4146–4157 *Nucleic Acids Research, 2012, Vol. 40, No. 9* *Published online 30 December 2011*
doi:10.1093/nar/gkr1281

Crystal structure of human polynucleotide phosphorylase: insights into its domain function in RNA binding and degradation

Chia Liang Lin[1,2], Yi-Ting Wang[2], Wei-Zen Yang[2], Yu-Yuan Hsiao[2] and Hanna S. Yuan[2,3,*]

[1]Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu 30013, [2]Institute of Molecular Biology, Academia Sinica, Taipei 11529 and [3]Graduate Institute of Biochemistry and Molecular Biology, National Taiwan University, Taipei 10051, Taiwan, R.O.C.

Received November 4, 2011; Revised December 8, 2011; Accepted December 9, 2011

ABSTRACT

Human polynucleotide phosphorylase (hPNPase) is a 3′-to-5′ exoribonuclease that degrades specific mRNA and miRNA, and imports RNA into mitochondria, and thus regulates diverse physiological processes, including cellular senescence and homeostasis. However, the RNA-processing mechanism by hPNPase, particularly how RNA is bound via its various domains, remains obscure. Here, we report the crystal structure of an S1 domain-truncated hPNPase at a resolution of 2.1 Å. The trimeric hPNPase has a hexameric ring-like structure formed by six RNase PH domains, capped with a trimeric KH pore. Our biochemical and mutagenesis studies suggest that the S1 domain is not critical for RNA binding, and conversely, that the conserved GXXG motif in the KH domain directly participates in RNA binding in hPNPase. Our studies thus provide structural and functional insights into hPNPase, which uses a KH pore to trap a long RNA 3′ tail that is further delivered into an RNase PH channel for the degradation process. Structural RNA with short 3′ tails are, on the other hand, transported but not digested by hPNPase.

INTRODUCTION

Polynucleotide phosphorylase (PNPase) is an evolutionary conserved 3′-to-5′ exoribonuclease that plays a key role in RNA processing and turnover in various species, ranging from bacteria, worms and plants to mammals (1–4). PNPase catalyzes the processive phosphorolysis of RNA by using an inorganic phosphate to cleave the phosphodiester linkage at the 3′-end of a RNA chain, thus generating nucleoside diphosphates as cleavage products. Bacterial PNPase participates not only in stable RNA maturation but also plays a pivotal role in mRNA turnover (5–7). The fact that plant PNPase contains an extra chloroplast signal peptide whereas mammalian PNPase possesses an extra mitochondrial signal peptide at the N-terminus suggests that the different PNPases target to chloroplasts and mitochondria, respectively (8).

Interestingly, human PNPase (hPNPase) has diversified functions not only in mitochondria but also in the cytoplasm (9). The gene encoding hPNPase (*hPNPase*$^{old\text{-}35}$) was first identified as one of the upregulated genes in the process of terminal differentiation and cellular senescence (10,11). hPNPase is ubiquitously expressed in normal tissues and is further induced by type I interferons (IFN-α and IFN-β) to degrade specifically *c-myc* mRNA, as well as a subset of microRNAs, and to regulate cell growth arrest and apoptosis (12–14). As a result, the over expression of this RNA-metabolizing enzyme is associated with growth inhibition and some pathological changes that occur during aging, such as inflammation (15,16). On the other hand, in mitochondria, hPNPase is located primarily in the intermembrane space and is involved in mitochondrial RNA processing and homeostasis (17). Moreover, hPNPase was demonstrated to have an unanticipated role in RNA translocation, as it imports RNAs, such as RNase P, 5S rRNA and mitochondrial RNA processing (MRP) RNAs, from the cytosol to mitochondria (18,19). hPNPase also interacts and forms a complex with SUV3 helicase located in the mitochondrial matrix to cooperatively degrade mitochondrial double-stranded RNA (20,21). Therefore, hPNPase has multiple functions in RNA degradation and translocation in the cytosol and mitochondria that are associated with its prominent role in regulating diverse physiological and pathological processes.

*To whom correspondence should be addressed. Tel: +886 2 27884151; Fax: +886 2 27826085; Email: hanna@sinica.edu.tw

All PNPase from different species share similar domain structures with one α-helical, one KH, one S1 and two RNase PH domains (Figure 1A) (22). The previously reported crystal structures of two bacterial PNPases show a trimeric architecture with six RNase PH domains assembled into a hexameric ring-like conformation containing a central channel for RNA binding and degradation (23–25). The S1 and KH domains, which are presumably responsible for RNA binding, are disordered in these structures with only partially visible backbones. It was noted that the architecture of these PNPases is similar to those of exosomes which are large protein complexes participating in RNA processing and degradation in archaea and eukaryotes (26–28). The archaeal and human exosome cores contain nine proteins with six RNase PH-like proteins that form a similar hexameric ring-like structure and three S1/KH domain-containing proteins that form an S1 pore associate on the top of the ring (29–31). Current models for RNA degradation by a PNPase propose that the RNA substrates are bound by the S1 domain and/or the KH domain and are further threaded into the central channel, within which the active site is situated in the second RNase PH domain (23,26). Previous mutational studies showed that the deletion of KH and S1 domains reduces the RNA binding and cleavage activities of bacterial PNPase (24,32,33). Nevertheless, the presence of the KH and S1 domains in hPNPase seems less critical for its activity in inducing the senescence phenotype (11).

To further characterize the molecular basis and define the domain function of PNPase for RNA binding and degradation, we analyzed the structure and biochemical properties of hPNPase. We have purified deletion mutants of hPNPase and determined the crystal structure of an S1 domain-truncated hPNPase (residues 46–669, see Figure 1A) at a resolution of 2.1 Å. We found that hPNPase assembled into a trimeric ring structure in which the three KH domains form an RNA-binding pore situated on the top of the structure. Our mutational and biochemical studies further confirmed that the KH domain in hPNPase, but not the S1 domain, is important for RNA binding and turnover. PNPase therefore has a unique arrangement with regard to RNA binding as compared to exosomes, their structural homologues. The pore formed by KH domain extends the length of the central RNA-binding channel in hPNPase, thus suggesting that only structured RNA with long 3′ tails can be bound and degraded in this channel. This study provides a solid structural model of hPNPase for in-depth understanding of this intriguing enzyme that contribute to RNA translocation and degradation in various cellular processes.

MATERIALS AND METHODS

Cloning, protein expression and purification

The genes of full-length hPNPase (residues 46–783) and ΔS1 hPNPase (residues 46–669) were amplified by PCR using Taq DNA polymerase (Stratagene) and subcloned into the NheI/SalI sites of expression vector pET28a (Novagen) to generate N-terminal His-tagged constructs. The G622D mutant of hPNPase was generated by using the QuickChange site-directed mutagenesis kit (Stratagene).

Single colonies of the *Escherichia coli* strain BL21-CodonPlus (DE3)-RIPL (Stratagene) transformed with full-length hPNPase or ΔS1 hPNPase plasmids were incubated at 37°C overnight in 10 ml LB medium supplemented with 35 μg/ml kanamycin, 35 μg/ml streptomycin and 25 μg/ml chloramphenicol. The cultures were

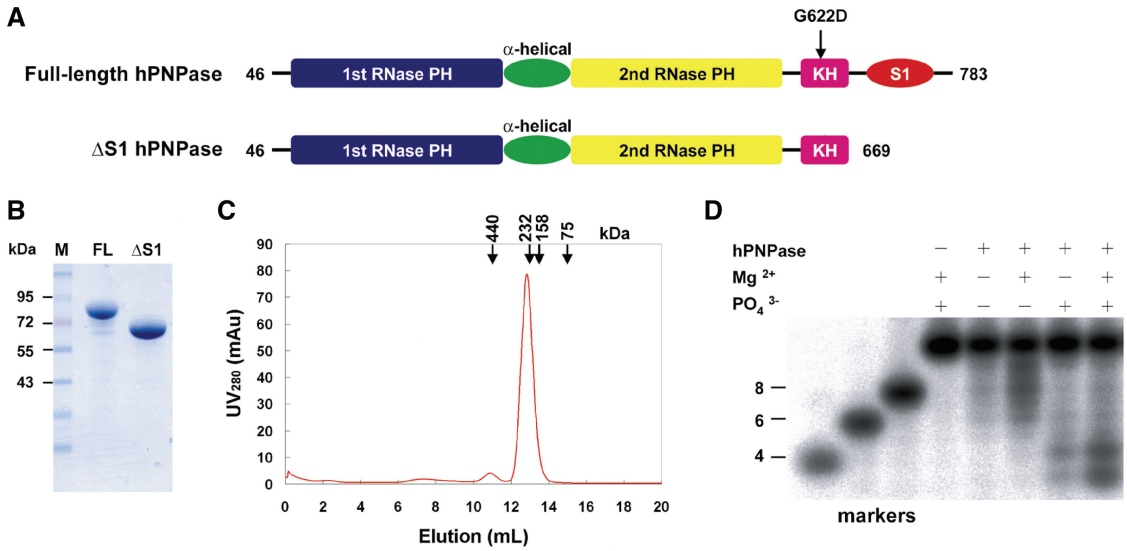

Figure 1. Recombinant human PNPase is a trimeric phosphorylase capable of digesting RNA. (**A**) Domain organization of full-length (FL) and S1 domain-truncated (ΔS1) hPNPase. (**B**) Purity of full-length and ΔS1 hPNPase, as analyzed by 10% SDS–PAGE. (**C**) The gel filtration (Superdex 200) profile of full-length hPNPase shows that the enzyme was eluted as a trimeric protein. Protein markers: ferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa) and conalbumin (75 kDa). (**D**) The full-length hPNPase (0.2 μM) was incubated for 30 min at 37°C with isotope-labeled 12-mer poly(A) ssRNA. In the presence of both magnesium and phosphate ions, hPNPase cleaved RNA most efficiently.

