ABSTRACT

Title of Dissertation:	DISCOVERYING THE ROLES OF SOLUBLE DHH-DHHA1 TYPE PHOSPHODIESTERASES IN RNA DEGRADATION AND CYCLIC DINUCLEOTIDE SIGNALING
	Tanner Matthew Myers, Doctor of Philosophy, 2023
Dissertation directed by:	Professor Wade C. Winkler, Department of Cell Biology and Molecular Genetics

The synthesis and degradation of RNA is a fundamental and essential process for all life forms. It is imperative that cells utilize multiple mechanisms to modulate the lifetimes of RNA molecules to ultimately control protein synthesis, to maintain homeostasis or adapt to environmental challenges. One mechanism to directly alter the stability or half-life of an RNA is through the direct enzymatic activity of ribonucleases (RNases). Bacterial organisms encode for many different RNases that possess distinct functions in RNA metabolism. It is through the actions of multiple cellular RNases that a long RNA polymer corresponding to an mRNA, rRNA, tRNA, or sRNA can be fully degraded back into nucleotide monophosphate precursors. The nucleoside monophosphate precursors are recognized by kinases that ultimately recycle these molecules for use in the synthesis of other long RNA polymers. The processing of RNA polymers by endo- and exoribonucleases generates nucleoside monophosphates as well as short RNA oligonucleotides ranging from 2-6 nucleotides in length. Previously, a subset of enzymes

broadly referred to as “nanoRNases” were found to process these short RNA fragments. Oligoribonuclease (Orn), NanoRNase A (NrnA), NanoRNase B (NrnB), and NanoRNase C (NrnC) were previously ascribed the function of indiscriminately processing nanoRNA substrates. However, recent analyses of the evolutionarily related DnaQ-fold containing proteins Orn and NrnC have provided compelling evidence that some nanoRNase protein families possess distinct dinucleotide substrate length preferences (Kim *et al.,* 2019; Lormand *et al.,* 2021). To determine whether all nanoRNases are diribonucleotide-specific enzymes, we utilized a combination of *in vitro* and *in vivo* assays to conclusively elucidate the substrate specificity and intracellular roles of the DHH-DHHA1 family proteins NrnA and NrnB. Through an *in vitro* biochemical survey of many NrnA and NrnB protein homologs, including from organisms of varying degrees of relatedness, we have determined that there are many functional dissimilarities contained within the DHH-DHHA1 protein family. Furthermore, we have conducted a rigorous investigation into the biological and biochemical functions of NrnA and NrnB in the Gram-positive model organism *Bacillus subtilis.* These analyses have shown that *B. subtilis* NrnA and NrnB are not redundant in biochemical activities or intracellular functions, as previously believed. In fact, we have found that *B. subtilis* NrnA is a 5’-3’ exoribonuclease that degrades short RNAs 2-4 nucleotides in length during vegetative growth, while *B. subtilis* NrnB is specifically expressed within the developing forespore and functions as a 3’-5’ exoribonuclease that processes short RNA in addition to longer RNA substrates (>40-mers). Our collective data provide a strong basis for the subdivision of the DHH-DHHA1 protein family on the basis on their diverse substrate preferences and intracellular functions.

DISCOVERYING THE ROLES OF SOLUBLE DHH-DHHA1 TYPE PHOSPHODIESTERASES IN RNA DEGRADATION AND CYCLIC DINUCLEOTIDE SIGNALING

by

Tanner Matthew Myers

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Dr. Wade Winkler, Chair
Dr. Myles Poulin
Dr. Douglas Julin
Dr. Jinwoo Lee
Dr. Daniel Nelson

Acknowledgements

Thank you to Dr. Wade Winkler for being an extremely supportive adviser who has allowed me the freedom to explore and build my skillset as an independent researcher.

To the past and present members of the Winkler lab, Cordelia, Amr, Madi, Conor, Thao, Daniel Trettel, and Daniel Wu. I appreciate all of the time that you have all sacrificed to listen to all of my practice talks, reading, and editing my papers, as well as my nonsensical ramblings about enzymes and bacterial spores.

Thank you to Dr. Mara Shainheit, Dr. Ana-Maria Soto, Dr. John Weldon, Dr. Katharine Bowdy, Dr. Elana Ehrlich, and all the chemistry and biology professors at Towson that have inspired me to pursue a graduate degree.

I would like to thank my advisory committee for all of their advice and guidance during my PhD.

Finally, thank you to my parents Dana Myers and James Webb. Without your support I could never have made it this far in my academic pursuits.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AMP: adenosine monophosphate

AS: ammonium sulfate

ATP: adenosine triphosphate

cAMP: 3’,5’-cyclic adenosine monophosphate

c-di-AMP: bis-(3'-5')-cyclic dimeric adenosine monophosphate

c-di-GMP: bis-(3'-5')-cyclic dimeric guanosine monophosphate

c-di-NMP: bis-(3'-5')-cyclic dimeric nucleotide monophosphate

DGC: diguanylate cyclase

DNA: deoxyribonucleic acid

FPLC: fast protein liquid chromatography

HPLC: high-performance liquid chromatography

6xHis: 6 sequential histidine residues

10xHis: 10 sequential histidine residues

k_{cat}: turnover number

K_d: dissociation constant

K_M: Michaelis constant

k_{cat}/K_M: specificity constant

LB: lysogeny broth

LC-MS/MS: liquid chromatography-mass spectrometry

MBP: maltose binding protein

PAGE: polyacrylamide gel electrophoresis

pAp: 3’-phosphoadenosine 5’-phosphate

pApA or AA: 5′-phosphonoadenylyl-(3′-5′)-adenosine

PCR: polymerase chain reaction

PDE-A: phosphodiesterase A

PDE-B: phosphodiesterase B

PEI: Polyethyleneimine

pGpG or GG: 5′-phosphonoguanylyl-(3′-5′)-guanosine

PMSF: phenylmethane sulfonyl fluoride or phenylmethylsulfonyl fluoride

RNA: ribonucleic acid

mRNA: messenger ribonucleic acid

rRNA: ribosomal ribonucleic acid

tRNA: transfer ribosomal nucleic acid

SDS: sodium dodecyl sulfate

SUMO: small ubiquitin-related modifier

TLC: thin layer chromatography

2xYT: 2x yest extract tryptone media

Chapter 1: Introduction

1.1 Copyright Notice

Some data in Chapter 1 was originally published by eLife as: 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.

1.2 Principles of RNA Degradation in Bacteria

Bacteria synthesize many different RNA polymers to grow and adapt to various cellular stresses. DNA dependent RNA polymerase synthesizes cellular RNA chains by sequentially joining ribonucleoside triphosphates and releasing pyrophosphate. Cellular RNA polymers are distinguishable from each other not only due to the differences in nucleobase composition, but also by the cellular functions that they serve. In the synthesis of a single polypeptide, cells utilize various species of RNA polymers. To accomplish this feat, a cell requires a messenger RNA (mRNA) which serves as a template that can be decoded by ribosomes. Ribosomes are ribonucleoproteins composed of ribosomal RNAs (rRNA) that serve as essential structural and functional components that are required to translate the mRNA code. In order to synthesize a polypeptide, a ribosome must transfer the amino acid from RNAs known as transfer RNAs (tRNA). Yet, all cellular RNAs, whether they are small regulatory RNA, mRNAs, rRNAs or tRNAs, must undergo processing and turnover by ribonucleases

Ribonucleases are enzymes that degrade RNA into smaller fragments. These enzymes can be further classified as endoribonucleases which cleave RNA molecules internally, or exoribonucleases that remove nucleotide monophosphates from either terminus of the RNA.

Exoribonucleases can be further subcategorized by the termini and direction in which they remove nucleoside monophosphates either acting in the 3'-5' or 5'-3' directions. Furthermore, exoribonucleases can be classified as using a processive or distributive cleavage mechanisms. Processive RNases bind to an RNA and make multiple successive cleavages prior to releasing the substrate. Distributive RNases are thought to make one cleavage event per binding interaction between substrates; these enzymes act on RNAs by binding, cleaving, releasing and rebinding RNA substrates for another base removal. In total, cells encode for various RNases that utilize these different enzymatic mechanisms for the processing of RNA.

Cells utilize different RNases to selectively modulate the folding and the half-lives of mRNAs, rRNAs, and tRNAs. After RNA molecules such as tRNAs or rRNAs are synthesized, many different endo- and exoribonucleases specifically process these RNA species to ensure proper folding, maturation and ultimately function (Hui *et al.,* 2014; Bechhofer and Deutscher 2019; Trinquier *et al.,* 2020). Also, endo- and exoribonucleases regulate the lifetimes of mRNA transcripts, tuning their half-lifes from seconds to dozens of minutes. Even more interesting is that the fragments of RNAs generated from the processing of mRNA processing can have regulatory implications. Small RNAs (sRNAs) are a class of non-coding RNA that can arise as independent transcriptional units or via mRNA processing; these regulatory RNAs can then influence the translation or degradation of target mRNAs through complementary base pairing interactions (Chao and Vogel 2016; Adams and Storz 2020). Combinations of RNases tune the stability of sRNAs according to their need, or in response to cellular stress conditions. Therefore, all bacterial species rely on a complex array of RNA polymers that can be regulated by a variety of different exo- and endoribonucleases.

While many of the fundamental RNA degradation pathways have been elucidated utilizing the Gram-negative model organisms *Escherichia coli*, one key complication in the field of RNA decay has been that no two organisms encode the same collection of RNases (Trinquier *et al.,* 2020). It was largely assumed that the fundamental principles governing RNA degradation would be conserved in all bacterial organisms. However, the subsequent analysis of the Gram-positive organism *Bacillus subtilis* provided an alternative RNA degradation pathway that is not present in *E. coli*. While much of our knowledge of bacterial RNA degradation has stemmed from research utilizing either *E. coli* or *B. subtilis,* researchers are now investigating whether these pathways are conserved in non-model organisms.

1.2.1 Endonucleolytic mediated RNA degradation in *E. coli* and *B. subtilis*

The fundamental principles and initial models for RNA degradation derived primarily from studies on the mechanisms used by the Gram-negative model organism *E. coli*. These mechanisms were extrapolated to the Gram-positive bacterium *B. subtilis* and while there are some striking similarities, there are also key differences in the ways these two historically significant model organisms degrade RNA. Furthermore, advances in genomic analyses revealed that different microbes encode for very different sets of RNases. This finding has led researchers to question whether the fundamental principles of *E. coli* RNA degradation will hold true in organisms that harbor vastly different collections of RNases. While research utilizing non-model organisms is certainly going to improve our collective understanding of RNA degradation, herein we will take a closer look at the major discoveries using *E. coli* and *B. subtilis.*

Early hypotheses proposed that RNAs could exist as unstable messengers for protein synthesis, and subsequent theories suggested that these intermediates possess structural elements that can exert a great influence on the lifetime of the RNA (Brenner, Jacob, and Meselson 1961;

Apirion 1973). Many mRNAs as well as various other forms of intracellular RNA (e.g., rRNA, tRNA, sRNA) contain secondary and tertiary structures that can impede degradation by exoribonucleases. Early observations with *E. coli* suggested that RNA can be destabilized and more accessible to degradation through the action of an endoribonuclease (Hui *et al.,* 2014). In fact, later experiments showed that the stability of most mRNA *E. coli* transcripts is dictated by the membrane-bound endonuclease RNase E, which is essential for growth under standard conditions (Babitski *et al.,* 1991; Cohen and McDowell 1997). RNase E bypasses RNA structural elements to make an internal ssRNA cleavage at AU rich sequences (Cohen and McDowell 1997). This is referred to as the direct access pathway (Figure 1.1). Importantly, RNase E also serves as the scaffold for a multiprotein complex referred to as the degradosome (Py *et al.,* 1996). This might be unsurprising because RNase E is a large 1061 amino acid protein that consists of many different subdomains that can serve as sites for the association for the degradosome proteins. RNase E copurifies with the 3'-5' exoribonuclease PNPase, the metabolic enzyme enolase, and the ATP dependent helicase RhlB (Py *et al.,* 1996). While researchers had known that RNase E and the degradosome were crucially important for *E. coli* RNA processing, it remained unknown whether a degradosome complex was a feature of RNA processing in *B. subtilis*, or if it used a direct access pathway akin to RNase E.

B. subtilis does not encode a protein homolog of RNase E. The search for a *B. subtilis* enzyme that acted as a functional homolog of RNase E led to the discovery of two very different enzymes – the heterodimeric exoribonuclease RNase J1/J2 and the endoribonuclease RNase Y (Even *et al.,* 2005; Shahbabian *et al.,* 2009). Subsequent data suggested that RNase Y may act as the key functional homolog of RNase E, based largely on the observation that RNase Y significantly affects global mRNA levels (Laalami *et al.,* 2013). Interestingly, although RNase Y

is a membrane bound endonuclease that also cleaves at AU rich sequences, it bears virtually no structural or shared sequence identity with RNase E (Shahbabian *et al.,* 2009; Khemici *et al.,* 2015; DeLoughery *et al.,* 2018). While the discovery of a direct access pathway in *B. subtilis* (Figure 1.1) provided researchers with a more generalizable understanding of RNA degradation, there were still many unanswered questions remaining. For example, does RNase Y form a multiprotein degradosome-like complex akin to RNase E?

