Structure and mechanism of *Mycobacterium smegmatis* polynucleotide phosphorylase

MIHAELA-CARMEN UNCIULEAC,1 SHREYA GHOSH,1 M. JASON DE LA CRUZ, YEHUDA GOLDGUR, and STEWART SHUMAN

Molecular Biology and Structural Biology Programs, Sloan-Kettering Institute, New York, New York 10065, USA

ABSTRACT

Polynucleotide phosphorylase (PNPase) catalyzes stepwise phosphorolysis of the 3′-terminal phosphodiesters of RNA chains to yield nucleoside diphosphate products. In the reverse reaction, PNPase acts as a polymerase, using NDPs as substrates to add NMPs to the 3′-OH terminus of RNA chains while expelling inorganic phosphate. The apparent essentiality of PNPase for growth of *M. tuberculosis* militates for mycobacterial PNPase as a potential drug target. A cryo-EM structure of *Mycobacterium smegmatis* PNPase (MsmPNPase) reveals a characteristic ring-shaped homotrimer in which each protomer consists of two RNase PH-like domains and an intervening α-helical module on the inferior surface of the ring. The carboxy-terminal KH and S1 domains, which impart RNA specificity to MsmPNPase, are on the opposite face of the core ring and are conformationally mobile. Single particle reconstructions of MsmPNPase in the act of poly(A) synthesis highlight a 3′-terminal (rA)₄ oligonucleotide and two magnesium ions in the active site and an adenine nucleobase in the central tunnel. We identify amino acids that engage the 3′ segment of the RNA chain (Phe68, Arg105, Arg112, Arg430, Arg431) and the two metal ions (Asp526, Asp532, Gln546, Asp548), and we infer those that bind inorganic phosphate (Thr470, Ser471, His435, Lys534). Alanine mutagenesis pinpointed RNA and phosphate contacts as essential (Arg105, Arg431, Lys534, Thr470 + Ser471), important (Arg112, Arg430), or unimportant (Phe68) for PNPase activity. Severe phosphorolysis and polymerase defects accompanying alanine mutations of the enzymic metal ligands suggest a two-metal mechanism of catalysis by MsmPNPase.

Keywords: 3′ polymerase; RNA phosphorylase; metal-dependent catalysis; single particle cryo-EM

INTRODUCTION

Polynucleotide phosphorylase (PNPase) is a versatile enzyme that plays synthetic and degradative roles in bacterial RNA metabolism (Mohanty and Kushner 2000). In its degradative mode, PNPase catalyzes processive phosphorolysis of the 3′-terminal phosphodiester of RNA chains to yield nucleoside diphosphate (NDP) products. In its polymerase mode, PNPase uses NDPs as substrates to add NMPs to the 3′-OH terminus of RNA chains while expelling phosphate. The polymerase activity of PNPase was a crucial enabling reagent in the synthesis of mRNAs that led to the breaking of the genetic code (Lengyel 2012). The RNA polymerase reaction is the microscopic reversal of the RNA phosphorolysis reaction and both require a divalent cation, which can be either magnesium or manganese. We reported previously that *Mycobacterium smegmatis* (Msm) PNPase also acts on DNA polynucleotides as a DNA 3′-phosphorylase and a dADP-dependent DNA polymerase (Unciuleac and Shuman 2013a). The DNA modification reactions of PNPase depend on manganese and are not sustained effectively by magnesium.

Insights into the structural organization and catalytic mechanism of bacterial PNPase have emerged from crystal structures of exemplary enzymes (Symmons et al. 2000; Shi et al. 2008; Nurmohamed et al. 2009; Hardwick et al. 2012; Wang et al. 2017). PNPases are ring-shaped homotrimers. Each protomer is composed of five domain modules. Two RNase PH-like domains comprise the core of the trimeric ring. A metal-binding site is located within the distal PH domain (Nurmohamed et al. 2009). An α-helical module separating the PH domains is disposed on the inferior surface of the ring. The carboxy-terminal KH and S1 domains are on the opposite face of the ring and are conformationally mobile and thus usually not visualized in many of the available PNPase crystal structures. In the *Caulobacter*
PNPase structure, the three KH modules of the trimer form a narrow aperture that contacts the RNA and through which the RNA chain is threaded into the central cavity of the ring (Hardwick et al. 2012).

M. smegmatis PNPase is a homotrimer of a 763-aa polypeptide (Unciuleac and Shuman 2013a). The Mn²⁺–DNA and Mg²⁺–RNA end modifying activities of MsmPNPase were coordinately ablated by mutating the metal ligand Asp526 to alanine, signifying that the catalytic magnesium and manganese ions bind to the same site on PNPase. Deletions of the carboxy-terminal S1 and KH domains of MsmPNPase exerted opposite effects on the RNA and DNA modifying activities. Subtracting the S1 domain diminished RNA phosphorylase and polymerase activity; simultaneous deletion of the S1 and KH domains further crippled the enzyme with respect to RNA substrates. In contrast, the S1 and KH domain deletions enhanced the DNA polymerase and phosphorylase activity of MsmPNPase (Unciuleac and Shuman 2013a). Two distinct modes of nucleic acid binding by MsmPNPase were observed: (i) metal-independent RNA-specific binding that required the S1 domain, and (ii) metal-dependent binding to RNA or DNA that was optimal when the S1 domain was deleted (Unciuleac and Shuman 2013a). The metal effect on RNA binding by the MsmPNPase core trimer is not dictated by the metal catalyst of phosphoryl transfer in the active site, insofar as alanine mutations of the predicted catalytic metal-binding residues Asp526 and Asp532 did not affect Mg/Mn-dependent RNA binding to the full-length and ΔS1 MsmPNPases (Unciuleac and Shuman 2013b). These findings added a new dimension to our understanding of PNPase specificity, whereby the carboxy-terminal S1 and KH modules serve a dual purpose: (i) to help capture an RNA polynucleotide substrate for processive 3′ end additions or resections, and (ii) to provide a specificity filter that selects against a DNA polynucleotide substrate.

