NrnA is a 5’-3’ exonuclease that processes short RNA substrates in vivo and in vitro

Cordelia A. Weiss1,†, Tanner M. Myers2,†, Chih Hao Wu1, Conor Jenkins2, Holger Sondermann3,4, Vincent T. Lee1,† and Wade C. Winkler1,2,†

1Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD 20742, USA, 2Department of Chemistry and Biochemistry, University of Maryland, College Park, MD 20742, USA, 3CSSB – Centre for Structural Systems Biology, Deutsches Elektronen-Synchrotron (DESY), 22607 Hamburg, Germany and 4Christian-Albrechts-Universität, 24118 Kiel, Germany

Received March 10, 2022; Revised October 25, 2022; Editorial Decision October 26, 2022; Accepted November 30, 2022

ABSTRACT

Bacterial RNases process RNAs until only short oligomers (2–5 nucleotides) remain, which are then processed by one or more specialized enzymes until only nucleoside monophosphates remain. Oligoribonuclease (Orn) is an essential enzyme that acts in this capacity. However, many bacteria do not encode for Orn and instead encode for NanoRNase A (NrnA). Yet, the catalytic mechanism, cellular roles and physiologically relevant substrates have not been fully resolved for NrnA proteins. We herein utilized a common set of reaction assays to directly compare substrate preferences exhibited by NrnA-like proteins from Bacillus subtilis, Enterococcus faecalis, Streptococcus pyogenes and Mycobacterium tuberculosis. While the M. tuberculosis protein specifically cleaved cyclic di-adenosine monophosphate, the B. subtilis, E. faecalis and S. pyogenes NrnA-like proteins uniformly exhibited striking preference for short RNAs between 2–4 nucleotides in length, all of which were processed from their 5′ terminus. Correspondingly, deletion of B. subtilis nraA led to accumulation of RNAs between 2 and 4 nucleotides in length in cellular extracts. Together, these data suggest that many Firmicutes NrnA-like proteins are likely to resemble B. subtilis NrnA to act as a housekeeping enzyme for processing of RNAs between 2 and 4 nucleotides in length.

INTRODUCTION

The general cycle of RNA metabolism involves polymerization of NTPs by DNA-dependent RNA polymerase, followed by fragmentation and hydrolysis of cellular RNAs by ribonucleolytic enzymes (RNases). The process of RNA degradation is essential, as it is required for replenishing NTP pools; therefore, RNA degradation pathways profoundly affect gene expression, environmental adaptation, and cellular viability. Moreover, defects in RNA degradation may lead to aberrant NTP pools that interfere with nucleotide signaling pathways. While the general principles of RNA processing and decay are conserved among bacteria, the individual strategies and classes of RNases differ from microbe to microbe. As such, there is no universally conserved set of RNases that all bacteria rely upon for mRNA turnover (1). Indeed, of the 40 RNases that have been identified for Escherichia coli and Bacillus subtilis, only a subset is shared between the two microorganisms (2). For example, while bulk E. coli mRNA degradation is typically initiated by endoribonuclease E (RNase E), which is essential in E. coli and other Gammaproteobacteria, RNase E is conspicuously absent in the genomes of B. subtilis and many Firmicutes. Instead, B. subtilis employs endoribonuclease Y (RNase Y) as a key enzyme to initiate global mRNA turnover (3). The action of endonucleolytic cleavage results in two mRNA decay products that are no longer protected at the 3′ (upstream product) or 5′ (downstream product) terminus, which can then be subject to degradation by exoribonucleases. Generally, the exoribonucleases are thought to act processively, cleaving in the 3′-5′ direction. However, for some bacteria, RNAs are also processed from the 5′ terminus by RNase J1 or RNase AM (YciV), which are the only 5′-3′ exoribonucleases to be discovered in bacteria to date (4,5). The combined action of RNase enzymes leads to processing of mRNAs until only short (< 5 nucleotides) RNA fragments remain. Then, a separate category of RNase enzymes is required for processing of these short RNAs.

In E. coli and other bacteria, the enzyme Oligoribonuclease (Orn) performs the crucial terminal step of RNA degra-
vation. Originally discovered in the 1960s, it was argued that Orn is a 3'-5' exoribonuclease that specifically cleaves short oligoribonucleotides from between 2 and 5 nucleotides in length (6–10). Interestingly, Orn is essential for viability in several species where it has been tested; however, the molecular basis of Orn’s essentiality is not yet known (9). This essentiality makes Orn unique among all other known 3'-5' exoribonucleases, which have been shown in some instances to be functionally redundant. Yet, a recent biochemical reassessment of Orn revealed that it binds and cleaves dinucleotides with greater specificity than other short RNAs (11). Similarly, an X-ray crystallographic analysis of RNA-bound Orn revealed a substrate binding pocket that appears to be remarkably restricted. Based on these combined data, it was proposed that Orn might function most of the time as a ‘diribonuclease’, an enzyme that preferentially acts upon diribonucleotides rather than longer RNA substrates (11). This discovery suggests that processing of very small RNA species may occur through discrete degradation of diribonucleotides and others short RNAs.

For degradative roles with Orn. These proteins were identified over exoribonucleases, which have been shown in some instances to preferentially cleave dinucleotides rather than longer RNA substrates. Based on this complementation phenotype, it was suggested that Orn might function most of the time as a ‘diribonuclease’, an enzyme that preferentially acts upon diribonucleotides rather than longer RNA substrates (11). This discovery suggests that processing of very small RNA species may occur through discrete steps and that cleavage of dinucleotides might represent a particularly crucial bottleneck in RNA degradation. Yet not all bacteria encode for Orn. In fact, the phylogenetic distribution of orn is limited to Proteobacteria and Actinobacteria, as well as Eukaryotes. Therefore, it is important to discover what alternative strategies bacteria utilize for degradation of diribonucleotides and other short RNAs (12–15).

A few candidate proteins have been discovered to have overlapping roles with Orn. These proteins were identified based on their ability to complement the growth defect elicited by a conditional E. coli orn mutant (14–16). Based on this complementation phenotype, it was suggested that these proteins participate in the terminal steps of RNA degradation, akin to Orn. Two of these RNases are encoded by B. subtilis: NanoRNase A (NrnA Bs) and NanoRNase B (NrnB Bs). Since their discovery, it has been largely assumed that NrnA Bs and NrnB Bs are likely to behave in a redundant manner to cleave oligoribonucleotides. While Orn’s active site includes a signature DEDD(h) amino acid motif, NrnA Bs and NrnB Bs are both members of the DHH superfamily family of proteins (17). This superfamily consists of enzymes that exhibit DNA or RNA phosphodiesterase activity that together encompass a broad range of substrate specificities. Proteins in this superfamily display a conserved N-terminal domain, responsible for their inclusion in the DHH superfamily (Pfam PF01368), while sequences of the C-terminal domain have been found to vary among family members. For example, a subset of proteins in the DHH superfamily contain an associated C-terminal DHHA1 (DHH-associated domain 1, Pfam PF02272) domain. This DHHA1 subfamily includes proteins of diverse functions, such as NrnA proteins, cyclic-diadenosine monophosphate (c-di-AMP) phosphodiesterase GdpP, alanyl-tRNA synthetase (AlaRS), and RecJ (18,19). RecJ is a 5'-3' exonuclease that processes single-stranded DNA (ssDNA) and is involved in homologous recombination and DNA repair (20). Therefore, not all proteins with DHH and DHHA1 motifs display the same nuclelease activity. Furthermore, the annotation of some subfamilies of DHH-DHHA1 proteins has been complicated by an overall lack of diagnostic sequence criteria to denote them. For example, many bacteria encode for proteins with DHH-DHHA1 domains, yet it is unclear how many should be annotated as NrnA proteins. In example, a subset of NrnA proteins evolved to cleave only certain targeted sequences? Are some NrnA proteins differentially regulated? Moreover, are there subclasses of NrnA proteins that specialize in processing of signaling nucleotides? The full range of intracellular roles and physiologically relevant substrates remains to be explored for NrnA and NrnA-like proteins.

Prior investigations of NrnA-like proteins led to several different opinions on NrnAs enzymatic activities and preferences. One possibility is that NrnA acts as a diribonuclease, akin to Orn. This speculation stems from the observation that overexpression of B. subtilis nrnA and nrnB in a Pseudomonas aeruginosa PA14 Δorn strain resulted in restoration of the wild-type PA14 phenotype, as compared to PA14 Δorn, which otherwise forms small colonies (13). Therefore, NrnA Bs and NrnB Bs either perform a similar function as Orn or exhibit functions that overlap with Orn. However, the published NrnA data are not uniform in their conclusions. For example, some studies have concluded that NrnA exhibits 3'-5' exoribonucleolytic activity (15), while others have reported 5'-3' exoribonucleolytic activity (19). Yet another study suggested that NrnA could exhibit both activities, preferentially cleaving longer RNA substrates from the 5' end while targeting shorter RNA substrates from their 3' terminus (21). If this were true, NrnA would represent a rare instance where an RNAse exhibits dual polarity on oligoribonucleotide substrates. Yet other data argue that NrnA can act on ssDNA substrates (19). Furthermore, some stand-alone DHH-DHHA1 domain-containing proteins (e.g. Staphylococcus aureus Pde2, Streptococcus pneumoniae Pde2) were proposed to hydrolyze c-di-AMP to AMP in a two-step process via the linear intermediate AA, although there is not uniform agreement on that either (18,22). Whether NrnA-like proteins directly hydrolyze cyclic dinucleotide signaling molecules or possess an expanded range of RNA substrates has been further complicated by analyses of a Mycobacterium tuberculosis ‘NrnA-like protein’ that has been recently renamed CnpB (or Rs2837c) (23–26). While some data have argued that CnpB acts directly on short RNA substrates 2 - 5 nucleotides in length in addition to degrading a 24-mer RNA substrate (23), other studies have argued that CnpB specifically hydrolyzes c-di-AMP and c-di-GMP to nucleoside monophosphates in addition to linearizing 2'3'-cGAMP to GA (2'-5') (24–26). Furthermore, cyclic nucleotide cleavage has also been reported for the Mycobacterium smegmatis ‘NrnA-like’ protein (27).

Yet other studies have linked ‘NrnA-like’ proteins to sulfur metabolism through their putative influence on pAp, which in some bacteria acts as an important carrier of sulfate. Specifically, NrnA Bs was shown to complement an E. coli cysQ mutant, which is auxotrophic for cysteine; in this strain, expression of B. subtilis nrnA restored cysteine prototrophy (15). Moreover, deletion of B. subtilis nrnA resulted in a slower growth rate in the absence of cysteine (15). Based on these observations, it was suggested that NrnA Bs may act as a CysQ-like phosphatase that dephosphorylates pAp to AMP. However, the reduction in B. subtilis growth rate for ΔnrnA was mild overall, hinting that there could
still be a different enzyme that functions as a pAp phosphatase in *B. subtilis*.