The existence of a degradosome in *B. subtilis* has been highly controversial given that a stable degradosome complex has not been isolated from this organism. Much of the current data suggesting that there might be interactions between RNase Y and other RNases, helicases, and glycolytic enzymes has resulted from bacterial two-hybrid experiments, which are known to show a high false positive rate (Commichau *et al.,* 2009; Lehnik-Habrink *et al.,* 2010). It was speculated from these data that RNase Y interacts with various other proteins *in vivo*, but this complex could be far more transient than the degradosome found in *E. coli.* Nevertheless, a recent transcriptomic analysis of *B. subtilis* RNase Y has provided strong evidence that there are at least three protein factors that can alter the specificity of RNase Y for the processing of different mRNA transcripts (Deloughery *et al.* 2018). It is remarkable that a direct access pathway fundamental for gene regulation has been found in both *E. coli* and *B. subtilis*, albeit these organisms evolved vastly different proteins to accomplish this important cellular process.

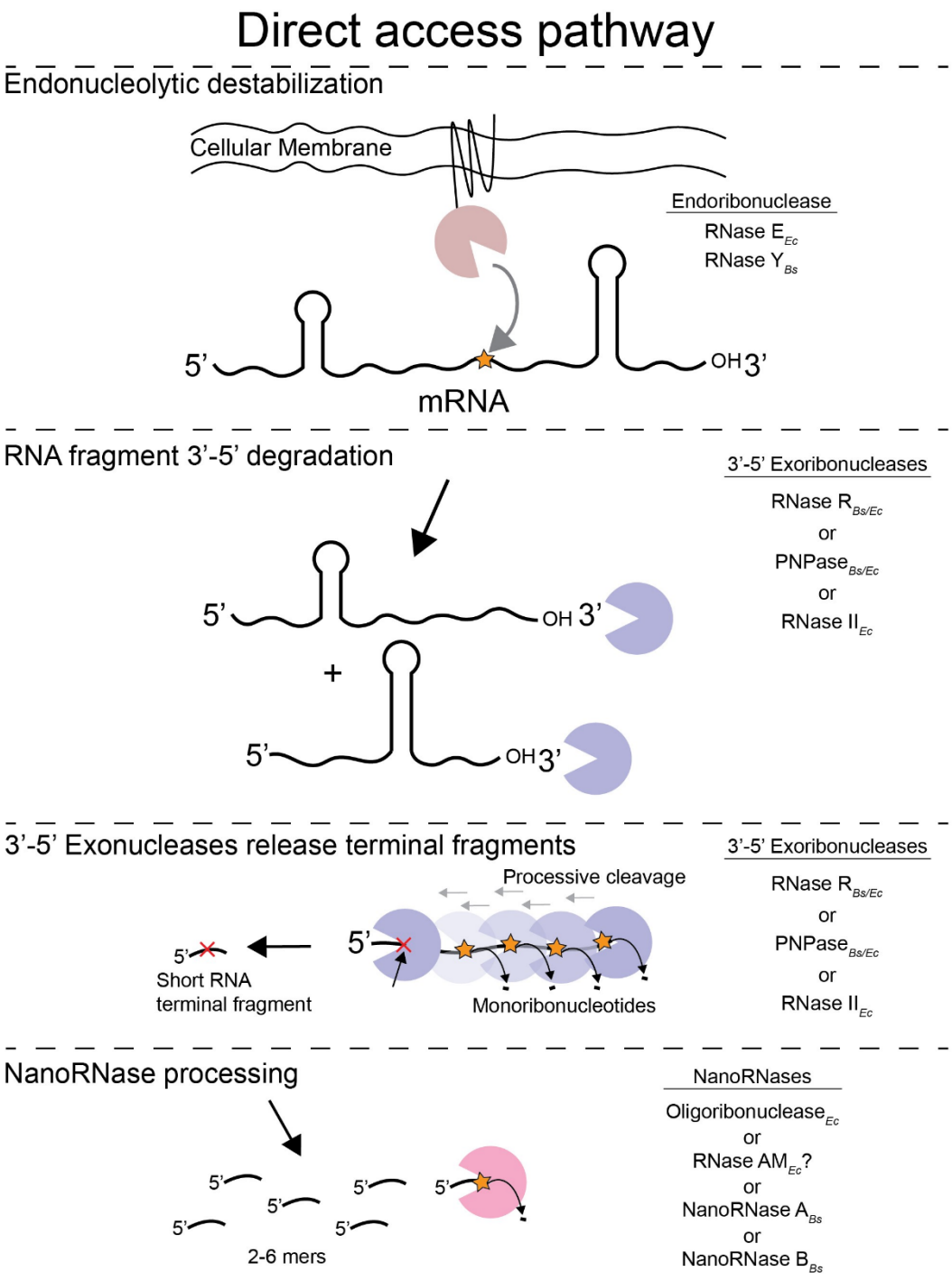

Figure 1.1. A schematic showing the endonucleolytic based direct access pathway utilized by *E. coli* and *B. subtilis*. The degradation of RNA polymers usually proceeds by an initial destabilization event in which an endonuclease makes an internal cleavage event. Next, processive 3'-5' exoribonucleases degrade the destabilized fragments generating monoribonucleotides and releasing short terminal RNA fragments. NanoRNases have been thought to indiscriminately process short RNA fragments that can accumulate from the previous steps of RNA degradation.

1.2.2 Exonucleolytic 3'-5' RNA degradation pathways in *E. coli* and *B. subtilis*

Exonucleolytic RNases degrade RNAs by sequentially removing mononucleotides from long RNA polymers. Interestingly, it is thought that the primary function of 3'-5' exoribonucleases is to degrade RNAs that were initially cleaved by an endoribonuclease in the direct access pathway. RNA transcripts oftentimes feature highly structured 3' or 5' untranslated regions and endonucleolytic cleavage can render a long RNA transcript more readily accessible to 3'-5' exoribonucleases (Figure 1.1). It is widely accepted that the primary 3'-5' exoribonucleases found in *E. coli* are PNPase, RNase R, and RNase II (Hui *et al.,* 2014; Bechhofer and Murray 2019). *B. subtilis* harbors PNPase and RNase R, but an RNase II homolog is conspicuously absent from this organism genome. Experiments using *E. coli* have shown that single deletion strains of *pnpA, rnb,* or *rnr* generally do not display major growth phenotypes (McMurry and Levy 1987; Bechhofer and Deutscher 2019). Interestingly though, double knockout strains of *pnpA* (PNPase) and *rnb* (RNase II) are inviable as well as strains that lack *pnpA* and *rnr* (Donovan and Kushner 1986; Cheng *et al.,* 1998). A key difference between *E. coli* and *B. subtilis* is that while double mutant strains for general 3'-5' exoribonucleases are inviable, this is not the case for *B. subtilis*. In fact, double mutant strains for *pnpA* and *rnr* still do not show any major growth defects (Oussenko *et al.,* 2002). Even triple mutant strains of *B. subtilis pnpA rnr* and another 3'-5' exoribonuclease encoding gene called *yhaM,* only display a modest growth defect in which the doubling time of the triple mutant is only twice that of the wild type cells (Oussenko *et al.,* 2002). These genetic experiments in *E. coli* and *B. subtilis* suggest that the general 3'-5' exoribonucleases have overlapping functions and are important for cell survival.

1.2.3 Exonucleolytic 5'-3' RNA degradation pathways in *E. coli* and *B. subtilis*

Experiments in eukaryotic cells provided evidence that RNA degradation could proceed by Xrn1-mediated exonucleolytic degradation from the 5' to 3' direction (Houseley and Tollervey 2009). It was previously thought that 5' to 3'-dependent RNA degradation pathways were not present in bacterial organisms. Yet, *B. subtilis* RNase J1 was discovered for this exact function (Even *et al.,* 2005; Mathy *et al.,* 2007), demonstrating why it is important to further investigate the paradigms that have been constructed from a few model organisms. RNase J1 is a 5'-3' exoribonuclease that is found in *B. subtilis* and other Gram-positive bacteria but is not encoded by *E. coli* (Hui *et al.,* 2014). RNase J1 forms a heterodimer with another protein with high shared sequence identity referred to as RNase J2 (Even *et al.,* 2005; Mathy *et al.,* 2007),. Together, RNase J1/J2 complexes degrade RNA from the 5'-3' direction with the help of the pyrophosphohydrolase RppH (Figure 1.2). Previous studies have suggested that the 5'-3' exonucleolytic decay pathway is highly dependent on the removal of pyrophosphate from the 5' ends of transcripts by RppH (Richards *et al.,* 2011). Once a pyrophosphate is removed from the 5' terminus, RNase J1/J2 can recognize and degrade the monophosphorylated RNA transcripts more efficiently. Interestingly, the effects of deleting the gene encoding RNase J1 (*rnjA*) have been contested. Yet it is now widely accepted that *rnjA* is an essential (or near essential) gene in *B. subtilis* (Kobayashi *et al.,* 2003; Figaro *et al.,* 2013). This

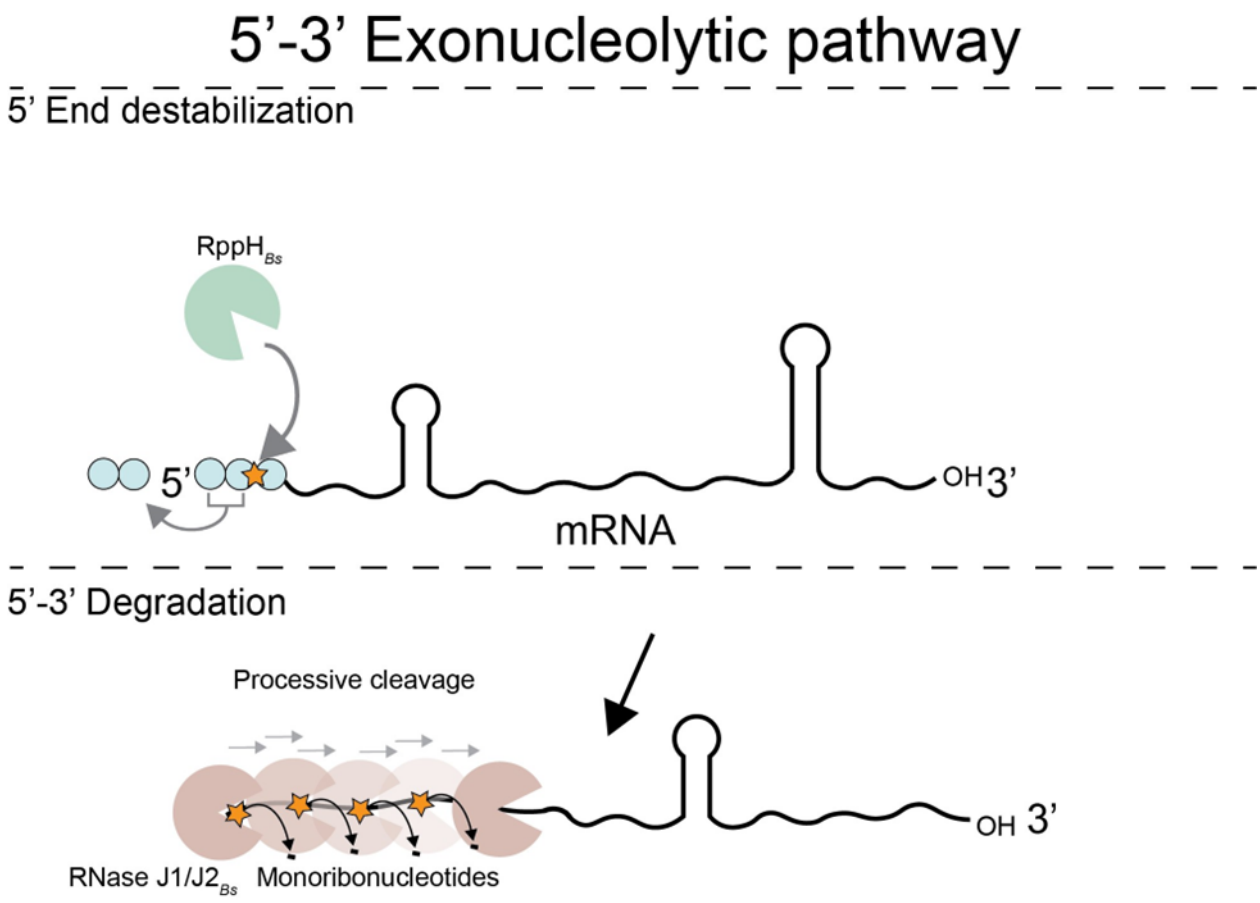

Figure 1.2. Diagram of the 5’-3’ exonucleolytic RNA degradation pathway in *B. subtilis*. RNA polymers are generally synthesized containing a 5’ triphosphate group which is represented by the blue circles. Degradation of RNA polymers can occur in the event that the pyrophosphohydrolase RppH destabilizes an RNA by removing pyrophosphate from the 5’ end of a transcript. The monophosphorylated RNA can be degraded by the 5’-3’ exoribonuclease heterodimer RNase J1/J2.

essentiality indicates the importance of the 5'-3' exonucleolytic decay pathway in *B. subtilis*. These reports also suggest that it is crucially important to investigate the many different RNases and potential RNA decay pathways present in different bacterial organisms investigate the many different RNases and potential RNA decay pathways present in different bacterial organisms.