Whereas MsmPNPase has been well characterized biochemically, the genetics and physiology of mycobacterial PNPase are under-developed. The upstream pepR stream (a lipoprotein), lppU (a lipoprotein), Rv2783c (PNPase), pepR (a zinc protease), and Rv2781c (a putative oxidoreductase). The upstream lppU gene and the downstream pepR and Rv2781c genes are deemed inessential for M. tuberculosis growth, as defined by the recovery of viable bacteria with transposon insertions within these respective open reading frames (Sassetti et al. 2003). In contrast, no transposon insertion within the PNPase gene was recovered, raising the prospect that PNPase might be essential for growth of M. tuberculosis under laboratory conditions (Sassetti et al. 2003). This appears to be the case insofar as direct attempts to delete the M. tuberculosis gene encoding PNPase were unsuccessful and a CRISPR/Cas9 knockdown approach that reduced PNPase mRNA levels by 20-fold resulted in slowed cell growth (Plociński et al. 2019). Accordingly, mycobacterial PNPase merits consideration as a target for inhibitor discovery. This idea is given credence by the reports that Asp67Asn and Lys702Gln mutations in MtuPNPase confer resistance of clinical M. tuberculosis isolates to the anti-TB pro-drug pyrazinamide (PZA) and that pyrazinoic acid (POA, the active form of the antibiotic) inhibits MtuPNPase activity in vitro (Njire et al. 2017; He et al. 2019).

In order to gain further insight into the mycobacterial PNPase mechanism, we set out to determine the structure of the enzyme. We grew large hexagonal crystals of recombinant full-length MsmPNPase but found that they diffracted X-rays to no better than 10 Å resolution. Alternatively, as described in this report, we obtained a structure of the MsmPNPase via single particle cryo-electron microscopy with an RNA 3′ terminus and two magnesium ions visualized in the active site. Guided by the structure, we conducted a mutational analysis of MsmPNPase that illuminates enzymic functional groups essential for its phosphorylase and polymerase activities.

RESULTS AND DISCUSSION

Cryo-EM structure of mycobacterial PNPase in the act of poly(A) synthesis

Poly(A) synthesis was initiated during a 3 min reaction of purified MsmPNPase with ADP, magnesium, and oligo(rA)₁₂ primer at 22°C. The mixture was then quenched on ice and immediately applied to a Superdex-200 analytical gel filtration column. Aliquots of the peak PNPase fractions were applied to EM grids and imaged with a Titan Krios G2 transmission electron microscope equipped with a Gatan K3 camera. Discrete MsmPNPase particles (Supplemental Fig. S1A) appeared as toroidal rings with apparent three-fold symmetry after 2D and 3D classification (Supplemental Fig. S1B, C). A population of 840,000 particles was used for ab initio 3D reconstruction. The workflow for refinement of the MsmPNPase cryo-EM map to 3.07 Å resolution is described in Materials and Methods and summarized in Supplemental Figure S1D–J. Two noncontiguous masses of density were evident in the cryo-EM maps at all stages of refinement (Supplemental Fig. S1E–H). The larger component of the map comprises the core PNPase trimer ring, into which we built three continuous polypeptide chains, spanning MsmPNPase residues Asp8 to Ile594. Model refinement statistics are compiled in Supplemental Table S1. We were unable to model the carboxy-terminal S1 domain or the preceding KH domain into the weaker cloud of density adjacent to the PNPase core. We surmise that the KH-S1 portion of MsmPNPase is conformationally mobile, as found previously in crystal structures of other PNPases.
The refined model of the MsmPNPase fold is shown in Figure 1A, colored by protomer and oriented so that the inferior/bottom surface of the trimer is facing forward. Figure 1B shows a surface model in the same orientation, thereby highlighting the central pore of the PNPase ring. The structure is rotated in Figure 1C to feature the component domain modules of the core MsmPNPase protomer, wherein the PH domains are colored cyan and blue and the intervening α-helical module is colored beige. (For the purpose of orientation, the missing KH-S1 module is shown as a trio of gray ovals hovering above the carboxyl termini of the modeled PNPase core structure.) Figure 2A depicts the core trimer with the top surface of the trimer (i.e., from which RNA exits the central tunnel) facing forward.

