



A distinct RNA recognition mechanism governs Np₄ decapping by RppH

Rose Levenson-Palmer^{a,b,1} , Daniel J. Luciano^{a,b,1} , Nikita Vasilyev^c, Ashok Nuthanakanti^c , Alexander Serganov^c, and Joel G. Belasco^{a,b,2}

^aSkirball Institute of Biomolecular Medicine, New York University School of Medicine, New York, NY 10016; ^bDepartment of Microbiology, New York University School of Medicine, New York, NY 10016; and ^cDepartment of Biochemistry and Molecular Pharmacology, New York University School of Medicine, New York, NY 10016

Edited by Gisela Storz, Division of Molecular and Cellular Biology, National Institute of Child Health and Human Development, Bethesda, MD; received September 20, 2021; accepted December 17, 2021

Dinucleoside tetraphosphates, often described as alarmones because their cellular concentration increases in response to stress, have recently been shown to function in bacteria as precursors to nucleoside tetraphosphate (Np₄) RNA caps. Removal of this cap is critical for initiating 5' end-dependent degradation of those RNAs, potentially affecting bacterial adaptability to stress; however, the predominant Np₄ decapping enzyme in proteobacteria, ApaH, is inactivated by the very conditions of disulfide stress that enable Np₄-capped RNAs to accumulate to high levels. Here, we show that, in *Escherichia coli* cells experiencing such stress, the RNA pyrophosphohydrolase RppH assumes a leading role in decapping those transcripts, preferring them as substrates over their triphosphorylated and diphosphorylated counterparts. Unexpectedly, this enzyme recognizes Np₄-capped 5' ends by a mechanism distinct from the one it uses to recognize other 5' termini, resulting in a one-nucleotide shift in substrate specificity. The unique manner in which capped substrates of this kind bind to the active site of RppH positions the δ-phosphate, rather than the β-phosphate, for hydrolytic attack, generating triphosphorylated RNA as the primary product of decapping. Consequently, a second RppH-catalyzed deprotection step is required to produce the monophosphorylated 5' terminus needed to stimulate rapid RNA decay. The unconventional manner in which RppH recognizes Np₄-capped 5' ends and its differential impact on the rates at which such termini are deprotected as a prelude to RNA degradation could have major consequences for reprogramming gene expression during disulfide stress.

diadenosine tetraphosphate | Ap₄A | Np₄A | noncanonical cap | X-ray structure

Messenger RNA (mRNA) degradation is an important mechanism for controlling gene expression in all organisms. mRNA lifetimes directly impact protein synthesis by limiting the number of times a transcript can be translated by ribosomes. The half-lives of distinct messages in the same cell can vary widely, from seconds to an hour in bacteria and from minutes to days in vertebrates, yet the mechanisms that govern these differences remain poorly understood.

In *Escherichia coli*, the essential endoribonuclease RNase E plays a central role in controlling the decay rates of most mRNAs (1). This regulatory enzyme gains access to its cleavage sites in RNA either directly or in a 5' end-dependent manner. The latter pathway requires prior conversion of the 5' triphosphate to a monophosphate by a two-step process involving a diphosphorylated intermediate whose β-phosphate is removed by the RNA pyrophosphohydrolase RppH, a member of the Nudix hydrolase superfamily (2–4). The resulting monophosphorylated 5' end accelerates endonucleolytic cleavage by binding selectively to a pocket on the surface of RNase E and facilitating access by this enzyme to internal cleavage sites (5, 6). Other Nudix hydrolases, many only distantly related to *E. coli* RppH, appear to play an analogous role in deprotecting mRNA 5' ends in disparate prokaryotic and eukaryotic organisms (7–12).

Hundreds of *E. coli* mRNAs are degraded by this 5' end-dependent mechanism (3, 13), their susceptibility being determined in large measure by the 5'-terminal substrate preferences of RppH and RNase E. Both require at least two unpaired nucleotides at the RNA 5' end and prefer three or more (14, 15). In addition, RppH, which can convert both triphosphorylated and diphosphorylated 5' ends to monophosphates, favors substrates whose second nucleotide is a purine rather than a pyrimidine (4, 14). This nucleotide binds in a cleft near the active site and helps to position the β-phosphate for hydrolytic attack in the catalytic center (16). By contrast, the 5'-terminal sequence dependence of RNase E is negligible (15). Both RppH and RNase E assemble with other *E. coli* proteins to form distinct multimeric complexes but remain highly active in the absence of their respective protein partners (DapF for RppH and polynucleotide phosphorylase, RhlB, and enolase for RNase E) (1, 17–19).

Recently, we have reported high levels of a novel 5'-terminal RNA modification, a nucleoside tetraphosphate (Np₄) cap (Fig. 1A), in *E. coli* cells experiencing stress (20). Acquired by incorporation of a dinucleoside 5', 5'''-P¹, P⁴-tetraphosphate (Ap₄A, Gp₄A, Cp₄A, or Up₄A) during transcription initiation by RNA polymerase (21), these caps directly affect 5' end-dependent

Significance

Dinucleoside tetraphosphate alarmones function in bacteria as precursors to 5'-terminal nucleoside tetraphosphate (Np₄) caps, becoming incorporated at high levels into RNA during stress and thereby influencing transcript lifetimes. However, little is known about how these noncanonical caps are removed as a prelude to RNA degradation. Here, we report that the RNA pyrophosphohydrolase RppH assumes a leading role in decapping those transcripts under conditions of disulfide stress and that it recognizes Np₄-capped 5' ends by an unexpected mechanism, generating a triphosphorylated RNA intermediate that must undergo further deprotection by RppH to trigger degradation. These findings help to explain the uneven distribution of Np₄ caps on bacterial transcripts and have important implications for how gene expression is reprogrammed in response to stress.

Author contributions: R.L.-P., D.J.L., N.V., A.N., A.S., and J.G.B. designed research; R.L.-P., D.J.L., N.V., and A.N. performed research; R.L.-P., D.J.L., N.V., A.N., A.S., and J.G.B. analyzed data; and R.L.-P., D.J.L., N.V., A.N., A.S., and J.G.B. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

¹R.L.-P. and D.J.L. contributed equally to this work.

²To whom correspondence may be addressed. Email: joel.belasco@med.nyu.edu.

This article contains supporting information online at <http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2117318119/-DCSupplemental>.

Published February 7, 2022.

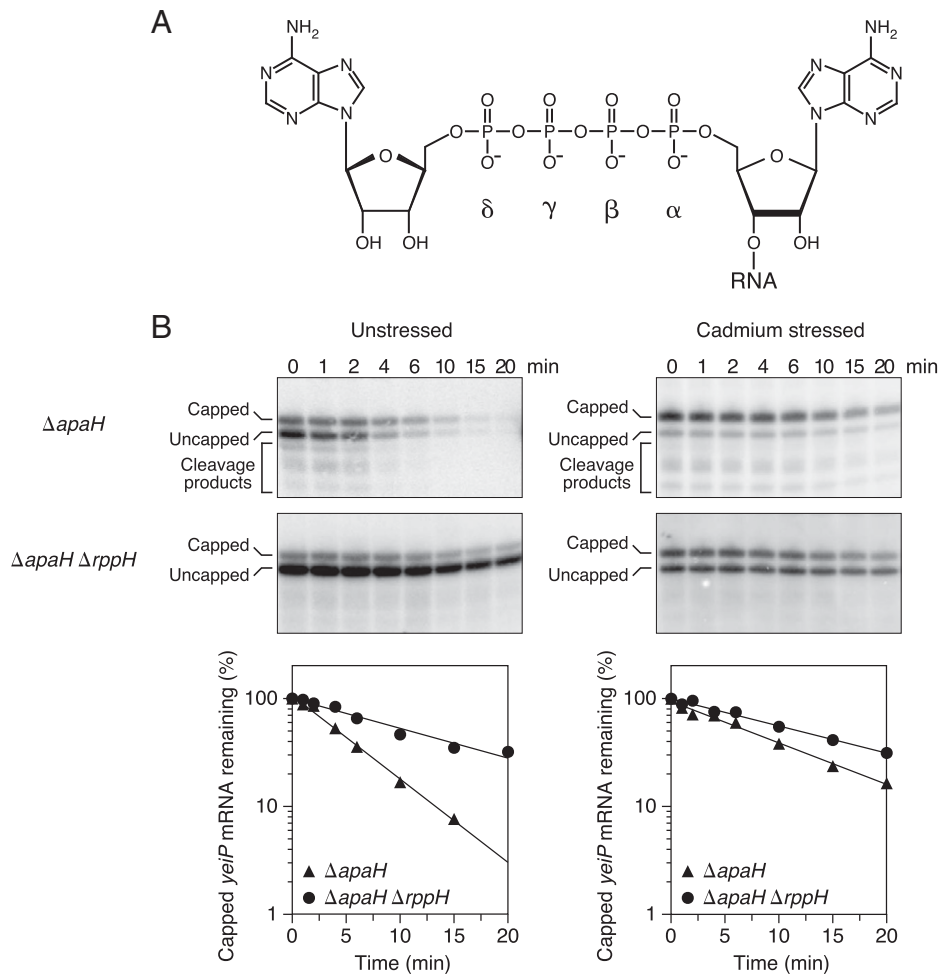


Fig. 1. Effect of disulfide stress on RppH activity in *E. coli*. (A) Structure of the 5' end of Ap₄-capped *yeiP* mRNA. Only the cap and the first transcribed nucleotide of this A-initiated RNA are shown. The α -, β -, γ -, and δ -phosphates of the cap are labeled. (B) Rate of loss of capped *yeiP* mRNA after arresting transcription in *E. coli* mutants lacking ApaH and/or RppH. (Top) Northern blots. Cultures of the indicated strains (Δ *apaH* or Δ *apaH* Δ *rppH*) growing in MOPS-glucose medium either were treated with cadmium chloride (cadmium stressed; Right) or were not (unstressed; Left), and equal amounts of total RNA extracted at time intervals after inhibiting transcription were cleaved site-specifically with a deoxyribozyme targeting *yeiP* (SI Appendix, Table S3) and analyzed by boronate gel electrophoresis and blotting with a radiolabeled *yeiP* probe. (Bottom) Graphs. The amount of capped *yeiP* mRNA that remained was plotted semilogarithmically as a function of time, and best-fit lines were calculated by linear regression. First-order rate constants for the disappearance of capped *yeiP* mRNA were determined from the slopes of these lines. Representative experiments are shown.