1.3 Cyclic di-nucleotide second messengers

1.3.1 Cyclic di-GMP

Bacterial organisms use secondary metabolites called second messenger signaling molecules, which communicate with protein and RNA effectors to control many aspects of cell physiology. While there is considerable knowledge on the mechanisms and applications of the second messenger cyclic AMP (cAMP), other second messenger molecules went unnoticed in bacteria until 1987, when it was discovered that cyclic diguanosine monophosphate (c-di-GMP) allosterically regulated cellulose synthesis in *Acetobacter xylinum* (Ross *et al.,* 1987). It is now generally accepted that c-di-GMP is widely used amongst bacterial organisms to regulate many different complex biological processes such as cell growth, biofilm formation, and virulence (Römling *et al.,* 2013; Jenal *et al.,* 2017).

The synthesis of c-di-GMP proceeds by joining two molecules of GTP by proteins called diguanylate cyclases (DGC) (Jenal *et al.,* 2017). DGCs are proteins that contain a catalytic GGDEF domain. The degradation of c-di-GMP proceeds by a two-enzyme mechanism in which the first enzyme, a phosphodiesterase A (PDE-A), which contains either an EAL or HD-GYP domain, cleaves the c-di-GMP molecule to generate the linear diribonucleotide pGpG. Next, a

second enzyme referred to as phosphodiesterase B (PDE-B) degrades pGpG to generate two molecules of GMP. Interestingly, this pathway was complicated by the observation that EAL or HD-GYP domain-containing proteins degrade pGpG *in vitro* (Tomayo *et al.,* 2005; Stelitano1 *et al.,* 2013). However, a more recent analysis using *Pseudomonas aeruginosa* provided evidence that EAL or HD-GYP domain-containing proteins do not meaningfully process pGpG in *ex vivo* lysate experiments, and that the RNase Oligoribonuclease (Orn) is the primary enzyme responsible for pGpG degradation in *P. aeruginosa* (Orr *et al.,* 2018). In fact, many subsequent studies have found that the linear diribonucleotide pGpG represents an intersection between c-di-GMP homeostasis and RNA degradation (Lee, *et al.,* 2022).

1.3.2 Cyclic di-AMP

The second messenger signaling molecule c-di-AMP was initially discovered in a structural analysis of the DNA integrity scanning protein DisA (Witte *et al.,* 2008). It was subsequently shown that c-di-AMP is a crucially important signaling molecule responsible for osmolyte regulation, DNA integrity, cell wall homeostasis, and sporulation (Corrigan and Gründling, 2013; Stülke and Krüger, 2020). Interestingly, the synthesis and degradation of c-di-AMP resembles that of c-di-GMP. Akin to c-di-GMP synthesis, c-di-AMP is generated by proteins containing a diadenylate cyclase (DAC) domain in which two ATPs are utilized to form the c-di-AMP molecule (Witte *et al.,* 2008). The degradation of c-di-AMP is also generally assumed to occur by way of a two-enzyme mechanism. The core PDE-A enzyme that is believed to linearize c-di-AMP is GdpP, which is comprised of a heme binding

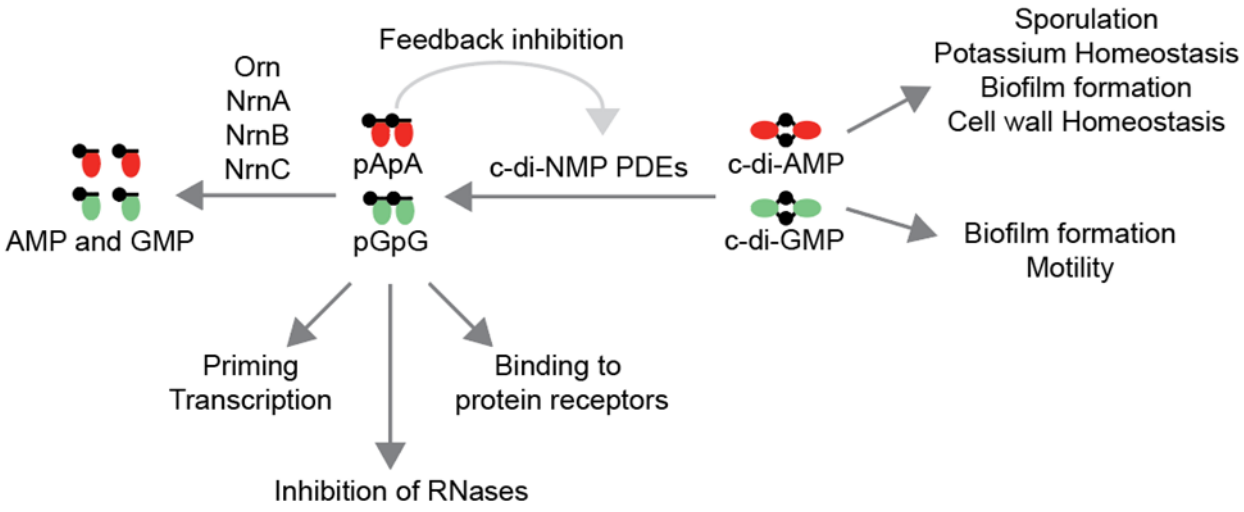

Figure 1.3. Schematic illustrating the functional significance of the diribonucleotides pGpG and pApA. C-di-GMP or c-di-AMP regulate a variety of different intracellular processes. Diribonucleotides have been shown to feedback inhibit the degradation of c-di-GMP or c-di-AMP through feedback inhibition. Additionally, diribonucleotides have been shown to exhibit various intracellular functions outside of c-di-GMP or c-di-AMP signaling. NanoRNases such as Orn, NrnA, NrnB, and NrnC have been shown to process the diribonucleotides pGpG and pApA *in vitro* and in *in vivo*.

transmembrane Per-Arnt-Sim (PAS) domain, a degenerate GGDEF domain and the catalytic DHH-DHHA1 domains. C-di-AMP is also linearized by the enzyme PgpH, which contains a transmembrane domain and a catalytic HD domain (Rao *et al.,* 2010; Huynh *et al.,* 2015). The PDE-B enzymes responsible for the processing of pApA include standalone DHH-DHHA1 proteins referred to as NanoRNase A (NrnA). Recent reports also suggested that some standalone DHH-DHHA1 proteins have the capacity to degrade c-di-AMP directly in addition to linear dinucleotides such as pApA (Manikandan *et al.,* 2014; He *et al.,* 2015, Tang *et al.,* 2015).

1.4 NanoRNases – enzymes at the intersection of c-di-NMP signaling and RNA degradation

Recent investigations into the processing of c-di-GMP and c-di-AMP have provided evidence that the action of two enzymes is required to completely process these molecules to nucleoside monophosphates. In fact, there is also substantial data that suggests that not only does c-di-NMP processing overlap with RNA degradation, but that diribonucleotides generated by PDE-A enzymes can result in feedback inhibition (Figure 1.3). In organisms where the genes encoding PDE-B enzymes have been deleted or interrupted, the concentrations of diribonucleotides (e.g., pGpG and/or pApA) increase and inhibit the processing of c-di-NMP's (Bai *et al.,* 2013; Orr *et al.,* 2015; Bowman *et al.,* 2016; Orr *et al.,* 2018; Fahmi *et al.,* 2019. For these reasons, the degradation pathways of cyclic di-NMPs overlaps with the global pathways for degradation of cellular dinucleotides. The enzymes that degrade dinucleotides and other short RNAs (i.e*.,* oligonucleotides less than 5 nucleotides in length) are generally called "nanoRNases" (Lee *et al.,* 2022). Given the overlap between cyclic di-NMP homeostasis and RNA degradation there has been general confusion about the RNA substrates preferred by PDE-B enzymes. Specifically, it has remained unclear whether there are subsets of PDE-B enzymes

that exhibit a physiologically relevant function in both the processing of linear dinucleotides and c-di-NMP's.

1.4.1 Oligoribonuclease

The protein Oligoribonuclease (Orn) was one of the earliest RNases isolated and identified (Stevens and Niyogi 1967; Niyogi and Datta 1975; Datta and Niyogi 1975; Lee *et al.,* 2022). Orn's initial activity was described from a biochemical fractionation of Δ*rna* Δ*pnpA E. coli* in which the hydrolysis of ^{14}C-pApA was utilized to measure enzyme activity (Stevens and Niyogi 1967). Additionally, the partially purified enzyme was shown to degrade short RNA substrates ranging from 2-5 nucleotides in length. Orn's amino acid composition and genomic location in the *E. coli* chromosome were identified many years after the initial reports of its activity (Yu and Deutscher 1995; Zhang *et al.,* 1998). It was also later discovered that the gene encoding Orn is in fact an essential gene in *E. coli*, which suggests that short RNA processing is a crucially important step in RNA degradation and metabolism (Ghosh and Deutscher 1999). While Orn certainly has a crucial role in *E. coli*, it has since been determined that many bacterial organisms do not encode for *orn* (Mechold *et al.,* 2007; Fang *et al.,* 2009; Liu *et al.,* 2012; Lee *et al.,* 2022). Bacterial organisms that encode for an *E. coli* Orn homolog include mostly Actinobacteria, Fibrobacteres, Betaproteobacteria, Deltaproteobacteria, and Gammaprotebacteria (Orr *et al.,* 2015; Lee *et al.,* 2022). Interestingly, the essentiality of *orn* in *E. coli* paved the way for genetic complementation experiments in which researchers were able to identify Orn functional homologs encoded by other organisms (Mechold *et al.,* 2007; Fang *et al.,* 2009; Lui *et al.,* 2012; Lee *et al.,* 2022).

Orn is a member of the DnaQ-fold containing protein family that harbors an active site DEDD motif and is thought to form a stable dimeric assembly (Kim *et al.,* 2019). Although

many previous studies have indicated that Orn is largely an indiscriminate nanoRNase the functions on RNase 2-5 nucleotides in length, a more recent analysis has conducted a rigorous structural and biochemical investigation into the substrate specificity of Orn from *Vibrio cholera* (Kim *et al.,* 2019). Surprisingly, Orn_{Vc} displays a stark preference for the binding and cleavage of diribonucleotides (Kim *et al.,* 2019). In fact, the reported K_d for Orn_{Vc} for pGpG is reported to be 90 nM. Furthermore, crystallographic structures of Orn_{Vc} bound to diribonucleotides has revealed an active site architecture that sterically precludes the binding of RNAs longer than 2-mers. Moreover, experiments comparing the degradation products of a radiolabeled 7-mer RNA in wild type or Δ*orn Pseudomonas aeruginosa* whole cell lysates show a clear accumulation of diribonucleotides in the absence of *orn* (Kim *et al.,* 2019). These findings provide evidence that some or all nanoRNases might specifically process diribonucleotides, thereby acting as 'diribonucleases', enzymes that specifically process dinucleotides (Lee *et al.,* 2022). Whether there are substrate length specificities contained in other classes of nanoRNases, which are composed of other domain architectures, remains to be seen. While these recent data demonstrate that some Orn proteins act as diribonucleases, it remains unclear whether this is a feature exhibited by all Orn homologs and by nanoRNases overall.