Given the range of different claims that have been made on the activities and functions of NrnA-like proteins, it is difficult to categorically summarize exactly what enzymatic activity should be expected for a ‘NrnA’ protein. Simply put, there is great uncertainty about the substrate preferences and cellular functions of NrnA-like proteins. Importantly, the experimental conditions and assay formats have varied widely among these different studies. Given that NrnA<sub>Bs</sub> (and NrnB<sub>Bs</sub>) can complement *orn* deletion mutants (13–15), the simplest hypothesis is that all three proteins share exactly overlapping substrate requirements and overall cellular roles. But as summarized above, the available data on NrnA (and NrnB) paint a more confusing picture of what exactly these proteins are likely to be doing inside cells. Is NrnA a ‘diribonuclease’ akin to Orn? Or is it an RNase that specializes in processing either ‘short’ or ‘long’ RNA substrates? Or is NrnA a nuclease that processes both RNA and DNA substrates, as well as cyclic dinucleotides? And does NrnA process RNAs only in one direction, as do most RNase enzymes, or does it exhibit bidirectional abilities?

To answer these questions, a variety of putative NrnA substrates must be assayed under a set of common reaction conditions. This could lead to a clearer definition of NrnA’s activity and could help discover distinct subclasses of NrnA-like proteins. This, in turn, could lead to identification of signature residues that guide substrate recognition for stand-alone DHH-DHHAA1 family proteins.

In this study, we utilized a common set of reaction assays that closely resembled the assays that were recently used to re-examine Orn’s diribonuclease activity (11). We reasoned that this approach would provide the best evidence for NrnA<sub>Bs</sub>’s substrate preferences, including the enzyme’s directionality. To examine whether NrnA<sub>Bs</sub>’s activity could impact dinucleotide pools in vivo, we investigated whether the loss of *nrnA* or *nrnB* could affect cyclic di-GMP signaling. To directly investigate whether NrnA<sub>Bs</sub> and NrnB<sub>Bs</sub> could affect the abundance of other short RNAs in *in vivo*, we examined the cleavage pattern of a radiolabeled 10-mer RNA in cellular extracts of wild-type, *B. subtilis* strains were grown in LB at 37°C, shaking with aeration (unless otherwise noted). When appropriate, *B. subtilis* strains were grown in the presence of 5 µg/ml chloramphenicol or 100 µg/ml spectinomycin. The methods for creating markerless deletions of *B. subtilis* ∆*nrnA*, ∆*nrnB*, and ∆*ornA∆ornB*, and the integration of fluorescent c-di-GMP riboswitch-yfp reporter constructs have been previously described (13,28). To construct complementation strains, *nrnA<sub>Bs</sub>*, *nrnB<sub>Bs</sub>*, *orn<sub>Ec</sub>*, and *cysQ<sub>Ec</sub>*, were each PCR amplified from chromosomal DNA preparations. Sequences were subcloned via Gibson assembly (29) into the *amyE* integration vector pDR111, which harbors an IPTG-inducible promoter upstream of the target gene. Transformation of *B. subtilis* was performed using a previously described protocol (30). To construct *E. coli* strains for overexpression of targeted proteins, *nrnA<sub>Bs</sub>*, *nrnA<sub>Ec</sub>*-pde2<sub>Sp</sub>, *gdpP<sub>Bs</sub>*<sup>82–659</sup>, *orn<sub>Vc</sub>*<sup>Δ</sup>, and *cysQ<sub>Ec</sub>* were each PCR amplified from chromosomal DNA preparations. The sequences encoding *cnpB<sub>Mt</sub>*<sup>Δ</sup>, *cynB<sub>Mt</sub>*<sup>Δ</sup>, *cnpB<sub>Bs</sub>*<sup>Δ</sup>, *disA<sub>Bs</sub>*<sup>Δ</sup>, and *ppd<sub>Cs</sub>* were codon optimized for expression in *E. coli* and purchased from Integrated DNA Technologies. The different gene sequences were subcloned via Gibson assembly (29) into IPTG-inducible expression vector pAmr30, to yield an N-terminal 10xHis-SUMO tag that is cleavable by bdSENP1 protease (31). The *ppd<sub>Cs</sub>* and *gdpP<sub>Bs</sub>*<sup>82–659</sup>*<sup>Δ</sup> sequences were subcloned via Gibson assembly (29) into the IPTG-inducible expression vector pHis-parallel, to yield an N-terminal 6xHis tag. To create the point mutants of *nrnA<sub>Bs</sub>* mutations were generated using the Q5 Site-Directed Mutagenesis Kit (NEB). *E. coli* XL10-Gold<sup>®</sup> (Agilent) was initially transformed with all plasmids, and sequences of all inserts were verified by Sanger sequencing. *E. coli* T7 Express (NEB) was transformed with all plasmids that were used for overexpression and purification of targeted proteins.

Fluorescence microscopy and quantification

Single colonies were used to inoculate liquid minimal salts glycerol glutamate (MSgg) medium (32) and grown at 37°C with shaking overnight. The following morning, cultures of each strain were inoculated 1:50 in fresh MSgg medium and cultured, shaking at 37°C until reaching an OD<sub>600</sub> of 1.0. Aliquots of each culture were placed on 1.5% low-melting-point agarose MSgg pads and allowed to dry for 10 min before inverting the pads and placing them on a glass bottom dish (Willco Wells). Cells were imaged at room temperature using a Zeiss Axio-Observer Z1 inverted fluorescence microscope equipped with a Rolera EM-C2 electron-multiplying charge-coupled (EMCC) camera, enclosed within a temperature-controlled environmental chamber. Quantification was performed with Fiji and FIJI software (33,34).

Protein overproduction and purification

*E. coli* strains harboring expression vectors for 10xHis-SUMO-tagged protein sequences were cultured, shaking, overnight at 37°C. The following morning, the cultures were diluted in 500 ml fresh 2xYT supplemented with
Preparation of whole cell lysates

Overnight cultures of *B. subtilis* strains were cultured shaking at 37°C. The following morning, cultures were diluted into 500 ml fresh LB and grown at 37°C with shaking to OD600 ~0.8. 250 μM IPTG was then added to the cultures and cells were grown for an additional 40 min. Cells were harvested by centrifugation and concentrated 10× in 25 mM Tris–HCl, 100 mM NaCl, pH 8.0. Following addition of 1 mM PMSF, cells were sonicated, and lysates were then aliquoted and stored at -80°C.

Oligoribonucleotide labeling

Synthetic RNAs (2–7-mers) were purchased from TriLink Biotechnologies or Sigma-Aldrich. Each RNA was subjected to radioactive end-labeling or non-radioactive phosphorylation by T4 Polynucleotide Kinase (NEB). Each RNA was subjected to phosphorylation with equimolar concentrations of either [γ-32P]-ATP or ATP. T4 PNK and 1× T4 PNK Reaction Buffer. Reactions comprising a final concentration of either 0.5 μM 5'-32P]-radiolabeled RNA or 2.0 μM phosphorylated RNA were incubated at 37°C for 40 min, followed by heat inactivation of T4 PNK at 65°C for 20 min. These labeling reactions were then utilized for subsequent RNase cleavage assays, as described in text. Due to the short length of RNA substrates, labeling efficiency was not less than complete; residual [γ-32P]-ATP can be observed near the solvent front of denaturing PAGE gels.

Oligoribonucleotide and c-di-AMP cleavage reactions

Phosphorylated RNA or c-di-AMP (1.0 μM), including trace amounts of the respective radiolabeled substrate, were subjected to cleavage in vitro at room temperature by 100 nM purified NrnA, NrnA(A80D D156N), NrnA(EF), NrnA(EF), CnpB, CnpB, CnpB, CnpB, CysQ, CysQ, and CysQ. Reactions containing either physiological buffer or Pde2Sp. These reactions were conducted in the presence of 25 mM Tris, pH 8.0, 300 mM NaCl and 200 μM MnCl2. At the appropriate times, aliquots of the reaction were removed and quenched in the presence of 150 mM EDTA on ice and heat inactivated at 95°C for 5 min. The reactions in Figure 2D were conducted by subjecting trace amounts of radiolabeled RNA molecules 2–7, and 10 nucleotides in length (GG, AGG, (A)2GG, (A)GG, (A)GG, (A)GG, (A)GG) to cleavage in vitro by 20 nM of purified NrnA, or CysQ. Reactions containing either physiological buffer comprised 25 mM Tris, pH 8.0, 100 mM NaCl, 1 mM MgCl2, or 50 μM MnCl2, or a high manganese containing buffer composed of 25 mM Tris, pH 8.0, 100 mM NaCl, and 5 mM MnCl2. For reactions with *B. subtilis* whole cell lysates, trace amounts of radiolabeled RNA were added to lysates in the presence of 25 mM Tris–HCl pH 8.0, 300 mM NaCl, 200 μM MnCl2, and 25 mM MgCl2. At the appropriate times, aliquots of the reaction were removed and quenched in the
presence of 150 mM EDTA on ice and heat inactivated at 95°C for 5 min. All cleavage reactions were separated on denaturing 20% PAGE containing 1× TBE and 4 M urea. The gels were imaged using Cytiva Amersham Typhoon™ laser scanner platform and analyzed for the appearance of truncated 32P-labeled products. The intensity of the radiolabeled nucleotides was quantified using ImageJ software (34).

Differential radial capillary action of ligand assay (DRA-CALA)

Apparent equilibrium binding reactions were performed by incubating trace amounts (~1 nM) 5′-[32P] dinucleotide with increasing concentrations of purified NrnA(D80N D156N)Bs in binding buffer (25 mM Tris–HCl pH 8.0, 300 mM NaCl, 200 μM MnCl2) for 30 min at room temperature. For competition assays, 1 μM purified NrnA(D80N D156N)Bs was incubated with trace (~1 nM) 5′-[32P]AA in the presence of 10 or 100 μM unlabeled competitor, in binding buffer. Aliquots of all reactions were spotted onto a nitrocellulose membrane (GE) using a fixed replicator pin tool and allowed to air dry prior to being imaged on a Cytiva Amersham Typhoon™ laser scanner platform. Images were quantified using ImageJ software, and fraction bound was calculated based on previously described methods (34–36).

Swimming motility assays

Plates of LB supplemented with 0.2% (w/v) agar with or without 100 μM IPTG were prepared and left to dry at room temperature overnight. Simultaneously, cells were grown overnight in LB with or without 100 μM IPTG. 4 μL of each stationary phase culture was stab-inoculated into the semi-solid agar plate and left to incubate at 30°C for 12 h. The diameter of the bacterial migration was measured with a ruler.