RNA degradation. Dinucleoside tetraphosphates (Np₄As) were first described over 50 y ago as byproducts of transfer RNA aminoacylation and are present in all realms of life (22–24), yet their biological function as cap precursors was only recently recognized. Increases in their abundance are correlated with many interesting bacterial phenotypes including several related to pathogenesis (25–28). Np₄As are often characterized as alarmones because their cellular concentration rises dramatically in response to certain stresses, especially disulfide stress induced by cysteine cross-linkers like cadmium and diamide (29, 30), which arrest growth. The same stress conditions greatly increase the cellular concentration of Np₄-capped RNAs, raising their abundance to such an extent that they come to represent 4 to 76% of most primary transcripts (20). As a result of molecular symmetry (Np₄A = Ap₄N) and the mechanism of cap acquisition, both the cap nucleotide and the first transcribed nucleotide of Np₄-capped RNAs can be A, G, C, or U.

Cap removal is a critical first step in the degradation of Np₄-capped transcripts via the 5' end-dependent pathway. Both in vitro and in *E. coli*, Np₄-capped RNA can be deprotected by either of two enzymes, ApaH or RppH (20, 31), which also

function as Np₄A hydrolases (32–34). Of the two, ApaH appears to be the principal Np₄-decapping enzyme in unstressed *E. coli* cells. Its inactivation by cysteine cross-linkers enables Np₄-capped transcripts to accumulate to a high cellular concentration during disulfide stress, a phenomenon recapitulated in unstressed *apaH* deletion mutants, which lack this enzyme altogether (20). By contrast, although RppH contributes to deprotecting Np₄-capped RNA in *E. coli* under ordinary growth conditions, deleting the *rppH* gene is not sufficient to cause Np₄-capped RNAs to rise to measurable levels. Despite the importance of ApaH and RppH for governing the lifetimes of Np₄-capped transcripts, nothing is known about their substrate specificity as decapping enzymes, which could have a major impact on gene expression during stress.

Because ApaH is inactivated in disulfide-stressed cells (20), we have now examined the importance of RppH for deprotecting Np₄-capped transcripts under those conditions and found that it remains active. Biochemical and crystallographic studies have revealed unexpected differences in the recognition of capped versus triphosphorylated RNAs by this enzyme, resulting in distinct substrate preferences and reaction products.

These studies suggest an explanation for the strikingly uneven distribution of Np_4 caps on cellular transcripts and provide key insights into the mechanism by which Np_4 -capped transcripts are degraded when such caps are abundant.

Results

Effect of Disulfide Stress on RppH Activity in *E. coli*. Two *E. coli* pyrophosphohydrolases, ApaH and RppH, have each been shown to remove Np_4 caps from RNA 5' ends, both in vitro and in vivo (20, 31). In *E. coli*, ApaH appears to be the predominant source of Np_4 decapping activity under ordinary growth conditions, where it is so efficient that Np_4 -capped RNAs are virtually undetectable. Np_4 -capped transcripts are able to accumulate to significant levels under conditions that inactivate ApaH, such as disulfide stress, or in mutant strains ($\Delta apaH$) that lack this enzyme altogether (20). To determine whether RppH remains active during disulfide stress, we induced this condition in isogenic *E. coli* mutants lacking either ApaH alone or both ApaH and RppH, blocked further RNA synthesis, and measured the rate at which Np_4 -capped transcripts were lost. In principle, any difference between the rates in these two strains can be attributed to the action of RppH.

As a model RNA, we chose the *E. coli* *yeiP* transcript, which encodes a paralog of the translation elongation factor EF-P and is highly capped in wild-type *E. coli* cells experiencing disulfide stress and in unstressed $\Delta apaH$ cells (20). We first determined the rate of *yeiP* decapping by RppH in the absence of stress. Log-phase cultures of $\Delta apaH$ and $\Delta apaH \Delta rppH$ cells were treated with rifampicin to arrest transcription, and total RNA was harvested at time intervals thereafter. Capped and uncapped cellular transcripts were then separated by electrophoresis through a polyacrylamide matrix modified with boronate side chains that selectively retard the migration of capped RNAs by transiently forming a covalent adduct with the vicinal diol of the cap nucleoside (35). Distinct bands representing capped and uncapped *yeiP* mRNA were subsequently detected by Northern blotting (Fig. 1B, Left). By subtracting the rate of disappearance of the capped *yeiP* transcript in the $\Delta apaH \Delta rppH$ strain from the rate in the isogenic $\Delta apaH$ strain, we calculate that the first-order rate constant for *yeiP* decapping by RppH is $0.115 \pm 0.008 \text{ min}^{-1}$ in unstressed *E. coli* cells (SI Appendix, Table S1).

We then repeated these measurements in the same *E. coli* strains after treatment with cadmium chloride to induce disulfide stress (Fig. 1B, Right). Under these conditions, the rate constant for *yeiP* decapping by RppH was calculated to be $0.031 \pm 0.004 \text{ min}^{-1}$ (SI Appendix, Table S1). Thus, RppH is $27 \pm 4\%$ as active during cadmium stress as it is in unstressed cells. We conclude that RppH retains the ability to deprotect Np_4 -capped transcripts in *E. coli* under disulfide stress conditions that completely inactivate ApaH. The presence of *yeiP* decay intermediates resulting from 5' monophosphate-dependent cleavage by RNase E (36) in $\Delta apaH$ but not $\Delta apaH \Delta rppH$ cells (Fig. 1B) shows that RNase E also remains active under these stress conditions.

Requirements for Unpaired Nucleotides at the 5' End. Because RppH not only contributes to the deprotection of Np_4 -capped RNA under normal growth conditions but also appears to be the primary decapping enzyme during disulfide stress, knowing the substrate preferences of this enzyme is crucial for understanding the diverse rates at which capped transcripts are degraded in *E. coli*. To examine those preferences, we employed a set of structurally unambiguous RNA substrates similar to those that had previously been used to investigate the specificity of RppH when removing pyrophosphate from triphosphorylated RNA (11, 14, 37). Synthesized in vitro transcription in the

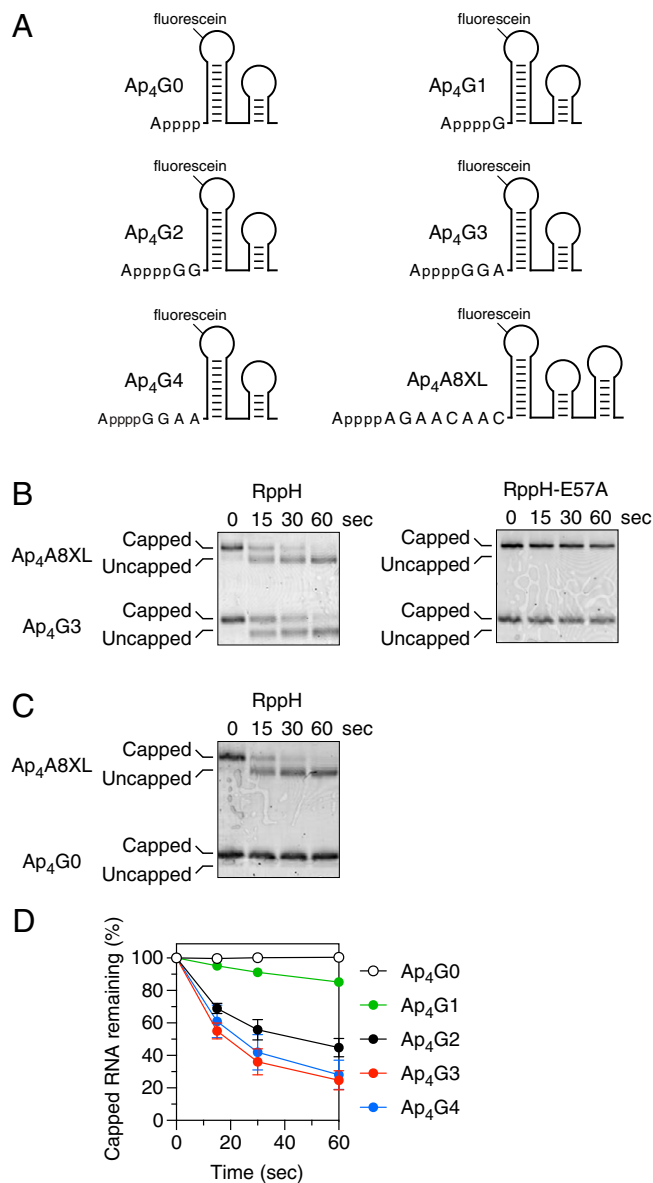


Fig. 2. Effect of the number of unpaired 5'-terminal nucleotides on decapping by RppH in vitro. (A) 5'-terminal sequence and expected secondary structure of Ap_4G_0 , Ap_4G_1 , Ap_4G_2 , Ap_4G_3 , Ap_4G_4 , and Ap_4A8XL . Each bore a 5'-terminal adenosine tetraphosphate cap and a fluorescein label at the top of the first stem-loop. In each RNA name, the first uppercase letter indicates the identity of the cap nucleotide, the second uppercase letter indicates the identity of the first transcribed nucleotide, and the second numeral indicates the number of unpaired transcribed nucleotides at the 5' end. Truncated derivatives of Ap_4G_4 (Ap_4G_3 , Ap_4G_2 , Ap_4G_1 , and Ap_4G_0) lacked one to four nucleotides from the 3' boundary of the unpaired 5'-terminal RNA segment. (B) Decapping of Ap_4G_3 . Ap_4G_3 was mixed with Ap_4A8XL and treated with either active RppH (30 nM; Left) or catalytically inactive RppH-E57A (30 nM; Right), and the decapping of each RNA was monitored as a function of time by boronate gel electrophoresis and fluorescence. (C) Decapping of Ap_4G_0 . Ap_4G_0 was mixed with Ap_4A8XL , and their decapping by active RppH (30 nM) was monitored as a function of time. (D) Rate of Ap_4G_4 , Ap_4G_3 , Ap_4G_2 , Ap_4G_1 , and Ap_4G_0 decapping. Cap removal by active RppH was monitored as in A and B, and the percentage of capped RNA remaining was calculated at each time point from the ratio of capped substrate to decapped product. Each time point is the average of three or more independent measurements. Error bars correspond to standard deviations. Time courses for decapping of the internal standard Ap_4A8XL by RppH are graphed in SI Appendix, Fig. S2A.

presence of a specific Np₄A (synonymous with Ap₄N due to molecular symmetry), each substrate comprised an Np₄-capped 5'-terminal RNA segment that was not base paired followed by two 3'-terminal stem-loops, one of which bore a fluorescein tag in its apical loop (Fig. 2A and *SI Appendix*, Fig. S1A). Another Np₄-capped RNA containing a third stem-loop (Ap₄A8XL; *SI Appendix*, Fig. S1B) was included in every reaction as an internal standard. To compare rates of decapping, each RNA substrate was combined with Ap₄A8XL and RppH, and reaction samples were quenched with ethylenediaminetetraacetate (EDTA) at various times thereafter. The substrates and reaction products were then separated by boronate gel electrophoresis and visualized by fluorescence.