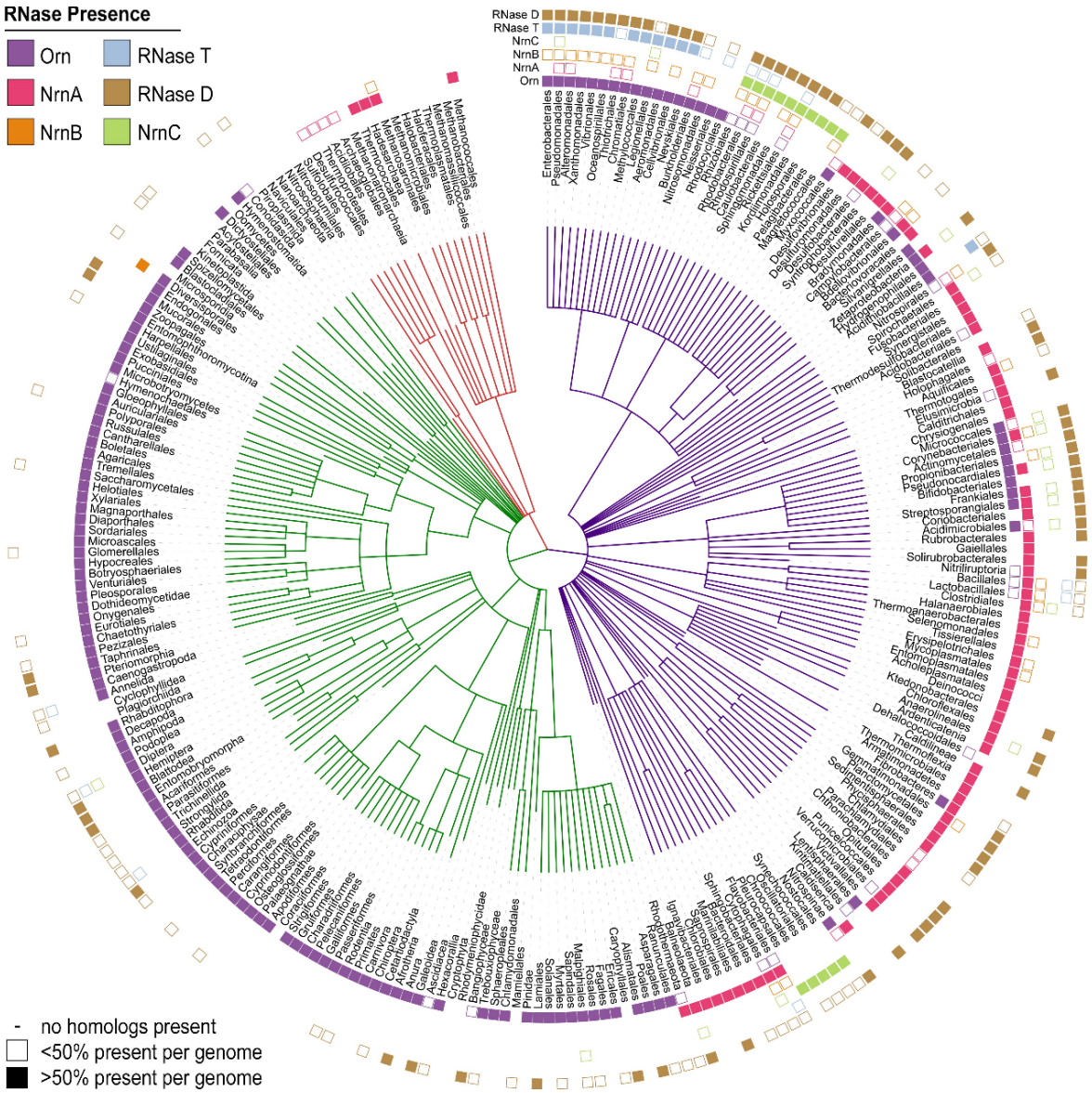

Figure 1.4. Presence of RNase homologs across sequenced organism classes. Shown is a 'Tree of Life' with all taxonomic groups at the class level with at least one substantially complete proteome available in the dataset. The tree is based on the structure of the NCBI Taxonomy database, with bacterial taxa shown with purple lines, eukaryotic taxa shown with green lines, and archaeal taxa shown with red lines. The presence of each RNase homolog as a proportion of the total proteins in that taxonomic group is shown as either a filled square (>50% presence of a homolog per genome) or an empty square (<50% presence of a homolog per genome). Lack of a square indicates no homologs for that family were present in genomes of that class. *Phylogenetic analysis was performed by Dr. Jonathan, Winkler laboratory, University of Maryland and was previously published in (Lormand et al., 2021).*

1.4.2 NanoRNase C

Orn is generally found in various classes of Proteobacteria, however, sequences corresponding to a functional Orn homolog are conspicuously absent in many members of the Alphaproteobacterial class of organisms. This observation led a group of researchers to screen a genomic library of *Bartonella birtlesii* genes for genes that would complement a conditional Δ*orn E. coli* strain (Lui *et al.,* 2012). This screen led to the discovery of the protein NanoRNase C (NrnC). NrnC is also a member of the DnaQ-fold containing protein family and also possesses a DEDD motif. However, NrnC adopts an octameric oligomeric state assembly which contrasts with the dimeric Orn (Yaun *et al.,* 2019; Kim *et al.,* 2019; Lormand *et al.,* 2021). Akin to *orn*, *nrnC* has been shown to be an essential gene in *Caulobacter crescentus* and *Brucella abortus* (Christen *et al.,* 2011; Sternon *et al.,* 2018). However, the intracellular activity of NrnC was still debated.

The first *in vitro* biochemical analysis of NrnC from *B. birtlesii* indicated that this enzyme exhibits exoribonucleolytic activity against both short and long RNAs (Lui *et al.,* 2012). Specifically, NrnC was shown to degrade a 3-mer and 5-mer RNA bearing a 5'Cy5 label as well as a 5'-radiolabeled 24-mer substrate (Lui *et al.,* 2012). Another activity described for NrnC from *Agrobacterium tumefaciens* is the capacity of this protein to processively degrade long ssRNA, ssDNA, and dsDNA substrates, but not dsRNA (Yuan *et al.,* 2019). Interestingly though, a very recent study utilized X-ray crystallography in addition to cryo-electron microscopy to investigate the structural basis for NrnC's substrate specificity (Lormand *et al.,* 2021). The structural analysis of $NrnC_{Bh}$ bound to short RNAs 2-5 nucleotides in length revealed an active site architecture that conceptually resembles Orn_{Vc} in which diribonucleotides are most likely the preferred substrate (Lormand *et al.,* 2021). Structures of $NrnC_{Bh}$ bound to 3-5 mer RNAs resulted

in poor substrate density maps except for the first two nucleotides, which were well coordinated within the active site (Lormand *et al.,* 2021). This data is corroborated by *in vitro* RNase activity assays in which the purified NrnC$_{Bh}$ preferentially degrades diribonucleotides as compared to RNAs of greater lengths. Additionally, this specificity is reflected in RNA binding experiments as the reported K_d for NrnC$_{Bh}$ for pGpG is 18 nM while the K_d for a 3-mer is 3.5 μM (Lormand *et al.,* 2021). Although Orn and NrnC proteins exhibit similarities in their general folding architecture and active site motifs, it was unclear whether these proteins convergently or divergently evolved. A recent phylogenetic analysis demonstrated that NrnC is likely to have evolved from two different families of DnaQ-fold containing RNases, RNase D and RNase T (Lormand *et al.,* 2021). These data suggest that NrnC more recently diverged from RNase D, while Orn exhibits an ancient divergence from RNase T (Lormand *et al.,*2021). Altogether, these data suggest that Orn and NrnC have likely convergently evolved their diribonucleotide substrate specificity. Yet, while strides have been made in understanding the properties of a few Orn and NrnC proteins, it is still an important challenge facing researchers going forward to determine if there are specialized paralogs of these proteins exhibiting alternate activities.

1.5 DHH-DHHA1 Enzymes

Proteins contained in the DHH family of phosphodiesterases are widely distributed amongst bacteria, archaea, and eukaryotes. The N-terminal domain of the DHH family of proteins is distinguished by the presence of the signature DHH sequence motif in addition to multiple aspartic acid residues which coordinate divalent metal ion cofactors (Aravind and Koonin 1998; Lee *et al.,* 2022). The superfamily of DHH family proteins includes multiple specialized subfamilies; the average organism encodes for several different types of DHH family protein homologs. This is most likely due to the prevalence of many different accessory domains

that co-occur with the N-terminal DHH domain. These different architectures in part dictate the protein's substrate preferences (Aravind and Koonin 1998; Lee *et al.,* 2022). Two C-terminal domains that can co-occur with the DHH domain are referred to as the DHHA1 (DHH-associated domain 1) or DHHA2 (DHH-associated domain 1) domain. The typical assembly of a standalone DHH protein consists of the N-terminal DHH domain connected via a linker region to either a DHHA1 or DHHA2 domain. Proteins composed of DHH-DHHA2 domains represent a group of different exopolyphosphatase enzymes. The DHHA1 can be recognized not only by its tertiary structured, but also by a highly conserved GGGH motif (Schmier *et al.,* 2017). For example, the DHH-DHHA1 protein family includes subclasses of proteins corresponding to but not limited to: NanoRNase A (NrnA), NanoRNase B (NrnB), RecJ, and GdpP (Lee *et al* 2022). Although NrnA, NrnB, RecJ, and GdpP are composed of contiguous DHH-DHHA1 assemblies, NrnA and NrnB contain no other discernible domain architectures other than the DHH and DHHA1 domains. For this reason, DHH/DHHA1 proteins are referred to in this project as standalone DHH-DHHA1 proteins. Additionally, NrnA and NrnB proteins have been shown to exhibit RNase activity, although it has remained largely unclear which factors delineate this subfamily of proteins (Mechold *et al.,* 2007; Fang *et al.,* 2009). This differs from RecJ and GdpP, which also include additional domains. For example, GdpP is a DHH-DHHA1 protein that also contains a transmembrane domain, a putative heme-sensing PAS domain, and a degenerate GGDEF domain that primarily functions as a c-di-AMP phosphodiesterase (Rao *et al.,* 2010; Wang *et al.,* 2018). RecJ contains the DHH-DHHA1 domains in addition to an OB fold domain and an extended C-terminal domain; RecJ functions as a 5’-3’ DNase involved in DNA repair (Cheng *et al.,* 2016). While functions have been identified for some of the major branches, it is less clear for many of the DHH family proteins.

1.5.1 NanoRNase A and NanoRNase B

NrnA and NrnB are two widely conserved DHH-DHHA1 proteins that were initially discovered in *B. subtilis* and were shown to exhibit RNase activity on short RNA substrates (Mechold *et al.,* 2007; Fang *et al.,* 2009). Interestingly, a recent phylogenetic analysis conducted by our laboratory examined the distribution of nanoRNases across various taxonomic lineages and revealed that NrnA is the most widely distributed nanoRNase amongst bacterial organisms (Figure 1.4) (Lormand *et al.,* 2021). The distribution of NrnB is far less common. At the time of this analysis there were virtually no criteria for distinguishing between proteins referred to as NrnA or NrnB. Some researchers have claimed that certain NrnA-like protein homologs exhibit c-di-AMP phosphodiesterase activity (Yang *et al.,* 2014; He *et al.,* 2016). When I began my graduate studies, I recognized that there were many claims regarding the activity a researcher might expect from an NrnA protein. For instance, the capacity of NrnA-like proteins to function as c-di-AMP phosphodiesterases has been heavily disputed. Also, the mechanism by which NrnA recognizes and degrades substrates has also been heavily contested. Furthermore, $NrnA_{Bs}$ and $NrnB_{Bs}$ have been thought to be essentially redundant proteins despite only sharing ~20% sequence identity with each other. I will elaborate further about the state of knowledge surrounding NrnA and NrnB proteins when I began my graduate studies in the respective chapters of this dissertation. However, some of the key questions that I hoped to answer with my research were as follows: 1. Are NrnA and NrnB redundant in enzymatic activity? 2. Do NrnA and NrnB both function as PDE-B enzymes in c-di-GMP processing? 3. Do NrnA and NrnB proteins exhibit diribonucleotide substrate specificity akin to Orn and NrnC? 4. Do specialized functional homologs of NrnA or NrnB exist? 5. Does *B. subtilis* harbor more nanoRNases outside of NrnA and NrnB?

1.6 RNase AM (YciV)

It was previously thought that *E. coli* does not harbor a 5'-3' exoribonuclease. However, it was recently discovered that protein RNase AM (YciV) acts on RNA substrates from the 5' end (Ghodge and Raushel 2015). RNase AM is a member of the polymerase and histidinol phosphatase (PHP) family of proteins. Researchers have previously known that histidinol phosphatases generally hydrolyze organophosphoesters. The discovery of RNase AM stems from a clustering analysis of PHP family proteins in which RNase AM was found to be a member of a distinct cluster of orthologous protein groups where the other protein COGs functions were known. In fact, one of the COGs was ascribed the function of removing the 3' phosphate from 3'-5'-bis-phosphonucleotides (Cummings *et al.,* 2014). A kinetic analysis of RNase AM_{Ec} revealed that pAp, 2'-deoxy-pAp, and pGp are the preferred substrates for this enzyme based on a comparison of specificity constants. NanoRNase activity was also reported for RNase AM_{Ec} as this protein was shown to degrade short purine containing RNAs 2-5 nucleotides in length (Ghodge and Raushel 2015). This *in vitro* analysis of RNase AM_{Ec} indicates that this protein could exert a dual function in the degradation of the sulfate carrying molecule pAp and in the degradation of short RNAs akin to Orn. It is currently unclear whether RNase AM_{Ec} functions as a nanoRNase *in vivo*, but there are reports that RNase AM_{Ec} matures the 5' ends of all three ribosomal RNAs (Jain, 2020). Still, another group has provided evidence that RNase AM proteins might also function to remove flavin adenine dinucleotide caps from RNA transcripts (Sharma *et al.,* 2022). More research will be required to unequivocally determine the functional roles of RNase AM in *E. coli* and other organisms that encode for different homologous proteins.

1.7 Outlook

Bacteria encode for and utilize many RNases, some of which possess functionally redundant functions, while others possess distinct specialized roles in RNA degradation. Oftentimes, bacteria will encode for an essential membrane bound endoribonuclease that controls the stability and half-lives of most mRNA transcripts in addition to multiple 3'-5' exoribonucleases that are partially redundant in function (Hui *et al.,* 2014; Bechhofer and Deutscher, 2019; Trinquier *et al.,* 2020). In some rarer instances, bacteria will also encode for a 5'-3 exoribonuclease such as RNase J that works in conjunction with a pyrophosphohydrolase RppH to control the half-lives of transcripts (Richards *et al.,* 2011; Hsieh *et al.,* 2013; Foley *et al.,* 2015). Yet, all organisms seem to encode for at least one often essential nanoRNase that degrades the small oligoribonucleotide fragments that are generated by the endo- and exoribonucleases (Lormand *et al.,* 2021). It is through the combined efforts of these RNases that cells can regulate gene expression by controlling the half-lives of mRNAs, thereby generating stable functional rRNAs, tRNAs, and sRNAs. Furthermore, nanoRNases have been directly implicated in c-di-NMP signaling pathways. This is because many PDE-A enzymes can be feedback inhibited by the linear diribonucleotides substrates of nanoRNases. This means that nanoRNases are PDE-Bs that can influence the intracellular concentrations of c-di-NMPs, which can in turn dramatically alter cellular phenotypes (Stülke *et al.,* 2020; Jenal *et al.,* 2017). Altogether, RNases play crucial roles in gene expression and RNA metabolism.