grown to an OD_{600} of 0.6 and then induced with 0.5 mM IPTG at 18°C for 20 h. The harvested cells were disrupted by a microfluidizer in buffer containing 50 mM Tris (pH 8.0), 300 mM NaCl and 10 mM imidazole. Crude cell extracts were first loaded onto a Ni-NTA resin affinity column (QIAGEN) followed by a HiTrap Heparin (GE Healthcare) and a Superdex 200 gel filtration column (GE Healthcare) in 50 mM Tris (pH 8.0) and 150 mM NaCl. Purified protein samples were concentrated to suitable concentrations and stored at −20°C until use.

Nuclease activity assays

The RNA substrate was first labeled at its 5′-end with [γ-^{32}P]ATP using T4 polynucleotide kinase and then purified on a Microspin G-25 column (GE Healthcare) to remove the free nucleotide. The RNA substrate (0.1 pmol) was then incubated for 30 min at 37°C with different concentrations of hPNPase (0.1–0.8 μM) in a reaction buffer containing 10 mM Tris (pH 8.0), 50 mM KCl, 1 mM DTT, 1 mM $MgCl_2$ and 0.5 mM NaH_2PO_4. After incubation, the reaction was stopped by adding 2× TBE–Urea sample buffer (Bio-Rad). Samples were then separated on 20% polyacrylamide/7 M urea gels, which were exposed to a phosphor imaging plate (Fujifilm) and analyzed by autoradiography (FLA-5000, Fujifilm).

RNA-binding assays

Electrophoretic mobility shift assays (EMSA) were performed by incubating 0.1 pmol of 5′-end ^{32}P-labeled ssRNA substrate with hPNPase at various concentrations ranging from 1 to 16 μM on ice for 30 min. The binding buffer contained 10 mM Tris (pH 8.0), 50 mM KCl, 1 mM DTT and 10 mM EDTA. After incubation, reaction products were resolved on 20% TBE gels (Invitrogen) and visualized by autoradiography (FLA-5000, Fujifilm).

Crystallization and crystal structure determination

ΔS1 hPNPase was concentrated to 10 mg/ml in buffer of 50 mM Tris (pH 8.0) and 150 mM NaCl. Crystals of ΔS1 hPNPase were grown by the hanging-drop vapor diffusion method at room temperature. The crystallization drop was made by mixing 1 μl of protein solution and 1 μl of reservoir solution containing 0.1 M citrate buffer (pH 5.0), 10% (v/v) 2-propanol and 26% (v/v) polyethylene glycol 400. X-ray diffraction data were collected at 100 K from beamline BL-13C1 at the National Synchrotron Radiation Research Center, Hsinchu, Taiwan, and processed and scaled by using the HKL2000 program. The structure was solved by the molecular replacement method using the crystal structure of *E. coli* PNPase (PDB entry: 3GCM) as the searching model in the program MOLREP of CCP4. The structure model was built using Coot and refined in Phenix. All diffraction and refinement statistics are listed in Table 1. Structural coordinates and diffraction structure factors of ΔS1 hPNPase have been deposited in the RCSB Protein Data Bank with the PDB ID code of 3U1K.

Table 1. X-ray data collection and refinement statistics for ΔS1 hPNPase

Data collection statistics	
Wavelength (Å)	1.0
Space group	R3
Cell dimensions (a, b, c) (Å)	289.8, 289.8, 92.8
Resolution (Å)	30–2.1 (2.2–2.1)*
Observed/unique reflections	520 863/162 293
Data redundancy	3.2 (3.1)
Completeness (%)	100 (100)
Rsym (%)	4.6 (31.1)
$I/\sigma(I)$	30.9 (4.7)
Refinement statistics	
Resolution range	30–2.1
Reflections (work/test)	162 282/8129
R-work/R-free (%)	18.0/22.3
Number of atoms (protein/water)	19 033/1153
Average B-factor (protein/solvent) (Å^2)	42.4/43.4
RMSD in bond length (Å)/bond angle (°)	0.007/1.071

*Values in parentheses refer to the highest resolution shell.

RESULTS

hPNPase is a trimeric exoribonuclease digesting ssRNA up to ∼4 nt

Full-length His-tagged hPNPase (residues 46–783) and an S1 domain-truncated (ΔS1) hPNPase (residues 46–669) without the N-terminal mitochondrial localization sequence (residues 1–45) were overexpressed in *E. coli* (Figure 1A). The recombinant proteins were purified by chromatographic methods using a Ni-NTA resin affinity column, followed by using a HiTrap Heparin and a Superdex 200 gel filtration column. The purified hPNPases had a high homogeneity, with only minute amounts of degraded fragments, as analyzed by 10% SDS–PAGE (Figure 1B). The gel filtration profile of the purified protein revealed one major peak with an estimated molecular weight of ∼240 kDa (Figure 1C). Because hPNPase has a theoretical molecular weight of 80.9 kDa, this result indicates a trimeric conformation of hPNPase.

To test whether the full-length recombinant hPNPase was functional, the protein was incubated with ^{32}P-labeled single-stranded RNA [poly(A)$_{12}$] in the presence or absence of phosphate and magnesium ions. The nuclease assays showed that hPNPase digested ssRNA most efficiently in the presence of both magnesium and phosphate ions and generated end products of 3–5 nt (Figure 1D), thereby confirming that the recombinant enzyme was a Mg^{2+}-dependent phosphorylase. Taken together, these results showed that the recombinant hPNPase was a functional trimeric phosphorylase capable of digesting single-stranded RNA to produce final products ∼4 nt in length.

Crystal structure of ΔS1 hPNPase

Both full-length and ΔS1 hPNPase were screened for crystallization conditions; however, only ΔS1 hPNPase yielded well-ordered crystals that diffracted X-rays to a resolution of 2.1 Å. The deletion mutant crystallized in

the rhombohedral space group R3 with four monomers per asymmetric unit. The structure was solved by molecular replacement using the crystal structure of *E. coli* PNPase (PDB entry: 3GCM) as the searching model. Two types of trimers were packed in the rhombohedral unit cell, one with a crystallographic 3-fold axis and one without. Because the two types of trimers had almost identical structures (average RMSD: 0.35 Å for 1672 Cα atoms), the trimer without the 3-fold crystallographic symmetry was used for structural analysis thereafter. The X-ray data collection and refinement statistics for ΔS1 hPNPase are listed in Table 1.

Each monomeric subunit of ΔS1 hPNPase in the crystal structure contained two visible RNase PH domains, one α-helical domain, one C-terminal KH domain and two citrate ions originating from the crystallization buffer (Figure 2A). In the trimeric ΔS1 hPNPase, the six RNase PH domains assembled into a ring-like structure with a central channel for RNA binding and cleavage (Figure 2B). The two citrates were bound in the RNA-binding channel near the active site in the second RNase PH domain of hPNPase, which displayed a geometry similar to those found in *E. coli* PNPase (25). The residues Arg446, Ser484, Asp538 and Asp544, which are critical for the phosphorylase activity of hPNPase as identified by previous site-directed mutagenesis studies (34), were located closely to the bound citrates, thus suggesting that the citrate ions were bound at the active site (see Figure 2D).