A DALI search (Holm 2020) of the Protein Data Base affirmed a high degree of tertiary structure homology between mycobacterial PNPase and the PNPase enzymes from six different bacterial genera representing three phyla (Supplemental Table S2). As one might expect, the “top hit” was to the PNPase from the fellow Actinobacteria species Streptomyces antibioticus (Symmons et al. 2000) with respect to overall fold (Z score 54.8) and amino acid identity (73%). Whereas the fold similarity was only slightly less than PNPases from Proteobacteria (Coxiella, Caulobacter, Escherichia, and Acinetobacter; Z scores 50.7 to 47.7) and the Firmicutes taxon Staphylococcus aureus (Z score 45.1), the amino acid identities were clearly lower (43%–45%). MsmPNPase was more distantly related to human PNPase (Lin et al. 2012), a mitochondrial protein, with which it shares just 36% amino acid identity (Supplemental Table S2).

The cryo-EM maps for the MsmPNPase core trimer indicated that RNA and magnesium were bound to the enzyme. By virtue of the threefold symmetry of the PNPase core, the densities (and ensuing models) for the RNA moieties are observed in association with each protomer (Fig. 2A), notwithstanding that only one protomer can be engaged in catalysis at any time and only one RNA chain can traverse the central pore. With respect to the pore, we observed clear density for an adenine nucleobase engaged in a π-stack on the Phe89 aromatic side chain (Fig. 2C). The adenine is positioned near the center of the RNA channel (Fig. 2A; ball and stick model). A side view of the core trimer highlights that the adenine is situated high up in the channel near the exit site (Fig. 2B; ball and stick model). The crystal structure of E. coli PNPase contained electron density for a lone purine nucleoside at the same position in the channel and making a π-stack on the equivalent Phe77 side chain (Nurmohamed et al. 2009; pdb 3GCM).

RNA density for an (rA)₄ tetranucleotide was apparent adjacent to the active site and in contact with several amino acid side chains of the two PH domains, for example, Phe68, Arg105, Arg112, and Arg430 (Fig. 3). The 5′-OH of the (rA)₄ is pointing toward the center of the pore (Fig. 2A; stick model), whereas the 3′-OH of the (rA)₄ is directed toward a pair of magnesium ions (Fig. 2B; green spheres) bound to the second PH domain of each protomer. The (rA)₄ 5′-OH is 19 Å away from the adenine near the tunnel exit, from which we construe that a poly(A) chain is present and threading through the tunnel but is either disordered, or obscured by the threefold symmetry, and hence not amenable to modeling. An ordered single nucleotide or nucleobase situated high up in the central channel, disconnected from a short oligonucleotide in the active site, has been noted previously in crystal structures of Sulfolobus and Pyrococcus exosome/PNPases (Lorentzen et al. 2007, pdb 2JEA; Navarro et al. 2008, pdb 2PO1).

![Figure 1. CryoEM structure of MsmPNPase.](image-url)
RNA and magnesium contacts in the MsmPNPase active site

Figure 4 shows a detailed stereo view of the active site, with the RNA—5′-A₄pA₃pA₂pA₁OH—depicted as a stick model. Two magnesium ions, separated by 4.7 Å, are rendered as green spheres and labeled in panel B. The adenine nucleobase (Ad) in the central channel is rendered as a ball-and-stick model. (C) Close-up stereo view of the central channel looking down from the top of the PNPase trimer, highlighting how the adenine nucleobase (stick model with gray carbons) makes a π-stack on Phe89 (stick model, colored by protomer). The cryo-EM densities for adenine (blue mesh) and Phe89 (beige mesh) are shown (contoured at 2 σ, with 2 Å carve radius). Arg95 projecting into the central channel below Phe89 is also shown as a stick model.
a catalytic role, insofar as a D526A mutation of MsmPNPase abolished its phosphorylase activity (Unciuac and Shuman 2013a). The M2 cation in the MsmPNPase structure is coordinated between Asp532, Gln546-O$_{\varepsilon}$, and Asp548. The latter two contacts might be water-bridged, based on interatomic distances, though waters are not apparent in the density map. (Note that no densities at the M1 and M2 sites were seen in a 2.8 Å cryo-EM map of MsmPNPase that had not been preincubated with ADP and magnesium prior to preparation of the EM grids; not shown.) The M2-binding residues Gln546 and Asp548 are conserved in the structures of E. coli PNPase (as Gln506 and Asp508), Caulobacter PNPase (as Gln508 and Asp510; pdb 4AIM; Hardwick et al. 2012), and human PNPase (as Gln558 and Asp560; pdb 3U1K; Lin et al. 2012). The equivalent of the M2-binding residues in the Sulfolobus exosome PNPase are Gln204 and Asn206 (Rrp41 subunit; pdb 4BA2; Lorentzen and Conti 2012).

There was no interpretable density for ADP substrate or PO$_4$ product in the MsmPNPase structure, which might reflect the heterogeneity of active site occupancy in a nonsynchronous population of PNPase enzymes engaged in poly(A) synthesis. (Alternatively, ADP or PO$_4$ might have dissociated from MsmPNPase during the gel filtration step.) As a proxy for the PO$_4$ product site (equivalent to the PO$_4$ substrate site in the phosphorylase reaction), we superimposed the RNA-bound MsmPNPase structure on the crystal structure of Caulobacter PNPase in a binary complex with inorganic phosphate (pdb 3AIM) and then imported the phosphate anion (rendered as a semitransparent stick model with a green phosphorus atom) into the MsmPNPase active site depicted in Figure 4. From such a position, the phosphate could be coordinated by MsmPNPase side chains Thr470, Ser471, His435, and Lys534.