In the case of nanoRNases, some of the biochemical analyses of these proteins have generated much confusion as to what the primary roles of these enzymes are in cells. Orn, NrnA, NrnB, and NrnC are enzymes that have been ascribed the function of indiscriminately degrading short 'nanoRNAs' 2-5 nucleotides in length (Lee *et al.,* 2022). However, referring to the proteins

collectively as “nanoRNases” has complicated the functional annotation of these proteins given that they are likely to have derived from convergent evolution and are composed of different protein folding architectures. Further complicating the use of the word “nanoRNase” to refer to these proteins are the recent studies from our laboratories, which provided substantial evidence showing that Orn and NrnC specifically exhibit diribonucleotide substrate specificities. That these enzymes are restricted to only cleaving diribonucleotides makes it more technically accurate to refer to them as “diribonucleases” as opposed to “nanoRNases”. Also, are all the known nanoRNases actually diribonucleases, or do some of them act against nanoRNAs in general? And what even is a “nanoRNA”? Proteins have been referred to as nanoRNases if the enzyme functions in the degradation of short RNAs; however, the actual substrate length of a “nanoRNA” has not been strictly defined and is therefore ambiguous. At what point does a short RNA oligomer stop being a nanoRNA and become a general RNA?

In my graduate studies, I have chosen to investigate the DHH-DHHA1 proteins NrnA and NrnB. My goal is to unequivocally determine the substrate specificity and intracellular functions of these enzymes. Of the known nanoRNases, the least is known about NrnA- and NrnB-type enzymes; therefore, I expect my work on these proteins to significantly improve their annotation. In Chapter 2, I describe non-radioactive methods that can be used to gain rapid insight into nanoRNase mechanisms and functions. In Chapter 3, we investigated the biochemical characteristics and intracellular functions of NrnA in *B. subtilis.* Also in Chapter 3, we conducted a small biochemical survey to investigate whether some NrnA-like proteins exhibit activity directly against c-di-AMP. In Chapter 4, we describe an approach for the purification of *B. subtilis* NrnB. Then, Chapter 5 investigates the intracellular function of *B. subtilis* NrnB and compares it with NrnA. Chapter 6 presents a biochemical survey that can be used to identify

features that could aid in the annotation of NrnB proteins. Finally, Chapter 7 outlines the challenges facing researchers investigating nanoRNase function in addition to providing theoretical methods for aiding in the analysis of nanoRNases. Improving the collective functional annotation of nanoRNases, and specifically the standalone DHH-DHHA1 proteins, is a major challenge that we have undertaken in this study. These efforts have significantly improved the collective understanding of nanoRNases and their functions in diribonucleotide processing, c-di-NMP metabolism, and general RNA processing.

Chapter 2: Non-radioactive methods for assessing phosphodiesterase activity

2.1 Copyright Notice

Some data in Chapter 2 was originally published by Nucleic Acids Research Journal as: Weiss,C.A. [#], Myers,T.M.[#], Wu,C.H., Jenkins,C., Sondermann,H., Lee,V.T. and Winkler,W.C. (2022) NrnA is a 5′-3′ exonuclease that processes short RNA substrates in vivo and in vitro. *Nucleic Acids Res*, 50, 12369–12388.

2.2 Introduction

Phosphodiesterases are enzymes of great biochemical interest due to their prevalence in all life forms and their often-central roles in cellular metabolism. Enzymes that are classified as phosphodiesterases utilize H_2O or inorganic phosphate PO_4^{3-} as nucleophiles to break one of the two bridging phosphodiester bonds found in phosphodiester linkages (Yang, 2010). Phosphodiesterases are subdivided into many different protein families based on their domain architecture and substrate preferences. Common candidate substrates for individual phosphodiesterase subgroups include 3’,5’-cyclic adenosine monophosphate (cAMP), 3’,5’-cyclic guanosine monophosphate (cGMP), bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-GMP), bis-(3'-5')-cyclic dimeric adenosine monophosphate (c-di-AMP), in addition to DNA and RNA. The cellular concentrations of the different second messenger signaling molecules regulate a myriad of different complex biological processes including carbon utilization, motility, osmolyte homeostasis, sporulation, or competence (Gomelsky, 2011). The processing of various forms of intracellular RNA is essential for controlling the synthesis of proteins, and the maintenance of DNA is important for ensuring genomic integrity. Given the

fundamental importance of these phosphodiesterases, it is important that researchers have access to an assortment of biochemical assays.

Many of the commonly used biochemical assays on phosphodiesterases have significant drawbacks associated with them. For example, how the RNA substrates are prepared can affect the efficiency of an assay. The biochemical analyses of cAMP and c-di-AMP phosphodiesterases typically starts with the generation of a ^{32}P-labeled substrate, which must be generated by an adenylate cyclase (to generate cAMP) or a diadenylate cyclase (to generate c-di-AMP). This process can be exceedingly laborious because these cyclases are not commercially available and must be individually purified and tested for specific activity. Additionally, the products for the phosphodiesterase reactions are typically analyzed using denaturing PAGE (polyacrylamide gel electrophoresis) or thin layer chromatography (TLC). These separation techniques are robust and can be easily adapted for resolution of different RNA products; however, they are also relatively low throughput and require a substantial time commitment. One small molecule separation technique that circumvents the need for radioactivity is through the use of an HPLC or FPLC. These analyses have the advantage of being free of radioactive labeling, yet often require a large amount of substrate for detection purposes. They may also require relatively expensive chemical standards for investigating elution profiles of the targeted molecules. Additionally, analyses of phosphodiesterase degradation products by HPLC or FPLC is often low throughput and limited to an endpoint assay format. Herein, we describe two different nonradioactivity-based methods for assessing phosphodiesterase activity. The first method is based on the release of the fluorescent nucleotide analog 2-aminopurine. The second method is a coupled enzyme assay that quantifies the release of AMP in order to monitor activity of phosphodiesterases that act on cAMP, c-di-AMP, or RNA. The 2-aminopurine fluorescence method was developed for other

RNAs previously (Roembke *et al.,* 2014; Zhou *et al.*, 2017), and a similar version of the luminescence-based coupled enzyme assay was also reported previously (Sturm and Schramm, 2009). We aim to draw attention to the usefulness of these methods while including describing our adaptations. In general, we highlight how these methods can be used to investigate the substrate preferences and mechanisms employed by the short RNA-degrading *Bacillus subtilis* NanoRNase A.

2.3 Results

2.3.1 The release of the nucleotide 2-aminopurine results in an increase in fluorescence that can be utilized to monitor RNase activity

Nucleotide analogs are frequently utilized by biochemists to investigate the mechanisms of different enzymes in addition to probing the conformational dynamics of DNA and RNA structures. One such nucleotide analog that is frequently used is 2-aminopurine. The nucleobase 2-aminopurine is a purine analog that exhibits intrinsic fluorescence properties, which contrasts with the canonical nucleobases that do not exhibit fluorescence (Jones and Neely, 2015). Additionally, since 2-aminopurine exhibits fluorescence, it can be selectively excited even when it is situated within large and small DNA or RNA fragments. Another interesting feature of 2-aminopurine is that the fluorescence emission is directly influenced by whether the nucleobase is contained in a stacked confirmation with other nucleobases, is unstacked or is free in solution (Somsen *et al*., 2005). When 2-aminopurine is contained in a stacked confirmation the fluorescence emission is low, but when 2-aminopurine is not in a stacked confirmation or is free in solution there is a substantial increase in fluorescence. Previous analyses have shown that 2-aminopurine exhibits reduced fluorescence even when in a stacked confirmation with just one other nucleotide in a very short single-stranded dinucleotide (Somsen *et al.,* 2005). The

differential fluorescence between a dinucleotide containing 2-aminopurine and that of free 2-aminopurine has been previously exploited in an assay format for monitoring the enzymatic activity of *Mycobacterium smegmatis* Orn (i.e., an enzyme that cleaves dinucleotides) (Zhou *et al.* 2017). To investigate the cleavage of diribonucleotides by NrnA$_{Bs}$ we adapted the method used for *M. smegmatis* Orn (Zhou *et al*., 2017). Specifically, we monitored the cleavage of a diribonucleotide where one of the nucleobases corresponded to 2-aminopurine (Figure 2.1A). For this analysis we subjected 10 μM of pAp(2AP) to cleavage by 10 nM of purified NrnA$_{Bs}$ and found that there was an appreciable change in fluorescence intensity over time, indicating that 2-aminopurine monophosphate (2AP) was being liberated (Figure 2.1B). As a control for background substrate fluorescence emission, we included a reaction that did not contain NrnA$_{Bs}$. We observed that excitation of pAp(2AP) resulted in a steady background fluorescence emission that did not change over time. From these data we speculated that this reaction progress curve could be utilized to estimate the kinetic efficiency of diribonucleotide cleavage by NrnA$_{Bs}$. Previous kinetic analyses of the NrnA-like protein homolog found in *Thermotoga maritima* reported that the Michaelis-Menten constant (K_M) for the diribonucleotide substrate pApA was 204 μM, while the specificity constant (k_{cat}/K_M) was 686 M^{-1} s^{-1}. We decided to attempt to fit the enzymatic progress curve in (Figure 2.1C) to a single exponential equation. Since 10 μM of the substrate pAp(2AP) is well below the reported K_M found in the analysis of the *T. maritima* protein homolog, we reasoned that this substrate concentration might be sufficiently low enough to represent a second order or bimolecular reaction where both the concentration of [S] and [E] dictate the rate of reaction progress. Under second order reaction conditions, the rate can be approximated using pseudo-first order kinetics to record an observed rate constant (k_{obs}). The rate constant k_{obs} is determined under these assay conditions where [S] < K_M is likely proportional to

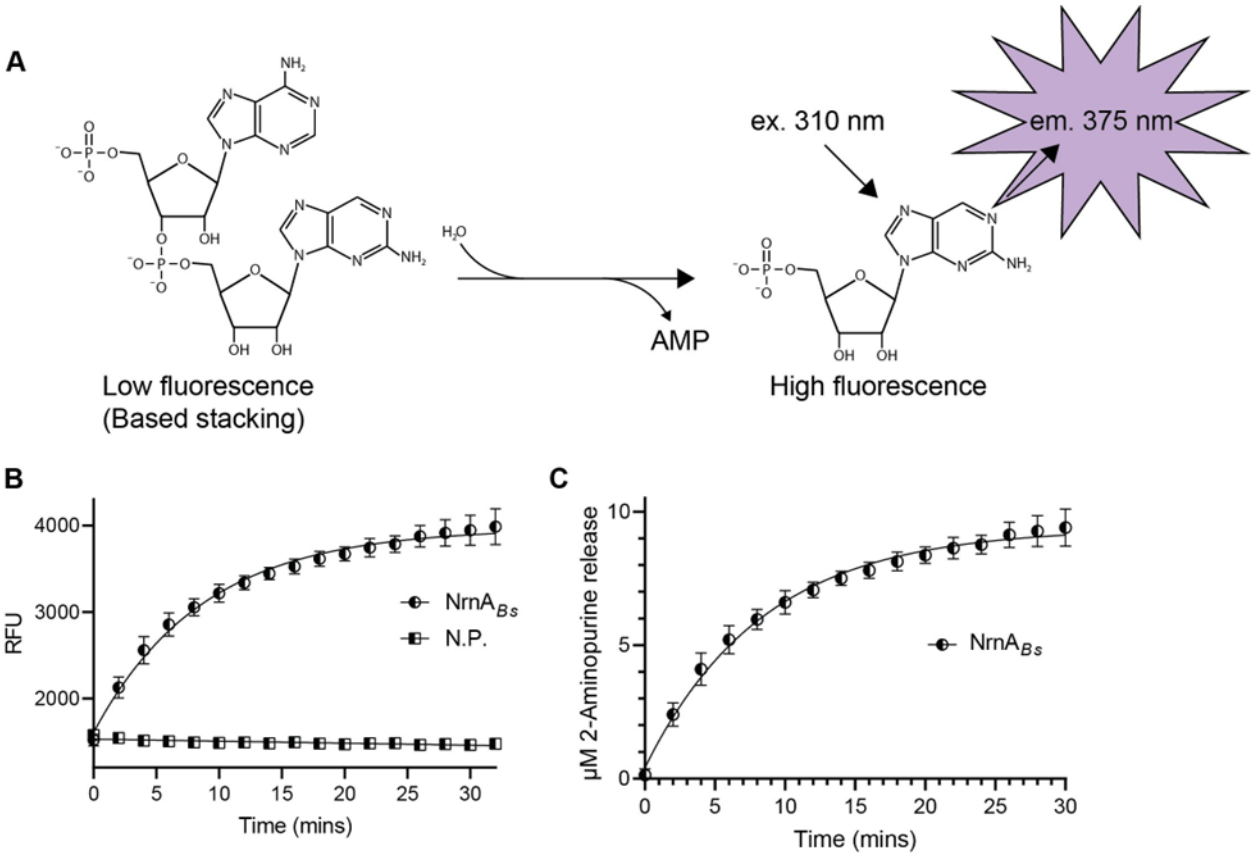

Figure 2.1. The release of the nucleotide 2-aminopurine results in an increase in fluorescence that can be utilized to monitor RNase activity. (A) Schematic showing the change in fluorescence emission for 2-aminopurine in a stacked confirmation with an adenosine nucleotide vs. free unstacked 2-aminopurine. (B) Time course fluorescence emission changes upon the hydrolysis of 10 μM pAp(2AP) by 10 nM of purified NrnA$_{Bs}$ and the fluorescence output for a reaction containing no protein (N.P.). (C) The fluorescence data in (B) normalized to fit the concentration of 2-aminopurine liberated fit to a first-order rate equation $R^2 = 0.98$. The equation used to fit the curve in (C) was Y=Y0 + (Plateau-Y0)*(1-exp(-K*x)). Measurements are representative of two or three individual replicates.