The three KH domains are located on the top of the ring forming a novel KH pore. The top view of the trimeric hPNPase shows clearly that the KH domain is associated with the first RNase PH domain of the adjacent monomer (Figure 2B). This result implies that the KH domain contributes to trimer formation and stabilization by interacting not only with the two other KH domains but also with the RNase PH domain in the neighboring molecule. In the crystal structure of *E. coli* PNPase, the C-terminal KH and S1 domains were not visible, whereas in the crystal structure of *Streptomyces antibioticus* PNPase, the KH and S1 domains were partially ordered and modeled as a poly-alanine chain. Superimposition of ΔS1 hPNPase with *S. antibioticus* PNPase (PDB entry 1E3P) and *E. coli* PNPase (PDB entry 3CDI) in the region of RNase PH domains gave an average RMSD of 1.38 Å for 374 Cα atoms and 1.16 Å for 382 Cα atoms, respectively (Figure 2C). This result shows that the N-terminal RNase PH domains of hPNPase have a similar overall structure to those of bacterial PNPase. However, the C-terminal KH domain of hPNPase only overlapped partially with the poly-alanine chain in *S. antibioticus* PNPase. We further constructed a structural model of the full-length hPNPase by superimposition of the structure of ΔS1 hPNPase (PDB entry 3U1K) with those of *S. antibioticus* PNPase (PDB entry 1E3P) and S1 domain of *E. coli* PNPase (PDB entry 1SRO) (35). The model shows that the S1 domain is located outside of the KH pore, suggesting that S1 domain may not be a part of the central pore (Figure 2E).

A KH pore is formed in hPNPase

Previous structural analysis revealed the presence of an S1 pore for RNA binding in human and archael exosomes (29–31). Three Rrp4 (or Csl4) subunits form the S1 pore bound on the top of the hexameric ring in archaeal exosomes, whereas three RNA-binding proteins, Rrp40, Rrp4 and Csl4, associate to an open S1 pore in human exosomes (see Figure 3). In contrast, in hPNPase, the three KH domains form a KH pore situated on the top of the hexameric ring-like structure. The KH pore extends the central channel formed by the RNase PH domains and therefore is likely involved in the binding of RNA substrates, which are further delivered to the active site located within the central channel.

The crystal structures of two proteins with a KH domain bound with RNA have been determined, neuro-oncological ventral antigen 2 (Nova2) KH3 (PDB entry 1EC6) and splicing factor 1 (SF1) KH (PDB entry 1K1G) (36,37). Superimposition of the KH domain of hPNPase with the Nova2 and SF1 KH–RNA complex structures gave an average RMSD of 1.45 Å for 48 Cα atoms and 3.12 Å for 30 Cα atoms, respectively (Figure 4A and B). This result suggests that the KH domain in hPNPase has a structure closely resembling those in Nova2 and SF1. Moreover, the RNA was bound in a similar manner in the KH domain on the edge of the α-helix and β-sheet in Nova2 and SF1. Therefore, a complex model of the hPNPase KH domain bound with an RNA molecule was built based on the structure of the Nova2–RNA complex. The hPNPase KH-RNA model had a 4-nt RNA bound in the RNA-binding cleft of the KH domain (Figure 4C). Superimposition of the bound RNA with the trimeric hPNPase further showed that the RNA substrate fitted right into the KH pore of the central channel. One of the three possible models is shown in Figure 4D and E. The sugar phosphate backbone of the 4-nt RNA in the KH pore was parallel to the channel direction, with the 3′-end pointing inward. This result implies that the RNA substrate threads through the KH pore to reach the central channel formed by the RNase PH domains for degradation.

The C-terminal S1 domain of hPNPase is not critical for RNA binding and cleavage

Previous structural analysis showed that the S1 domain in exosomes is involved in RNA binding. However, our crystal structure analysis suggests that the KH domain likely plays a dominant role for RNA binding in hPNPase through formation of a KH pore. To determine if the S1 domain in hPNPase is involved in RNA binding, full-length and ΔS1 hPNPase were incubated with $poly(A)_{12}$ and $poly(U)_{12}$ RNA in the absence of phosphate and magnesium ions. The electrophoretic mobility shift assay showed that the full-length and ΔS1 hPNPase bound $poly(A)_{12}$ and $poly(U)_{12}$ RNA with similar affinities. Moreover, ΔS1 hPNPase bound RNA almost as tightly as full-length hPNPase, thus suggesting that the S1 domain is not of critical importance for RNA binding (Figure 5A).

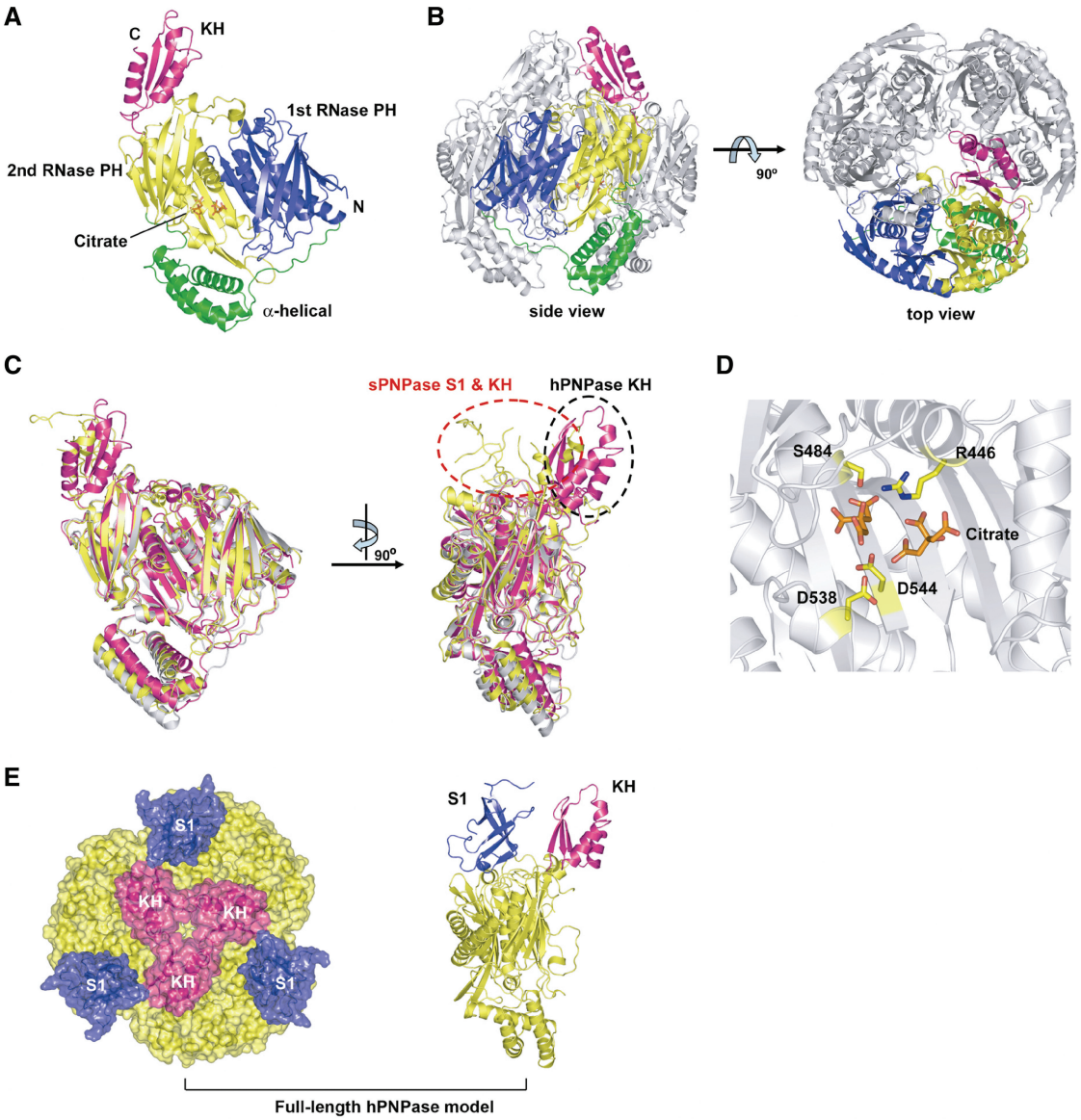

Figure 2. Crystal structure of ΔS1 hPNPase. (**A**) The monomeric structure of ΔS1 hPNPase comprises two RNase PH domains (blue and yellow), an α-helical domain (green) and a KH domain (magenta). Two citrate ions (orange) bound near the active site are displayed in stick models. (**B**) Side view and top view of the trimeric ΔS1 hPNPase which assembled into a ring-like structure with a central channel. The three KH domains are located on the top of the ring, forming a novel RNA-binding KH pore. (**C**) Superimposition of the structure of ΔS1 hPNPase (PDB entry 3U1K, magenta) with those of *S. antibioticus* PNPase (PDB entry 1E3P, yellow) and *E. coli* PNPase (PDB entry 3CDI, gray). The RNase PH and α-helical domains matched well; however, the KH domain in hPNPase (marked by a dashed black oval) does not superimpose with the poly-alanine chain in *S. antibioticus* PNPase (marked by a dashed red oval). (**D**) Two citrate ions (stick models in orange) bind at the active site in the second RNase PH domain in hPNPase. (**E**) Trimeric and monomeric full-length structural models of hPNPase generated by superimposition of the structure of ΔS1 hPNPase (PDB entry 3U1K) with those of *S. antibioticus* PNPase (PDB entry 1E3P) and S1 domain of *E. coli* PNPase (PDB entry 1SRO). Blue, magenta and yellow represent S1 domain, KH domain and RNase PH domain, respectively.