2-Aminopurine hydrolysis assays

To determine the polarity of NrnA Bs, we used a fluorescence-based assay as described previously (37). This assay is based on the differential fluorescence output of the nucleotide analog 2-aminopurine. 2-Aminopurine shows reduced fluorescence when base-stacked with other nucleobases; however, free 2-aminopurine nucleotides exhibit increased fluorescence output. We monitored the fluorescence of 2-aminopurine generated from phosphodiester hydrolysis of synthetic 4-mer RNA substrates that were purchased from GE Healthcare Dharmaco. For this analysis, 100 nM of NrnA Bs was incubated with 10 μM of the RNAs containing a 2-aminopurine and a specific phosphorothioate modification. These reactions also contained 25 mM Tris–HCl, pH 8.0, 300 mM NaCl, 200 μM MnCl2. Reactions were conducted in a black 384-well plate using a Spectramax M5 plate reader, and fluorescence was measured every two min using an excitation wavelength of 310 and an emission wavelength of 375.

pAp hydrolysis assays

NrnA-like proteins were assayed for pAp phosphatase activity using the Sigma-Aldrich® Malachite Green Phosphate Assay Kit (MAK307). This assay was modeled on a previously described method, with slight modifications (38). A standard curve was generated using 1 mM free phosphate standard. All reactions were composed of 25 mM Tris–HCl pH 8.0, 300 mM NaCl, 100 μM pAp (Sigma), and either 200 μM of MnCl2 (for NrnA-like proteins), or 1 mM MgCl2 (for CysQ) and conducted at room temperature. The indicated protein was added to its respective reaction and allowed to incubate for the indicated reaction time of either 5 or 30 min. Reactions were then quenched with one fourth of a volume of acidic malachite green dye solution and incubated for 30 min to allow for color development. Absorbance values at 620 nm (A620) were measured using an Agilent Cary 60 UV–Vis spectrophotometer and correlated against the standard curve to determine the concentration of inorganic phosphate released in each reaction. Several reactions were repeated under the same assay conditions but with higher concentration of NrnA, as indicated in Table 1.

| Protein | Concentration of phosphate liberated from pAp (μM) |
|---------|----------------------------------------------------|
|          | Concentration of phosphate liberated from pAp (mean ± SD) (μM) |
|          | 5 min | 30 min |
| B. subtilis NrnA – 5 nM | ND | ND |
| B. subtilis NrnA – 50 nM | ND | 9.7 ± 0.2 |
| B. subtilis NrnA – 500 nM | 9.3 ± 0.9 | 43.7 ± 1.7 |
| E. faecalis NrnA – 5 nM | ND | ND |
| S. pyogenes Pde2 – 5 nM | ND | ND |
| M. tuberculosis CnpB – 5 nM | ND | ND |
| E. coli CysQ – 5 nM | 6.3 ± 0.7 | 31.4 ± 1.2 |

ND, not detected

Luciferase coupled enzyme assay: expression and purification of pyruvate phosphate dikinase

T7Express competent cells expressing His6-PPDK Cs were cultured to an OD600 of 0.6 at 37°C and induced with 1 mM IPTG for 2 h. Cells were harvested by centrifugation at 5000 rpm for 20 min. The cell pellets were weighed and resuspended in 1 g cell pellet per 10 mL of 25 mM Tris–HCl, pH 8.0, 1 M NaCl, 5% glycerol. To this mixture, PMSF was added to reach 1 mM and lysozyme to 0.25 mg/ml. Lysis proceeded by sonication on ice, then a 10% solution of polyethyleneimine was added to the sample to reach a final concentration of 0.5%. Lysates were clarified by two 15-minute rounds of centrifugation at 12 000 rpm. Next, PPDKCs was precipitated with ammonium sulfate. Solid ammonium sulfate was gradually added to the clarified lysate to reach a final concentration of 40% saturation using the 0-degree Celsius convention. The lysate was then left to rock on ice for 30 min, after which it was centrifuged at 10 000 rpm for 10 min. After centrifugation, more solid ammonium sulfate was added to a final concentration of 50% saturation. The lysate was left to rock on ice for 30 min, and then centrifuged at 10 000 rpm for 10 min. At this point, the supernatant was decanted, and the precipitated PPDKCs was resuspended in 25 mM Tris–HCl, pH 8.0, 300
mM NaCl, 5% glycerol. The protein sample was further purified by immobilization metal ion affinity chromatography using cOmplete™ His-Tag Purification Resin (Roche). The protein was incubated with equilibrated nickel resin for 1 h shaking gently on ice. Post incubation, the resin was washed with 10 column volumes of 25 mM Tris–HCl, pH 8.0, 300 mM NaCl, 5% glycerol 10 mM imidazole. Next, the resin was washed with 10 column volumes of 25 mM Tris–HCl, pH 8.0, 300 mM NaCl, 5% glycerol, 25 mM imidazole. Finally, PPDK<sub>Cs</sub> was eluted from the column using 5 column volumes of 25 mM Tris–HCl, pH 8.0, 300 mM NaCl, 5% glycerol, 250 mM imidazole. Purified PPDK<sub>Cs</sub> was dialyzed to remove excess imidazole. Then, purified PPDK<sub>Cs</sub> was concentrated using Corning® Spin-X™ UF concentrators, and small aliquots of protein were flash frozen in liquid nitrogen and stored at −80°C.

**Luciferase coupled enzyme assay: determining the enzyme units of purified pyruvate phosphate dikinase**

The enzymatic activity of NrnA<sub>Bs</sub> was assessed using a slightly modified continuous kinetic assay as described previously (39). For this analysis, AMP generated from phosphodiester hydrolysis of RNA substrates by NrnA<sub>Bs</sub> was converted to ATP by Pyruvate Phosphate Dikinase (PPDK<sub>Cs</sub>). In these reactions, ATP was then used by Luciferase enzyme to generate a luminescence signal. A 2× coupling buffer was used for these assays and was comprised of 100 mM Tris–HCl, pH 8.0, 200 mM NaCl, 2 mM phosphoenolpyruvic acid, 2 mM sodium pyrophosphate, 15 mM (NH₄)₂SO₄. The 2× coupling buffer was filter sterilized and kept frozen. The enzymatic activity of dilutions of PPDK<sub>Cs</sub> were determined by monitoring the conversion of 100 μM AMP into ATP, using luminescence, and correlating ATP generation to a standard curve. Prior to the enzyme assay, 200 μL of fresh ATPlite substrate mix was added to 1 mL of the thawed 2X coupling buffer. The enzyme assays were conducted in the presence of the 1× coupling buffer containing the ATPlite mix, 10 mM MgCl₂, 0.5 mM MnCl₂, >50 units of PPDK<sub>Cs</sub> and the respective concentrations of the RNA substrates and purified NrnA<sub>Bs</sub> at 25°C. To ensure that there was a stable luminescence signal, we monitored the reactions containing all assay components except for the purified protein for 1–2 min prior to the addition of purified protein. For each concentration a blank was used that included all assay components except NrnA<sub>Bs</sub>, to control for background luminescence signal. Luminescence signals were measured continuously in a white 384-well plate using a Spectramax M5 plate reader. The luminescence signals were fit to standard curves to determine the concentrations of AMP released. For the degradation of AA, we divided the rates of AMP release by two since two molecules of AMP are liberated for each reaction.

**Mass spectrometry proteomics analysis**

Wt and Δnra strains were grown overnight and then subcultured followed by growth to an OD of 0.8. Cells were pelleted, supernatant removed and then resuspended in 5% SDS with 100 mM triethylammonium bicarbonate (TEAB). Protein abundance was acquired by using a Thermo Scientific Nanodrop 8000 spectrophotometer for sample normalization. Sample reduction in 200 mM DTT was performed for 1 hour at 57°C. Samples were then alkylated with 500 mM iodoacetamide for 45 min in the dark at RT and quenched by addition of 200 mM DTT. The S-Trap™ Micro Spin Columns (ProtiFi) preparation was performed according to manufacturer’s protocols on 200 μg of protein (1:25, Trypsin/Lys-Cl-Protein Promega Cat.# VS073). Digested peptides were dried in a cold-trap vacuum centrifuge, and resuspended in 95.5 water:acetonitrile, 0.1% v/v formic acid was added to the samples and stored at −80°C until analysis. Samples were analyzed on a Thermo Orbitrap Eclipse Tribrid mass spectrometer equipped with a Thermo Ultimate 2500 UPLC and FAIMS unit operating under 3 different CVs (−50, −65 and −85). The liquid chromatography buffers used were mass spectrometry grade reagents (H₂O and 0.1% formic acid for buffer A and 95% acetonitrile with 5% H₂O and 0.1% formic acid for buffer B). A 75 cm Thermo Fisher Easy-Spray column was utilized for separation. A 200-minute gradient was used for separation at 0.3 μl/min flow rate (0 min 5% B, 10 min 10% B, 150 min 70% B, 158 min 80% B, 161 min 90% B, 171 90% B, 172 min 5% B, 200 min 5% B). MS1 was collected in the orbitrap at a resolution of 240 000 scanning from 250–2000 m/z with a standard AGC target and MIPS active. An intensity threshold of 2.04e⁴ was utilized along with charge state inclusion between 2 and 7. Dynamic exclusion was half to 60 s with a mass tolerance of 5 ppm. Data dependent MS2 were acquired in the ion trap, operating under the turbo scan rate with HCD activation at a 28% collision energy.

**Mass spectrometry data analysis**

Data analysis was performed with Metamorpheous Version 0.0.320 (40). The database search was performed against the *B. subtilis* proteome from Uniprot (41) (Downloaded 5/11/2021). Search parameters included full Trypsin digestion, 2 maximum miss cleavage events, 10 ppm precursor mass tolerance and 0.6 Da fragment tolerance. Dynamic modifications of oxidation on methionine and acetylation on the protein N-terminus were allowed. A static modification of carboxymethylation was added to cystines. The FlashLFQ algorithm was utilized for relative label free quantification (42). Q-values and false discovery rate (FDR) were calculated using Percolator 3.0 (43). Results were filtered by requiring >0 unique peptides per protein, >1 PSMs per protein, and protein FDR ≤0.01.

**Collection of DHH-DHHA1 protein sequences and subsets based on taxonomic class**

Within the DHH phosphoesterase superfamily (Pfam: PF01368), there is a vast collection of protein subfamilies whose members can be distinguished by the presence of at least one additional (to the DHH domain), subfamily-specific domain. We were interested in instances where the DHH domain co-occurs with an adjacent DHH-associated domain, DHHA1 (Pfam: PF02272). A set of sequences whose domain architecture included the ordered combination of DHH and DHHA1 domains was collected by searching against the UniProtKB database using hhm-
search with a relaxed inclusion threshold, $E$-value $\leq 1E-1$, and passing pre-trained profile hidden Markov models (HMMs) for each domain as the input. The resultant set consisted of only sequences in which there are subsequences that correspond to both DHH and DHHA1 domains. While the presence of these two domains were required, we permitted sequences that contained additional domains. Given that two subsequences are common across all sequences in the set, these subsequences can be concatenated for each sequence. Let the set of chimeric sequences be denoted by $S = \{ S_1, S_2, \ldots, S_n \}$. In $S (n = 63,390)$, we examined the organisms in which DHH/DHHA1 sequences are encoded and determined their prevalence within various taxonomic groups (classes). The DHH/DHHA1 proteins are phylogenetically pervasive. We selected two subsets of $S$ from specific taxa—Bacilli ($n = 11,068$), and Actinomycetia ($n = 1,627$)—to subsequently perform comparisons within subsets of sequences to assess sequence diversity that is driven by functional specialization rather than resulting from phylogeny (shared ancestry).