The first capped substrate tested was Ap₄G3, which contained three unpaired 5'-terminal nucleotides, the first of which was G (Fig. 2A). RppH was able to efficiently decap both Ap₄G3 and Ap₄A8XL, whereas catalytically inactive RppH-E57A was unable to remove the cap from either substrate, as expected (Fig. 2B). Because a 5'-terminal stem-loop is known to prevent RppH-catalyzed pyrophosphate removal from triphosphorylated RNA (3, 14), we next examined whether such a structure also inhibits Np₄ cap removal by this enzyme. A substrate with no unpaired nucleotides at the 5' end, Ap₄G0, was completely resistant to RppH, while Ap₄A8XL in the same reaction mixture was readily decapped (Fig. 2C). To determine the number of unpaired 5'-terminal nucleotides required for decapping by RppH, we tested three more Ap₄-capped substrates containing one, two, or four unpaired nucleotides there (Ap₄G1, Ap₄G2, and Ap₄G4, respectively). RppH reacted slowly with Ap₄G1, substantially faster with Ap₄G2, and even more rapidly with Ap₄G4, whose reactivity resembled that of Ap₄G3 (Fig. 2D). The internal standard Ap₄A8XL was efficiently decapped in every reaction (*SI Appendix*, Fig. S2A). We conclude that RppH requires at least one unpaired nucleotide at the 5' end for Np₄ cap removal and prefers two or more. These requirements differ from those observed for pyrophosphate removal from triphosphorylated RNA, where RppH was found to require at least two unpaired 5'-terminal nucleotides and to prefer three or more (14).

Effect of Nucleotide Identity on Decapping. The requirement for at least one unpaired nucleotide at the 5' end of capped substrates raised the possibility that RppH might be sensitive to the identity of that nucleotide. We therefore compared the reactivity of four Ap₄-capped substrates (Ap₄A4, Ap₄G4, Ap₄C4, and Ap₄U4; *SI Appendix*, Fig. S1C) that differed in sequence only at the first of four unpaired 5'-terminal nucleotides (i.e., at the first transcribed nucleotide). RppH showed a marked preference for decapping substrates that bore a purine rather than a pyrimidine there (Fig. 3A and *SI Appendix*, Fig. S2B). Little discrimination was observed between A and G or between U and C at that position. The preference for a purine at the 5' terminus of Np₄-capped substrates is reminiscent of the nucleobase preference of this enzyme at the second nucleotide of triphosphorylated substrates (14).

Because Ap₄, Gp₄, Cp₄, and likely Up₄ caps have been observed in *E. coli* (20), we also investigated whether RppH discriminates between different types of caps (Ap₄A8, Gp₄A8, Cp₄A8, and Up₄A8; *SI Appendix*, Fig. S1D). No significant disparities were observed in its reactivity with substrates that differed only in the identity of the cap nucleobase (Fig. 3B and *SI Appendix*, Fig. S2C).

Influence of the 5'-Terminal Transcribed Nucleotide in *E. coli*. To determine whether the 5'-terminal nucleotide preference of RppH observed in vitro is also evident in vivo, we examined the influence of the first nucleotide of *efp* mRNA on its rate of

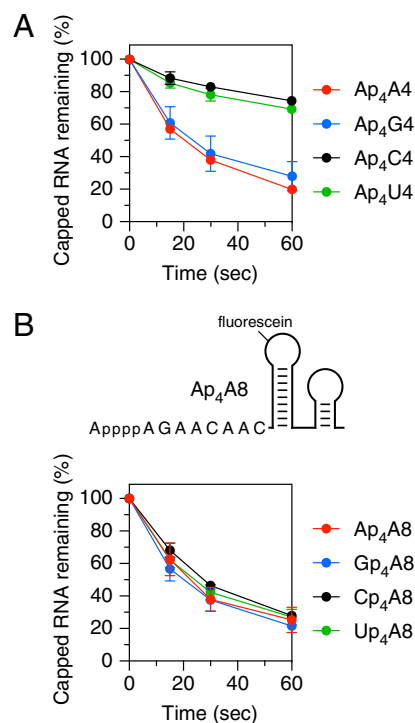


Fig. 3. Effect of nucleotide identity on decapping by RppH in vitro. (A) Reactivity of substrates differing in the first transcribed nucleotide. The decapping of Ap₄A4, Ap₄G4, Ap₄C4, and Ap₄U4 by RppH was monitored as in Fig. 2. (B) Reactivity of substrates differing in the cap nucleotide. (Top) 5'-terminal sequence and expected secondary structure of Ap₄A8. (Bottom) Decapping of Ap₄A8, Gp₄A8, Cp₄A8, and Up₄A8 by RppH. The 5'-terminal sequences of these substrates are shown in *SI Appendix*, Fig. S1. Corresponding time courses for decapping of the internal standard Ap₄A8XL by RppH are graphed in *SI Appendix*, Fig. S2 B and C. Each time point is the average of three or more independent measurements. Error bars correspond to SDs.

decapping by RppH in *E. coli*. This transcript, which encodes the translation elongation factor EF-P, naturally begins with C at the 5' end, as does its capped counterpart in Δ *apaH* cells (20). Plasmids encoding wild-type *efp* mRNA or either of two *efp* point mutants in which the 5'-terminal pyrimidine had been replaced with a purine (*efp*-C1A or *efp*-C1G) were introduced into an isogenic pair of Δ *apaH* and Δ *apaH* Δ *rppH* strains lacking the chromosomal copy of the *efp* gene. Log-phase cultures growing without stress were then treated with rifampicin, and the rate constant for RppH-mediated decapping of *efp* mRNA was calculated from the difference in the rate of disappearance of the Np₄-capped transcript in the two strains, as determined by boronate gel electrophoresis and Northern blotting.

The three *efp* transcripts were Np₄-capped to varying degrees in Δ *apaH* cells (Fig. 4A), and their rates of deprotection by RppH also differed substantially. Both of the 5'-terminal purine substitution mutants were decapped much faster than the wild-type, C-initiated transcript (Fig. 4B–D). The rate constant for decapping by RppH in *E. coli* was $0.099 \pm 0.009 \text{ min}^{-1}$ for *efp*-C1A and $0.089 \pm 0.012 \text{ min}^{-1}$ for *efp*-C1G versus only $0.034 \pm 0.005 \text{ min}^{-1}$ for wild-type *efp*, a nearly threefold acceleration (Fig. 4E and *SI Appendix*, Table S1). The influence of the first transcribed nucleotide on the rate of *efp* decapping by RppH in vivo, where the enzyme forms a heteromeric complex with the diaminopimelate epimerase DapF (17–19), was in full agreement with the 5'-terminal sequence preference observed for decapping by monomeric RppH in vitro, evidence that DapF does not alter the substrate specificity of decapping by RppH.