the specificity constant k_{cat}/K_M.From the normalized reaction progress curve where we correlated the RFUs with the molar concentration of 2-aminopurine release in (Figure 2.1C), we obtained a k_{obs} value of 0.12 min^{-1}. The specificity constant for the cleavage of pAp(2AP) by $NrnA_{Bs}$ was determined by converting the k_{obs} to seconds which is 0.002 s^{-1}. Using the equation $k_{obs} = V\text{max}/K_M$ (which is also $k_{obs} = ([E_T] \bullet k_{cat})/K_M$) we can divide the k_{obs} value by the protein concentration which is $10 \bullet 10^{-9}$ M to obtain k_{cat}/K_M. From this analysis we estimate that the specificity constant k_{cat}/K_M is $2 \bullet 10^5$ M^{-1} s^{-1}, which is significantly greater than what was previously reported for other NrnA-like homologs (Drexler *et al.,* 2017). We speculate that this method could be further utilized in the analysis of short RNA-degrading enzymes to reduce costs associated with buying short RNA substrates.

2.3.2 Determining RNase polarity by monitoring the release of 2-aminopurine from modified oligoribonucleotides

In the original biochemical analysis of $NrnA_{Bs}$ the polarity of RNA cleavage was not fully established. Yet, it was largely assumed that this enzyme utilized a 3'-5' polarity based on the substrate degradation products seen in the cleavage of a 5-mer RNA that featured a 5'-Cy5 label (Mechold *et al.,* 2007). A subsequent mass spectrometry survey of the degradation products of $NrnA_{Bs}$ suggested that $NrnA_{Bs}$ degraded an 11-mer DNA substrate with 5'-3' polarity (Wakamatsu *et al.,* 2011). Yet another subsequent analysis argued that $NrnA_{Bs}$ was a unique bidirectional exoribonuclease that acted on short RNAs with 3'-5' exonucleolytic activity but processed long RNA substrates twelve nucleotides and longer by exonucleolytically cleaving from the 5' terminus (Schmier *et al.,* 2017). Given the lack of clarity and contradictory claims in the published literature, we sought to unequivocally determine the polarity that $NrnA_{Bs}$ utilized to degrade RNAs.

In this analysis, we subjected 5'-radiolabeled 4-mer RNAs with different phosphodiester backbone linkages to cleavage by purified NrnA$_{Bs}$. The degradation products were then resolved by denaturing 20% PAGE. Incubation of NrnA$_{Bs}$ with a short RNA 4-mer (^{32}P-AAGG) that contained only canonical phosphodiester linkages resulted in the release of 5'-^{32}P-AMP, without appearance of any other degradation intermediates (Figure 2.2A). The lack of degradation intermediates is an indication that NrnA$_{Bs}$ could be removing the 5' residue, which in this instance contains a radiolabeled phosphate group. This observation agreed with a prior investigation of NrnA$_{Bs}$ that showed that the cleavage of a 5'-^{32}P-AAA resulted in the accumulation of ^{32}P-AMP with no observable dinucleotide intermediate (Schmier *et al.*, 2017). This would be the expected outcome if NrnA$_{Bs}$ utilized a 5'-3' exonucleolytic mechanism. However, one other possibility could be that NrnA$_{Bs}$ recognizes short RNAs and processes them very quickly in the 3'-5' direction resulting in no radiolabeled intermediates. To test this hypothesis, we utilized 4-mer RNA substrates that contained a potentially non-hydrolyzable phosphorothioate linkage. One of the RNAs contained a phosphorothioate linkage to theoretically block the cleavage of the 3' residue, but not the 5' residue (^{32}P-AAG*ps*G). When we subjected the (^{32}P-AAG*ps*G) RNA to cleavage by NrnA$_{Bs}$ we found that the primary degradation product was ^{32}P-AMP (Figure 2.2A). These data are still consistent with the removal of the 5' residue. To bolster our analysis, we also tested the cleavage of (^{32}P-AA*ps*GG) to cleavage by NrnA$_{Bs}$. When we incubated ^{32}P-AA*ps*GG with NrnA$_{Bs}$, ^{32}P-AMP was still the only degradation fragment to accumulate. If NrnA$_{Bs}$ was able to cleave the 3' residue we would anticipate that the intermediate ^{32}P-AA*ps*G would accumulate. together, these data suggest that NrnA$_{Bs}$ utilizes a 5'-3' exonucleolytic mechanism.

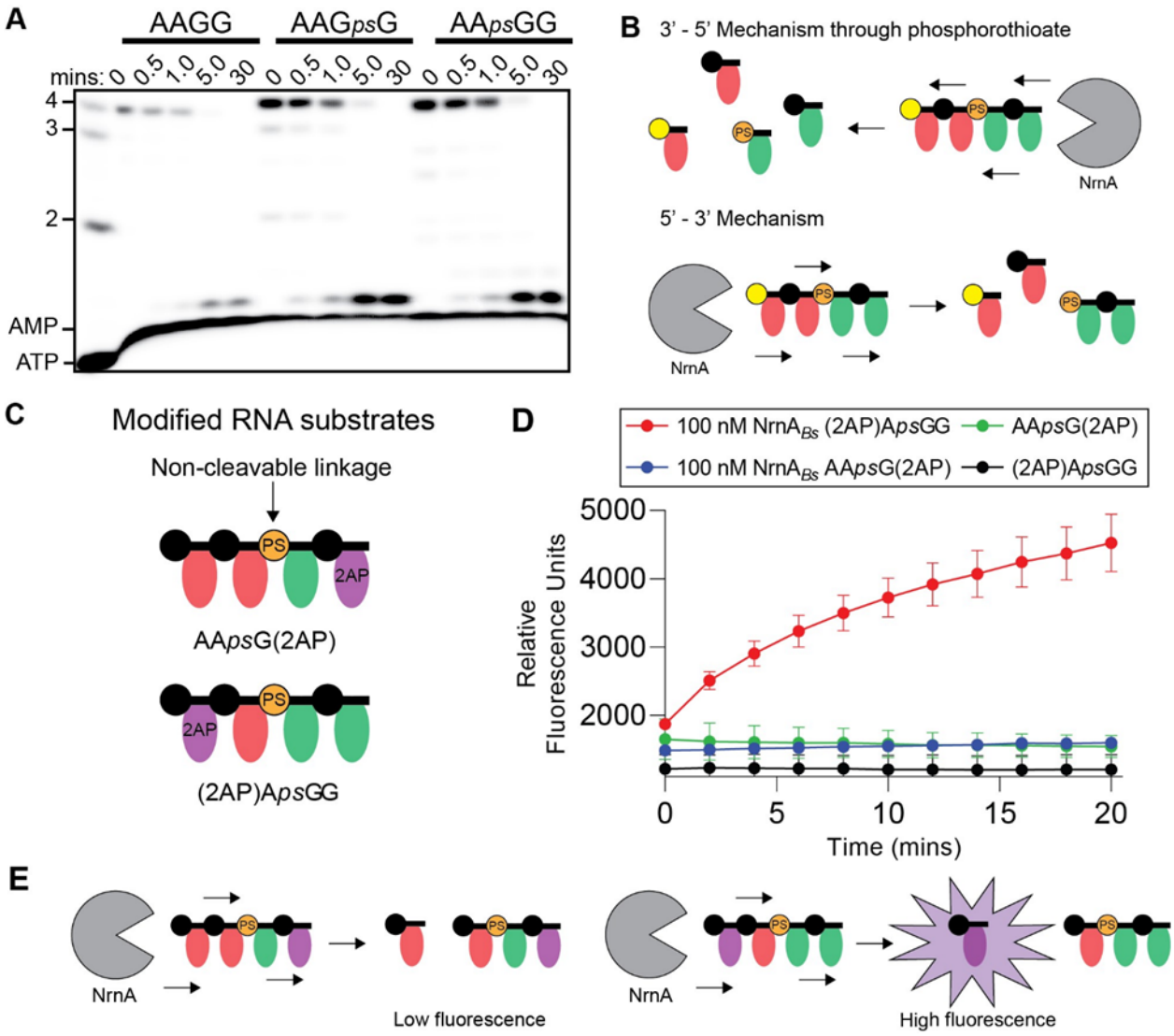

Figure 2.2. $NrnA_{Bs}$ degrades RNAs with 5'-3' polarity. (A) In these reactions, some RNA substrates contained a non-hydrolyzable, phosphorothioate-modified backbone (placement of the linkage represented by *ps*). 1 µM of RNA was incubated with trace amounts of ^{32}P-radiolabeled RNA and 100 nM of purified $NrnA_{Bs}$. Aliquots were removed at various time points and resolved by denaturing PAGE (B) A schematic showing the possible degradation products that could arise in the cleavage of the ^{32}P-AA*ps*GG substrate. (C) An illustration of the 2-aminopurine oligoribonucleotides utilized in (D). (D) 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 $NrnA_{Bs}$. (E) A schematic showing the results observed in (C).

RNase polarity is generally determined by analyzing the degradation products that arise in the processing of radiolabeled RNAs. Researchers frequently label RNA molecules at their 5' terminus using T4 polynucleotide kinase, which transfers the ^{32}P-γ-phosphate from ATP to the 5' hydroxyl of an RNA. Otherwise, researchers have the option to label the 3' end of RNAs by first generating ^{32}P-Cp by again transferring the ^{32}P- γ -phosphate from ATP to the 5' hydroxyl of cytosine 3' phosphate using a modified T4 PNK that lacks 3' phosphatase activity. Next, ^{32}pCp can be added to the 3' end of an RNA molecule by T4 RNA ligase 1. While we have routinely been able to radiolabel the 5' ends of short RNA oligonucleotides using T4 PNK as seen in (Figure 2.2A), we were unsuccessful in our attempts to ligate ^{32}pCp to the short RNAs to generate 3' end labeled short RNAs, which could have been used to more conclusively determine the polarity of NrnA$_{Bs}$ (Nilsen, 2014). To this end, we sought to adapt a non-radioactive method for assessing the removal of the 3' nucleotide from RNAs.

One complication in our analysis of NrnA$_{Bs}$'s polarity is that it was possible that NrnA$_{Bs}$ could process through the phosphorothioate linkage, which could directly hinder our ability to interpret the data in which the 5' residues of RNAs are the only traceable degradation products (Figure 2.2B). In order to investigate this we determined whether NrnA$_{Bs}$ would cleave off 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*ps*GG or AA*ps*G(2AP)) (Figure 2.2C). 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 (Somsen *et al.,* 2005; Zhou *et al.,* 2017). When NrnA$_{Bs}$ was incubated with (2AP)A*ps*GG, an increase in 2AP fluorescence was observed over time (Figure 2.2D). In contrast, no fluorescence increase was observed upon incubation of NrnA$_{Bs}$ with AA*ps*G(2AP), which indicates that NrnA$_{Bs}$ cannot

process through phosphorothioate linkages. Taken altogether, these assays suggest that NrnA$_{Bs}$ is exclusively a 5′-3′exoribonuclease for short RNA substrates and NrnA$_{Bs}$ cannot hydrolyze nucleotides connected by phosphorothioate linkages (Figure 2.2E).

2.3.3 Adapting a continuous luminescence-based coupled assay to measure the substrate specificity of NrnA$_{Bs}$

Phosphodiesterases that release adenosine monophosphate are widespread amongst all domains of life. AMP is a degradation product that arises in the breakdown of RNAs, c-di-AMP, and cAMP. To quantitatively determine whether NrnA$_{Bs}$ exhibits substrate length specificities we adapted a previously reported continuous coupled luminescence-based assay that couples AMP release to a series of different enzymatic reactions to produce a luminescent signal (Sturm and Schramm, 2017). Originally, this assay format was developed to measure the depurination kinetics of ribosomes by ricin toxin A-chain, however, we noticed that the assay format could be utilized to monitor the release of AMP from various enzyme reactions (Figure 2.3A). Furthermore, we speculated that this coupled enzyme assay format could be adapted for the analysis of other RNases, c-di-AMP phosphodiesterase, and cAMP phosphodiesterases that release AMP as a product (Figure 2.3B).