**CNYM method-based approach to detecting sequence clusters**

Suppose a subset of sequences $T = \{ n \in S \mid n$ is encoded in a specific taxon (class) $\}$; additionally, the sequences in $T$ are filtered using CD-Hit such that there is at most 80% pairwise sequence identity. We performed an all-against-all, pairwise BLAST search to quantify the similarity between pairs of aligned sequences, denoted by $T_i$ and $T_j$. The $E$-value of high scoring segment pairs from BLASTP were tabulated in a weighted adjacency matrix $I$; here, $I_{ij}$ is the $E$-value associated to the comparison of sequences $T_i$ and $T_j$ (for instance, $I_{ij} = 0$ indicates $T_i$ and $T_j$ are identical). This measure of pairwise similarity is not symmetric, therefore we took the lower $E$-value between pairs such that $I$ is symmetric. The matrix $I$ was used to construct a weighted, undirected graph which layout is based on assigning weights to the edges that correspond to the distance between sequences (nodes), then implemented stress majorization to find an ideal configuration of nodes. Since values in $I$ are the $E$-values, the edges between more similar sequences (indicated by lower $E$-values) are shorter. The neat utility from the Graphviz package was used for drawing graphs. Communities, or clusters, of similar sequences in the graph are detected by Clauset- Newman–Moore (CNM) modularity maximization. The greedy modularity communities function from the NetworkX Package was used for hierarchical and agglomerative partitioning of the sequences into distinct clusters, which we labeled based on the annotation of characterized representatives.

**RESULTS**

**Deletion of B. subtilis nrnA results in accumulation of the dinucleotide GG and cyclic di-GMP**

The recent re-evaluation of Orn's function as primarily acting as a diribonuclease (11) led us to reconsider the individual roles NrnA$_{Bs}$ and NrnB$_{Bs}$, have in vivo. Orn's role as a diribonuclease has implications in cyclic di-guanosine monophosphate (c-di-GMP) signaling, as the dinucleotide GG is an intermediate in the two-step conversion of c-di-GMP to two GMP molecules. The ability of B. subtilis nrnA and nrnB to complement the E. coli and P. aeruginosa $\Delta$orn mutants suggested that these enzymes are responsible for GG cleavage in B. subtilis in a manner that is analogous to Orn function (13–15). It has been shown previously that an increase in GG leads to a concomitant increase in c-di-GMP through feedback inhibition. Previously, a double $\Delta$nrnA$\Delta$nrnB mutant was assayed for c-di-GMP levels using a fluorescent c-di-GMP riboswitch reporter (13,28,44). This revealed that levels of c-di-GMP, and consequently GG, were higher in this strain compared to wild-type. However, in this study, we sought to determine the relative contributions of GG turnover from NrnA$_{Bs}$ and NrnB$_{Bs}$ individually. Therefore, c-di-GMP levels were assayed using the same c-di-GMP-responsive riboswitch-yfp reporter but for B. subtilis wild-type, $\Delta$nrnA, or $\Delta$nrnB cellular backgrounds (Figure 1A). The c-di-GMP-responsive riboswitch decreases downstream gene expression in response to an increase in c-di-GMP; therefore, YFP expression is inversely coupled to c-di-GMP abundance in this strain. Assessment of a constitutive yfp control reporter showed mild to no differences in fluorescence between the wild-type strain and the single $\Delta$nrnA or $\Delta$nrnB mutants (Figure 1B, D). In agreement with our prior observations (28), assessment of the c-di-GMP reporter in a wild-type B. subtilis population showed a bimodal distribution of fluorescence where one population exhibited high c-di-GMP (low YFP fluorescence) and the second exhibited low c-di-GMP (high YFP fluorescence) (Figure 1C, E). In the $\Delta$nrnA background, the riboswitch reporter exhibited very low fluorescence compared to the wild-type, indicating that cyclic di-GMP levels, and, by extension, GG levels, are higher in this strain. This observation is consistent with a role in turnover of GG for NrnA$_{Bs}$. However, a bimodal distribution of fluorescence was observed from the riboswitch-yfp reporter in the $\Delta$nrnB strain, similar to wild-type (Figure 1C, E). This suggests that the NrnB$_{Bs}$ does cleave GG during late exponential phase of growth (OD$_{600}$ $\sim$1.0), accumulation of the dinucleotide in the absence of NrnB$_{Bs}$ is not enough to promote feedback inhibition of enzymes that linearize c-di-GMP. Alternatively, NrnB$_{Bs}$ might not be expressed during exponential growth phase.

To further investigate how NrnA can affect B. subtilis c-di-GMP pools, we assessed motility, which is regulated by c-di-GMP in many bacteria (45). Prior experiments in B. subtilis have shown that elevated intracellular c-di-GMP levels lead to an inhibition of swarming motility—a social form of surface migration (46,47). B. subtilis uses flagella for swimming through liquid as well (47). Fortifying media with 0.2% agar allows the pores in the agar to be sufficient large enough to discourage swimming over the surface, but rather, permit swimming. The $\Delta$nrnA mutant exhibited a clear swimming defect compared to wild-type, further suggesting that intracellular c-di-GMP levels are elevated in this strain (Figure 1F). In contrast, the $\Delta$nrnB mutant had a phenotype similar to wild-type. Together, these data show that NrnA$_{Bs}$ but not NrnB$_{Bs}$ is important for maintaining cellular GGP pools during exponential phase growth.
NrnA<sub>Bs</sub> cleaves RNA substrates 2-4 nucleotides in length <em>in vitro</em>

To fully understand its length preferences, purified NrnA<sub>Bs</sub> was tested for cleavage of different RNA substrates. NrnA<sub>Bs</sub> was incubated with 5′-3′P-radiolabeled oligoribonucleotides of varying lengths and at substrate concentrations that exceeded enzyme concentration (10:1). These reactions also contained divalent cations for supporting catalysis. The products of these reactions were analyzed by urea-denaturing 20% polyacrylamide gel electrophoresis (PAGE). Analysis of NrnA<sub>Bs</sub> showed that it fully processed the dinucleotide AA into nucleoside monophosphates in 5 min (Figure 2A, B). Importantly, NrnA<sub>Bs</sub> also showed similar activity against a 3-mer RNA (AGG) within the same timeframe. While some processing of a 4-mer (AAGG) was observed in the first 5 min, the rate of cleavage appeared to be slower than that of the 3-mer or dinucleotide. This contrasts with a 5-mer RNA, which remained generally unprocessed after the first 5 min but was processed over longer time points.

NrnA<sub>Bs</sub> does not cleave longer RNA substrates <em>in vitro</em>

Using the same reaction conditions that were used to test short RNA substrates, we incubated NrnA<sub>Bs</sub> with 6-mer, 7-mer, 10-mer and 15-mer substrates. This revealed no detectable cleavage activity against the longer substrates (Figure 2B, C). Indeed, when NrnA<sub>Bs</sub> was simultaneously incubated with radiolabeled RNAs of 2-7, and 10 nucleotides in length, only the RNAs from 2-4 nucleotides in length were processed (Figure 2D). These data mildly conflict with prior reports that NrnA<sub>Bs</sub> degrades longer nucleic acid substrates greater than four nucleotides in length (15,19,21).
Figure 2. *B. subtilis* NrnA has a preference for hydrolyzing short RNAs 2–4 nucleotides in length, and acts with a 5′-3′ polarity. (A–C) 1 μM of RNA substrates containing trace amounts of 32P-radiolabeled RNA were incubated with 100 nM of purified NrnABs. The RNA substrates ranged from 2-7, 10 or 15 nucleotides in length. Aliquots were removed at various time points and resolved by denaturing PAGE. The cleavage data were graphed as the normalized intensity of the initial substrate, plotted as the average and SD of three independent experiments in (A and C). (D) Trace amounts of different 32P-radiolabeled RNA substrates that ranged in size (including 2–7, and 10-mer RNAs) were simultaneously incubated with 20 nM of purified NrnABs or OrnEc. These reactions also contained either low or high concentrations of manganese. Aliquots were removed at time intervals and quenched in 150 mM EDTA and 4 M urea. Degradation products were resolved by denaturing PAGE. (E) In these reactions, some RNA substrates contained a non-hydrolyzable, phosphorothioate-modified backbone (placement of the linkage represented by ps). 1 μM of RNA was incubated with trace amounts of 32P-radiolabeled RNA and 100 nM of purified NrnABs. Aliquots were removed at various time points and resolved by denaturing PAGE. (F) Fluorescence emission changes of 10 μM RNA substrates containing both a 2-aminopurine and an internal phosphorothioate linkage, upon incubation in the presence or absence of 100 nM NrnA.
Our data strongly suggest that NrnA<sub>Bs</sub> preferentially degrades short RNAs four nucleotides and smaller in length, and that the previously reported activities against longer substrates may have arisen due to the stochiometric excess of enzyme and/or long incubation times used in the prior assays. Therefore, we speculate that long RNA processing is not likely to be a biologically relevant function of NrnA<sub>Bs</sub>. Additionally, these data indicate that the processing of 2–4 mer RNAs by NrnA<sub>Bs</sub> is only modestly affected by increasing the concentration of Mn<sup>2+</sup>. Interestingly, when Orn<sub>Ec</sub> was simultaneously incubated with radiolabeled RNAs 2–7, and 10 nucleotides in length, dinucleotides are rapidly processed, and only minimal trinucleotide processing was seen over the longer time points (Figure 2D). In total, these data demonstrate that under reaction conditions where Orn<sub>Ec</sub> displays diribonuclease activity, NrnA<sub>Bs</sub> instead displays a specific preference for short RNAs between 2–4 nucleotides in length.

### NrnA<sub>Bs</sub> cleaves RNA substrates from their 5′ terminus

Interestingly, processing of short RNAs by NrnA<sub>Bs</sub> resulted in release of nucleoside monophosphates but did not result in the appearance of degradation intermediates. This suggests that cleavage activity does not occur in a 3′–5′ direction as intermediates would be expected for RNAs containing a radiolabeled phosphate at the 5′ terminus. This observation agrees with certain prior investigations on NrnA<sub>Bs</sub> that showed that cleavage of 5′-32P-AAA rapidly resulted in release of 32P-AMP (21), which would be the expected outcome if NrnA<sub>Bs</sub> were removing the terminal phosphate from the 5′ direction. Yet, NrnA proteins have also been described as a 3′ to 5′ exonuclease (15,21,23,48). Therefore, NrnA<sub>Bs</sub>′s directionality remains unclear. To test the directionality of NrnA<sub>Bs</sub>, cleavage under our standard assay conditions, we assessed NrnA<sub>Bs</sub> activity against 5′-32P-radiolabeled 4-mer RNAs that contained a putatively non-hydrolyzable phosphorothioate linkage at different locations. As anticipated, cleavage of a native 4-mer resulted in release of 32P-AMP (Figure 2E). However, NrnA<sub>Bs</sub> also released 32P-AMP upon incubation with a 4-mer containing a phosphorothioate linkage at its 3′ terminus. Furthermore, cleavage of a 4-mer with a phosphorothioate linkage in the middle bond of the RNA (32P-AA<sub>Ps</sub>GG) did not result in 32P-AMP, which would have been expected if NrnA<sub>Bs</sub> were cleaving from the 3′ direction, but instead again resulted in release of 32P-AMP (Figure 2E). The 5′-32P-AGG substrate could not be assessed in this analysis, as it was not successfully radiolabeled by polynucleotide kinase. However, since this assay did not rule out the possibility that NrnA could process through the phosphorothioate linkage, we utilized a second approach to assess termini selectivity. Specifically, we determined whether NrnA<sub>Bs</sub> would release a 2-aminopurine (2AP) nucleotide from either the 5′ terminus or 3′ terminus for 4-mer RNA substrates containing an internal phosphorothioate linkage (i.e. (2AP)A<sub>Ps</sub>GG or A<sub>Ps</sub>GG(2AP)). Free 2AP is known to exhibit an increased level of intrinsic fluorescence relative to 2AP contained within an RNA polymer, due to base stacking of the latter. When NrnA<sub>Bs</sub> was incubated with (2AP)A<sub>Ps</sub>GG, an increase in 2AP fluorescence was observed over time (Figure 2F). In contrast, no fluorescence increase was observed upon incubation of NrnA<sub>Bs</sub> with A<sub>Ps</sub>GsG(2AP). Taken together, these assays suggest that NrnA<sub>Bs</sub> is exclusively a 5′-3′ exoribonuclease for short RNA substrates.