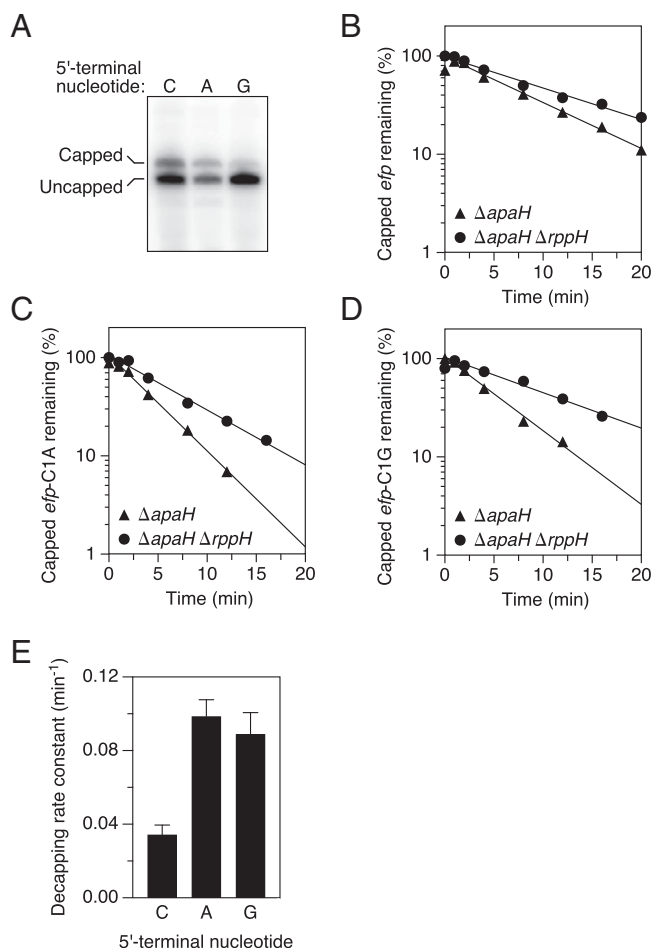


Fig. 4. Effect of the identity of the first transcribed nucleotide on decapping by RppH in *E. coli*. (A) Presence of capped *efp*, *efp*-C1A, and *efp*-C1G mRNA in unstressed cells lacking ApaH. Total RNA extracted from unstressed *E. coli* cells lacking the chromosomal *apaH* and *efp* genes and containing a plasmid that encoded *efp*, *efp*-C1A, or *efp*-C1G mRNA was cleaved site-specifically with a deoxyribozyme targeting *efp* (SI Appendix, Table S3), and capped plasmid-encoded transcripts were detected by boronate gel electrophoresis and blotting. (B–D) Rate of loss of capped *efp*, *efp*-C1A, and *efp*-C1G mRNA, respectively, after arresting transcription in *E. coli* mutants lacking ApaH and/or RppH as well as the chromosomal *efp* gene. Cultures of the indicated strains growing without stress in MOPS-glucose medium were treated with rifampicin, and equal amounts of total RNA extracted at time intervals thereafter were analyzed as in A. The amount of capped *efp*, *efp*-C1A, or *efp*-C1G mRNA that remained was plotted semilogarithmically as a function of time, and best-fit lines were calculated by linear regression. Representative experiments are shown. (E) Rate constants for decapping of *efp*, *efp*-C1A, and *efp*-C1G mRNA by RppH in *E. coli*. First-order rate constants for the disappearance of capped *efp*, *efp*-C1A, and *efp*-C1G mRNA were determined from the slopes of the best-fit lines in B–D, and the rate constant for RppH-mediated decapping of each mRNA was calculated by subtracting the rate constant for the disappearance of the capped transcript in $\Delta apaH \Delta rppH$ cells from the corresponding rate constant in $\Delta apaH$ cells.

Structure of RppH Bound to Ap₄A. By analogy to the regioselectivity of Ap₄A hydrolysis by RppH, which cuts dinucleoside tetraphosphates asymmetrically (34), RppH would be predicted to release the cap of Np₄-capped RNAs as either a nucleoside triphosphate or a nucleoside monophosphate. A nucleoside triphosphate (NTP) product would be expected if Np₄-capped RNA binds to RppH in the same manner as triphosphorylated RNA, which is cleaved between the α - and β -phosphates to release pyrophosphate (3). However, the substrate preferences

of RppH as a decapping enzyme differ from those observed for its reaction with triphosphorylated and diphosphorylated RNAs (4, 14); in particular, cap removal requires one fewer unpaired nucleotide at the 5' end and prefers a purine as the first transcribed nucleotide rather than as the second transcribed nucleotide. These findings suggest that the positioning of capped RNA bound to RppH may be shifted by one nucleotide relative to the positioning of triphosphorylated and diphosphorylated RNA substrates.

To visualize the interaction of RppH with the 5' end of Np₄-capped RNA, we determined the X-ray crystal structure of this enzyme bound to Ap₄A at 1.6 Å resolution (SI Appendix, Table S2). The structure was determined in the precleavage state with a catalytically competent enzyme, whose activity was inhibited by adding fluoride ions that replace the reactive water molecule in the catalytic center. In the structure, the tetraphosphate moiety of Ap₄A splays out on a positively charged surface of the protein (Fig. 5A), extending from the catalytic center where three Mg²⁺ ions are present to a semiopen cleft where one of the adenines (Ade1) is sandwiched between R27 and Q30 above and L83, V137, and F139 below (Fig. 5B). The side chain of R27 forms a cation- π interaction with the nucleobase of Ade1 but does not hold it tightly in the cleft, as evidenced by its partial electron density map (Fig. 5C). The other adenosine (Ade0) is disordered in the structure. The tetraphosphate bridge connecting the two nucleosides accepts 10 hydrogen bonds from seven amino-acid residues that line the binding path, including three each formed by the γ - and δ -phosphates and two each formed by the α - and β -phosphates (Fig. 5B). In addition, the δ -phosphate coordinates all three Mg²⁺ ions, and the γ -phosphate contacts Mg1 (SI Appendix, Fig. S3). Together, these interactions position the substrate optimally for catalysis.

The structure readily explains the catalytic mechanism for asymmetrical Ap₄A cleavage. The electron density peak positioned between Mg2 and Mg3 and assigned to a fluoride ion corresponds to a water molecule in the noninhibited structure (16). This nucleophile is located 2.9 Å from the phosphorus atom of the δ -phosphate at an angle (175°) that is ideal for in-line attack to sever the δ -phosphate from the γ -phosphate (Fig. 5B). Cleavage there by water would yield AMP and ATP as products, as previously reported (34).

Comparison of this structure to that of RppH bound to the 5'-triphosphorylated RNA oligonucleotide pppAGU (16) revealed striking similarities, including the absence of any appreciable change in protein conformation, as well as important differences. Specifically, Ade1 of Ap₄A replaces the second nucleoside (Gua2_{AGU}) of pppAGU in the cleft, the α -phosphate of Ap₄A replaces the phosphodiester linkage between Gua2_{AGU} and the preceding nucleoside, and the γ - and δ -phosphates of Ap₄A replace the α - and β -phosphates of the RNA ligand (Fig. 5D and E). As a result, the 5'-terminal nucleoside of pppAGU (Ade1_{AGU}) is replaced nonisosterically by the β -phosphate of Ap₄A. This phosphate forms two new hydrogen bonds with RppH residues R27 and Q95 that could not be formed by the ribose of Ade1_{AGU}. Thus, the four phosphates of Ap₄A maximize the number of interactions with RppH and, together with Ade1, effectively replace the two 5'-terminal nucleotides of the RNA ligand, thereby positioning the δ -phosphate of Ap₄A in the catalytic center, where the β -phosphate of pppAGU would have been.

Products of Decapping by RppH. Because Ap₄A is symmetrical, its complex with RppH does not alone reveal the bound orientation of Np₄-capped RNA. However, the preference of RppH for Np₄-capped substrates that begin with a purine suggests that the first transcribed nucleotide of capped RNA replaces Ade1 of Ap₄A in the cleft, where R27 could engage in a stronger cation- π interaction with a purine than with a pyrimidine (Fig. 5B). Likewise, the

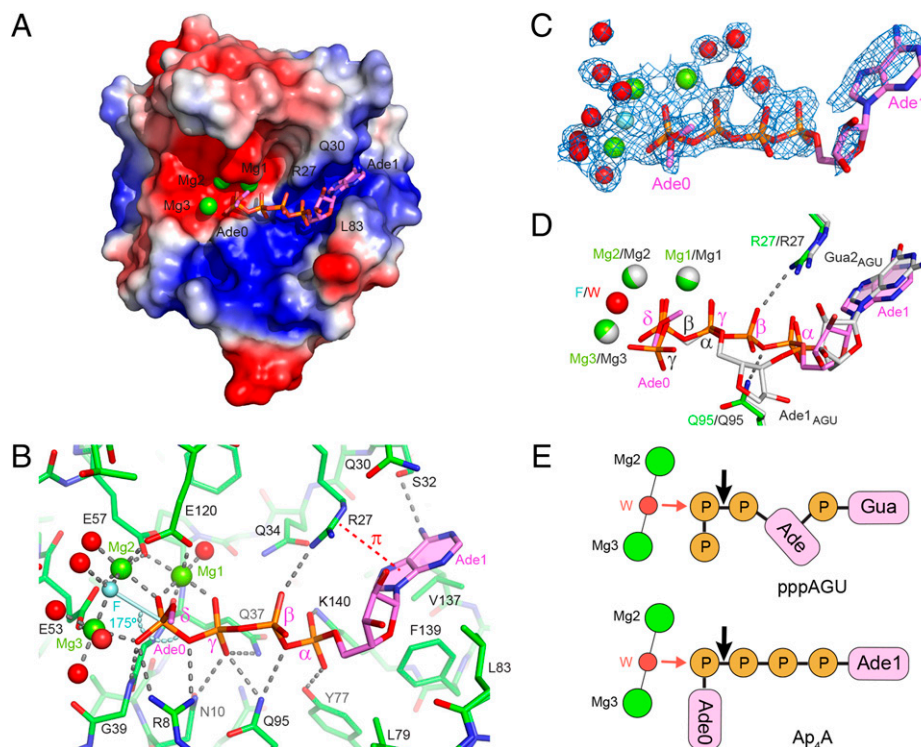


Fig. 5. Structure of RppH bound to Ap₄A. (A) Overall structure of the complex, with RppH in electrostatic surface representation and Ap₄A in violet sticks. Mg²⁺ ions are shown as green spheres. Note the alternative conformations of Ade0 as represented by the location of its C5' carbon atom. (B) Details of Ap₄A binding. The four phosphates of Ap₄A are labeled with Greek letters. Green spheres, Mg²⁺. Cyan sphere (F), fluoride. Red spheres, water molecules that coordinate Mg²⁺ or fluoride. Cyan stick, direction of in-line attack on the δ -phosphate. Curved cyan dashed line, angle of in-line attack by fluoride on the δ -phosphorus atom to displace the bridging oxygen atom. Gray dashed lines, hydrogen bonds and coordination bonds. Red dashed line, cation- π interaction. Only one of the possible locations of the Ade0 C5' carbon atom is depicted. Note that the spheres are not intended to represent the actual sizes of ions and water molecules. (C) Refined 2F_o-F_c electron density map (1 σ level, blue mesh) shown with the refined model of Ap₄A and bound Mg²⁺ ions. (D) View of the complexes formed by RppH with Ap₄A (green and violet) and the triphosphorylated RNA oligonucleotide pppAGU (gray) after all-atom superposition of the corresponding structures [PDB identifiers: 7SP3 (current work) and 4S2Y (16)]. The substrates, magnesium ions, and respective nucleophiles (F, fluoride; W, water) are shown along with the RppH residues (R27 and Q95) that form hydrogen bonds (dashed lines) with the β -phosphate of Ap₄A. (E) Schematic representation of substrates bound to RppH, vertically aligned according to the structural superposition in D. (Top) pppAGU; (Bottom) Ap₄A. Mg²⁺ ions (green circles) coordinate (thin black lines) a nucleophilic water molecule (W, red circle). Phosphates (P) are depicted as brown circles and nucleosides as magenta rectangles. Red and black arrows represent in-line nucleophilic attack and the site of phosphoanhydride cleavage, respectively. In the bottom schematic, the fluoride ion used to trap the crystallized enzyme-substrate complex in a catalytically active conformation is replaced by a water molecule normally present there.

indifference of RppH to the identity of the cap nucleoside suggests that the cap replaces Ade0, whose disordered structure in the complex of Ap₄A with RppH implies a lack of stable base-specific interactions with the protein. Binding Np₄-capped RNA in this manner would position its δ -phosphate for hydrolytic attack in the catalytic center. If so, the deprotection of such substrates by RppH should release the cap nucleotide as a nucleoside monophosphate (NMP), not an NTP, so as to generate a triphosphorylated RNA product that could then react further with RppH to produce a monophosphorylated RNA susceptible to rapid degradation by RNase E.