We next tested the nuclease activity of full-length and ΔS1 hPNPase by incubating poly(A)$_{12}$ and poly(U)$_{12}$ RNA with different concentrations of the recombinant hPNPase in the presence of 1 mM Mg^{2+} and 0.5 mM phosphate. Similar to the binding assays, full-length and ΔS1 hPNPase cleaved the poly(A)$_{12}$ and poly(U)$_{12}$ RNA with similar activities and ΔS1 hPNPase cleaved ssRNA substrate almost as efficiently as full-length PNPase (Figure 5B). Together with the binding assays, these results confirmed that the S1 domain in hPNPase is not critically important for RNA binding and cleavage. Moreover, this result is consistent with the earlier report showing that human PNPase, different from bacterial and plant chloroplast PNPase, does not bind and degrade polyadenylated RNA preferentially (34).

hPNPase digests the long 3′ tail of structured RNA

The KH pore is constricted and extends the length of the central RNA-binding channel, suggesting that hPNPase

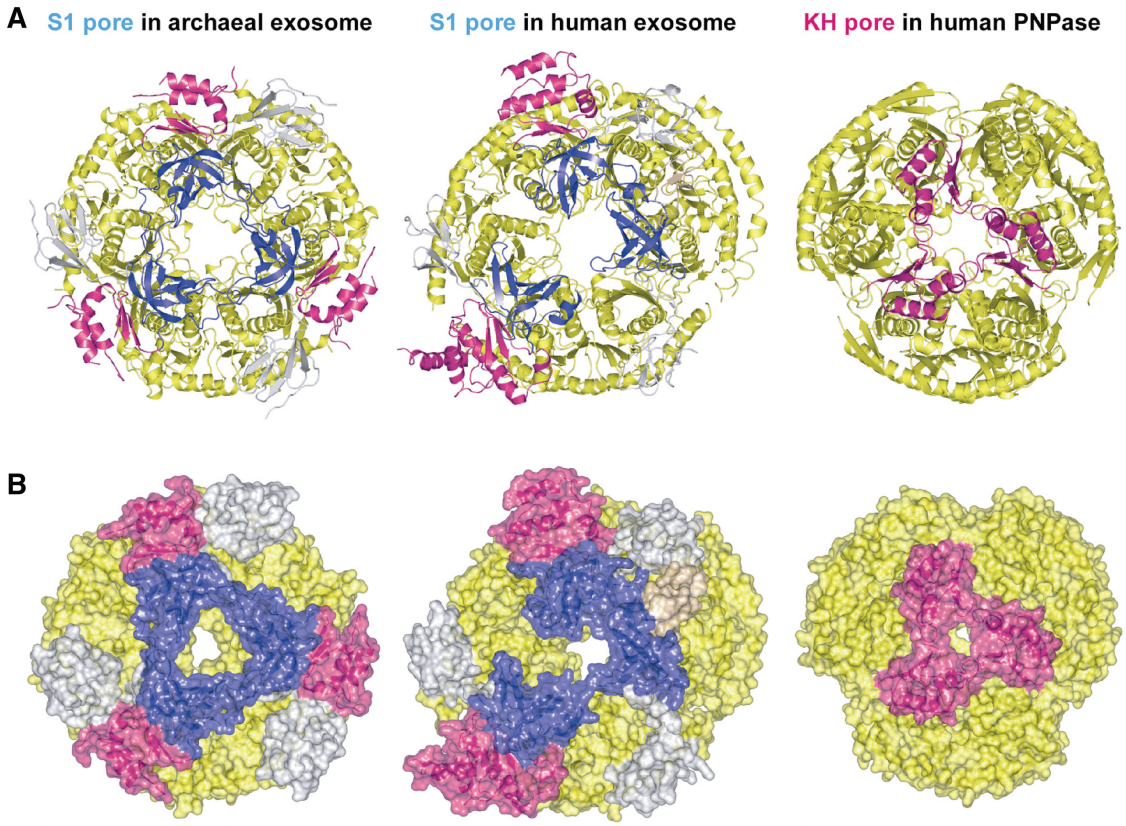

Figure 3. Comparison of the S1 pore in exosomes and the KH pore in hPNPase. (**A**) The archaeal exosome (PDB entry 2JE6) has an S1 pore formed by the three S1 domains of Rrp4 (blue). The human exosome (PDB entry 2NN6) has a more open S1 pore formed by the S1 domains of Rrp4, Rrp40 and Csl4 (blue). In contrast, human PNPase has a KH pore formed by three KH domains (magenta). (**B**) Molecular surface representation of the archaeal exosome, human exosome and human PNPase structures show clearly differences in the arrangement of the domains that form the S1 pore and KH pore in exosomes and hPNPase, respectively. The color code is the same as in panel A: S1 domain in blue, KH domain in magenta, RNase PH domain in yellow, NTD domain (in Rrp40, Rrp4 and Csl4) in gray and CTD domain (in Csl4) in beige.

likely degrades only the long 3′ tail of structured RNA. To test the substrate specificity of hPNPase, stem–loop RNA with different 3′ overhangs ranging from 7 to 20 nt were incubated with the enzyme. The duplex RNA with a long 3′ overhang of 15 and 20 nt were digested by the full-length and ΔS1 hPNPase (see Figure 5C). It can be seen clearly that ΔS1 hPNPase generated the major end products with a 3′ overhang of 11–14 nt (Figure 5C). Conversely, the stem–loop RNA with a short 3′ overhang of 7 and 10 nt were more resistant for digestion. This result shows that hPNPase can digest the long 3′ overhang of a structured RNA up to ~12 nt.

The KH pore is responsible for RNA binding in hPNPase

A common feature of the KH domain is a GXXG motif which is located between two α-helices and responsible for nucleic acid binding (38). Structure-based sequence alignment of the KH domain of PNPases showed that the GXXG motif is highly conserved in PNPase from different species (Figure 6A). However, the GXXG motif is not conserved in Rrp40 and Rrp4 exosome subunits; instead, they contain another motif, GXNG, with unknown function located between two β-sheets (Figure 6B). The GXXG motif of the KH domain is involved in RNA binding in several proteins, such as NusA and ERA (39,40). In hPNPase, this GXXG motif is located in a β-turn region interacting directly with the bound RNA substrate in the hPNPase–RNA complex model (Figure 4C). We therefore hypothesized that the GXXG motif located inside the KH pore of PNPase is involved in RNA binding.

To test this hypothesis, site-directed mutagenesis was used to mutate the first glycine residue to aspartate (G622D) in the putative RNA-binding GXXG motif in hPNPase. The first glycine in the GXXG motif has special backbone dihedral angles ($\phi = 74.3^\circ$ and $\psi = 170.5^\circ$) that are not within the energy-allowed regions for non-Gly residues. The replacement of Gly622 to Asp may disrupt the β-turn conformation and therefore interfere with the hydrophobic interactions between hPNPase and RNA. The gel shift assays showed that the RNA binding affinity of wild-type full-length hPNPase was almost completely abolished in the G622D mutant (Figure 6C). Moreover, the G622D mutant lost most of its RNA degradation activity as compared to full-length hPNPase (Figure 6D). Taken together, these results confirmed the hypothesis that the KH pore in hPNPase is directly involved in RNA binding for RNA degradation.