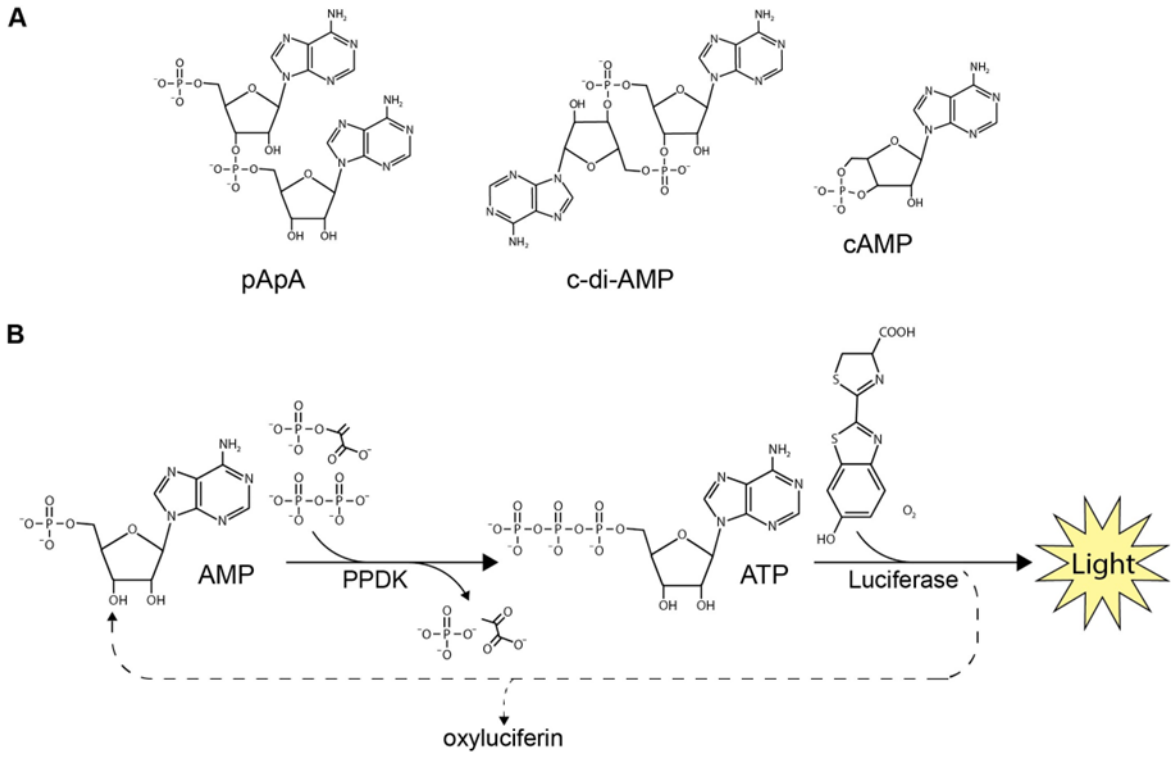

Figure 2.3. Diagram of coupled luminescent assay. (A) Possible substrates for enzymes that could be monitored utilizing the coupled luminescence assay. (B) A schematic showing the series of reactions for converting AMP into ATP by PPDK$_{Cs}$ and the generation of a traceable luminescence signal through the ATP dependent conversion of luciferin to oxyluciferin.

2.3.4 A method for the rapid purification of Clostridium symbiosum Pyruvate Phosphate Dikinase.

There are many different commercially available luciferase-based kits for the detection of ATP. The first step to this coupled assay is in the purification of a surplus of the coupling enzyme pyruvate phosphate dikinase ($PPDK_{Cs}$). Previously published methods for the purification of $PPDK_{Cs}$ involve a lengthy procedure encompassing multiple purification columns. To this end, we sought to simplify the purification by expressing a codon optimized $ppdk_{cs}$ with an N-terminal hexahistidine tag in T7 Express Competent *E. coli*. In our initial attempt to purify His_6-$PPDK_{Cs}$, we grew a small culture of T7 express *E. coli* cells harboring the recombinant plasmid in 2xYT at 37˚C to an $OD_{600} \sim 0.6$, and then induced expression by adding IPTG to a final concentration of 1 mM. Induction at 37˚C continued while shaking for 2 hours. After the induction, *E. coli* cells were harvested by centrifugation and resuspended in a buffer composed of 25 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10 mM imidazole, 5% glycerol. The cells were lysed by sonication, and the cell debris was removed by two rounds of centrifugation at 12,000 rpm for 20 minutes. The clarified extract was incubated with Roche nickel resin for 1 hour. The nickel resin was washed with 10 column volumes of 25 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole, 5% glycerol. Then, $PPDK_{Cs}$ was eluted from the resin with 5 column volumes of 25 mM Tris-HCl, pH 8.0, 300 mM NaCl, 250 mM imidazole, 5% glycerol. In the analysis of this purification by SDS-PAGE, we found that $PPDK_{Cs}$ could be rapidly purified to near homogeneity using this method (Figure 2.4A). However, we observed that there were contaminating nucleic acids as seen by the A260/280 ratio which was > 1.0. Since it was possible that contaminating nucleic acids could negatively influence the coupled enzyme reaction, we decided to utilize polyethyleneimine precipitation prior to running the lysate over the nickel

column to remove nucleic acids (Burgess *et al.* 2012). Additionally, we included an ammonium sulfate precipitation fractionation as a precaution to remove any remaining polyethyleneimine in the protein sample. Following precipitation by ammonium sulfate, $PPDK_{Cs}$ was resuspended in 25 mM Tris-HCl, pH 8.0, 300 mM NaCl, 5% glycerol, and purified by IMAC using Roche nickel resin as described in the small-scale purification protocol, except, a 10-column volume wash with 25 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10 mM imidazole, 5% glycerol, was added prior to the 20 mM imidazole wash. Analysis of the protein purification fractions by SDS-PAGE as seen in (Figure 2.4B), indicates that $PPDK_{Cs}$ could be rapidly purified to near homogeneity after a much-simplified purification procedure.

2.3.5 Testing the activity of purified *Clostridium symbiosum* Pyruvate Phosphate Dikinase

To implement a coupled enzyme assay format, it is crucially important to ensure that the coupling reaction rates never become rate limiting. One method for determining whether the coupled enzymatic reactions are rate limiting is to double the concentration of the initiating enzyme and obtain initial rates. If the initial rates double with the doubling of the initiating enzyme concentration in the assay, this indicates that the coupling enzymatic reactions are not rate limiting. This is because the initiating enzyme velocity should scale proportional to the enzyme's concentration. However, another method for determining that the coupling enzymes are in great excess to this initiator enzyme is to measure the units of enzyme activity and ensure that the initiator enzymes initial rates never exceed the units of activity of the coupling enzymes. Enzyme units are sometimes arbitrarily defined, but most commonly it can be assumed that 1 enzyme unit (U) is the amount of enzyme required to process 1 μM of substrate per minute.

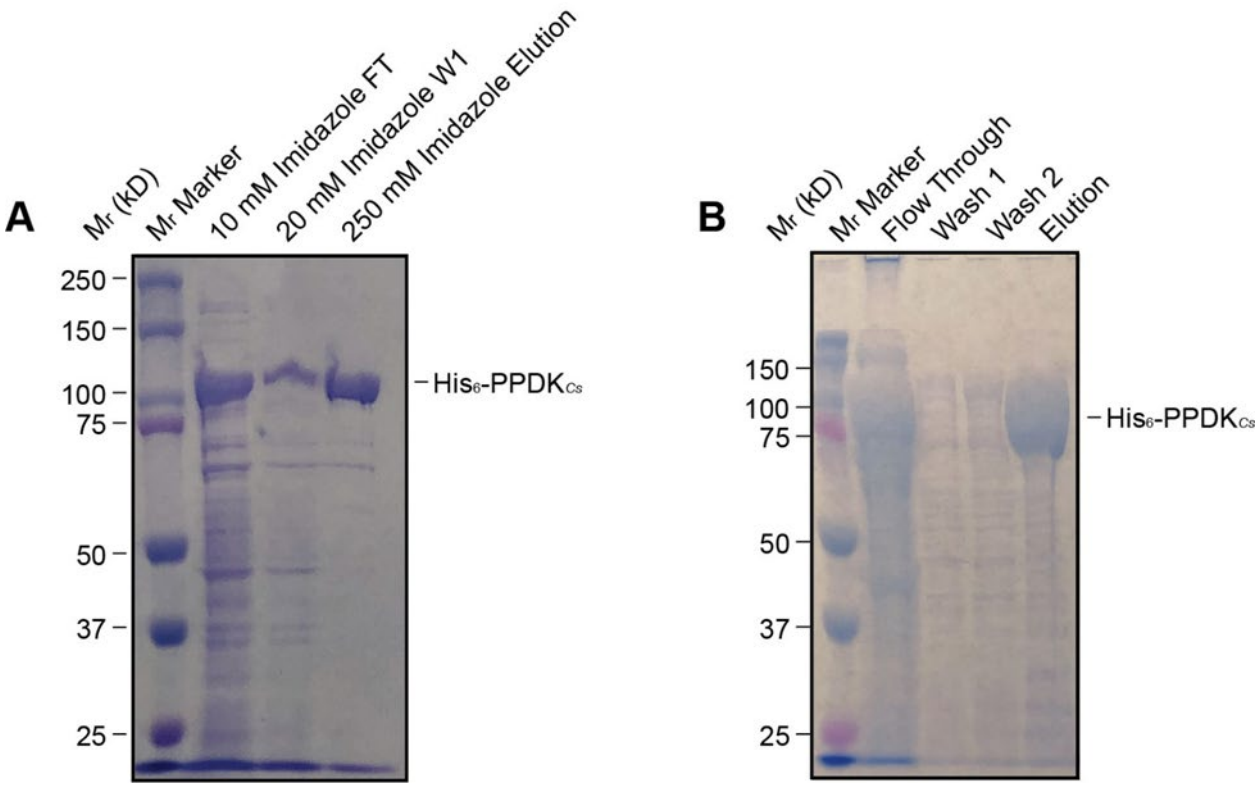

Figure 2.4. Purification of His_6-PPDK. (A-B) SDS-PAGE of fractions of $PPDK_{Cs}$ taken from (A) a small scale purification using IMAC or (B) a large scale purification using PEI precipitation, ammonium sulfate precipitation, and IMAC.

Determining the enzyme units in a purified enzyme preparation can be particularly useful in this scenario because you can ensure that you are not wasting excess purified protein samples in each reaction.

In our analysis of our purified PPDK$_{Cs}$ sample, we opted to determine the units of activity per µL of our purified protein sample. Many previous publications have measured the activity of PPDK$_{Cs}$ spectrophotometrically by coupling pyruvate formation to lactate dehydrogenase and monitoring the change in the oxidation state of NADH by the corresponding decrease in absorbance at 340 nm. In our preliminary tests of PPDK$_{Cs}$'s activity using this method yielded gross underestimations of the enzyme activity of our protein preparations, which may have been due to the limited sensitivity of this assay format. The enzyme units for PPDK$_{Cs}$ can be determined by monitoring the conversion of AMP to ATP and coupling this reaction with luciferase to generate a luminescence signal. To accomplish this, we generated a standard curve by incubating a range of concentrations of ATP with luciferase in the reaction buffer (Figure 2.5A). To assess the activity of PPDK$_{Cs}$ we included 75 µM of AMP to the reaction buffer and monitored the conversion of AMP to ATP by the luminescence output. For this analysis, it is advantageous to make multiple dilutions of your enzyme preparation to ensure that you are obtaining the initial rate of product formation by PPDK$_{Cs}$. If your PPDK$_{Cs}$ preparation is too active, the AMP will be converted to ATP prior to obtaining the initial rate of product formation. For our analysis we found that a 25X dilution of our PPDK$_{cs}$ stock was sufficient to give us a signal corresponding to <10% AMP (i.e., 7.5 µM AMP) to ATP conversion (Figure 2.5B). The initial rate of AMP to ATP conversion indicates that the enzymatic units for our PPDK$_{Cs}$ preparation are 160 U per µL of protein stock.

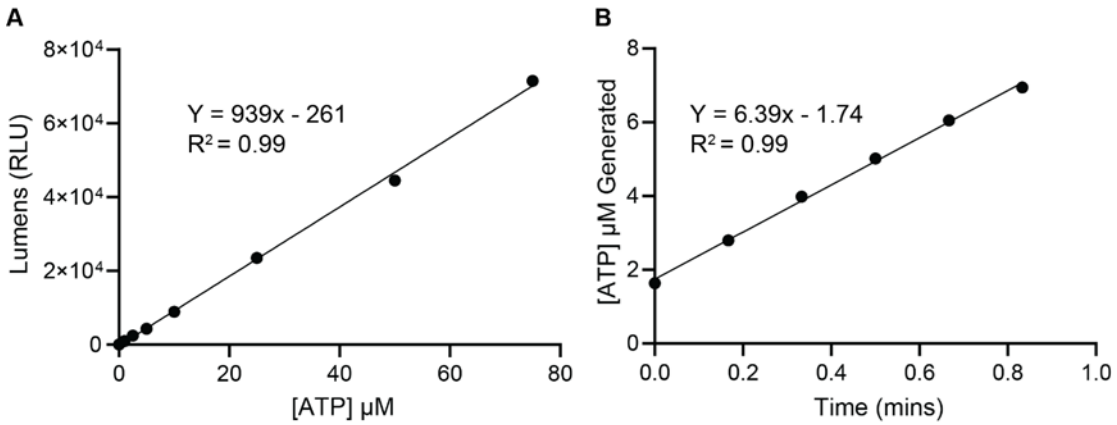

Figure 2.5. Determining the enzymatic activity of purified His_6-PPDK. (A) ATP standard curve (B) Initial rate of AMP to ATP conversion for 1 µL of 25X diluted His_6-PPDK sample incubated with 75 µM of AMP.