To test this hypothesis, we examined the RppH reaction products of the Np₄-capped RNA tetramer Ap₄GUAA. We chose a substrate of this length to ensure that the substrate and each of the possible reaction products would have distinct chromatographic mobilities on phosphoethyleneimine (PEI)-cellulose. As determined by thin-layer chromatography and ultraviolet (UV) shadowing, RppH released the cap of Ap₄GUAA primarily as AMP while generating an initial RNA product that comigrated with triphosphorylated GUAA and that was subsequently converted by RppH to monophosphorylated GUAA (Fig. 6A). Little (<20%) of the cap was released as ATP. To rule out the possibility that much of the cap released

as ATP was quickly converted to AMP, we tested the reactivity of ATP with RppH under the same conditions and found it to be largely inert (Fig. 6B), consistent with prior observations (38). Moreover, what little ATP did react was converted primarily to ADP rather than AMP. As expected, Ap₄GUAA did not react with catalytically inactive RppH-E57A.

To ensure that the positioning of Ap₄GUAA on the surface of RppH was not biased by its nucleotide sequence, which might favor binding of the 5'-terminal purine in the enzyme cleft, we tested another capped substrate, Ap₄GAUA, in which the first two transcribed nucleotides were both purines. Despite the presence of a purine at both of these positions, this substrate also yielded AMP as the principal mononucleotide product (Fig. 6C and D). Together, these findings suggest that RppH deprotects all Np₄-capped transcripts by releasing the cap primarily as a nucleoside monophosphate, thereby generating triphosphorylated RNA as the initial reaction product regardless of the 5'-terminal RNA sequence.

Finally, the reactivity of Ap₄GUAA and Ap₄GAUA with RppH was compared to that of their triphosphorylated and diphosphorylated counterparts by monitoring the reaction of each as a function of time. The Ap₄-capped substrates were the most reactive of all, irrespective of the RNA sequence (SI Appendix,

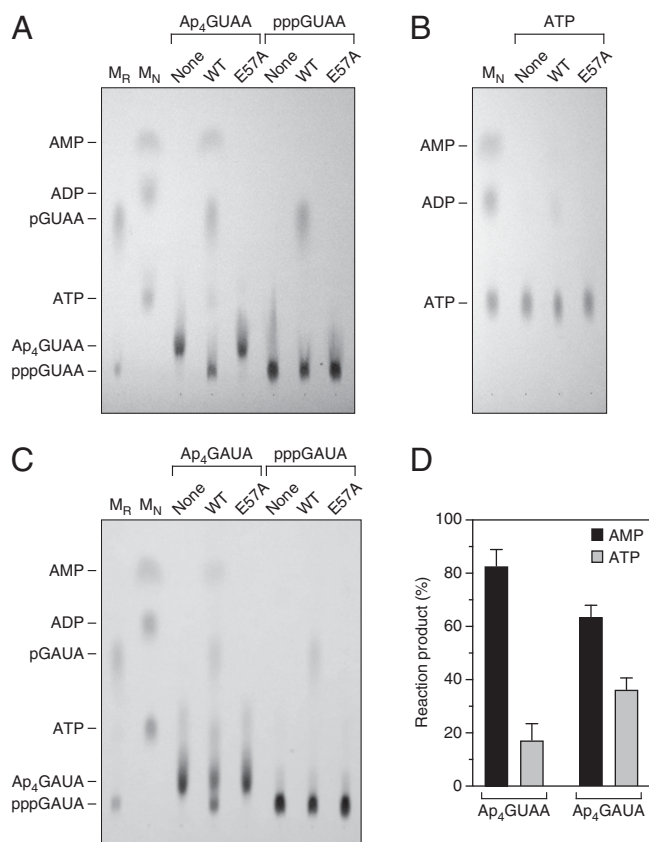


Fig. 6. Products of RppH-mediated decapping in vitro. (A) Ap₄GUAU. Ap₄GUAU and pppGUAU were examined by thin-layer chromatography on fluorescent PEI-cellulose plates before (None) and after treatment with wild-type RppH (WT) or catalytically inactive RppH (E57A). An equimolar mixture (M_R) of ATP, ADP, and AMP and a separate mixture (M_N) of pGUAU and pppGUAU were included as markers. Substrates and products were detected by UV shadowing. (B) ATP. The slow reaction of ATP with RppH was examined by thin-layer chromatography as in A. (C) Ap₄GAUA. The products of the reaction of Ap₄GAUA and pppGAUA with RppH were examined by thin-layer chromatography as in A, except that pGAUA and pppGAUA were substituted as RNA markers (M_R). (D) Products of decapping. The percentage of the Ap₄GUAU and Ap₄GAUA cap released by RppH as AMP or ATP was calculated by quantifying the AMP and ATP spots and determining their ratio. Each value is the average of two to three measurements. Error bars correspond to standard deviations. $P = 0.0002$ for Ap₄GUAU and 0.02 for Ap₄GAUA.

Fig. S4). Among the less reactive substrates, the diphosphorylated RNAs reacted faster than the triphosphorylated RNAs, a finding consistent with our previous observations (4, 18).

Discussion

It was long believed that only eukaryotic RNAs are capped at the 5' end. However, the recent discovery of 5' caps on bacterial transcripts and their impact on RNA degradation has transformed thinking about the 5'-terminal regulatory events that govern RNA function in bacteria (10, 20, 39). The most abundant types of bacterial caps yet reported are the Np₄ caps observed in *E. coli* during disulfide stress, when the Np₄A hydrolase and decapping enzyme ApaH becomes inactivated (20). Our findings now show that, under these conditions, the RNA pyrophosphohydrolase RppH assumes a leading role in RNA decapping, triggering the degradation of myriad Np₄-capped transcripts in that organism. By a combination of enzyme kinetics and X-ray crystallography, we have conducted a systematic investigation of capped substrate recognition by

this enzyme in vitro and in vivo and discovered that it binds Np₄-capped 5' ends in an unexpected manner to generate RNA products whose 5' phosphorylation state (primarily triphosphorylated) is unlike that of any product yet described for any decapping enzyme. Np₄-capped transcripts are the most reactive RNA substrates of RppH yet identified, exceeding the reactivity of their triphosphorylated (31) and even their diphosphorylated counterparts. These insights are key to understanding how bacterial mRNAs are degraded and their lifetimes are regulated during stress.

Previously determined structures of RppH bound to triphosphorylated and diphosphorylated RNA have revealed that the enzyme binds the second nucleotide of these uncapped substrates in a cleft, where that nucleobase is held in place by cation- π and hydrophobic interactions on either side that favor a purine there (14, 16). By doing so, RppH positions the β -phosphate for hydrolytic attack promoted by magnesium ions in the catalytic center, generating monophosphorylated RNA and either pyrophosphate or orthophosphate, respectively, as reaction products. By analogy, one might have expected an Np₄-capped substrate to be bound in a similar manner. However, the distinct substrate preferences of RppH when catalyzing the deprotection of Np₄-capped rather than uncapped substrates suggest otherwise. In particular, instead of requiring at least two unpaired 5'-terminal nucleotides, the second of which should ideally be a purine (14), RppH requires only one unpaired nucleotide at the 5' end of Np₄-capped substrates and prefers for that 5'-terminal nucleotide to be a purine. This one-nucleotide shift in specificity suggests that RppH binds the first transcribed nucleotide rather than the second transcribed nucleotide of Np₄-capped RNA in the cleft to form a complex that resembles RppH bound to Ap₄A. As a result, the δ -phosphate, instead of the β -phosphate, is positioned for hydrolytic attack. Corroborating this conclusion, the reaction of Ap₄-capped RNAs with RppH releases the cap primarily as AMP rather than ATP, even when the capped substrate is designed to increase the potential for the second nucleotide to bind in the cleft. A third possible mode of interaction, in which the cap nucleotide binds in the cleft, is ruled out by the indifference of the enzyme to the identity of that nucleotide and by the release of the cap principally as AMP.

What explains the distinct binding modes of Np₄-capped and uncapped substrates? For one thing, triphosphorylated and diphosphorylated RNAs would be unreactive if the 5'-terminal nucleotide, rather than the second nucleotide, were to bind in the cleft, as their phosphate chains would then be too short to reach the nucleophilic water molecule in the catalytic center. By contrast, although Np₄-capped RNAs could, in principle, bind productively in any of three ways, their decapping by RppH appears to rely primarily on only one of these binding modes, in which the first transcribed nucleotide docks in the cleft. The structure of RppH bound to Ap₄A and the influence of the number of unpaired 5'-terminal nucleotides on reactivity suggest likely explanations for this preference. When bound to RppH, an adenylylated nucleotide and the γ - and δ -phosphates of Ap₄A (a structural analog of Np₄-capped RNA) isosterically replace the second nucleotide and the α - and β -phosphates of triphosphorylated RNA and make similar contacts with the enzyme and magnesium ions (Fig. 5D). The key difference between these enzyme-substrate complexes is the β -phosphate of Ap₄A, which replaces the first nucleoside of triphosphorylated RNA and is uniquely able to accept hydrogen bonds from the positively charged side chain of R27 to a negatively charged nonbridging oxygen and from the side chain of Q95 to a bridging oxygen. These energetically favorable interactions would be lost if the second transcribed nucleotide of Np₄-capped RNA were to bind in the cleft. These interactions likely also explain the greater reactivity of Np₄-capped RNA as an RppH substrate compared to both triphosphorylated (31) and diphosphorylated RNA. The failure of

Np₄-capped substrates to bind RppH in the opposite orientation with the cap in the cleft may be related to the substantially greater reactivity of substrates bearing additional unpaired nucleotides downstream of the nucleotide in the cleft, which implies favorable interactions of those nucleotides with the enzyme that would be sacrificed if the cap were to occupy the cleft.