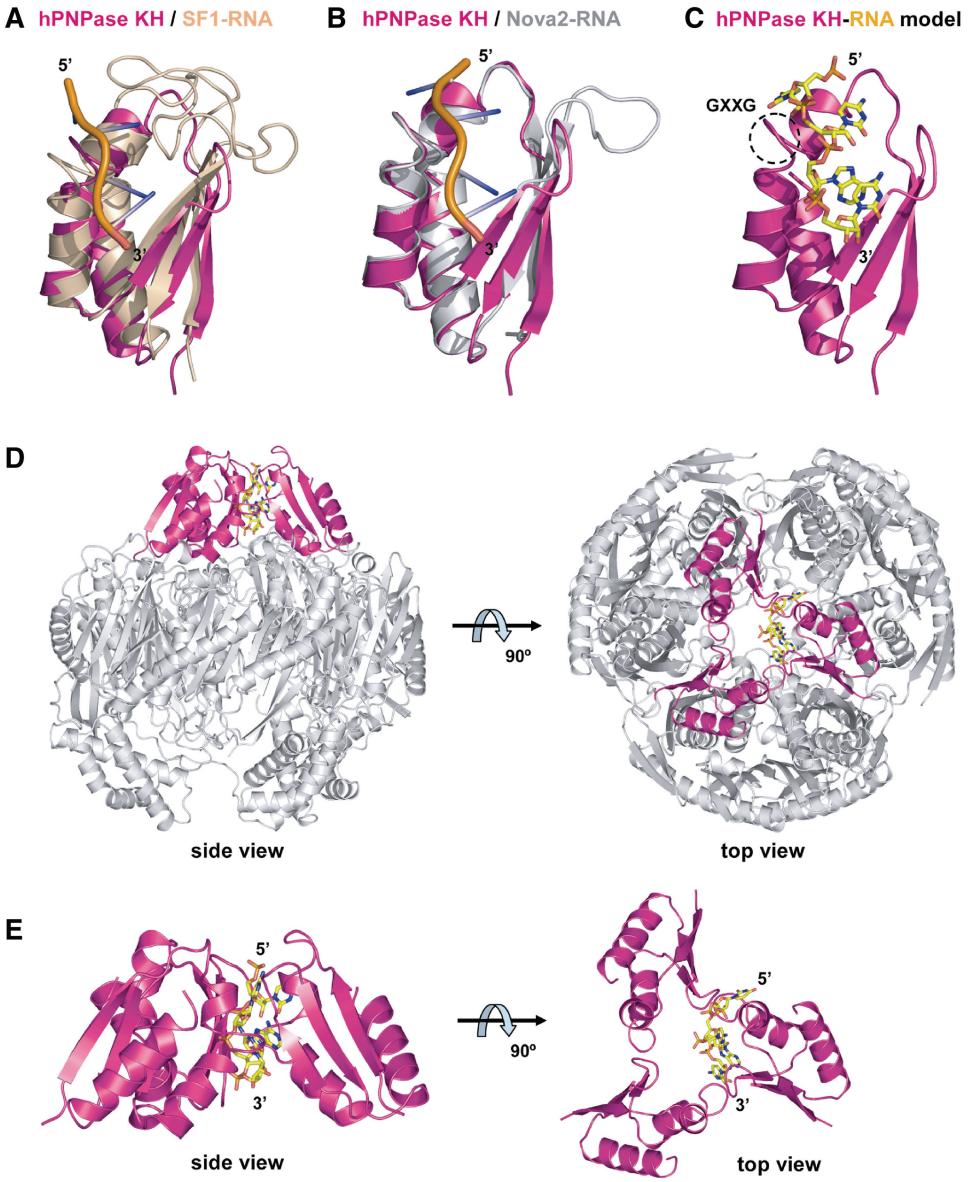

Figure 4. Structural model of RNA bound at the KH pore of hPNPase. (**A**) Superimposition of the hPNPase KH domain (magenta) with the KH domain in the SF1–RNA complex structure (PDB entry 1K1G, beige for SF1 and orange for RNA) gave an average RMSD of 3.12 Å for 30 Cα atoms. (**B**) Superimposition of the hPNPase KH domain (magenta) with the KH domain in the Nova2–RNA complex structure (PDB entry 1EC6, gray for Nova2 and orange for RNA) gave an average RMSD of 1.45 Å for 48 Cα atoms. (**C**) Structural model of the KH domain (magenta) of hPNPase in complex with a single-stranded, 4-nt RNA adapted from Nova2-RNA. The GXXG motif is marked by a dashed circle. (**D**) Structural model of the hPNPase–RNA complex. The 4-nt RNA is bound inside the KH pore. (**E**) A close-up view of the trimeric KH pore in hPNPase bound with RNA. The 3′-end of the RNA points inward into the channel.

DISCUSSION

KH pore in PNPase versus S1 pore in exosomes

Here, we report the first crystal structure of a PNPase with a well-ordered KH domain. The crystal structure of hPNPase shows a typical hexameric RNase PH ring capped with a novel trimeric KH pore. Our mutational and biochemical studies further confirmed that the GXXG motif within the KH domain participates directly in RNA binding. The KH domain was first described in human heterogeneous nuclear ribonucleoprotein K (hnRNP K), which contains three copies of the KH domain (41). Subsequently, the KH domain was found in various RNA-binding proteins, such as fragile X mental retardation protein (FMRP), splicing factor 1 (SF1), neuro-oncological ventral antigen 2 (Nova2), and NusA transcription elongation factor (36–39,42). Similar to the KH domain, the S1 domain is also considered to be an RNA-binding domain, with a five-stranded β-barrel structure originally identified in ribosomal protein S1 (43). Many RNA-binding proteins contain an S1 domain, such as RNase II, RNase E, RNase R and NusA transcription elongation factor (44–46).

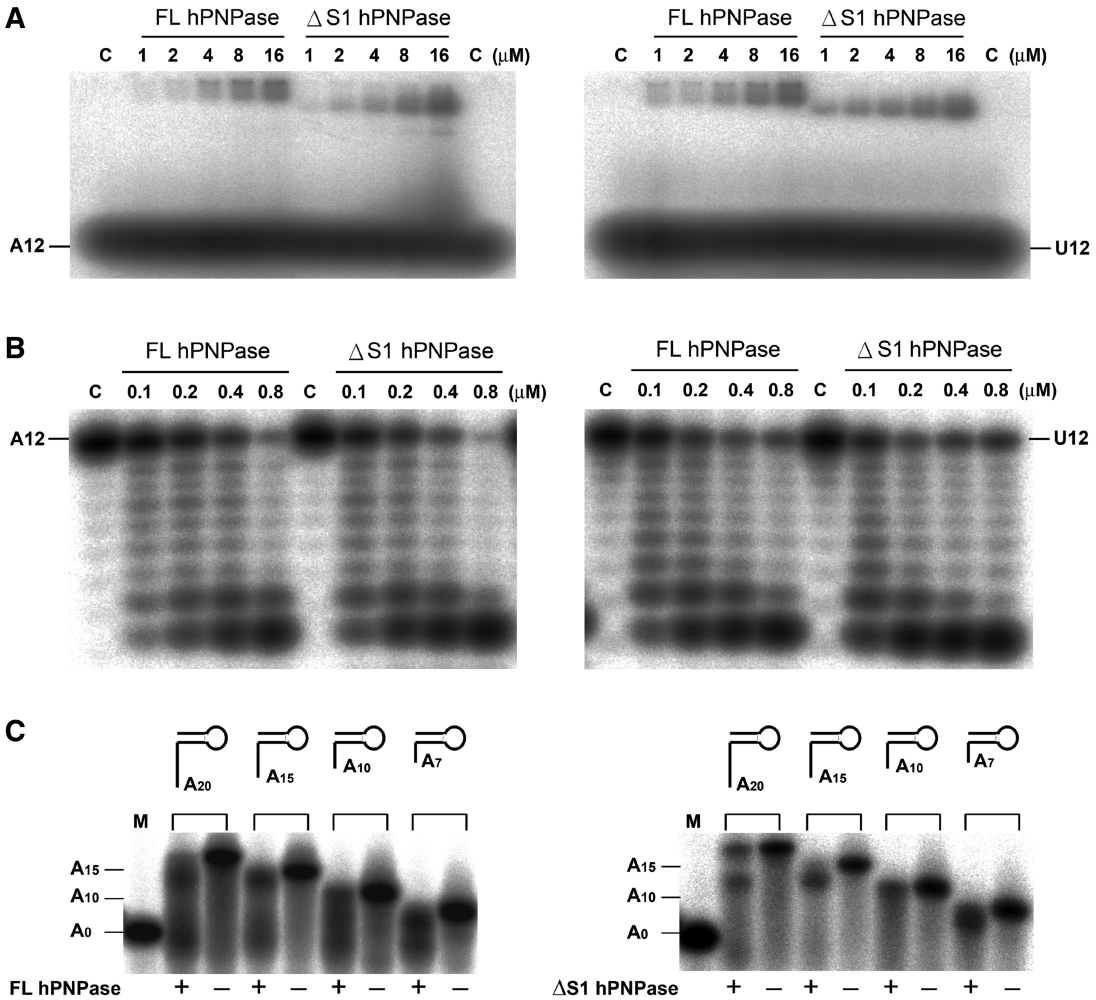

Figure 5. RNA binding and cleavage assays for full-length (FL) and ΔS1 hPNPase. (**A**) EMSA assay showing that the RNA substrate binds with a similar affinity to full-length and ΔS1 hPNPase. The 12-mer poly(A) and poly(U) ssRNA substrate (0.1 pmol) was incubated, respectively, with hPNPase at various concentrations ranging from 1 to 16 μM in the absence of phosphate and Mg^{2+} ions. (**B**) The RNase activity of full-length and ΔS1 hPNPase was examined by incubation of the ssRNA with different concentrations of protein (0.1–0.8 μM) for 30 min at 37°C. Full-length and ΔS1 hPNPase exhibited a similar RNase activity. (**C**) Full-length and ΔS1 hPNPase can digest the long 3′ overhang (15 and 20 nt) of stem–loop RNA and generate major products with an overhang of 11–14 nt. The stem–loop RNA with a short 3′ overhang (7 and 10 nt) were more resistant for digestion. The marker (M) of 20-nt RNA corresponds to the stem–loop region of RNA with an 8-bp duplex and a 4-nt loop. The A_0, A_{10} and A_{15} indicate the overhang size of 0, 10 and 15 nt, respectively.

