions, water, and other molecules were all built manually using COOT. The complex structures were refined using the Refmac5 program of the CCP4 suite or the phenix.refine program of Phenix suit. The structural refinement statistics were summarized in Supplementary Table 1.

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

Data availability

The data supporting the findings of this study are available from the corresponding authors upon reasonable request. Structural factors and coordinates have been deposited in the Protein Data Bank under accession codes 7E8J, 7E8K, and 7E8O for the apo-form

TcHARP, apo-form *PbHARP*, and *PbHARP*/pre-tRNA^{His} complex, respectively. Source data are provided with this paper.

Received: 6 June 2021; Accepted: 30 March 2022;

Published online: 28 April 2022

References

- Tuorto, F. et al. RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. *Nat. Struct. Mol. Biol.* **19**, 900–905 (2012).
- Krutyholowa, R., Zakrzewski, K. & Glatt, S. Charging the code - tRNA modification complexes. *Curr. Opin. Struct. Biol.* **55**, 138–146 (2019).
- Nagao, A. et al. Hydroxylation of a conserved tRNA modification establishes non-universal genetic code in echinoderm mitochondria. *Nat. Struct. Mol. Biol.* **24**, 778–782 (2017).
- Ranjan, N. & Rodnina, M. V. Thio-Modification of tRNA at the Wobble Position as Regulator of the Kinetics of Decoding and Translocation on the Ribosome. *J. Am. Chem. Soc.* **139**, 5857–5864 (2017).
- Popow, J. et al. HSPC117 is the essential subunit of a human tRNA splicing ligase complex. *Science* **331**, 760–764 (2011).
- Popow, J., Jurkin, J., Schleiffer, A. & Martinez, J. Analysis of orthologous groups reveals archaease and DDX1 as tRNA splicing factors. *Nature* **511**, 104–107 (2014).
- Reyes, V. M. & Abelson, J. Substrate recognition and splice site determination in yeast tRNA splicing. *Cell* **55**, 719–730 (1988).
- Harris, M. E. Theme and Variation in tRNA 5' End Processing Enzymes: Comparative analysis of protein versus ribonucleoprotein RNase P. *J. Mol. Biol.* **428**, 5–9 (2016).
- Karasik, A., Fierke, C. A. & Koutmos, M. Interplay between substrate recognition, 5' end tRNA processing and methylation activity of human mitochondrial RNase P. *RNA* **25**, 1646–1660 (2019).
- Marszalkowski, M., Willkomm, D. K. & Hartmann, R. K. 5'-end maturation of tRNA in aquifex aeolicus. *Biol. Chem.* **389**, 395–403 (2008).
- Paushkin, S. V., Patel, M., Furia, B. S., Peltz, S. W. & Trotta, C. R. Identification of a human endonuclease complex reveals a link between tRNA splicing and pre-mRNA 3' end formation. *Cell* **117**, 311–321 (2004).
- Saoura, M. et al. Mutations in ELAC2 associated with hypertrophic cardiomyopathy impair mitochondrial tRNA 3'-end processing. *Hum. Mutat.* **40**, 1731–1748 (2019).
- Wilusz, J. E., Freier, S. M. & Spector, D. L. 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. *Cell* **135**, 919–932 (2008).
- Xiong, Y. & Steitz, T. A. A story with a good ending: tRNA 3'-end maturation by CCA-adding enzymes. *Curr. Opin. Struct. Biol.* **16**, 12–17 (2006).
- Arrondel, C. et al. Defects in t(6)A tRNA modification due to GON7 and YRDC mutations lead to Galloway-Mowat syndrome. *Nat. Commun.* **10**, 3967 (2019).
- Yasukawa, T., Suzuki, T., Ishii, N., Ohta, S. & Watanabe, K. Wobble modification defect in tRNA disturbs codon-anticodon interaction in a mitochondrial disease. *EMBO J.* **20**, 4794–4802 (2001).
- Budde, B. S. et al. tRNA splicing endonuclease mutations cause pontocerebellar hypoplasia. *Nat. Genet.* **40**, 1113–1118 (2008).
- Schaffer, A. E. et al. CLP1 founder mutation links tRNA splicing and maturation to cerebellar development and neurodegeneration. *Cell* **157**, 651–663 (2014).
- Levinger, L., Morl, M. & Florentz, C. Mitochondrial tRNA 3' end metabolism and human disease. *Nucleic Acids Res.* **32**, 5430–5441 (2004).
- Shang, J. et al. Overexpression of *Schizosaccharomyces pombe* tRNA 3'-end processing enzyme Trz2 leads to an increased cellular iron level and apoptotic cell death. *Fungal Genet Biol.* **122**, 11–20 (2019).
- Lin, H. et al. CO2-sensitive tRNA modification associated with human mitochondrial disease. *Nat. Commun.* **9**, 1875 (2018).
- Guerriertakada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. The RNA Moiety of Ribonuclease-P Is the catalytic subunit of the enzyme. *Cell* **35**, 849–857 (1983).
- Guerrier-Takada, C. & Altman, S. Catalytic activity of an RNA molecule prepared by transcription in vitro. *Science* **223**, 285–286 (1984).
- Jarrous, N. & Gopalan, V. Archaeal/eukaryal RNase P: Subunits, functions and RNA diversification. *Nucleic Acids Res.* **38**, 7885–7894 (2010).
- Hartmann, E. & Hartmann, R. K. The enigma of ribonuclease P evolution. *Trends Genet.* **19**, 561–569 (2003).
- Walker, S. C. & Engelke, D. R. A protein-only RNase P in human mitochondria. *Cell* **135**, 412–414 (2008).
- Rackham, O. et al. Hierarchical RNA Processing Is Required for Mitochondrial Ribosome Assembly. *Cell Rep.* **16**, 1874–1890 (2016).