It was instructive to compare the structure of RNA-bound MsmPNPase to the crystal structure of a homologous Sulfolobus exosome PNPase enzyme in complex with a (rA)$_4$ RNA and a phosphate anion that approximates a substrate/product state of the phosphorylase/polymerase reactions, albeit absent a metal cofactor (Lorentzen and Conti 2012). Figure 5A,B depicts superimposed and horizontally offset images of the archaeal and mycobacterial PNPase active sites and their enzymic contacts to substrates. The core ring of the archaeal exosome PNPase is composed of distinct subunits Rrp42 and Rrp41 that are

![FIGURE 3. Active site RNA ligand modeled into cryo-EM density. Stereo view of the (rA)$_4$ tetranucleotide (stick model with gray carbons and yellow phosphorus atoms) modeled into surrounding density (green mesh, contoured at 1.25σ, with 2 Å carve radius). Adjacent amino acids Phe68, Arg105, Arg112, and Arg430 are depicted as stick models with beige carbons modeled into their density (blue mesh, contoured at 2σ, with 2 Å carve radius).](image-url)

![FIGURE 4. Stereo view of the MsmPNPase active site. PNPase protomer A is depicted as a cartoon model colored according to secondary structure (cyan α helices; magenta β strands). The (rA)$_4$ tetranucleotide is rendered as a stick model with gray carbons and yellow phosphorus atoms. Magnesium ions M1 and M2 are depicted as green spheres within their density (green mesh, contoured at 2σ, with 2 Å carve radius). A phosphate anion imported from the superimposed structure of E. coli PNPase is included as a semitransparent stick model with a green phosphorus atom, situated to the left of the metal ions. Amino acid side chains of protomer A that interact with the active site ligands are shown as stick models with beige carbons.](image-url)
homologous to the tandem PH domains of bacterial PNPases. The three amino acids engage the PO₄ anion in the active site of the archaeal exosome, two of which (Arg99 and Ser138) are conserved in MsmPNPase (as Arg431 and Thr470). The archaeal Arg139 equivalent in MsmPNPase is Ser471, which is poised to make a hydrogen bond to the phosphate. In MsmPNPase, the Lys534 and His435 side chains (which are not conserved in the *Sulfolobus* enzyme but are conserved in *E. coli* and Caulobacter PNPases) are likely to make additional contacts to the phosphate (Fig. 5B).

**Structure-guided alanine scan of MsmPNPase**

We constructed seven alanine mutants of MsmPNPase targeting active site amino acids imputed to coordinate the phosphate anion (T470A-S471A, K534A, R431A) and amino acids that engage the (rA)₄ RNA (R105A, R112A, R430A, R431A, F68A). The recombinant His-tagged MsmPNPase-Ala mutants and wild-type MsmPNPase were produced in *E. coli* and purified by sequential Ni-affinity, DEAE-Sephacel anion exchange, and Superdex-200 gel filtration chromatography steps. SDS-PAGE analysis of the peak homotrimer fractions from the gel filtration step affirmed similar purity of the wild-type and mutant MsmPNPase polypeptides (Supplemental Fig. S2A).

As an initial test of the MsmPNPases for RNA phosphorolysis activity, we reacted them for 15 min with a 5’-³²P-labeled 24-mer RNA (1 pmol) in the presence of 5 mM Mg²⁺ and 0.5 mM phosphate. The reaction products were analyzed by urea-PAGE and the labeled RNAs visualized by autoradiography. We saw that 1 pmol of wild-type MsmPNPase trimer sufficed to convert most of the input 24-mer RNA to a mixture of trinucleotide and dinucleotide end-products (Fig. 6A). The T470A-S471A, K534A, R431A, and R105A proteins were severely defective with respect to RNA end resection (Fig. 6A). These results signify that the contacts of these amino acids with the phosphate nucleophile and the 3’-terminal RNA segment are essential for the phosphorolysis reaction. Whereas mutants R112A and R430A resected nearly all of the input 24-mer RNA, the size distribution of the products seen after 15 min was longer and more heterogeneous than that of wild-type PNPase (Fig. 6A). In contrast, the products formed in 15 min by the F68A mutant were similar to that of the wild-type enzyme.

We then examined the kinetic profiles of RNA resection by the “partly active” MsmPNPase mutants R430 and
R112A in parallel with wild-type PNPase controls (Fig. 6B, C). Wild-type PNPase rapidly resected the 24-mer RNA to a predominant trinucleotide product within the first minute of the reaction and then slowly degraded the trinucleotide during a 15 min incubation (Fig. 6B). This pattern is indicative of a biphasic reaction in which initial resection is processive, after which the short product is prone to dissociate and only slowly reengages with the enzyme for further processing. A switch from processive to distributive action when the RNA chain is short accords with “classic” enzymology of bacterial PNPase, whereby Maxine Singer’s laboratory showed that long RNAs are resected processively, whereas phosphorolysis of short RNA oligonucleotides is not processive (Klee and Singer 1968; Chou and Singer 1970a). Comparing a series of short oligo(A) substrates, they found that the \( K_m \) for RNA increased by a factor of 39 in going from a pentanucleotide \( \text{p(Ap)}_4\text{A} \) to a trinucleotide \( \text{p(Ap)}_2\text{A} \) (Chou and Singer 1970b).