### NrnA<sub>Bs</sub> preferentially binds short RNAs

We next sought to correlate NrnA<sub>Bs</sub> cleavage preferences with binding affinities by using Differential Radial Capillary Action of Ligand Assay (DRaCALA) (35). Structural analyses of NrnA<sub>Bs</sub> and other NrnA-like proteins have demonstrated that the NrnA active site incorporates four aspartate residues (NrnA<sub>Bs</sub>D24, D26, D80, D156), and three histidine residues (NrnA<sub>Bs</sub>H20, H103, H104). Catalysis by NrnA and other DHH domain-containing proteins are coordinated by a two-ion metal mechanism, in which manganese is the preferred ion for catalysis (19,21,48), although magnesium can also support catalysis in vitro. Herein, we incubated NrnA<sub>Bs</sub> with 5′-radiolabeled AA in the presence of magnesium or manganese and assayed for protein-RNA complexes by DRaCALA (Figure 3A, B). We did not observe evidence of binding activity, most likely because the RNAse enzyme is active under these conditions, thereby processing the dinucleotides into nucleoside monophosphates. We also did not observe binding activity of 5′-radiolabeled AA with NrnA<sub>Bs</sub> that had been treated with the divalent cation chelator EDTA (Figure 3A). NrnA<sub>Bs</sub> is not catalytically active in the presence of excess EDTA; therefore, the lack of RNA-binding activity under these conditions may suggest that divalent ions are required for substrate binding. We therefore sought to investigate binding of AA to a catalytically inactive mutant. To that end, we mutated two of the highly conserved aspartate residues to asparagine (D80N D156N) (49). This mutant protein exhibited no detectable RNase cleavage activity (Supplementary Figure S1B), although it demonstrated binding activity to AA when incubated with supplemental divalent metals (Figure 3A). A survey of varying dinucleotide concentrations revealed that NrnA<sub>Bs</sub>D80N D156N bound GG best, with an equilibrium dissociation constant (K<sub>d</sub>) of 210 nM, while other dinucleotides (GA, AA, UA and CA) exhibited modestly poorer binding affinities (Figure 3A, B). The poorest binding affinities were observed for UA and CA, which were roughly an order of magnitude higher than GG. We were surprised at the differences in these apparent equilibrium binding affinities; therefore, we tested for their cleavage by wild-type NrnA<sub>Bs</sub> in vitro (Figure 3C). This showed that the dinucleotides were all equally good substrates for NrnA<sub>Bs</sub>, with only minor differences in their rates of cleavage. As it was previously reported that NrnA<sub>Bs</sub> could exhibit activity against single-stranded DNAs (19), we also compared cleavage of a DNA dinucleotide alongside the RNA dinucleotides. This revealed that NrnA<sub>Bs</sub> is indeed fully active against a DNA dinucleotide substrate (Figure 3C). From these aggregate data, we conclude that NrnA<sub>Bs</sub> generally cleaves dinucleotides but with a modest preference for purine-containing substrates.

We then exploited the binding activity of the double aspartate mutant to survey for recognition of different putative substrates. Specifically, we determined the fraction...
Figure 3. *B. subtilis* NrnA possesses a dinucleotide substrate length preference with a slight preference for 5′ purine residues. (A, B) DRaCALA, as described elsewhere (35), was used to measure binding of different radiolabeled dinucleotides to the catalytically inactive mutant NrnA<sub>Bs</sub> D80N D156N. (C) For cleavage assays, 1 μM of the indicated dinucleotide substrate was incubated with trace amounts of <sup>32</sup>P-radiolabeled RNA and 100 nM of NrnA Products were removed at time intervals and resolved by denaturing PAGE. Quantification of the normalized radioactive intensity of the initial substrate was plotted as the average and SD of three independent experiments. (D) Substrate binding competitions were conducted by DRaCALA. 1 μM of purified NrnA<sub>Bs</sub> D80N D156N was bound to ∼1 nM of <sup>32</sup>P-labeled AA, and subsequently incubated with either 10 or 100 μM of the indicated unlabeled competitor molecules. The fraction bound was normalized to the maximum binding exhibited by NrnA<sub>Bs</sub> D80N D156N to <sup>32</sup>P-radiolabeled AA. (D, E) Kinetic analysis of *B. subtilis* NrnA was performed using a previously published continuous kinetic assay (39), with slight modifications. The continuous coupled enzyme assay depended on the liberation of AMP from the RNA substrate by NrnA<sub>Bs</sub>, followed by the conversion of AMP to ATP by PPDK<sub>Cc</sub>, which was then used by luciferase to generate a pulse of light. Since the hydrolysis of AA results in the generation of 2x AMP molecules, the initial rates for AA hydrolysis were divided by 2. Substrate degradation was restricted to 10% or less, and a blank containing all reaction materials except for NrnA<sub>Bs</sub> was included as a control for background luminescence. (B–F) All data was plotted as the average and SD of three independent experiments.
bound of radiolabeled AA when co-incubated in the presence of unlabeled (10 or 10 µM) putative substrates (Figure 3D). As anticipated, the unlabeled dinucleotides AA and GG competed well against 5′-radiolabeled AA. However, ApA and GpG did not compete for binding, suggesting that a 5′ phosphate is likely to be critical for binding. Indeed, an adenosine with phosphates at its 5′ and 3′ termini (pAp) was unable to compete at 10 µM but could partially compete against AA when supplied at 100 µM, suggesting that pAp exhibits reduced but measurable binding affinity for NrnA<sub>Bs</sub>. pAp is a cytoplasmic carrier of sulfate when converted to pApS (50); however, pApS was unable to compete for binding to NrnA<sub>Bs</sub>. Other phosphorylated mononucleotides (e.g. ATP, AMP) were also unable to compete for binding. Finally, cyclic di-nucleoside monophosphates (c-di-AMP and c-di-GMP) were unable to compete against binding of AA. These data reveal that the NrnA<sub>Bs</sub> binding pocket is best suited for canonical dinucleotides and is unlikely to act directly on nucleotide signals.

**Kinetic analysis of dinucleotide and trinucleotide substrates against NrnA<sub>Bs</sub>**

While NanoRNase enzymes have been argued to cleave a range of short RNA substrates, the Orn enzyme is suspected in some bacteria to strongly prefer cleavage of dinucleotides under physiological conditions. When correspondingly tested the cleavage of NrnA<sub>Bs</sub> against 2-5 mer RNA substrates at potentially physiologically relevant concentrations the rate of cleavage for 2-4 mers was similar (Supplementary Figure S2). The concentrations used in this analysis likely represent a second order rate equation which could be useful for determining the substrate specificity of NrnA<sub>Bs</sub>. From these data, we speculated that NrnA<sub>Bs</sub> might possess a slight dinucleotide specificity. Previous reports regarding the specificity of NrnA-like proteins have led to different conclusions as to whether NrnA-like proteins are specific for dinucleotides or other short RNA substrates. NrnA<sub>Bs</sub> was originally reported to possess a preference for 3-mer substrates in cleavage assays using 5′-Cy5 modified oligonucleotides (15). However, our data in Figure 2D strongly suggests that the 5′ phosphate plays a crucial role in ligand binding. Therefore, these data led us to believe that the 5′-Cy5 modified RNAs could have influenced the reported enzyme activity. Also, the NrnA-like protein from *Thermatoga maritima* was previously reported to possess a dinucleotide specificity based on the protein’s reduced activity in cleavage assays against the trinucleotide AAA (49). Hence, we wanted to acquire additional quantitative data directly comparing the kinetics of cleavage for a dinucleotide and trinucleotide substrate. To that end, we modified a luminescence-based, coupled enzyme assay for measuring release of AMP from either an AA or ACC substrate (see Materials and Methods). These data showed that NrnA<sub>Bs</sub> displayed a Michaelis–Menten kinetic profile (Figure 3E, F). This revealed that the Km values differed modestly between the two RNAs (0.090 versus 0.325 mM). However, a direct comparison of the specificity constants (k<sub>cat</sub>/K<sub>m</sub>) for these reactions did not differ significantly. Therefore, NrnA<sub>Bs</sub> acts efficiently against both di- and trinucleotide RNA substrates.

**Substrate preferences of NrnA<sub>Ef</sub> and Pde2<sub>Sp</sub> closely resemble NrnA<sub>Bs</sub>, but not CnpB<sub>Mt</sub>**

Several different catalytic functions have been attributed to the NrnA-like *Mycobacterium tuberculosis* CnpB protein, including but not limited to cleavage of short and long RNA substrates and processing of c-di-AMP (23, 25, 26, 51). The streptococcal Pde2 protein has been implicated in processing of c-di-AMP but has not yet been assessed for cleavage of linear RNAs (52). Yet, the sequence identity between these proteins and NrnA<sub>Bs</sub> is reasonably high at 45% for *S. pyogenes* Pde2 and 25% for the *M. tuberculosis* CnpB. Therefore, we wanted to directly compare the substrate preferences of these enzymes (CnpB<sub>Mt</sub>, Pde2<sub>Sp</sub>) against NrnA<sub>Bs</sub> using the same substrates and reaction conditions. To further extend these analyses, we also included an NrnA-like proteins from *Enterococcus faecalis*, which has not yet been investigated. Herein, we incubated purified NrnA<sub>Bs</sub>, CnpB<sub>Mt</sub>, NrnA<sub>Ef</sub> and Pde2<sub>Sp</sub> with a variety of short RNA substrates. Each protein was incubated with 5′-radiolabeled oligoribonucleotides between 2-5 nucleotides in length and the reaction products were resolved by 20% denaturing PAGE (Figure 4). From these data, it is apparent that NrnA<sub>Ef</sub> and Pde2<sub>Sp</sub> process short RNAs between 2-4 nucleotides in length, remarkably like NrnA<sub>Bs</sub>. Moreover, both proteins fully processed the dinucleotide AA into nucleoside monophosphates within the same approximate time frame as NrnA<sub>Bs</sub> (Figure 4A–D). A trinucleotide RNA substrate was also fully processed by both proteins within a time frame that resembles dinucleotide substrate. Interestingly, while CnpB<sub>Mt</sub> was able to fully process AA to NMPs within 30 min, the overall rate of cleavage appeared to be slower than that of the other proteins (Figure 4E, F). Also, an assessment of longer substrates showed essentially no cleavage activity for CnpB<sub>Mt</sub> against them (Supplementary Figure S3). Taken together, the NrnA<sub>Ef</sub> and Pde2<sub>Sp</sub> proteins appear to exhibit substrate preferences that closely resemble NrnA<sub>Bs</sub>, while CnpB<sub>Mt</sub> exhibits lowered activity against linear RNAs.