Why does hPNPase have a KH pore for nucleic acid binding rather than an S1 pore as that found in exosomes? First, the structure-based sequence alignment of the PNPase KH domain shows that the GXXG motif is evolutionarily conserved in PNPase. On the contrary, the GXXG motif is not conserved in exosome subunits, such as Rrp40 and Rrp4. This observation indicates that the KH domain in exosome subunits probably does not participate in RNA binding; accordingly, the GXXG sequence appears to have degenerated during evolution. Second, the human exosome contains six RNase PH domain-like proteins, Rrp41, Rrp46, Mtr3, Rrp42, Rrp43 and Rrp45, which form a pseudo-hexameric ring structure with three S1/KH RNA-binding proteins, Rrp40, Rrp4 and Csl4, associate on the top of the ring. Both Rrp40 and Rrp4 have a S1 domain and a KH domain, whereas Csl4 has only a S1 domain. Similarly, in archaeal exosomes, three Rrp41 and three Rrp42 form a hexameric ring structure capped with three Rrp4 or Csl4. The RNA-binding protein Rrp4 has an S1 domain and a KH domain, whereas Csl4 has only an S1 domain. Therefore, not all eukaryotic exosome subunits have a KH domain capable of forming a KH pore for nucleic acid binding. Hence, exosomes have evolved to use an S1 pore, whereas PNPase has evolved to use a KH pore to trap RNA for degradation.

A constricted channel in hPNPase for efficient RNA binding

In PNPase and exosomes, RNA is bound in the central channel of the ring-like structures. It is intriguing to ponder what kind of channel might be more efficient for RNA binding. A previous study suggests that the KH/S1 domains are not only involved in RNA binding but also in

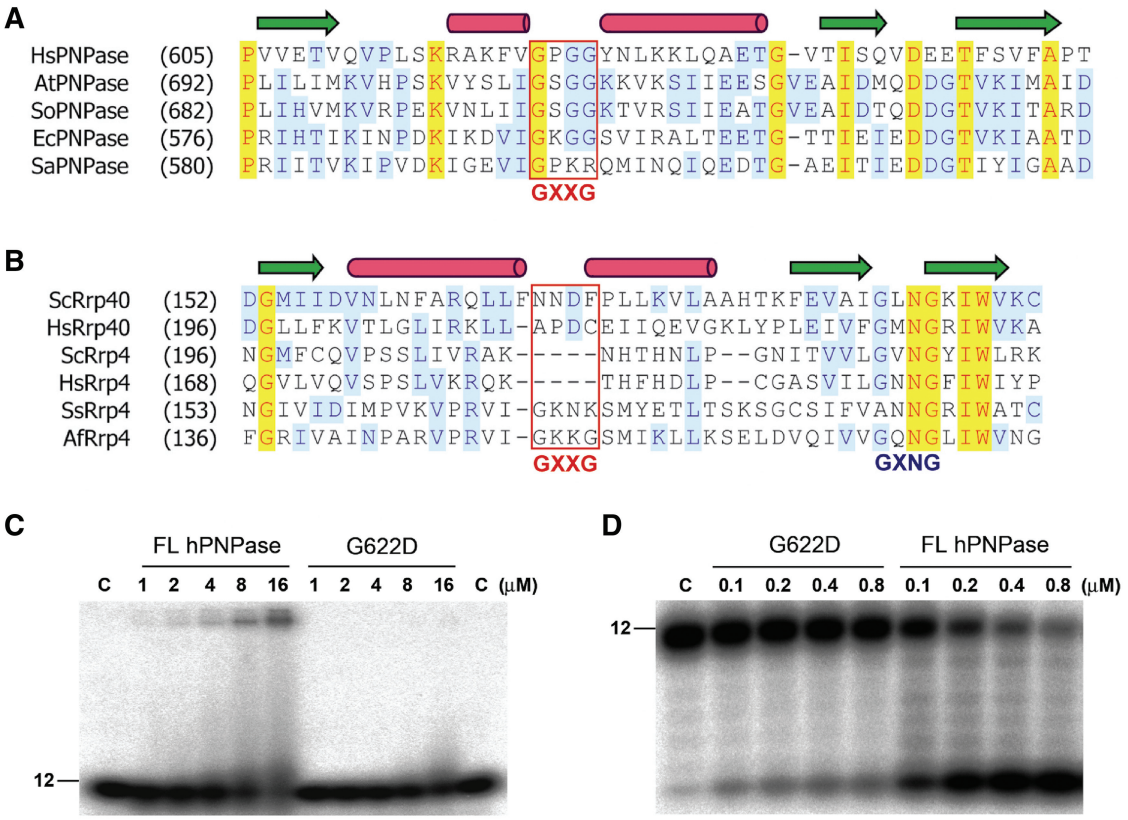

Figure 6. The GXXG motif in the KH domain of PNPase participates in RNA binding. (**A**) Sequence alignment of the KH domain of PNPase from different species shows that the GXXG motif is highly conserved. Secondary structure elements are shown on the top, with arrows indicating β-sheets and cylinders indicating α-helices. The PNPase-specific GXXG sequences are marked with a box. Sequences listed here include PNPase from *Homo sapiens* (Hs), *Arabidopsis thaliana* (At), *Spinacia oleracea* (So), *E. coli* (Ec) and *S. antibioticus* (Sa). (**B**) Sequence alignment of the KH domain in Rrp40 and Rrp4 exosome components shows that the GXXG motif was not conserved. Sequences listed here include components from the following species: *Saccharomyces cerevisiae* (Sc), *H. sapiens* (Hs), *S. solfataricus* (Ss) and *Archaeoglobus fulgidus* (Af). (**C**) EMSA assay showing that the hPNPase mutant G622D lost its RNA-binding activity in binding a 12-mer poly(A) ssRNA. (**D**) RNase activity assay showing that the hPNPase mutant G622D digested the 12-mer poly(A) RNA poorly as compared to wild-type PNPase.

the formation of a compact trimer with a constricted central channel (24). For example, the *Sulfolobus solfataricus* exosome Rrp41/Rrp42 core complex has a wider central channel than the holoenzyme Rrp4/Rrp41/Rrp42. The central channel of KH/S1 domain-truncated *E. coli* PNPase is also larger than that of full-length protein. Comparing the central channel of three PNPase structures, we found that *E. coli* PNPase has the widest central channel, likely because the KH/S1 domains are not present, while *S. antibioticus* PNPase has a more constricted central channel because it has a partially ordered KH/S1 RNA-binding domain (Figure 7A). Remarkably, hPNPase has an even more constricted central channel with two neck regions, one within the KH domain and one in a region on top of the RNase PH ring (displayed in red and pink in Figure 7A). We also noticed that the inactive human exosome has a wide channel, whereas the active *S. solfataricus* exosome has a narrow channel with a constricted neck region (Figure 7B). This result suggests that the RNA-binding KH domain in PNPase regulates the size of the channel and therefore in turn regulates the RNA-binding activity as well as the RNA cleavage activity of PNPase. A more constricted central channel appears more efficient for RNA binding and cleavage.