28. Gobert, A. et al. A single Arabidopsis organellar protein has RNase P activity. *Nat. Struct. Mol. Biol.* **17**, 740–U113 (2010).
29. Taschner, A. et al. Nuclear RNase P of *Trypanosoma brucei*: a single protein in place of the multicomponent RNA-protein complex. *Cell Rep.* **2**, 19–25 (2012).
30. Nickel, A. I. et al. Minimal and RNA-free RNase P in *Aquifex aeolicus*. *P. Natl. Acad. Sci. USA* **114**, 11121–11126 (2017).
31. Daniels, C. J., Lai, L. B., Chen, T. H. & Gopalan, V. Both kinds of RNase P in all domains of life: Surprises galore. *RNA*. **25**, 286–291 (2019).
32. Matelska, D., Steczkiewicz, K. & Ginalski, K. Comprehensive classification of the PIN domain-like superfamily. *Nucleic Acids Res.* **45**, 6995–7020 (2017).
33. Schwarz, T. S. et al. Homologs of *aquifex aeolicus* protein-only RNase P are not the major RNase P activities in the archaea *haloferax volcanii* and *methanosarcina mazei*. *IUBMB Life* **71**, 1109–1116 (2019).
34. Howard, M. J., Lim, W. H., Fierke, C. A. & Koutmos, M. Mitochondrial ribonuclease P structure provides insight into the evolution of catalytic strategies for precursor-tRNA 5' processing. *P Natl Acad. Sci. USA* **109**, 16149–16154 (2012).
35. Court, D. L. et al. RNase III: Genetics and Function; Structure and Mechanism. *Annu. Rev. Genet.* **47**, 405–40 (2013).
36. Gan, J. H. et al. Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. *Cell* **124**, 355–366 (2006).
37. Gao, Y. Q. et al. Structural basis for guide RNA trimming by RNase D ribonuclease in *Trypanosoma brucei*. *Nucleic Acids Res.* **49**, 568–583 (2021).
38. Nowotny, M., Gaidamakov, S. A., Crouch, R. J. & Yang, W. Crystal structures of RNase H bound to an RNA/DNA hybrid: Substrate specificity and metal-dependent catalysis. *Cell* **121**, 1005–1016 (2005).
39. Li, F., Liu, X., Zhou, W., Yang, X. & Shen, Y. Auto-inhibitory mechanism of the human mitochondrial RNase P protein complex. *Sci. Rep.* **5**, 9878 (2015).
40. Teramoto, T. et al. Pentatricopeptide repeats of protein-only RNase P use a distinct mode to recognize conserved bases and structural elements of pre-tRNA. *Nucleic Acids Res.* **48**, 11815–11826 (2020).
41. Ye, X. et al. Structure of C3PO and mechanism of human RISC activation. *Nat. Struct. Mol. Biol.* **18**, 650–657 (2011).
42. Liu, Y. et al. C3PO, an endoribonuclease that promotes RNAi by facilitating RISC activation. *Science* **325**, 750–753 (2009).
43. Zhang, J. et al. Structural basis for single-stranded RNA recognition and cleavage by C3PO. *Nucleic Acids Res.* **44**, 9494–9504 (2016).
44. Teramoto, T. et al. Minimal protein-only RNase P structure reveals insights into tRNA precursor recognition and catalysis. *J. Biol. Chem.* **297**, 101028 (2021).