The initial phase of resection by the R430A mutant is clearly slowed and multiple cleavage intermediates of 6 to 23 nt in length are evident at the 1 min and 2 min time-points that are slowly shortened to 5 nt and then 4 nt during the 15 min reaction, with only scant formation of a trinucleotide by 15 min (Fig. 6B). A similar kinetic pattern is seen for the R112A mutant (Fig. 6C). These altered kinetics suggest that the inflection point for product dissociation is shifted to a tetranucleotide to hexanucleotide range by the R430A and R112A mutations, in keeping with the structural data that Arg430 and Arg112 contact the third, fourth, and fifth nucleotides from the 3′-terminal nucleotide that undergoes phosphorolysis. The kinetic profile of the “active” MsmPNPase mutant F68A was interrogated during a 1 min reaction in parallel with the wild-type PNPase (Fig. 6D). Both enzymes rapidly resected the 24-mer to a trinucleotide after 5 sec, which was followed by gradual resection of the trinucleotide over 60 sec (Fig. 6D). Thus, the loss of the Phe68 aromatic side chain had little discernable impact on phosphorolysis under the conditions tested.

We estimated the magnitudes of the kinetic defects of the partly active R430A and R112A PNPase mutants by comparing the time required for the wild-type enzyme to resect the 24-mer to a trinucleotide (5 sec) to the times taken by the mutants to convert a significant fraction of the input 24-mer to a tetranucleotide, which were 5 min and 10 min for the R430A and R112A enzymes, respectively (Fig. 6). By this criterion, we surmise that Arg430 and Arg112 enhance the rate of phosphorolysis by \( \sim 60\times \) and \( \sim 120\times \) respectively.

The seven alanine mutants were also assayed for RNA polymerase activity by incubation for 15 min with the same 5′\(^{32}\)P-labeled 24-mer RNA (1 pmol) in the presence of 5 mM Mg\(^{2+}\) and 2 mM ADP. The reaction products were analyzed by urea-PAGE and the labeled RNAs visualized by autoradiography. The wild-type enzyme catalyzed extensive AMP transfer from ADP to the labeled primer to generate high molecular weight RNA chains that migrated near the top of the gel (Supplemental Fig. S2B). The mutational effects on polymerase activity generally echoed...
those seen for the phosphorylase reaction. To wit: (i) the R105A, K534A, R431A, and T470A-S471A proteins were inert or severely impaired in primer extension; (ii) the R112A and R430A proteins were less active than wild-type and formed shorter products during the 15 min polymerase reaction; and (ii) F68 was as active as wild-type PNPase (Supplemental Fig. S2B).

Alanine mutations of metal ligands

Alanine mutations D532A and Q546A-D548A were targeted to amino acids that either bridge the M1 and M2 magnesium ions (Asp532) or uniquely coordinate the M2 magnesium ion (Gln546 and Asp548). The purified recombinant mutant PNPase homotrimers (Supplemental Fig. S3A) were assayed for RNA phosphorylase activity in parallel with wild-type PNPase (Fig. 7). Loss of the Asp532 metal ligand drastically slowed the rate and extent of RNA resection, such that conversion of 24-mer substrate to tetranucleotide product by the D532A mutant was achieved after 15 min, at which time there was still a residual fraction of unresected 24-mer (Fig. 7). A similarly slowed kinetic profile was seen for the Q546A-D548A mutant (Fig. 7). The -180-fold rate decrement vis-à-vis wild-type elicited by loss of Gln546 and Asp548 contacts to the M2 metal ion suggests that the second metal plays an essential role in phosphorylase activity. The metal ligand mutations exerted a similar effect on the polymerase activity of MsmPNPase, whereby the fraction of the input radiolabeled 24-mer primer RNA that was extended and the rate of AMP addition to the extended primer by the D532A and Q546A-D548A enzymes during a 15 min reaction was markedly less than that of wild-type PNPase (Supplemental Fig. S3B).

Conclusions and mechanistic implications

The cryo-EM structure of MsmPNPase in the act of poly(A) synthesis provides new insights into the organization of the active site, especially the amino acids that contact the 3′ segment of the RNA chain (Phe68, Arg105, Arg112, Arg430, Arg431) and those that are inferred to contact inorganic phosphate (Thr470, Ser471, His435, and Lys534). Structure-guided mutational analysis identified particular RNA and phosphate contacts as essential (Arg105, Arg431, Lys534, Thr470+Ser471), important (Arg112, Arg430), or unimportant (Phe68) for PNPase activity. The structure underscores shared and divergent features of the RNA and phosphate interfaces of bacterial and archaeal RNA phosphorylases. The detection of two magnesium ions in the MsmPNPase active site is a novel finding and mutational analyses of the amino acid metal ligands (here for Asp532, Gln546, and Asp548 and previously for Asp526) is consistent with a two-metal mechanism in which both metals are critical for PNPase activity. (Although this is the most parsimonious view of our results, we cannot definitely exclude, based on the structural data, a scenario in which a single metal can occupy either of two positions in the PNPase active site.) We speculate that the metals promote catalysis in two ways: (i) by stabilizing the transition state on the terminal phosphodiester during the phosphorolysis reaction and on the ADP α-phosphate during the polymerase reaction; and (ii) by interaction with the RNA 3′-O leaving group in the phosphorylase reaction and the 3′-O nucleophile in the polymerase reaction. Validation of this model will hinge on capturing structures of Michaelis complexes of metal- and RNA-bound PNPase with either ADP or PO₄ occupying the substrate site. The latter ligands are not visualized in the cryo-EM structure reported here, presumably because of heterogeneity of occupancy under conditions permissive for polymerization. The challenge will be to freeze PNPase in a single state along the reaction pathway or as a minimally perturbed on-pathway mimic.