RNA degradation or translocation by hPNPase

It is intriguing why some RNAs, such as RNase P, 5S rRNA and MRP RNA, are bound and translocated by hPNPase, whereas some other RNAs, such as *c-myc* mRNA, are digested by hPNPase (15,16,18,19). How does hPNPase distinguish different RNAs and select some for degradation and others for transportation? The length of the central channel of hPNPase from the KH domain to the active site is ~60 Å, corresponding roughly to a length of 8 nt (Figure 7C). A previous crystal structure of archaeal exosome in complex with RNA shows that the active site cavity can harbor 4 nt (48,49). Therefore, a 3′ overhang of a duplex RNA is likely to be digested by hPNPase up to 12 nt. Indeed, this result is consistent with our biochemical data which reveal that hPNPase digests only the long 3′ overhang of a duplex RNA up to ~12 nt. The narrowest region of the central channel of hPNPase is only ~4 Å in diameter, not wide enough for accommodating single-stranded RNA, hence structural conformational changes are expected upon RNA binding. We suggest here that a structured RNA with a long 3′ tail of >12 nt is first bound to the KH pore and that binding induces a conformational change in the narrow neck regions. The RNA is then further threaded

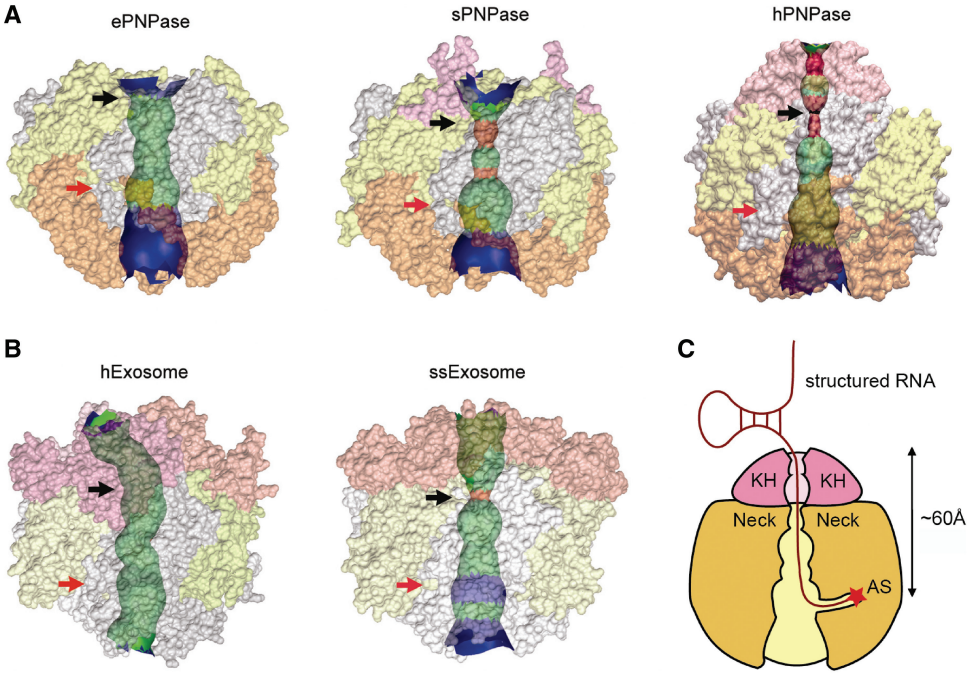

Figure 7. The RNA-binding central channel in PNPase and exosomes. (**A**) Comparison between the central channel in *E. coli* PNPase (ePNPase, PDB: 3CDI), *S. antibioticus* PNPase (sPNPase, PDB: 1E3P) and hPNPase (PDB: 3U1K) shows that hPNPase has the most constricted channel for RNA binding among these PNPases [calculated by using the program HOLE (47)]. The channel regions with a diameter of <5 Å are displayed in red and pink. The neck regions are marked by black arrows and the active sites are marked by red arrows. (**B**) The crystal structures of human exosome (PDB: 2NN6) and *S. solfataricus* exosome (PDB: 2JE6) show that the active *S. solfataricus* exosome has a more constricted channel as compared to the inactive human exosome. (**C**) Structural model of hPNPase bound with a structured RNA.

into the channel to reach the active site in the second RNase PH domain for degradation (Figure 7C).

Interestingly, the RNAs transported by hPNPase either have short or no 3′ overhangs: 1 nt in RNase P RNA (50), 0 nt in 5S rRNA (51) and 2 nt in MRP RNA (52). Because their 3′ overhangs are too short for being threaded into the RNA binding pore and cleavage channel, these RNAs with folded secondary structures are transported but not digested by hPNPase. How hPNPase binds to these structural RNA remains unclear. Further structural studies on PNPase–RNA and PNPase–helicase complexes may reveal detailed information regarding the interactions of PNPase with RNA and its partner proteins during RNA transportation, processing and degradation.

ACCESSION NUMBER

3U1K.

FUNDING

Academia Sinica and the National Science Council, Taiwan, R.O.C and National Synchrotron Radiation Research Center (BL-13B1 and BL-13C1, partial), a national user facility supported by the National Science Council of Taiwan, ROC. Funding for open access charge: Academia Sinica, Taiwan.

Conflict of interest statement. None declared.

REFERENCES

1. Mohanty,B.K. and Kushner,S.R. (2000) Polynucleotide phosphorylase functions both as a 3′→5′ exonuclease and a poly(A) polymerase in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **97**, 11966–11971.
2. Deutscher,M.P. (2006) Degradation of RNA in bacteria: comparison of mRNA and stable RNA. *Nucleic Acids Res.*, **34**, 659–666.
3. Andrade,J.M., Pobre,V., Silva,I.J., Domingues,S. and Arraiano,C.M. (2009) The role of 3′–5′ exoribonucleases in RNA degradation. *Prog. Mol. Biol. Trans. Sci.*, **85**, 187–229.
4. Das,S.K., Bhutia,S.K., Stokhi,U.K., Dash,R., Azab,B., Sarkar,D. and Fisher,P.B. (2011) Human polynucleotide phosphorylase (*hPNPase*$^{old\text{-}35}$): an evolutionary conserved gene with an expanding repertoire of RNA degradation functions. *Oncogene*, **30**, 1733–1743.
5. Marcaida,M.J., DePristo,M.A., Chandran,V., Carpousis,A.J. and Luisi,B.F. (2006) The RNA degradosome: life in the fast lane of adaptive molecular evolution. *Trends Biochem. Sci.*, **31**, 359–365.
6. Carpousis,A.J. (2007) The RNA degradosome of *Escherichia coli*: an mRNA-degrading machine assembled on RNase E. *Annu. Rev. Microbiol.*, **61**, 71–87.
7. Carpousis,A.J., Luisi,B.F. and McDowall,K.J. (2009) Endonucleolytic initiation of mRNA decay in *Escherichia coli*. *Prog. Mol. Biol. Trans. Sci.*, **85**, 91–135.
8. Schuster,G. and Stern,D. (2009) RNA polyadenylation and decay in mitochondria and chloroplasts. *Prog. Mol. Biol. Trans. Sci.*, **85**, 393–422.
9. Chen,H.-W., Koehler,C.M. and Teitell,M.A. (2007) Human polynucleotide phosphorylase: location matters. *Trends Cell Biol.*, **17**, 600–608.
10. Leszczyniecka,M., Kang,D.C., Sarkar,D., Su,Z.Z., Holmes,M., Valerie,K. and Fisher,P.B. (2002) Identification and cloning of human polynucleotide phosphorylase, *hPNPase*$^{old\text{-}35}$, in the