**NrnA<sub>Bs</sub> does not act on cyclic di-AMP**

While our binding analysis of the NrnA<sub>Bs</sub>D80N D156N mutant suggested that the protein was not likely to bind c-di-NMPs, other DHH-DHHHA1 domain-containing proteins have been observed to specifically hydrolyze c-di-AMP; therefore, we wanted to directly test for this activity using the same reaction conditions as for cleavage of small linear RNA substrates. GdpP is one of the DHH-DHHHA1 domain-containing proteins that has been implicated in cleavage of c-di-AMP. Unlike NrnA, which only contains the DHH-DHHHA1 domains, GdpP also includes an N-terminal transmembrane domain, a heme-binding PAS domain, and a degenerate GGDEF domain (53). However, in contrast to the architectural complexity of GdpP, *M. tuberculosis* CnpB only contains the DHH-DHHHA1 domains but has also been reported to exhibit c-di-NMP phosphodiesterase activity (25, 26). Yet, this observation has been debated and, akin to NrnA<sub>Bs</sub>, multiple other activities have been attributed to CnpB<sub>Mt</sub> such as cleavage of RNAs from 2-5 nucleotides in length, with an Orn-like preference for dinucleotides (23). Furthermore, despite show-
ing c-di-AMP hydrolysis activity in vitro, it was recently reported that expression of cnpB in vivo is entirely independent of c-di-AMP (54). In addition to CnpB, Pde2 has also been linked to homeostasis of cyclic dI-NMPs. Several studies have suggested that Pde2 can process c-di-AMP in vitro and can preferentially hydrolyze AA to AMP (18,52).

To determine whether NrnA-like proteins had the ability to cleave c-di-AMP, NrnA<sub>Br</sub>, CnpB<sub>Mt</sub>, NrnA<sub>EF</sub> and Pde2<sub>Sp</sub> were purified and then tested for cleavage of <sup>32</sup>P-radiolabeled c-di-AMP. All four proteins were assayed under the same conditions, which resembled the reaction conditions employed for analysis of dinucleotides. When the products of these reactions were analyzed by urea-denaturing 20% polyacrylamide gel electrophoresis (PAGE), only CnpB demonstrated an ability to directly hydrolyze c-di-AMP to AMP (Figure 5A, C); no cleavage activity was detected for NrnA<sub>Br</sub>, NrnA<sub>EF</sub> and Pde2<sub>Sp</sub>. To ascertain whether this activity may be representative of CnpB-like proteins, we also purified CnpB homologs from Mycobacterium smegmatis, Mycobacterium avium and Rhodococcus ruber and then assayed for cleavage of c-di-AMP. As a positive control, we purified and analyzed the cytoplasmic portion of B. subtilis GdpP, which is known to hydrolyze c-di-AMP (53). All CnpB-like proteins exhibited robust hydrolysis of c-di-AMP (Figure 5B). Together, these data suggest that CnpB<sub>Mt</sub> may broadly feature active site residues that make it functionally different from

Figure 4. Degradation profiles of short RNAs incubated with NrnA-like proteins. (A, C, E) 100 nM of the respective protein was incubated with 1 μM of RNA substrate, 2–5 nucleotides in length, containing trace amounts of <sup>32</sup>P-radiolabeled substrates. Aliquots from reactions were removed at the indicated times and quenched in 150 mM EDTA and 4 M urea, and degradation products were resolved by denaturing PAGE. (B, D, F) The fraction of initial substrate was plotted over time as the average and SD of three independent experiments.
Figure 5. NrnA-like proteins differ in their capacities to process c-di-AMP. (A, B) 100 nM of the respective protein was assayed for c-di-AMP phosphodiesterase activity against 1 μM of substrate containing trace amounts of 32P-labeled c-di-AMP. Aliquots were removed from the individual reactions and were quenched in 150 mM EDTA and 4 M urea. Degradation products from the reactions were resolved by denaturing 20% PAGE. (C) The fraction of initial substrate was plotted over time as the average and SD of three independent experiments.
other NrnA-like proteins, which do not directly process c-di-AMP.

An assessment of 3’phosphoadenosine-5’-phosphate (pAp) cleavage activity

It has also been reported that NrnA<sub>Ec</sub> and CnpB<sub>Mt</sub> might act as pAp phosphatases, which affect sulfur metabolism by altering homeostasis of 3’-phosphoadenosine-5’-phosphosulfate (pApS) (15,23). However, we found that pAp and pApS were only moderate and poor binding competitors for the NrnA<sub>Ec</sub> enzyme, respectively (Figure 3D). Therefore, to directly investigate this question, we surveyed each of the NrnA-like proteins for pAp hydrolysis under the same reaction conditions and alongside a positive control. Specifically, we purified <i>E. coli</i> CysQ, a known pAp phosphatase, as a positive control for these assays (10). Then, purified NrnA<sub>Ec</sub>, CnpB<sub>Mt</sub>, NrnA<sub>EF</sub> and Pde2<sub>Sp</sub> were each assayed for pAp phosphatase activity in the presence of divalent cations. Each protein was incubated with 100 μM pAp and the concentration of free phosphate generated in the reaction was determined via a malachite green detection assay. Under these conditions, 5 nM CysQ<sub>EF</sub> generated 6.3 and 31.4 μM free phosphate in 5 and 30 min, respectively (Table 1). However, free phosphate was not detectable following incubation with pAp for any of the NrnA-like proteins at these enzyme concentrations (NrnA<sub>Ec</sub>, CnpB<sub>Mt</sub>, NrnA<sub>EF</sub> and Pde2<sub>Sp</sub>). However, when the concentration of NrnA<sub>Ec</sub> was increased 10× and 100×, phosphate release increased. This observation may explain previously published reports on pAp hydrolysis (15,23).

Based on the general lack of robust phosphatase activity coupled with the inability of pAp to compete for binding against linear AA (as seen in Figure 3D), we speculate that NrnA-like proteins are unlikely to preferentially act on pAp under physiological conditions.

NrnA<sub>Ec</sub> is required for processing of short RNAs <i>ex vivo</i>

Deletion of <i>nrnA</i> does not appear to affect growth in rich medium (Supplementary Figure S4); however, our in vitro data suggested that NrnA preferentially cleaves nuclear monophosphates by wild-type lysates. It also showed that Δ<i>nrnB</i> lysates generated a degradation profile that was identical to wild-type (Figure 6A). In contrast, there was specific accumulation of 2–4-mers when the 10-mer RNA was incubated with <i>nrnA</i> lysates, concurrent with reduced levels of nucleoside monophosphates (Figure 6A). This observation suggests that NrnA<sub>Ec</sub> is the primary RNase responsible for cleaving 2–4-mers during late exponential phase and agrees well with our biochemical assessment of the purified proteins. We then integrated an inducible copy of <i>nrnA<sub>Ec</sub></i>, <i>nrnB<sub>Ec</sub></i>, <i>orn</i> or <i>cysQ<sub>Ec</sub></i> into a nonessential locus of the Δ<i>nrnA</i> strain and prepared lysates for RNA degradation (Figure 6B). An empty vector control was included in this analysis. Δ<i>nrnA</i> complemented with an empty vector resembled the Δ<i>nrnA</i> strain, showing again an accumulation of 2–4-mers. In contrast, bands corresponding to 2- to 4-mers decreased upon complementation with <i>nrnA</i>, resembling lysates from the wild-type strain. Although Δ<i>nrnB</i> lysates did not show any change in degradation of the 10-mer RNA, complementation of the Δ<i>nrnA</i> strain with <i>nrnB</i> surprisingly resulted in complete processing of the RNA substrate, including all intermediates. Consistent with its role in processing dinucleotides, the complementation with <i>orn</i> led to a marked decrease in 2-mers, alongside a moderate decrease in 3-mers. However, complementation of Δ<i>nrnA</i> with <i>cysQ</i> showed the same RNA accumulation as Δ<i>nrnA</i>, confirming that CysQ does not play a role in RNA degradation.

To address whether in vitro hydrolysis of a deoxydinucleotide substrate NrnA<sub>Ec</sub> was possibly a physiologically relevant substrate for NrnA<sub>Ec</sub>, we conducted <i>ex vivo</i> experiments wherein a 5’-radiolabeled 20-mer DNA was incubated in cellular lysates extracted from wild-type, Δ<i>nrnA</i>, or a Δ<i>nrnA</i> strain containing an inducible copy of <i>nrnA</i> (Supplementary Figure S5). These data suggested that the wild-type lysate degraded the 20-mer DNA to near completion, resulting in accumulation of deoxynucleotides. However, when we incubated the 20-mer DNA in the Δ<i>nrnA</i> cellular lysate, we observed an accumulation of deoxydinucleotides. In contrast, these deoxydinucleotides did not accumulate for the Δ<i>nrnA</i> strain expressing an inducible copy of <i>nrnA</i>. These data suggest that DNA dinucleotides may in theory represent viable substrates for NrnA<sub>Ec</sub>.

Since addition of the 10-mer RNA led to accumulation of short RNAs in lysates from the Δ<i>nrnA</i> mutant, we reasoned that this strain should also exhibit phenotypes that result from an overabundance of linear dinucleotides. Specifically, increased levels of dinucleotides are known to lead to an increase in c-di-NMPs (13,18,52). C-di-GMP abundance is increased in <i>B. subtilis</i> motile cells (28); therefore, elevated dinucleotides could in theory alter the proportion of motile cells, by affecting c-di-GMP abundance. To test this, we inoculated the various complementation strains into soft agar and measured the radius of swimming motility in an endpoint assay (Figure 6C). This revealed that the Δ<i>nrnA</i> mutation led to a significant decrease in swimming motility, suggesting that this strain exhibits elevated dinucleotides (both linear and cyclic). Correspondingly, the complementation with <i>nrnA</i> fully restored swimming motility. A swimming defect was not observed in the Δ<i>nrnB</i> mutant (Figure 6C), suggesting that NrnA<sub>Ec</sub> is primarily responsible for maintaining cellular GG pools during exponential phase growth; however, complementation of the Δ<i>nrnA</i> strain with <i>nrnB</i> resulted in full restoration of swimming motility. Therefore, while NrnB is not required during exponential phase growth for degradation of short RNAs, our data suggests that it is proficient in this ability if forcibly expressed. Finally, complementation with <i>orn</i> on only partially restored swimming motility, further hinting that the range of RNA substrates processed by NrnA might be broader than that of Orn, an enzyme that may specialize as a diribonuclease.
NrnA<sub>B</sub><sub>B</sub> is expressed during exponential growth

To directly determine whether NrnA<sub>B</sub><sub>B</sub> or NrnB<sub>B</sub><sub>B</sub> are produced during exponential growth (OD<sub>600</sub> = 0.8), an untargeted mass spectrometry proteomic analysis was performed on wild-type, Δ<sub>nrnA</sub> or Δ<sub>nrnB</sub> strains of <i>B. subtilis</i>. Peptide spectral matches (PSMs) associated with unique tryptic peptides from NrnB<sub>B</sub> versus the reference <i>B. subtilis</i> proteome were identified in the wild-type and Δ<sub>nrnB</sub> strains (Figure 6D and Supplementary Figure S6). A single PSM identified as a unique NrnB<sub>B</sub> peptide (Supplementary Figure S7) was found in the Δ<sub>nrnA</sub> strain. This match was determined to be a low confidence match by manually observation of the PSM as the identifying peaks are commiserate with peaks from instrument noise, thus it can be surmised that NrnB<sub>B</sub> could not be detected under these growth conditions. Therefore, we conclude that NrnA<sub>B</sub><sub>B</sub> is produced during vegetative growth, whereas NrnB<sub>B</sub><sub>B</sub> is not. Perhaps the latter is produced under different cellular conditions.