FIGURE 7. Effect of alanine mutations of enzymic metal ligands. Reaction mixtures (70 µL) containing 20 mM Tris-HCl (pH 7.5), 5 mM MgCl₂, 0.5 mM (NH₄)₂PO₄, 0.1 µM 5′-¹⁴C-labeled 24-mer RNA, and 0.1 µM wild-type or mutant PNPase homotrimer as specified above the gels were incubated at 37°C. Aliquots (10 µL) were removed at the times specified above the lanes and quenched immediately with formamide/EDTA. The reaction products were analyzed by urea-PAGE and visualized by autoradiography.
MATERIALS AND METHODS

M. smegmatis PNPase

The pET-MsmPNPase plasmids encoding wild-type or mutant PNPase proteins (Unculeac and Shuman 2013a) were transformed into Escherichia coli BL21(DE3). Cultures (1 L) of E. coli BL21(DE3)/pET16b-MsmPNPase were grown at 37°C in LB medium containing 0.1 mg/mL ampicillin until the A600 reached 0.6. The cultures were chilled on ice, adjusted to 0.5 mM isopropyl-β-D-thiogalactopyranoside and 2% (v/v) ethanol, and then incubated at 17°C for 20 h with constant shaking. Cells were harvested by centrifugation and the pellets were stored at −80°C. All subsequent procedures were performed at 4°C. Thawed bacteria were resuspended in 25 mL of lysis buffer (50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 10% sucrose) and then sonicated for 10 min (alternating 1 min on and 1 min off) to achieve lysis. Insoluble material was removed by centrifugation. The supernatants were incubated for 1 h under continuous rotation on a nutator with 10 mL of Ni2+-nitrilotriacetic acid-agarose resin (Qiagen) that had been equilibrated with lysis buffer. The resins were then recovered into columns and bound PNPase was step-eluted with 300 mM imidazole in buffer A (50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 10% glycerol) containing 20 mM imidazole. The resins were then poured into columns and bound PNPase was step-eluted with 300 mM imidazole in buffer A. The imidazole eluates were adjusted immediately to 0.5 mM EDTA and then dialyzed overnight against buffer B (50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 0.5 mM EDTA, 10% glycerol). The dialytes were applied to 2-mL DEAE-Sephacel columns that had been equilibrated with buffer B, and the PNPase proteins were recovered in the flow-through fractions. The PNPase preparations were gel-filtered through a column of Superdex-200 equilibrated in 20 mM Tris pH 8.0, 100 mM NaCl, 10% glycerol. The PNPase elution profiles were monitored by A280 and affirmed by SDS-PAGE of the column fractions. The peak Superdex-200 fractions of wild-type PNPase and PNPase-Ala mutants were used for biochemical assays. The peak fraction was concentrated to a flow rate of 0.3 mL/min. The peak fraction was concentrated to 2 mg/mL PNPase. Aliquots (3 µL) were applied to glow-discharged UltrAuFoil 300 mesh R1.2/1.3 grids (Quantifoil). Grids were blotted for 1.5 sec at 100% humidity and 5.2°C, then plunge-frozen in liquid ethane using a Vitrobot Mark IV (FEI/Thermo Fisher Scientific). All data were collected on a Titan Krios G2 transmission electron microscope (FEI/Thermo Fisher Scientific) operating at 300 kV with a K3 camera (Gatan) at the Richard Rifkind Center for Cryo-EM at the Memorial Sloan Kettering Cancer Center. Using a set defocus range of −0.8 µm to −1.2 µm, 2890 raw movies were recorded at a super-resolution pixel size of 0.532 Å (physical pixel size of 1.064 Å). Exposures were acquired at an electron dose of 20 e−/pixel−1 s−1 with a total exposure time of 3 sec, for an accumulated electron dose of 53 e−Å−2 per exposure. Intermediate frames were recorded every 0.05 sec for a total of 60 frames.

Dose-weighted drift correction of the movie frames was performed with MotionCor2 (Zheng et al. 2017). Drift-corrected sums were used for CTF determination and defocus estimation via CTFFIND-4.1 (Rohou and Grigorieff 2015). From these images, 3,057,420 particles were automatically picked using the Laplacian-of-Gaussian filter and extracted using a particle box size of 300 pixels in RELION-3 (Zivanov et al. 2018). Particles were transferred to cryoSPARC v2 (Punjani et al. 2017) for 2D classification, where data was initially split into 100 averaged classes. Of these, 13 classes constituting 840,466 particles with high-resolution particle features were selected for generating an ab initio 3D model. Excluded particles from the 2D classification were used to generate five classes of “bad” 3D models in another ab initio run. These “bad” 3D models were used, in conjunction with the first ab initio model of the selected particles and the original particle stack from RELION, as input in a 3D heterogeneous refinement job for the purpose of filtering out particles not associated with the sample of interest. This process yielded 890,249 particles for subsequent 3D refinement, or 5.9% more particles than the 13 original 2D classes. Due to apparent threefold symmetry in the 2D and 3D classes, threefold (C3) symmetry was applied in subsequent refinements. The filtered particles were then subjected to one round of 3D heterogeneous refinement, followed by one round of nonuniform refinement, resulting in a map at an overall resolution of 3.07 Å.