The mechanism by which *E. coli* RppH recognizes Np₄ caps is likely to be replicated in nearly all α -, β -, γ -, and ϵ -proteobacteria, including many important pathogens, as the active-site residues of RppH that interact with such caps and with Mg²⁺ are highly conserved in those species (14). It remains to be determined whether nonhomologous RNA pyrophosphohydrolases from δ -proteobacteria like *Bdellovibrio bacteriovorus* (8) and Firmicutes like *Bacillus subtilis* (9) can recognize Np₄ caps.

Besides its crucial role in Np₄ cap removal during stress, *E. coli* RppH has been shown to react in vitro with certain additional types of RNA caps and cap analogs but not with others. Our insights into the various modes of substrate binding by this enzyme now make it possible to explain those differences in reactivity and to predict the effect of RNA sequence and structure on the rate at which those caps are removed. For example, diadenosine polyphosphates bearing three to six bridging phosphates have been tested as substrates for RppH, and all but Ap₃A are reactive (34). In the case of Ap₄A, Ap₅A, and Ap₆A, ATP is always one of the reaction products, a finding consistent with the structure reported here for Ap₄A bound to RppH, in which adenine binding in the cleft positions the δ -phosphate beside a nucleophile poised for attack. Lacking a δ -phosphate, Ap₃A would be too short to react. Similarly, the eukaryotic cap analog m⁷Gp₃G is unreactive, whereas m⁷Gp₃-capped RNA reacts with RppH to produce m⁷GDP and monophosphorylated RNA (40), presumably because the second transcribed nucleotide of this capped RNA can bind in the cleft and position the β -phosphate for hydrolytic attack. The same rationale undoubtedly explains why RppH reacts slowly with RNA bearing an NAD or NADH cap to yield nicotinamide mononucleotide and monophosphorylated RNA but does not react with NAD or NADH themselves (10, 12, 34, 41), which each contain only two bridging phosphates. Because transcripts that are m⁷Gp₃ capped or NAD(H) capped are expected to resemble triphosphorylated RNA in their productive mode of binding, we predict that their efficient decapping by RppH requires at least two unpaired 5'-terminal nucleotides and prefers a purine at the second position. Therefore, the mechanism by which RppH recognizes Np₄-capped RNAs and positions them in the active site is probably unique among both capped and uncapped transcripts.

The 5' end-dependent degradation pathway is a key contributor to RNA turnover in *E. coli* (3). The rate at which transcripts are degraded in this manner depends both on whether or not they are capped and on the activity and specificity of each enzyme in the pathway. In unstressed wild-type cells, where the Np₄A hydrolase and decapping enzyme ApaH is active, Np₄A levels are very low and Np₄-capped transcripts are undetectable (20). Under these conditions, the RNA substrates available for 5'-terminal deprotection are predominantly triphosphorylated and diphosphorylated, and RppH acts not as a decapping enzyme but as an RNA pyrophosphohydrolase that converts their 5' ends to monophosphates (3, 4). Uncapped RppH substrates of this kind are most reactive if they contain at least two, and preferably three or more, unpaired 5'-terminal nucleotides and bear a purine at the second position, characteristics that also enable ready cleavage of their monophosphorylated reaction products by the endonuclease RNase E (14, 15). What little Np₄-capped RNA is made undergoes rapid decapping by ApaH to produce a diphosphorylated decay intermediate that must then react with RppH to generate the

monophosphorylated 5' end favored by RNase E (20). During disulfide stress, ApaH inactivation allows the cellular concentration of Np₄As and Np₄-capped transcripts to increase markedly (20) and forces RppH to assume responsibility for Np₄ cap removal. Therefore, the initial RNA product of decapping is primarily triphosphorylated under these conditions. As a result of the one-nucleotide shift in RppH specificity, decapping of these transcripts by this enzyme requires only one unpaired 5'-terminal nucleotide and prefers two or more while favoring RNAs that begin with a purine (Figs. 2–4). Nevertheless, to attain a 5' phosphorylation state vulnerable to rapid cleavage by RNase E, which remains active under these conditions, the triphosphorylated products of decapping must react a second time with RppH to convert their 5' ends to monophosphates, a reaction with distinct substrate preferences (14). As a consequence, the identities of both the first and second transcribed nucleotide appear to be critical for governing rates of RppH-mediated deprotection of Np₄-capped transcripts, whose susceptibility to 5' end-dependent degradation during disulfide stress is expected to be maximized by the presence of two unpaired 5'-terminal purines followed by at least one more unpaired nucleotide of any kind.

A recent survey of 14 *E. coli* RNAs showed that the steady-state percentage of each that was Np₄ capped when ApaH was inactivated by cadmium stress was 33 to 76% for A-initiated transcripts, 36 to 72% for C-initiated transcripts, and 26% for the lone U-initiated transcript examined yet was only 4 to 8% for G-initiated transcripts (20). Moreover, changing the first transcribed nucleotide of *yepP* mRNA from A to G reduced the capped fraction of that transcript from 76 to 5%. Consistent with those observations, we now find in Δ *apaH* cells that the steady-state percentage of Np₄-capped *efp* mRNA is only slightly affected by changing the 5'-terminal transcribed nucleotide from C (25 \pm 3%) to A (21 \pm 1%) but falls to just 4 \pm 1% when that nucleotide is changed to G (Fig. 4A). The percentage of an RNA that is capped at steady state depends on the combined effects of multiple cellular processes, especially the ratio of the rates of cap acquisition and cap removal. Rates of cap acquisition should depend on the availability of cap precursors (Np₄As = Ap₄Ns) relative to the NTP with which they compete for 5'-terminal incorporation during transcription initiation by RNA polymerase (21). In cadmium-stressed *E. coli*, the molar ratio of Ap₄G to GTP (0.26) is lower than the ratio of Ap₄C to CTP (0.73) and Ap₄U to UTP (0.68) and much lower than the ratio of Np₄A to ATP (2.6, where N = A, G, C, or U) (42). Furthermore, our findings now show that RppH decaps G- and A-initiated transcripts three to five times faster than C- and U-initiated transcripts (Figs. 3A and 4E). The combination of a low cap-acquisition rate and a high decapping rate probably explains the small percentage of G-initiated transcripts that are Ap₄-capped at steady state. By contrast, the much higher rate of cap acquisition by A-initiated RNAs, whose synthesis can begin with any Np₄A (21), likely counterbalances their rapid decapping by RppH, allowing a substantial percentage of A-initiated transcripts to be Np₄-capped at steady state. Similarly, the ample fraction of C- and U-initiated transcripts that are Ap₄-capped at steady state can be attributed to their intermediate rate of synthesis and slow rate of decapping.

The removal of Np₄ caps appears to be crucial for governing *E. coli* mRNA lifetimes during disulfide stress. By controlling rates of RppH-dependent decapping and degradation, the substrate recognition mechanism described here may help cells to recover from this stress by increasing the production of key stress-response proteins and reducing the synthesis of other proteins of less immediate importance. The ability to reset levels of gene expression under these perilous conditions may be critical for survival.

Materials and Methods

Strains and Plasmids. Measurements of 5' cap levels and decapping rates in *E. coli* were performed in isogenic derivatives of the K-12 strain BW25113 (43) bearing in-frame deletions of the *apaH* and *rppH* coding regions, either individually or in combination (20). The strains used to analyze *yeiP* mRNA contained plasmid pYeiP1 to facilitate detection of that mRNA by increasing its cellular concentration (36). The strains used to analyze *efp*, *efp*-C1A, and *efp*-C1G mRNA contained plasmid pEfp1, pEfp1-C1A, or pEfp1-C1G [derivatives of plasmid pBR322fd (13) encoding *efp* mRNA or a variant thereof] and lacked the chromosomal *efp* gene, which was deleted by P1 transduction from the *efp::kan* strain of the Keio collection (JW4107) followed by excision of the *kan* gene (44).

Rates of Decapping by RppH in *E. coli*. *E. coli* cells were grown to mid-log phase (an optical density of 0.3 at 650 nm) at 37 °C in MOPS-glucose medium (45) before arresting transcription with rifampicin (200 µg/mL). For measurements made during cadmium stress, the mid-log phase cells were treated with CdCl₂ (0.2 mM) for 90 min before adding rifampicin. Total RNA was extracted at time intervals after inhibiting transcription, and equal amounts of RNA (10 µg) were cut with a 10–23 deoxyribozyme (DZyeiP69 or DZefp87; *SI Appendix, Table S3*) specific for *yeiP* or *efp* mRNA, subjected to electrophoresis on a borate gel to separate capped from uncapped RNA, and analyzed by Northern blotting (20). Band intensities were graphed semilogarithmically as a function of time, and first-order rate constants for the disappearance of capped *yeiP* or *efp* mRNA were obtained from the slope of the best-fit line as determined by linear regression. The rate constant for RppH-mediated decapping in *E. coli* was then calculated by subtracting the rate constant for the disappearance of the capped transcript in Δ *apaH* Δ *rppH* cells from the corresponding rate constant in Δ *apaH* cells.