NrnA<sub>B</sub><sub>B</sub> is widespread in Firmicutes

Our aggregate data have shown that, in our hands, purified NrnA<sub>B</sub><sub>B</sub> acts preferentially on short RNAs from two to four nucleotides in length and does not act on c-di-AMP or pAp. These biochemical observations are further bolstered
by the nearly identical substrate preferences exhibited by Pde2<sub>Sp</sub> and NrnA<sub>EF</sub>, suggesting that all three proteins might be members of a common subclass of standalone DHH-DHHA1 proteins. By extension, the differing substrate preferences of CnpB<sub>Mt</sub> suggests it represents a different subclass of DHH-DHHA1 proteins. However, it is not clear what sequence and structural features delineate these two subclasses of DHH-DHHA1 proteins. Nor is it clear whether yet more subclasses still await discovery. Therefore, the inability to correctly annotate DHH-DHHA1 proteins is a significant problem. Since our data showed that NrnA<sub>Br</sub>, Pde2<sub>Sp</sub> and NrnA<sub>EF</sub> closely resembled one another in their biochemical preferences, we sought to determine whether these proteins are members of a defined subclass of DHH-DHHA1 proteins.

The DHH domain of NrnA<sub>Br</sub> and other DHH family proteins includes several distinct sequence motifs, each including a conserved aspartate residue (17). The C-terminal DHHA1 domain contains a GGGH-x-x-ASG motif that is likely to be adjacent to residues involved in substrate recognition. Also, an R-x-R-x-R motif (R262, 264 and 266 for NrnA<sub>Br</sub>) may be conserved in NrnA-like proteins and has also been suggested to participate in substrate recognition (17, 21). However, the range of residues and sequence motifs involved in selection of substrates has not yet been established. Nor has it been determined whether different subclasses of NrnA-like proteins may exhibit different substrate preferences and cellular purposes. Interestingly, a recent study demonstrated that a DHH-DHHA1 homolog from Vibrio cholerae acts specifically on GG, at the exclusion of other dinucleotides and other short RNA species (55). This important publication demonstrates proof-of-principle that certain bacteria may have evolved specialized versions of DHH-DHHA1 proteins, and that still more cellular roles for DHH-DHHA1 proteins are likely to await discovery.

Phylogenetic analysis of all proteins with a DHH-DHHA1 domain show that these domains can be standalone (‘NrnA-like’) or occur in tandem with other domains. For example, GdpP proteins, which specifically process c-di-AMP, utilize several domains alongside their main domains. For example, GdpP proteins, which specifically process c-di-AMP, utilize several domains alongside their main domains. However, it is not clear what sequence and structural features delineate these two subclasses of DHH-DHHA1 proteins. Nor is it clear whether yet more subclasses still await discovery. Therefore, the inability to correctly annotate DHH-DHHA1 proteins is a significant problem. Since our data showed that NrnA<sub>Br</sub>, Pde2<sub>Sp</sub> and NrnA<sub>EF</sub> closely resembled one another in their biochemical preferences, we sought to determine whether these proteins are members of a defined subclass of DHH-DHHA1 proteins.

Our experiments on purified NrnA<sub>Br</sub> revealed a marked preference for cleavage of very short RNAs, from two to five nucleotides in length, which were specifically processed from their 5’ terminus. Moreover, our analysis showed that NrnA<sub>Br</sub> preferentially hydrolyzes RNAs 2–4 nucleotides in length even when other longer RNA substrates are included in the same reaction. This observation was further bolstered by experiments wherein a radiolabeled oligonucleotide was added to cellular extracts and the RNase products were analyzed by denaturing electrophoresis. This showed that depletion of NrnA<sub>Br</sub> resulted in accumulation of 2-mers, 3-mers and 4-mers, while, in contrast, deletion of NrnB<sub>R</sub> did not result in accumulation of these short RNAs under normal vegetative growth conditions. And while deletion of NrnB<sub>R</sub> led to accumulation of RNAs between 2 and 4 nucleotides in length, it also affected cellular signaling. Deletion of nrnA resulted in an increase in cyclic di-guanosine monophosphate (c-di-GMP) with a corresponding decrease in swimming motility. Therefore, the processing of dinucleotides by NrnA<sub>Br</sub> is required for maintaining proper homeostasis of cyclic dinucleotide signaling molecules. We also purified NrnA-like proteins from several related species (Streptococcus pyogenes Pde2, Enterococcus faecalis NrnA, Mycobacterium tuberculosis CnpB) and assayed them using the same reaction conditions as with NrnA<sub>Br</sub>. Pde2<sub>Sp</sub> and NrnA<sub>EF</sub> exhibited a strong preference for short RNAs and closely resembled NrnA<sub>Br</sub> overall. Moreover, our bioinformatic analysis of DHH-DHHA1 proteins showed that these proteins may be closely related to NrnA<sub>Br</sub>. Therefore, together, our data demonstrate that a cohesive group of Firmicutes NrnA proteins are required for housekeeping processing of RNAs between two and four nucleotides in length. In contrast, CnpB-like proteins from M. tuberculosis, M. smegmatis, M. avium and R. ruder exhibited dif-
different substrate preferences when assayed under the same reaction conditions. While they were also capable of cleaving very short RNAs, the CnpB proteins were able to directly process c-di-AMP to NMPs, as compared to NrnA<sub>Bs</sub>, Pde2<sub>Sp</sub> and NrnA<sub>Sp</sub>, which were unable to cleave c-di-AMP under our assay conditions. Lastly, none of the proteins exhibited a significant level of pAp hydrolysis activity in vitro under our assay conditions. Interestingly, a recent biochemical analysis of the *Lysteria monocytogenes* NrnA showed virtually no activity when assayed for pAp hydrolysis (56). Therefore, our aggregate data show that NrnA proteins exhibit broader substrate preferences as compared to Orn. While Orn specifically processes dinucleotides, NrnA acts on RNAs between 2 and 4 nucleotides in length and processes them from their 5′ terminus. We also show that NrnA<sub>Bs</sub> is fully capable of processing DNA dinucleotides, although the physiological significance of this is still currently unclear and further work is needed to determine whether NrnA plays a meaningful role in processing cellular DNA dinucleotides.

NrnA<sub>Bs</sub> and NrnB<sub>Bs</sub> were previously assumed to be effectively synonymous in their biochemical activity and intracellular functions (14). In fact, NrnA<sub>Bs</sub> and NrnB<sub>Bs</sub> were initially identified based on their ability to complement a conditional *E. coli orn* mutant strain (14,15). Encoding for both NrnA and NrnB proteins does not appear to be a widespread arrangement amongst bacterial genomes and based on the data shown herein. NrnB<sub>Bs</sub> is not expressed during exponential growth, nor does it appear to process RNAs during exponential growth, in contrast to NrnA<sub>Bs</sub>. Therefore, we speculate that the presence of both an NrnA and NrnB protein has allowed *B. subtilis* the opportunity to specialize the functional role(s) of the latter protein. Interestingly, NrnB-like sequences appeared to cluster together into a separate but cohesive group of sequences; perhaps this indicates that they will share biochemical properties with each other, but distinct from NrnA. Therefore, in total, our studies argue that NrnA<sub>Bs</sub> is the housekeeping enzyme for degradation of short RNAs in *B. subtilis*, while NrnB<sub>Bs</sub> is likely to be expressed under a different phase of cellular growth. However, even with comprehensive analysis of *B. subtilis* NrnA, there are still gaps in our understanding of the degradation pathway for short RNAs. While deletion of *orn* genes is either lethal or results in severe growth defects for the gammaproteobacterial species in which it has been mutated, deletion of *nrrA* does not present *B. subtilis* with a severe growth defect (Supplementary Figure S4). This may suggest that there is redundancy in this step of the RNA degradation pathway or that another RNase can accommodate removal of dinucleotides upon depletion of
NrnA, assuming that accrual of dinucleotides is the basis of the growth phenotype observed for orn mutant strains. This once again highlights how the RNA degradation pathway differs dramatically between E. coli and B. subtilis. And as a corollary to that statement, the complexity and variation between the suite of E. coli and B. subtilis RNases is likely to broaden even further when other, non-model bacteria are considered.

Standalone DHH-DHHA1 domain containing proteins are broadly distributed amongst many taxonomic groups of bacteria, and the members of this protein family are commonly annotated as either NrnA or NrnB proteins (12). A great challenge going forward will be to annotate appropriately the specialized members of this protein family. There have been multiple reports suggesting that NrnA-like proteins are functionally diverse, although some of these prior published data have been debated extensively. For example, in a biochemical analysis of the NrnA-like proteins from T. maritima and S. pneumoniae, it was surmised that previously reported c-di-AMP hydrolysis activity of the S. pneumoniae and M. tuberculosis NrnA-like proteins might be an artifact of assay conditions (49). Yet in another analysis, it was suggested that the S. aureus NrnA-like protein preferred cleavage of the dinucleotide AA, but that c-di-AMP could also be an intracellular substrate for this protein (18). While our data agrees that a large subset of NrnA-like proteins is unlikely to hydrolyze c-di-AMP, our biochemical assays suggested that c-di-AMP is indeed a primary substrate for the M. tuberculosis, M. smegmatis, M. avium and R. ruber CnpB proteins. It is therefore tempting to speculate that c-di-AMP hydrolysis activity is the primary function of the NrnA-like/CnpB proteins encoded by many actinobacterial organisms. In contrast, our sequence clustering analysis suggests that the NrnA-like proteins found in Bacilli function primarily to hydrolyze 2–4 mer RNAs. The current annotation convention for the standalone DHH-DHHA1 protein family will undoubtedly need to be updated to account for the functionally distinct members of this protein family. Indeed, this will be exceedingly difficult in the absence of characteristic sequence motifs that predict the enzyme’s substrate preference(s). However, rigorous enzymological analysis will also be essential in determining the substrate specificity of these enzymes. In the study herein, we find great value in directly comparing the biochemical properties of different purified enzymes under a common set of reaction conditions. In future experiments, these efforts can be expanded for examination of many other proteins to improve the functional annotation of different subclasses of DHH-DHHA1 protein sequences.

**DATA AVAILABILITY**

Mendeley Data is an open-source depository of primary data (https://data.mendeley.com/my-data/). Upon publication, all data will be accessible with a public accession number (doi: 10.17632/rtkt5t6yt9.1).

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online.

**ACKNOWLEDGEMENTS**

We thank Myles Poulin for assistance with analysis of kinetic assays.

**FUNDING**

National Institutes of Health [AI142400 to W.C.W and V.T.L]. Funding for open access charge: National Institutes of Health; C.J. is funded by a SMART Scholarship (funded by OUSD/R&E) (the Under Secretary of Defense-Research and Engineering); National Defense Education Program (NDEP)/BA-1, Basic Research. Conflict of interest statement. None declared.