An initial MsmPNPase model generated in Phyre2 (Kelley et al. 2015) was fit to the experimental map and real space refinement was performed in Phenix (Adams et al. 2010). Iterative model building was performed using O (Jones et al. 1991). The final PNPase model contains 1761 protein residues, 15 nucleotides, and six magnesium ions. RMSDs for bond lengths and angles are 0.006 Å and 0.65°, respectively. 96.4% of the protein residues fall in the most favored regions of the Ramachandran plot. The correlation between the map and the model is 0.85.

Cryo-EM of MsmPNPase in the act of poly(A) synthesis

A reaction mixture (740 µL) containing 3.58 mg of the DEAE fraction of wild-type PNPase (in 650 µL buffer B), 4.9 mM MgCl2, 1.9 mM ADP, and 0.1 mM oligo(a)12 was incubated for 3 min at room temperature to allow poly(A) synthesis, then quenched on ice and loaded immediately onto a Superdex 200 Increase analytical gel filtration column in 20 mM Tris-HCl, pH 8.0, 100 mM NaCl at a flow rate of 0.3 mL/min. The peak fraction was concentrated to 2 mg/mL PNPase. Aliquots (3 µL) were applied to glow-discharged UltrAuFoil 300 mesh R1.2/1.3 grids (Quantifoil). Grids were blotted for 1.5 sec at 100% humidity and 5.2°C, then plunge-frozen in liquid ethane using a Vitrobot Mark IV (FEI/Thermo Fisher Scientific). All data were collected on a Titan Krios G2 transmission electron microscope (FEI/Thermo Fisher Scientific) operating at 300 kV with a K3 camera (Gatan) at the

RNA phosphorylase assay

A synthetic 24-mer RNA was 5′-radiolabeled by using T4 polynucleotide kinase and [γ32P]ATP and then purified by electrophoresis through a native 20% polyacrylamide gel. RNA phosphorylase reaction mixtures containing 20 mM Tris-HCl (pH 7.5), 5 mM MgCl2, 0.5 mM (NH4)2PO4, 0.1 µM 5′-32P-labeled 24-mer RNA and 0.1 µM recombinant MsmPNPase homotrimer as specified were incubated at 37°C. The reactions were quenched at the times specified by adding 10 µL of 90% formamide, 50 mM EDTA to 10 µL of phosphorylase reaction mixture. The samples were analyzed by electrophoresis through a 40-cm 20% polyacrylamide gel containing 7.5 M urea in 44.5 mM Tris-borate (pH 8.3), 1 mM EDTA. The radiolabeled RNAs were visualized by autoradiography.
RNA polymerase assay

Reaction mixtures containing 20 mM Tris-HCl (pH 7.5), 5 mM MgCl₂, 2 mM ADP, 1 pmol (0.1 µM) primers, and 10 pmol (0.1 µM) recombinant MsmPNPase homotrimer were incubated at 37°C. The reactions were quenched with an equal volume of 90% formamide, 50 mM EDTA. The samples were analyzed by electrophoresis through a 40-cm 12% polyacrylamide gel containing 7.5 M urea in 44.5 mM Tris-borate (pH 8.3), 1 mM EDTA. The radiolabeled RNAs were visualized by autoradiography.

DATA DEPOSITION

Atomic structure models of MsmPNPase have been deposited in the PDB and EMDB databases under ID codes 7LD5 and EMD-23282, respectively.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

This research was supported by National Institutes of Health (NIH) grants AI064693 (National Institute of Allergy and Infectious Diseases) and GM126945 (National Institute of General Medical Sciences). The MSKCC Cryo-EM core facility is supported by National Cancer Institute (NCI) grant P30-CA008748.

Received May 2, 2021; accepted June 1, 2021.

REFERENCES

Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, et al. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol. Crystallogr. 66: 213–221. doi: 10.1107/S0907444909052925

Chou JY, Singer MF. 1970a. A kinetic analysis of the phosphorylase of oligonucleotides by polynucleotide phosphorylase. J Biol Chem 254: 995–1004. doi: 10.1016/S0021-9258(18)63281-8

Chou JY, Singer MF. 1970b. The effect of chain length on the phosphorylase of oligonucleotides by polynucleotide phosphorylase. J Biol Chem 254: 1005–1011. doi: 10.1016/S0021-9258(18)63282-X

Hardwick SW, Gubbey T, Hug I, Jenal U, Luisi BF. 2012. Crystal structure of Caulobacter crescentus polynucleotide phosphorylase reveals a mechanism of RNA substrate channeling and RNA degradation assembly. Open Biol 2: 120028. doi: 10.1098/rsob.120028

He L, Cui P, Shi W, Li Q, Zhang W, Li M, Zhang Y. 2019. Pyrazinoc acid inhibits the bifunctional enzyme (Rv2783) in Mycobacterium tuberculosis by competing with tmRNA. Pathogens 8: 230. doi: 10.3390/pathogens8040230