RNA Synthesis by In Vitro Transcription. To prepare double-stranded DNA templates for synthesizing Ap₄G4 and related substrates by in vitro transcription, pairs of complementary oligodeoxynucleotides (200 pmol each; *SI Appendix, Table S3*) were annealed and then extended with the Klenow fragment of DNA polymerase (five units; New England Biolabs) and deoxynucleoside triphosphates (0.5 mM each) in a solution (20 µL) containing Tris-HCl, pH 7.9 (10 mM), NaCl (50 mM), MgCl₂ (10 mM), and dithiothreitol (1 mM) (37). The fully double-stranded products were phenol extracted, ethanol precipitated, and dissolved in water. Templates for the synthesis of Ap₄A8XL, Ap₄A4, and the four Np₄A8 RNAs contained a T7 ϕ 2.5 promoter; the other templates all contained a T7 ϕ 6.5 promoter (46). Capped RNAs were synthesized by in vitro transcription of the double-stranded DNA template (0.25 pmol/µL) for 8 h at 37 °C with T7 RNA polymerase (5 units/µL, New England Biolabs) in the presence of Ap₄A, Gp₄A, Cp₄A, or Up₄A at a concentration (1 to 2 mM) \geq 5-fold higher than that of the NTP with which it was competing for 5'-terminal incorporation (0.1 to 0.4 mM). Each reaction mixture also contained other NTPs (ATP, GTP, and/or CTP) for internal incorporation (1 mM), fluorescein-12-UTP (0.1 mM), Tris-HCl, pH 7.9 (40 mM), MgCl₂ (6 mM), spermidine (2 mM), dithiothreitol (10 mM), and rNasin (0.25 units/µL, Promega) in a final volume of 40 µL. The resulting transcripts were purified by gel electrophoresis and elution of the band of interest.

Ap₄GUA and Ap₄GUA were synthesized by in vitro transcription of the double-stranded DNA template (2 pmol/µL; *SI Appendix, Table S3*) for 3 h at 37 °C with T7 RNA polymerase (0.1 µg/µL) in a solution containing Tris-HCl, pH 8.0 (100 mM), MgCl₂ (20 mM), spermidine (2 mM), dithiothreitol (40 mM), Ap₄G (3 mM), ATP (6 mM), and UTP (3 mM) (47). Their triphosphorylated, diphosphorylated, and monophosphorylated counterparts (pppGUA, ppGUA, pGUA, pppGUA, ppGUA, and pGUA) were prepared in the same manner, except that Ap₄G was replaced by GTP, GDP, or GMP (3 mM), respectively. The products of transcription were purified by anion-exchange chromatography on a 1-mL MonoQ (5/50) column (GE Healthcare). RNA was eluted with a 25-mL 0 to 50% gradient of 2 M triethylammonium bicarbonate buffer, pH 8.5, at a flow rate of 1 mL/min, and the desired RNA fractions were frozen and lyophilized. The dried RNAs were dissolved in water, and their concentration was determined spectrophotometrically. The identity of each RNA product was confirmed by mass spectrometry (Bruker UltraFlex MALDI-TOF). Ap₄GUA, pppGUA, ppGUA, pGUA, Ap₄GUA, pppGUA, ppGUA, and pGUA have calculated masses of 1,816.97, 1,487.76, 1,407.79, 1,327.81, 1,816.97, 1,487.76, 1,407.79, and 1,327.81 Da, respectively. Their observed masses were 1,818.00 ([M+H]⁺), 1,488.11, 1,407.68, 1,327.50, 1,818.51 ([M+H]⁺), 1,489.32 ([M+H]⁺), 1,407.66, and 1,327.70, respectively.

RppH Purification. For the in vitro decapping assays, *E. coli* RppH and RppH-E57A, each bearing an amino-terminal hexahistidine tag, were produced in *E. coli* from plasmids pLacRppH6 and pLacRppH6-E57A, respectively, purified by

affinity chromatography on TALON beads, and stored at –80 °C in a buffer containing HEPES, pH 7.5 (10 mM), NaCl (300 mM), and glycerol (50% vol/vol) (3, 14).

Structure determination was performed with a truncated variant of *E. coli* RppH that bore an additional serine residue at the amino terminus and lacked carboxyl-terminal residues 159 through 176, which were replaced by two alanines (16, 18). This variant, RppH_t, retained enzymatic activity but yielded better crystals than the full-length protein. It was produced as a fusion with amino-terminal tandem decahistidine and SUMO tags by using a T7 RNA polymerase-based expression system in *E. coli* strain BL21(DE3). The recombinant protein was purified by affinity chromatography on a HisTrap FF column (GE Healthcare), and the tags were cleaved off by a His-tagged variant of ULP1 protease, leaving an extra serine residue at the amino terminus. The affinity tags and the protease were removed by passage through a HisTrap FF column, and RppH_t was further purified by ion-exchange chromatography on a HiTrap SP column and gel filtration on Superdex 75 (GE Healthcare).

Rates of Decapping by RppH In Vitro. To compare rates of RppH-mediated decapping in vitro, each Np₄-capped transcript to be tested (750 fmol) was combined with Ap₄A8XL (500 fmol) and prewarmed to 37 °C for 3 min in a buffer containing HEPES, pH 7.5 (20 mM), MgCl₂ (10 mM), NaCl (50 mM), dithiothreitol (1 mM), rNasin (10 units; Promega), and glycerol (1% vol/vol) in a volume of 34.1 µL. An unreacted sample (7 µL) was removed and mixed with loading buffer (7 µL; 95% formamide, 20 mM EDTA, pH 8, containing bromophenol blue). RppH was freshly diluted for each reaction into a high-salt buffer containing HEPES, pH 7.5 (20 mM), MgCl₂ (10 mM), NaCl (300 mM), dithiothreitol (1 mM), and glycerol (10% vol/vol), and the diluted enzyme (7.9 µL) was added to the preheated reaction mixture to achieve a final concentration of 30 nM. Reaction samples (8.5 µL) were quenched with loading buffer (8.5 µL) after 15, 30, and 60 s, and all four samples were subjected to electrophoresis on a denaturing gel containing 13.7% acrylamide:bisacrylamide (19:1), 0.3% 3-acrylamidophenylboronic acid, 7 M urea, and 0.1 M Tris-acetate, pH 9.0 (35, 48). The fluorescein-labeled RNAs were then detected with a Typhoon Trio imager (GE Healthcare) and quantified with ImageQuant TL software. The percentage of each RNA that was capped at each time point was calculated from the ratio of band intensities (capped versus uncapped) and then normalized to the percentage that was capped at time 0.

Structure Determination by X-ray Crystallography. RppH_t was prepared in a solution of 20 mM sodium acetate, pH 5.0, 100 mM NaCl, and 1 mM dithiothreitol. Crystals of RppH_t were grown in sitting drop format, typically against 0.4 mL of reservoir solution at 20 °C for 4 d. Crystals were grown from a mixture comprising 1 µL of 1 mM protein and 2 µL of reservoir solution [0.4 M (NH₄)₂SO₄, 10% (wt/vol) PEG3350, and 10% glycerol]. To soak Ap₄A into the crystals, they were transferred to 1 µL of soaking solution containing 0.1 M sodium cacodylate, pH 6.0, 0.2 M Na₂SO₄, 25 mM NaF, 15% PEG3350, 10% glycerol, and 1 mM Ap₄A. After incubating the crystals for 20 min, they were transferred to a fresh 1-µL drop of the same solution and cross-linked by incubation over a 2-µL drop of 50% glutaraldehyde for 30 min. The crystals were transferred to a 2-µL drop of 50 mM MOPS, pH 7.0, 25 mM MgCl₂, 25 mM NaF, 1 mM Ap₄A, 30% (vol/vol) pentaerythritol propoxylate 5/4 PO/OH, and 15% PEG3350 for 20 min and then mounted and frozen in a cryostream without additional cryoprotector.

Diffraction data were collected at 100 K at the home Rigaku X-ray source and on beamline 24ID-C of the Advanced Photon Source (Argonne National Laboratory). Data were processed by using the X-ray Detector Software (XDS) suite (49). The crystal structure was solved by molecular replacement using the structure of *E. coli* RppH (Protein Data Bank [PDB] code 452W) as a search model and the Phaser-MR implementation in PHENIX (50). The models were rebuilt manually in COOT (51) and refined in PHENIX. The ligand, water molecules, and ions were added at the late stages of refinement based on the $F_o - F_c$ and $2F_o - F_c$ electron density maps. Three density map peaks in the catalytic center were assigned to Mg²⁺ ions based on their octahedral coordination geometry and coordination distances in the range of 1.9 to 2.2 Å, which are characteristic of these cations.

Products of Decapping by RppH. To characterize the products of RppH-mediated decapping, either Ap₄GUA, pppGUA, Ap₄GUA, or pppGUA (6 nmol) was prewarmed to 37 °C for 2 min in a buffer containing HEPES, pH 7.5 (20 mM), MgCl₂ (10 mM), NaCl (50 mM), and dithiothreitol (1 mM), resulting in a total volume of 6 µL. For each reaction, RppH or catalytically inactive RppH-E57A was freshly diluted into a high-salt buffer containing HEPES, pH 7.5 (20 mM), MgCl₂ (10 mM), NaCl (300 mM), and dithiothreitol (1 mM) to achieve an enzyme concentration of 14 µM, and the diluted enzyme (1 µL) was added to the prewarmed substrate, resulting in a final enzyme concentration

of 2 μM . The reactions were quenched after 8 min by adding 0.7 μL EDTA (220 mM), and the products were analyzed by thin-layer chromatography on fluorescent PEI-cellulose plates (Millipore Sigma) developed with 0.4 M ammonium sulfate, pH 5.5. Spots visualized by UV shadowing were photographed and quantified by using ImageJ software.

Comparisons of the reactivity of Ap₄GUAA, pppGUAA, ppGUAA, Ap₄GAUA, pppGAUA, and ppGAUA were performed identically, except that the reaction mixtures were scaled up 5.5-fold and samples (7 μL) were quenched after 0, 2, 4, 8, and 16 min. The extent of reaction at each time point was calculated from the molar ratio of substrate to RNA product(s), with correction for the difference in the extinction coefficients of the capped and uncapped RNAs.