**REFERENCES**

1. Hui, M.P., Foley, P.L. and Belasco, J.G. (2014) Messenger RNA degradation in bacterial cells. *Annu. Rev. Genet.*, **48**, 537–559.
2. Trinquier, A., Durand, S., Braun, F. and Condon, C. (2020) Regulation of RNA processing and degradation in bacteria. *Biochim. Biophys. Acta Gene Regul. Mech.*, **1863**, 194505.
3. Shahbabian, K., Jamalli, A., Zieg, L. and Putzer, H. (2009) RNase Y, a novel endoribonuclease, initiates riboswitch turnover in *Bacillus subtilis*. *EMBO J.*, **28**, 3523–3533.
4. Mathy, N., Bénard, L., Pellegrini, O., Daou, R., Wen, T. and Condon, C. (2007) 5′-to-3′ exonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5′ stability of mRNA. *Cell*, **129**, 681–692.
5. Ghodge, S.V. and Raushel, F.M. (2015) Discovery of a previously unrecognized ribonuclease from *Escherichia coli* that hydrolyzes 5′-phosphorylated fragments of RNA. *Biochemistry-us*, **54**, 2911–2918.
6. Stevens, A. and Niyoji, S.K. (1967) Hydrolysis of oligoribonucleotides by an enzyme fraction from *Escherichia coli*. *Biochem. Biophys. Res. Commun.*, **29**, 550–555.
7. Niyoji, S.K. and Datta, A.K. (1975) A novel oligoribonuclease of *Escherichia coli*. I. Isolation and properties. *J. Biological Chem.*, **250**, 7307–7312.
8. Datta, A.K. and Niyoji, K. (1975) A novel oligoribonuclease of *Escherichia coli*. II. Mechanism of action. *J. Biological Chem.*, **250**, 313–319.
9. Ghosh, S. and Deutscher, M.P. (1999) Oligoribonuclease is an essential component of the mRNA decay pathway. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 4372–4377.
10. Mechold, U., Ogryzko, V., Ngo, S. and Danchin, A. (2006) Oligoribonuclease is a common downstream target of lithium-induced pAp accumulation in *Escherichia coli* and human cells. *Nucleic Acids Res.*, **34**, 2364–2373.
11. Kim, S.-K., Lormand, J.D., Weiss, C.A., Eger, K.A., Turdiev, H., Turdiev, A., Winkler, W.C., Sondermann, H. and Lee, V.T. (2019) A dedicated diribonucleotidase resolves a key bottleneck for the terminal step of RNA degradation. *Elife*, **8**, e46313.
12. Lormand, J.D., Kim, S.-K., Walters-Marrah, G.A., Brownfield, B.A., Fromme, J.C., Winkler, W.C., Goodson, J.R., Lee, V.T. and Sondermann, H. (2021) Structural characterization of NrnC identifies unifying features of dinucleotidases. *Elife*, **10**, e70146.
13. Orr, M.W., Weiss, C.A., Severin, G.B., Turdiev, H., Kim, S.-K., Turdiev, A., Liu, K., Tu, B.P., Waters, C.M., Winkler, W.C. et al. (2018) A subset of exoribonucleases serve as degradative enzymes for pGpG in c-di-GMP signaling. *J. Bacteriol.*, **200**, e00300-18.
14. Fang, M., Zeisberg, L., Condon, C., Ogryzko, V., Danchin, A. and Mechold, U. (2009) Degradation of nanorNA is performed by multiple redundant RNases in *Bacillus subtilis*. *Nucleic Acids Res.*, **37**, 5114–5125.
15. Mechold, U., Fang, G., Ngo, S., Ogryzko, V. and Danchin, A. (2007) YtfJ from *Bacillus subtilis* has both oligoribonuclease and pAp-phosphatase activity. *Nucleic Acids Res.*, **35**, 4552–4561.
16. Liu, M.F., Cescau, S., Mechold, U., Wang, J., Cohen, D., Danchin, A., Boulouis, H.-J. and Biville, F. (2012) Identification of a novel nanorNase in *Bartonella* *Microbiology*, **158**, 886–895.
Aravind, L. and Koonin, E.V. (1998) A novel family of predicted phosphoesterases includes Drosophila prune protein and bacterial RecJ exonuclease. Trends Biochem. Sci., 23, 17–19.

Boman, L., Zedun, M.S., Scudder, C.F., Kaever, V. and Gründling, A. (2016) New insights into the cyclic di-adenosine monophosphate (c-di-AMP) degradation pathway and the requirement of the cyclic dinucleotide for acid stress resistance in Staphylococcus aureus. J. Biol. Chem., 291, 26970–26986.

Wakamatsu, T., Kim, K., Uemura, Y., Nakagawa, N., Kuramitsu, S. and Masui, R. (2011) Role of RecJ-like protein with 5′-3′ exonuclease activity in oligo(deoxy)xynucleotide degradation. J. Biol. Chem., 286, 2807–2816.

Dianov, G., Sedgwick, B., Daly, G., Olsson, M., Lovett, S. and Lindahl, T. (1994) Release of 5′-terminal deoxyribose-phosphate residues from incised abasic sites in DNA by the Escherichia coli RecJ protein. Nucleic Acids Res., 22, 993–998.

Schmir, B.J., Nelerka, C.M. and Malhotra, A. (2017) Structural basis for the bidirectional activity of Bacillus nanorNase NrnA. Sci. Rep., 7, 11065.

Bai, Y., Yang, J., Eisele, L.E., Underwood, A.J., Koester, B.L., Waters, C.M., Metzger, D.W. and Bai, G. (2013) Two DHF subfamily 1 proteins in Streptococcus pneumoniae possess cyclic di-AMP phosphodiesterase activity and affect bacterial growth and virulence. J. Bacteriol., 195, 5123–5132.

Postic, G., Danchin, A. and Mechold, U. (2012) Characterization of NrnA homologs from Mycobacterium tuberculosis and Mycoplasma pneumoniae. RNA, 18, 155–165.

Manikanand, K., Sabareesh, V., Singh, N., Saigal, K., Mechold, U. and Sinha, K.K.M. (2014) Two-step synthesis and hydrolysis of cyclic di-AMP in Mycobacterium tuberculosis. PLoS One, 9, e86996.

He, Q., Wang, F., Liu, S., Zhu, D., Cong, H., Gao, F., Li, B., Wang, H., Lin, Z., Liao, J. et al. (2016) Structural and biochemical insight into the mechanism of Rv2837c from Mycobacterium tuberculosis as a c-di-NMP phosphodiesterase. J. Biol. Chem., 291, 3668–3681.

Dey, R.J., Dey, B., Zheng, Y., Cheung, L.S., Zhou, J., Sayre, D., Kumar, P., Guo, H., Lammichane, G., Sintim, H.O. et al. (2017) Inhibition of innate immune cytosolic surveillance by an M. tuberculosis phosphodiesterase. Nat. Chem. Biol., 13, 210–217.

Tang, Q., Luo, Y., Zheng, C., Yin, K., All, M.K., LiX. and He, J. (2015) Functional analysis of a c-di-AMP-specific phosphodiesterase MsPDE from Mycobacterium smegmatis. Int. J. Biol. Sci., 11, 813–824.

Weiss, C.A., Hoberg, J.A., Liu, K., Tu, B.P. and Winkler, W.C. (2019) Siphonophilus saurinus, a new species level ciliate containing a cyclic di-GMP phosphodiesterase. J. Eukaryot. Microbiol., 66, 396–401.

Sturin, M.B. and Schramm, V.L. (2009) Detecting ricin: sensitive luminescent assay for ricin. A-chain ribosome depurination kinetics. Anal. Chem., 81, 2847–2853.

Solntsev, S.K., Shortreed, M.R., Frey, B.L. and Smith, L.M. (2018) Enhanced global post-translational modification discovery with MetaMorphes. J. Proteome Res., 17, 1844–1851.

The UniProt Consortium, Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Agirre, E., Ahmad, S., Alpi, E., Bowler-Barnett, E.H., Britto, R. et al. (2020) UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res., 49, D480–D489.

Millikan, R.J., Solntsev, S.K., Shortreed, M.R. and Smith, L.M. (2018) Ultrafast peptide label-free quantification with FlashLFQ. J. Proteome Res., 17, 386–391.

The, M., MacCoss, M.J., Noble, W.S. and Kall, L. (2016) Fast and accurate protein false discovery rates on large-scale proteomics data sets with percolator 3.0. J. Am. Soc. Mass Spectrom., 27, 1719–1727.

Weiss, C.A. and Winkler, W.C. (2021) Riboswitch-mediated detection of metabolite fluctuations during live cell imaging of bacteria. Methods Mol. Biol., 2323, 153–170.

Jenal, U., Reinders, A. and Lori, C. (2017) Cyclic di-GMP: second messenger extraordinaire. Nat. Rev. Microbiol., 15, 271–284.

Chen, Y., Chai, Y., Guo, J. and Losick, R. (2012) Evidence for cyclic di-GMP-mediated signaling in Bacillus subtilis. J. Bacteriol., 194, 5080–5090.

Subramanian, S., Gao, X., Dann, C.E. and Kearns, D.B. (2017) MotI (DgrA) acts as a molecular clutch on the flagellar star protein MotA in Bacillus subtilis. J. Bacteriol., 199, 2715–2726.

Chang, T.-J., Kao, T.-W., Chen, C. and Chen, J.-C. (2014) Functional analysis of 20-cyclic di-AMP-specific phosphodiesterase. Proteome Res., 9, 566–576.

Bai, G., Sun, X., Taneja, B. and Kall, L. (2014) New insights into the cyclic di-AMP phosphodiesterase in Mycobacterium tuberculosis. J. Biol. Chem., 289, 15397–15406.

Ishida, C., Taneja, B. and Bai, G. (2014) Deletion of the cyclic di-AMP phosphodiesterase gene (crpB) in Mycobacterium tuberculosis leads to reduced virulence in a mouse model of infection. Mol. Microbiol., 93, 65–79.

Fahimi, T., Faiz, S., Port, G.C. and Cho, K.H. (2019) The second messenger c-di-AMP regulates diverse cellular pathways involved in stress response, biofilm formation, cell wall homeostasis, Spb expression, and virulence in Streptococcus pyogenes. Infect. Immun., 87, e00147-19.

Rao, F., See, R.Y., Zhang, D., Toh, D.C., Ji, Q. and Liang, X.-Z. (2010) YybT is a signaling protein that contains a cyclic dinucleotide phosphodiesterase domain and a GGDEF domain with ATPase activity. J. Biol. Chem., 285, 473–482.

Zhang, Y., Yang, J. and Bai, G. (2018) Regulation of the CRISPR-associated genes by Rv2837c (CnpB) via an Orn-like activity in Tuberculosis complex Mycobacteria. J. Bacteriol., 200, e00743-17.

Heo, K., Lee, J.-W., Jung, H., Kwon, S., Lee, J., Seok, C.K., Ha, N.-C. and Seok, Y.-J. (2020) A PnpG-specific phosphodiesterase regulates cyclic di-GMP signaling in Vibrio cholerae. J. Biol. Chem., 298, 101626.

Gall, A.R., Haseh, B.Y., Siletti, C., Waters, C.M. and Huynh, T.N. (2021) NraA is a linear dinucleotide phosphodiesterase with limited function in cyclic dinucleotide metabolism in Listeria monocytogenes. J. Bacteriol., 204, e00206-21.