Holm L. 2020. DALI and the persistence of protein shape. Protein Sci 29: 128–140. doi: 10.1002/pro.3749

Jones TA, Zou JY, Cowan SW, Kjeldgaard M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr A 47: 110–119. doi: 10.1107/S0108767390010224

Kelley LA, Mezulis S, Yates CM, Wasm MN, Stemberg MJ. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10: 845–858. doi: 10.1038/nprot.2015.053

Klee CB, Singer MF. 1968. The processive degradation of individual polynucleotide chains: Micrococcus lysodeikticus polynucleotide phosphorylase. J Biol Chem 243: 923–927. doi: 10.1016/S0021-9258(18)63284-4

Lengyel P. 2012. Memories of a senior scientist: on passing the fiftieth anniversary of the beginning of deciphering the genetic code. Annu Rev Microbiol 66: 27–38. doi: 10.1146/annurev-micro-010312-100615

Lin CL, Wang YT, Yang WZ, Hsiao YY, Yuan HS. 2012. Crystal structure of human polynucleotide phosphorylase: insights into its domain function in RNA binding and degradation. Nucleic Acids Res 40: 4146–4157. doi: 10.1093/nar/gkr1281

Lorentzen E, Conti E. 2012. Crystal structure of a 9-subunit archaeal exosome in pre-catalytic states of the phosphorolytic reaction. Archaea 2012: 721869. doi: 10.1155/2012/721869

Lorentzen E, Dziembowska A, Lindner D, Seraphin B, Conti E. 2007. RNA channelling by the archaeal exosome. EMBO Rep 8: 470–476. doi: 10.1038/sj.embor.7400945

Mohanty BK, Kushner SR. 2000. Polynucleotide phosphorylase functions both as a 3′-to-5′ exonuclease and a poly(A) polymerase in Escherichia coli. Proc. Natl Acad. Sci 97: 11966–11971. doi: 10.1073/pnas.202295997

Navarro MV, Oliveira CC, Zanchin NI, Guimarães BG. 2008. Insights into the mechanism of progressive RNA degradation by the archaean exosome. J Biol Chem 283: 14120–14131. doi: 10.1074/jbc.M801005200

Njire M, Wang N, Wang B, Tan Y, Cai X, Liu Y, Mugweru J, Guo J, Hameed HMA, Tan S, et al. 2017. Pyrazinoic acid inhibits a bifunctional enzyme in Mycobacterium tuberculosis. Antimicrob Agents Chemother 61: e00070-17. doi: 10.1128/AAC.00070-17

Nurmoshmed S, Vaidialingam B, Callaghan AJ, Luisi BF. 2009. Crystal structure of Escherichia coli polynucleotide phosphorylase core bound to RNase E, RNA and manganese: implications for catalytic mechanism and RNA degradation assembly. J Mol Biol 389: 17–33. doi: 10.1016/j.jmb.2009.03.051

Plocrińska P, Miacios M, Houghton J, Niemiec E, Plocrińska R, Brzostek A, Słomka M, Dziadek J, Young D, Dziembowski A. 2019. Proteomic and transcriptomic experiments reveal an essential role of RNA degradosome complexes in shaping the transcriptome of Mycobacterium tuberculosis. Nucleic Acids Res 47: 5892–5905. doi: 10.1093/nar/gkz251

Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. 2017. CryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods 14: 290–296. doi: 10.1038/nmeth.4169

Rohou A, Grigorieff N. 2015. CTFIND4: fast and accurate defocus estimation from electron micrographs. J Struct Biol 192: 216–221. doi: 10.1016/j.jsb.2015.08.008

Sassetti CM, Boyd DH, Rubin EJ. 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol 48: 77–84. doi: 10.1046/j.1365-2958.2003.03425.x

Shi Z, Yang WZ, Lin-Chao S, Chak KF, Yuan HS. 2008. Crystal structure of Escherichia coli PNPase: central channel residues are involved in processive RNA degradation. RNA 14: 2361–2371. doi: 10.1261/rna.1244308

Symmons MF, Jones GH, Luisi BF. 2000. A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. Structure 8: 1215–1226. doi: 10.1016/S0969-2126(00)00521-9

Unciuleac M, Shuman S. 2013a. Distinctive effects of domain deletions on the manganese-dependent DNA polymerase and DNA phosphorylase activities of Mycobacterium smegmatis.
polynucleotide phosphorylase. *Biochemistry* **52**: 2967–2981. doi:10.1021/bi400281w
Unciuleac M, Shuman S. 2013b. Discrimination of RNA from DNA by polynucleotide phosphorylase. *Biochemistry* **52**: 6702–6711. doi:10.1021/bi401041v
Wang X, Wang C, Wu M, Tian T, Cheng T, Zhang X, Zang J. 2017. Enolase binds to RspA in competition with PNPase in *Staphylococcus aureus*. *FEBS Lett* **591**: 3523–3535. doi:10.1002/1873-3468.12859
Zheng SQ, Palovcak E, Armache J-P, Verba KA, Cheng Y, Agard DA. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**: 331–332. doi:10.1038/nmeth.4193
Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJH, Lindahl E, Scheres SHW. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**: e42166. doi:10.7554/eLife.42166