45. Feyh, R. et al. Structure and mechanistic features of the prokaryotic minimal RNase P. *Elife* **10** <https://doi.org/10.7554/eLife.70160> (2021).
46. Reiter, N. J. et al. Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature* **468**, 784–789 (2010).
47. Wu, J. et al. Cryo-EM structure of the human ribonuclease P holoenzyme. *Cell* **175**, 1393–1404 e1311 (2018).
48. Klemm, B. P. et al. The diversity of ribonuclease P: Protein and RNA catalysts with analogous biological functions. *Biomolecules* **6** <https://doi.org/10.3390/biom6020027> (2016).
49. Schencking, I., Rossmann, W., Hartmann, R. K. *Diversity and evolution of RNase P*. (Springer, Cham, 2020).
50. Wan, F. et al. Cryo-electron microscopy structure of an archaeal ribonuclease P holoenzyme. *Nat. Commun.* **10**, 2617 (2019).
51. Wang, M. J., Davis, N. W. & Gegenheimer, P. Novel mechanisms for maturation of chloroplast transfer RNA precursors. *EMBO J.* **7**, 1567–1574 (1988).
52. Rossmann, W. & Karwan, R. M. Characterization of human mitochondrial RNase P: Novel aspects in tRNA processing. *Biochem. Biophys. Res. Co.* **247**, 234–241 (1998).
53. Holzmann, J. et al. RNase P without RNA: Identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell* **135**, 462–474 (2008).
54. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7** <https://doi.org/10.7554/eLife.42166> (2018).
55. Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: The integration of data reduction and structure solution-from diffraction images to an initial model in minutes. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **62**, 859–866 (2006).
56. Adams, P. D. et al. PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **58**, 1948–1954 (2002).
57. Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **59**, 1131–1137 (2003).
58. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **60**, 2126–2132 (2004).

Acknowledgements

We thank the staff of BL17U1 beamline at the Shanghai Synchrotron Radiation Facility (SSRF), the staff of BL18U1, BL19U1 beamlines of National Facility for Protein Science Shanghai for help during data collection. This work was supported by the National Natural Science Foundation of China (32171197 and 31870721 to J.G.) and the Key Research and Development Project of China (2016YFA0500600 to J.G.).

Author contributions

Y.Y.L. performed the crystallization and biochemical studies with the help of H.H.L., X.C., Z.W.S., Y.X.Z., Q.Y.S., X.Z., J. Y. and C.L.C. Y.Y.L. and Y.Q.G. collected and processed the crystal diffraction data. S.C.S., Y.Y.L. and G.L.L. performed the electron microscopy study. Y.Y.L., J.Z.L., J.B.M. and J.H.G. analyzed the data. Y.Y.L. and J.H.G. wrote 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/s41467-022-30072-6>.

Correspondence and requests for materials should be addressed to Jianhua Gan.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022