context of terminal differentiation and cellular senescence. *Proc. Natl Acad. Sci. USA*, **99**, 16636–16641.
11. Sarkar,D., Park,E.S., Emdad,L., Randolph,A., Valerie,K. and Fisher,P.B. (2005) Defining the domains of human polynucleotide phosphorylase (hPNPase^{OLD-35}) mediating cellular senescence. *Mol. Cell. Biol.*, **25**, 7333–7343.
12. Sarkar,D., Leszczyniecka,M., Kang,D.C., Lebedeva,I.V., Valerie,K., Dhar,S., Pandita,T.K. and Fisher,P.B. (2003) Down-regulation of myc as a potential target for growth arrest induced by human polynucleotide phosphorylase (*hPNPase^{old-35}*) in human melanoma cells. *J. Biol. Chem.*, **278**, 24542–24551.
13. Sarkar,D., Park,E.S. and Fisher,P.B. (2006) Defining the mechanism by which IFN-β dowregulates c-myc expression in human melanoma cells: pivotal role for human polynucleotide phosphorylase (*hPNPase^{old-35}*). *Cell Death Differ.*, **13**, 1541–1553.
14. Das,S.K., Sokhi,U.K., Bhutia,S.K., Azab,B., Su,Z.Z., Sarkar,D. and Fisher,P.B. (2010) Human polynucleotide phosphorylase selectively and preferentially degrades microRNA-221 in human melanoma cells. *Proc. Natl Acad. Sci. USA*, **107**, 11948–11953.
15. Sarkar,D., Park,E.S., Barber,G.N. and Fisher,P.B. (2007) Activation of double-stranded RNA-dependent protein kinase, a new pathway by which human polynucleotide phosphorylase (*hPNPase^{old-35}*) induces apoptosis. *Cancer Res*, **67**, 7948–7953.
16. Sarkar,D. and Fisher,P.B. (2006) Molecular mechanisms of aging-associated inflammation. *Cancer Lett.*, **236**, 13–23.
17. Slomovic,S. and Schuster,G. (2008) Stable PNPase RNAi silencing: its effect on the processing and adenylation of human mitochondrial RNA. *RNA*, **14**, 310–323.
18. Chen,H.-W., Rainey,R.N., Balatoni,C.E., Dawson,D.W., Troke,J.J., Wasiak,S., Hong,J.S., McBride,H.M., Koehler,C.M., Teitell,M.A. *et al.* (2006) Mammalian polynucleotide phosphorylase is an intermembrane space RNase that maintains mitrochondrial homeostasis. *Mol. Cell. Biol.*, **26**, 8475–8487.
19. Wang,G., Chen,H.W., Oktay,Y., Zhang,J., Allen,E.L., Smith,G.M., Fan,K.C., Hong,J.S., French,S.W., McCaffery,J.M. *et al.* (2010) PNPASE regulates RNA import into mitochondria. *Cell*, **142**, 456–467.
20. Wang,D.D.-H., Shu,Z., Lieser,S.A., Chen,P.-L. and Lee,W.-H. (2009) Human mitochondrial SUV3 and polynucleotide phosphorylase form a 330-kDa heteropentamer to cooperatively degrade double-stranded RNA with a 3′-to-5′directionality. *J. Biol. Chem.*, **284**, 20812–20821.
21. Minczuk,M., Piwowarski,J., Papworth,M.A., Awiszus,K., Schalinski,S., Dziembowski,A., Dmochowska,A., Bartnik,E., Tokatlidis,K., Stepien,P.P. *et al.* (2002) Localisation of the human hSuv3p helicase in the mitochondrial matrix and its preferential unwinding of dsDNA. *Nucleic Acids Res.*, **30**, 5074–5086.
22. Bermudez-Cruz,R.M., Ramirez,F., Kameyama-Kawabe,L. and Montanez,C. (2005) Conserved domains in polynucleotide phosphorylase among eubacteria. *Biochimie*, **87**, 737–745.
23. Symmons,M.F., Jones,G.H. and Luisi,B.F. (2000) A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure*, **8**, 1215–1226.
24. Shi,Z., Yang,W.-Z., Lin-Chao,S., Chak,K.-F. and Yuan,H.S. (2008) Crystal structure of *Escherichia coli* PNPase: central channel residues are involved in processive RNA degradation. *RNA*, **14**, 2361–2371.
25. Nurmohamed,S., Vaidialingam,B., Callaghan,A.J. and Luisi,B.F. (2009) Crystal structure of *Escherichia coli* polynucleotide phosphorylase core bound to RNase E, RNA and manganese: implications for catalytic mechanism and RNA degradosome assembly. *J. Mol. Biol.*, **389**, 17–33.
26. Januszyk,K. and Lima,C.D. (2010) Structural components and architectures of RNA exosomes. *Adv. Exp. Med. Biol.*, **702**, 9–28.
27. Lin-Chao,S., Chiou,N.-T. and Schuster,G. (2007) The PNPase, exosome and RNA helicases as the building components of evolutionarily-conserved RNA degradation machines. *J. Biomed. Sci.*, **14**, 523–532.
28. Houseley,J., LaCava,J. and Tollervey,D. (2006) RNA-quality control by the exosome. *Nat. Rev. Mol. Cell. Biol.*, **7**, 529–539.
29. Liu,Q., Greimann,J.C. and Lima,C.D. (2006) Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell*, **127**, 1223–1237.
30. Buttner,K., Wenig,K. and Hopfner,K.-P. (2005) Structural framework for the mechanism of archaeal exosomes in RNA processing. *Mol. Cell*, **20**, 461–471.
31. Lorentzen,E., Dziembowski,A., Lindner,D., Seraphin,B. and Conti,E. (2007) RNA channelling by the archaeal exosome. *EMBO Rep.*, **8**, 470–476.
32. Jarrige,A.-C., Brechemier-Baey,D., Mathy,N., Cuche,O. and Portier,C. (2002) Mutational analysis of polynucleotide phosphorylase form *Escherichia coli*. *J. Mol. Biol.*, **321**, 397–409.
33. Stickney,L.M., Hankins,J.S., Miao,S. and Mackie,G.A. (2005) Function of the conserved S1 and KH domains in polynucelotide phosphorylase. *J. Bacteriol.*, **187**, 7214–7221.
34. Portnoy,V., Palnizky,G., Yehudai-Resheff,S., Glaser,F. and Schuster,G. (2008) Analysis of the human polynucleotide phosphorylase (PNPase) reveals differences in RNA binding and response to phosphate compared to its bacterial and chloroplast counterparts. *RNA*, **14**, 297–309.
35. Bycroft,M., Hubbard,T.J.P., Proctor,M., Freund,S.M.V. and Murzin,A.G. (1997) The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell*, **88**, 235–242.
36. Lewis,H.A., Musunuru,K., Jensen,K.B., Edo,C., Chen,H., Darnell,R.B. and Burley,S.K. (2000) Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. *Cell*, **100**, 323–332.
37. Liu,Z.H., Luyten,I., Bottomley,M.J., Messias,A.C., Houngninou-Molango,S., Sprangers,R., Zanier,K., Kramer,A. and Sattler,M. (2001) Structural basis for recognition of the intron branch site RNA by splicing factor 1. *Science*, **294**, 1098–1102.
38. Regan,L., Valverde,R. and Edwards,L. (2008) Structure and function of KH domains. *FEBS J.*, **275**, 2712–2726.
39. Beuth,B., Pennell,S., Arnvig,K.B., Martin,S.R. and Taylor,I.A. (2005) Structure of a *Mycobacterium tuberculosis* NusA-RNA complex. *EMBO J.*, **24**, 3576–3587.
40. Tu,C., Zhou,X.M., Tropea,J.E., Austin,B.P., Waugh,D.S., Court,D.L. and Ji,X.H. (2009) Structure of ERA in complex with the 3′ end of 16S rRNA: implications for ribosome biogenesis. *Proc. Natl Acad. Sci. USA*, **106**, 14843–14848.
41. Siomi,H., Matunis,M.J., Michael,W.M. and Dreyfuss,G. (1993) The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. *Nucleic Acids Res.*, **21**, 1193–1198.
42. Valverde,R., Pozdnyakova,I., Kajander,T., Venkatrarnan,J. and Regan,L. (2007) Fragile X mental retardation syndrome: structure of the KH1-KH2 domains of fragile X mental retardation protein. *Structure*, **15**, 1090–1098.
43. Bycroft,M., Hubbard,T.J.P., Proctor,M., Freund,S.M.V. and Murzin,A.G. (1997) The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell*, **88**, 235–242.
44. Frazao,C., Mcvey,C.E., Amblar,M., Barbas,A., Vonrhein,C., Arraiano,C.M. and Carrondo,M.A. (2006) Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature*, **443**, 110–114.
45. Schubert,M., Edge,R.E., Lario,P., Cook,M.A., Strynadka,N.C.J., Mackie,G.A. and McIntosh,L.P. (2004) Structural characterization of the RNase E S1 domain and identification of its oligonucleotide-binding and dimerization interfaces. *J. Mol. Biol.*, **341**, 37–54.
46. Gopal,B., Haire,L.F., Gamblin,S.J., Dodson,E.J., Lane,A.N., Papavinasasundaram,K.G., Colston,M.J. and Dodson,G. (2001) Crystal structure of the transcription elongation/antitermination factor NusA from *Mycobacterium tuberculosis* at 1.7 angstrom resolution. *J. Mol. Biol.*, **314**, 1087–1095.
47. Smart,O.S., Goodfellow,J.M. and Wallace,B.A. (1993) The pore dimensions of gramicidin A. *Biophy. J.*, **65**, 2455–2460.
48. Lorentzen,E. and Conti,E. (2005) Structural basis of 3′ end RNA recognition and exoribonucleolytic cleavage by an exosome RNase PH domain. *Mol. Cell*, **20**, 473–481.

49. Navarro,M.V.A.S., Oliveira,C.C., Zanchin,N.I.T. and Guimara,B.G. (2008) Insights into the mechanism of progressive RNA degradation by the archaeal exosome. *J. Biol. Chem.*, **283**, 14120–14131.
50. Kikovska,E., Svard,S.G. and Kirsebom,L.A. (2007) Eukaryotic RNase P RNA mediates cleavage in the absence of protein. *Proc. Natl Acad. Sci. USA*, **104**, 2062–2067.
51. Luehrsen,K.R. and Fox,G.E. (1981) Secondary structure of eukaryotic cytoplasmic 5S ribosomal RNA. *Proc. Natl Acad. Sci. USA*, **78**, 2150–2154.
52. Schmitt,M.E., Bennett,J.L., Dairaghi,D.J. and Clayton,D.A. (1993) Secondary structure of RNase MRP as predicted by phylogenetic comparison. *FASEB J.*, **7**, 208–213.