- G. A. Mackie, RNase E: At the interface of bacterial RNA processing and decay. *Nat. Rev. Microbiol.* **11**, 45–57 (2013).
- H. Celesnik, A. Deana, J. G. Belasco, Initiation of RNA decay in *Escherichia coli* by 5' pyrophosphate removal. *Mol. Cell* **27**, 79–90 (2007).
- A. Deana, H. Celesnik, J. G. Belasco, The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature* **451**, 355–358 (2008).
- D. J. Luciano, N. Vasilyev, J. Richards, A. Serganov, J. G. Belasco, A novel RNA phosphorylation state enables 5' end-dependent degradation in *Escherichia coli*. *Mol. Cell* **67**, 44–54 (2017).
- G. A. Mackie, Ribonuclease E is a 5'-end-dependent endonuclease. *Nature* **395**, 720–723 (1998).
- A. J. Callaghan *et al.*, Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover. *Nature* **437**, 1187–1191 (2005).
- Z. Wang, X. Jiao, A. Carr-Schmid, M. Kiledjian, The hDcp2 protein is a mammalian mRNA decapping enzyme. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12663–12668 (2002).
- S. A. Messing *et al.*, Structure and biological function of the RNA pyrophosphohydrolase BdRppH from *Bdellovibrio bacteriovorus*. *Structure* **17**, 472–481 (2009).
- J. Richards *et al.*, An RNA pyrophosphohydrolase triggers 5'-exonucleolytic degradation of mRNA in *Bacillus subtilis*. *Mol. Cell* **43**, 940–949 (2011).
- H. Cahová, M. L. Winz, K. Höfer, G. Nübel, A. Jäschke, NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs. *Nature* **519**, 374–377 (2015).
- T. Bischler *et al.*, Identification of the RNA pyrophosphohydrolase RppH of *Helicobacter pylori* and global analysis of its RNA targets. *J. Biol. Chem.* **292**, 1934–1950 (2017).
- E. Grudzien-Nogalska *et al.*, Structural and mechanistic basis of mammalian Nudt12 RNA deNADding. *Nat. Chem. Biol.* **15**, 575–582 (2019).
- D. J. Luciano *et al.*, Differential control of the rate of 5'-end-dependent mRNA degradation in *Escherichia coli*. *J. Bacteriol.* **194**, 6233–6239 (2012).
- P. L. Foley, P. K. Hsieh, D. J. Luciano, J. G. Belasco, Specificity and evolutionary conservation of the *Escherichia coli* RNA pyrophosphohydrolase RppH. *J. Biol. Chem.* **290**, 9478–9486 (2015).
- J. Richards, J. G. Belasco, Distinct requirements for 5'-monophosphate-assisted RNA cleavage by *Escherichia coli* RNase E and RNase G. *J. Biol. Chem.* **291**, 5038–5048 (2016).
- N. Vasilyev, A. Serganov, Structures of RNA complexes with the *Escherichia coli* RNA pyrophosphohydrolase RppH unveil the basis for specific 5'-end-dependent mRNA decay. *J. Biol. Chem.* **290**, 9487–9499 (2015).
- C. R. Lee, M. Kim, Y. H. Park, Y. R. Kim, Y. J. Seok, RppH-dependent pyrophosphohydrolysis of mRNAs is regulated by direct interaction with DapF in *Escherichia coli*. *Nucleic Acids Res.* **42**, 12746–12757 (2014).
- A. Gao *et al.*, Structural and kinetic insights into stimulation of RppH-dependent RNA degradation by the metabolic enzyme DapF. *Nucleic Acids Res.* **46**, 6841–6856 (2018).
- Q. Wang *et al.*, DapF stabilizes the substrate-favoring conformation of RppH to stimulate its RNA-pyrophosphohydrolase activity in *Escherichia coli*. *Nucleic Acids Res.* **46**, 6880–6892 (2018).
- D. J. Luciano, R. Levenson-Palmer, J. G. Belasco, Stresses that raise Np4A levels induce protective nucleoside tetraphosphate capping of bacterial RNA. *Mol. Cell* **75**, 957–966 (2019).
- D. J. Luciano, J. G. Belasco, Np4A alarmones function in bacteria as precursors to RNA caps. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 3560–3567 (2020).
- F. J. Finamore, A. H. Warner, The occurrence of P1, P4-diguanosine 5'-tetraphosphate in brine shrimp eggs. *J. Biol. Chem.* **238**, 344–348 (1963).
- P. C. Zamecnik, M. L. Stephenson, C. M. Janeway, K. Randerath, Enzymatic synthesis of diadenosine tetraphosphate and diadenosine triphosphate with a purified lysyl-sRNA synthetase. *Biochem. Biophys. Res. Commun.* **24**, 91–97 (1966).
- P. C. Lee, B. R. Bochner, B. N. Ames, Diadenosine 5',5''-P1,P4-tetraphosphate and related adenylylated nucleotides in *Salmonella typhimurium*. *J. Biol. Chem.* **258**, 6827–6834 (1983).
- T. M. Ismail, C. A. Hart, A. G. McLennan, Regulation of dinucleoside polyphosphate pools by the YgdP and ApaH hydrolases is essential for the ability of *Salmonella enterica* serovar typhimurium to invade cultured mammalian cells. *J. Biol. Chem.* **278**, 32602–32607 (2003).
- S. Hansen, K. Lewis, M. Vulić, Role of global regulators and nucleotide metabolism in antibiotic tolerance in *Escherichia coli*. *Antimicrob. Agents Chemother.* **52**, 2718–2726 (2008).
- R. D. Monds *et al.*, Diadenosine tetraphosphate (Ap4A) metabolism impacts biofilm formation by *Pseudomonas fluorescens* via modulation of c-di-GMP-dependent pathways. *J. Bacteriol.* **192**, 3011–3023 (2010).
- X. Ji *et al.*, Alarmone Ap4A is elevated by aminoglycoside antibiotics and enhances their bactericidal activity. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 9578–9585 (2019).
- A. Varshavsky, Diadenosine 5', 5''-P1, P4-tetraphosphate: A pleiotropically acting alarmone? *Cell* **34**, 711–712 (1983).
- B. R. Bochner, P. C. Lee, S. W. Wilson, C. W. Cutler, B. N. Ames, AppppA and related adenylylated nucleotides are synthesized as a consequence of oxidation stress. *Cell* **37**, 225–232 (1984).
- O. Hudeček *et al.*, Dinucleoside polyphosphates act as 5'-RNA caps in bacteria. *Nat. Commun.* **11**, 1052 (2020).
- A. Guranowski, H. Jakubowski, E. Holler, Catabolism of diadenosine 5',5''-P1,P4-tetraphosphate in prokaryotes. Purification and properties of diadenosine 5',5''-P1,P4-tetraphosphate (symmetrical) pyrophosphohydrolase from *Escherichia coli* K12. *J. Biol. Chem.* **258**, 14784–14789 (1983).
- P. Plateau, M. Fromant, A. Brevet, A. Gesquière, S. Blanquet, Catabolism of bis(5'-nucleosidyl) oligophosphates in *Escherichia coli*: Metal requirements and substrate specificity of homogeneous diadenosine-5',5''-P1,P4-tetraphosphate pyrophosphohydrolase. *Biochemistry* **24**, 914–922 (1985).
- M. J. Bessman *et al.*, The gene *ygdP*, associated with the invasiveness of *Escherichia coli* K1, designates a Nudix hydrolase, Orf176, active on adenosine (5')-pentaphosphate-(5')-adenosine (Ap5A). *J. Biol. Chem.* **276**, 37834–37838 (2001).
- G. L. Igloi, H. Kössel, Affinity electrophoresis for monitoring terminal phosphorylation and the presence of queuosine in RNA. Application of polyacrylamide containing a covalently bound boronic acid. *Nucleic Acids Res.* **13**, 6881–6898 (1985).
- J. Richards, D. J. Luciano, J. G. Belasco, Influence of translation on RppH-dependent mRNA degradation in *Escherichia coli*. *Mol. Microbiol.* **86**, 1063–1072 (2012).
- P. K. Hsieh, J. Richards, Q. Liu, J. G. Belasco, Specificity of RppH-dependent RNA degradation in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 8864–8869 (2013).
- A. Gao, N. Vasilyev, A. Kaushik, W. Duan, A. Serganov, Principles of RNA and nucleotide discrimination by the RNA processing enzyme RppH. *Nucleic Acids Res.* **48**, 3776–3788 (2020).
- Y. G. Chen, W. E. Kowtoniuk, I. Agarwal, Y. Shen, D. R. Liu, LC/MS analysis of cellular RNA reveals NAD-linked RNA. *Nat. Chem. Biol.* **5**, 879–881 (2009).
- M. G. Song, S. Bail, M. Kiledjian, Multiple Nudix family proteins possess mRNA decapping activity. *RNA* **19**, 390–399 (2013).
- G. B. Conyers, M. J. Bessman, The gene, *jalA*, associated with the invasion of human erythrocytes by *Bartonella bacilliformis*, designates a nudix hydrolase active on dinucleoside 5'-polyphosphates. *J. Biol. Chem.* **274**, 1203–1206 (1999).
- H. Coste, A. Brevet, P. Plateau, S. Blanquet, Non-adenylylated bis(5'-nucleosidyl) tetraphosphates occur in *Saccharomyces cerevisiae* and in *Escherichia coli* and accumulate upon temperature shift or exposure to cadmium. *J. Biol. Chem.* **262**, 12096–12103 (1987).
- K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6640–6645 (2000).
- T. Baba *et al.*, Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
- F. C. Neidhardt, P. L. Bloch, D. F. Smith, Culture medium for enterobacteria. *J. Bacteriol.* **119**, 736–747 (1974).
- T. M. Coleman, G. Wang, F. Huang, Superior 5' homogeneity of RNA from ATP-initiated transcription under the T7 ϕ 2.5 promoter. *Nucleic Acids Res.* **32**, e14 (2004).
- N. Vasilyev, A. Serganov, Preparation of short 5'-triphosphorylated oligoribonucleotides for crystallographic and biochemical studies. *Methods Mol. Biol.* **1320**, 11–20 (2016).
- A. E. Ivanov, H. Larsson, I. Y. Galaev, B. Mattiasson, Synthesis of boronate-containing copolymers of N,N-dimethylacrylamide, their interaction with poly(vinyl alcohol) and rheological behaviour of the gels. *Polymer (Guildf.)* **45**, 2495–2505 (2004).
- W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
- P. D. Adams *et al.*, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
- P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).