2.3.6 Kinetic analysis of NrnA$_{Bs}$

NanoRNase enzymes have been argued to cleave a range of different short RNAs of different lengths. Recent work on from our laboratories on Oligoribonuclease (Orn) and NanoRNase C (NrnC) strongly suggest that these enzymes possess active sites that specifically accommodate dinucleotides (Kim *et al.,* 2019; Lormand *et al*, 2021). 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$_{Bs}$ was originally reported to possess a preference for 3-mer over a 5-mer in cleavage assays using 5′-Cy5 modified oligonucleotides (Mechold *et al.,* 2007). In the degradation of the 5'-Cy5 3-mer RNA, the 3-mer was rapidly processed to a 2-mer prior to much slower processing to monoribonucleotide. Although diribonucleotides were never directly assayed, it remains a possibility that the 5'-Cy5 modification could have influenced these reactions. Also, the NrnA-like protein from *T. maritima* was previously reported to possess a dinucleotide specificity based on the protein's reduced activity in cleavage assays against the trinucleotide pApApA. Hence, we wanted to acquire additional quantitative data directly comparing the kinetics of cleavage for a dinucleotide and trinucleotide substrate. To that end, we utilized a modified luminescence-based coupled enzyme assay to monitor the release of AMP from the cleavage of pApA or pApCpC. First, we generated a standard curve consisting of a range of free AMP concentrations (Figure 2.6A). Next, we incubated purified NrnA$_{Bs}$ with a range of different pApA concentrations and measured the initial rates of AMP release (Figure 2.6B). Substrate degradation was restricted to 10% or less, and a blank containing all reaction materials except for NrnA$_{Bs}$ was included as a control for background luminescence. Since the cleavage of pApA

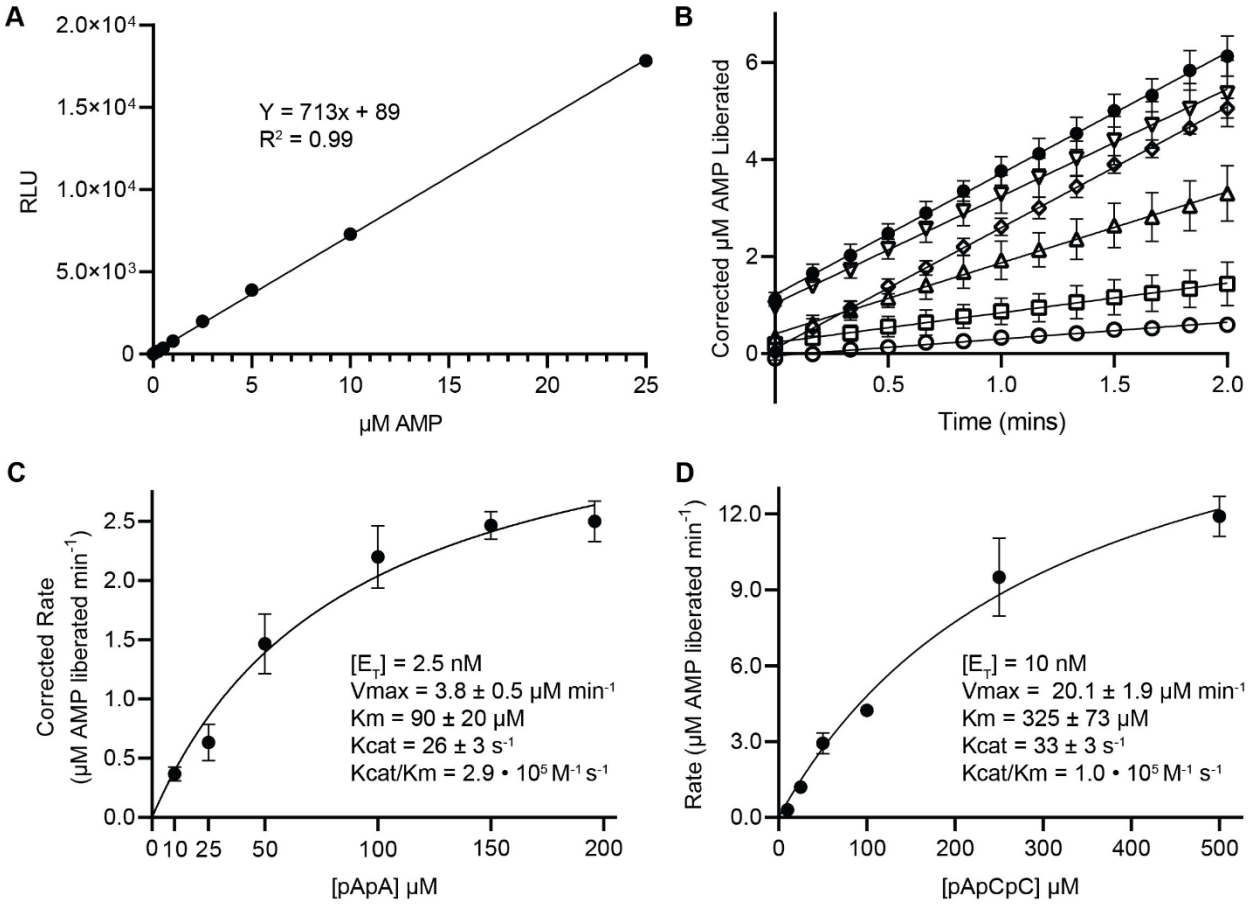

Figure 2.6. Kinetic analysis of *B. subtilis* NrnA. (A) A standard plot of luminescence versus AMP. (B) Initial rate slopes for the cleavage of increasing concentrations of diribonucleotide by 2.5 nM $NrnA_{Bs}$. (C, D) Kinetic curve fits for the cleavage of pApA by 2.5 nM of $NrnA_{Bs}$ (C) or pApApC by 10 nM of $NrnA_{Bs}$. (B-D) All data was plotted as the average and SD of three independent experiments.

liberates 2 molecules of AMP, we corrected the initial rates accordingly. The cleavage of pApA and pApCpC by NrnA$_{Bs}$ both displayed a Michaelis-Menten kinetic profile (Figure 2.6C, D). These data revealed that the K_M values differed modestly between the two RNAs (0.090 versus 0.325 mM). A direct comparison of the specificity constants (k_{cat}/K_M) for these reactions did not differ significantly. Therefore, NrnA$_{Bs}$ acts efficiently against both di- and trinucleotides, but exhibits a slight preference for 2-mers.

2.4 Discussion

Many conventional methods for analyzing RNases and other phosphodiesterases have many drawbacks including the reliance on radioactive labeling and chromatographic separation of small molecules in a discontinuous assay format. Additionally, the cost of radioactive nucleotides is skyrocketing. To avoid these complications, we have adapted fluorescent and luminescence-based assays to quickly investigate the mechanism and kinetic properties of NrnA$_{Bs}$. While we show that these assay formats can be utilized in the analysis of an RNase that preferentially cleaves short RNA substrates, we anticipate that these methods could be further adapted to investigate other RNases, c-di-AMP phosphodiesterases, or cAMP phosphodiesterases..

2.5 Materials and Methods

2.5.1 2-Aminopurine hydrolysis assay

To determine the polarity of NrnA$_{Bs}$ we used a fluorescence-based assay as described previously (Zhou et al. 2017). 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 2-mer RNA (pAp(2AP) or 4-mer RNA substrates (AA*ps*G(2AP) or ((2AP)A*ps*GG) that were purchased from GE Healthcare Dharmacon. 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. In the analysis of $NrnA_{Bs}$ cleavage of the diribonucleotide (pAp(2AP)) the final concentration of $NrnA_{Bs}$ was 10 nM, the concentration of diribonucleotide was 10 µM and the reaction contained 25 mM Tris-HCl, pH 8.0, 300 mM NaCl, 1 mM $MgCl_2$, 50 µM $MnCl_2$. These reactions also contained 25 mM Tris-HCl, pH 8.0, 300 mM NaCl, 200 µM $MnCl_2$. Reactions were conducted in a black 394-well plate using a Spectramax M5 plate reader, and fluorescence was measured every two minutes using an excitation wavelength of 310 and an emission wavelength of 375.

2.5.2 Expression and Purification of Pyruvate Phosphate Dikinase

T7 Express 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 hours. Cells were harvested by centrifugation at 5,000 rpm for 20 minutes. 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, $PPDK_{Cs}$ 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 minutes, after which it was centrifuged at 10,000 rpm for 10 minutes. 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 minutes, and then centrifuged at 10,000 rpm for 10 minutes. At this point, the supernatant was decanted, and the precipitated PPDK$_{Cs}$ 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 cOmpleteTM His-Tag Purification Resin (Roche). The protein was incubated with equilibrated nickel resin for 1 hour 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, PPDKCs 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$_{Cs}$ was dialyzed to remove excess imidazole. Then, purified PPDK$_{Cs}$ was concentrated using Corning® Spin-X® UF concentrators, and small aliquots of protein were flash frozen in liquid nitrogen and stored at -80°C.

2.5.3 Luciferase coupled enzyme assay

The enzymatic activity of NrnA$_{Bs}$ was assessed using a slightly modified continuous kinetic assay as described previously (Sturm and Schramm, 2009). For this analysis, AMP generated from phosphodiester hydrolysis of RNA substrates by NrnA$_{Bs}$ was converted to ATP by pyruvate phosphate dikinase (PPDK$_{Cs}$). In these reactions, ATP was then used by luciferase enzyme to generate a luminescence signal. A 2X 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, 15mM $(NH_4)_2SO_4$. The 2X coupling buffer was filter sterilized and kept frozen. The enzymatic activity of dilutions of PPDK$_{Cs}$ were determined by monitoring the conversion of 75 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 1X coupling buffer containing the ATPlite mix, 10 mM $MgCl_2$, 0.5 mM $MnCl_2$, >50 units of $PPDK_{Cs}$ and the respective concentrations of the RNA substrates and purified $NrnA_{Bs}$ 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 $NrnA_{Bs}$, 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 fitted to standard curves to determine the concentrations of AMP released. For the degradation of pApA, we divided the rates of AMP release by two since two molecules of AMP are liberated for each reaction.

Chapter 3: NrnA is a 5'-3' exonuclease that processes short RNA substrates *in vivo* and *in vitro*

3.1 Copyright Notice

Chapter 3 was originally published by Nucleic Acids Research Journal as: Weiss C.A. [#], Myers T.M. [#], Wu C.H., Jenkins C., Sondermann H., Lee V.T., Winkler W.C. NrnA is a 5′-3′ exonuclease that processes short RNA substrates in vivo and in vitro. Nucleic Acids Res. 2022; 50:12369–12388.

Author contributions: Weiss C.A. [#] and Myers T.M.[#] equally contributed to this work and performed the research. Additionally, Wu C.H. conducted the clustering analysis of the DHH-DHHA1 proteins, Jenkins C. conducted the proteomic analysis of *B. subtilis*, Sondermann H., Lee V.T., aided in the conceptualization of the research and the data analysis, Winkler W.C. oversaw the project, and Weiss C.A. [#], Myers T.M.[#], and Winkler W.C. wrote the paper.

3.2 Introduction

The iterative cycle of polymerizing NTPs into long RNA molecules by DNA-dependent RNA polymerase transcripts and the subsequent breakdown of these molecules by ribonucleases (RNases) is a fundamental process utilized by all organisms to regulate gene expression and to recycle nucleotide precursors. Perturbations in mRNA degradation pathways have been shown to elicit a profound impact on gene expression, thus impacting the ability of a cell to adapt and maintain homeostasis. Additionally, other intracellular species of intracellular RNA such as tRNA, rRNA, and sRNAs undergo RNase mediated processing events that are critical for the growth and development of organisms. It is largely assumed that the general principles of RNA processing and decay are conserved among bacteria, however, the collection of RNases that are encoded for in bacteria differ from microbe to microbe (Hui *et al.,* 2014; Bechhofer and

Deutscher, 2019; Tinquier *et al*., 2020). Although there are some RNases that are broadly conserved such as PNPase, there is no universal set of RNases conserved amongst all bacteria. This is exemplified in the comparison of RNases in the two most widely studied bacterial organisms *E. coli* and *B. subtilis*. Of the 40 RNases that have been identified in *E. coli* and *B. subtilis* only nine of these gene products are shared between these organisms (Trinquier et al., 2020). For both *E. coli* and *B. subtilis*, RNA degradation is initiated by endoribonucleolytic cleavage of long RNA transcripts. Interestingly, despite this mechanistic similarity, *E. coli* utilizes the protein RNase E while *B. subtilis* utilizes an unrelated protein RNase Y (Hui *et al.,* 2014). The internal cleavage by endoribonucleases generates destabilized RNA fragments that can be degraded by exoribonucleases. Generally, exoribonucleases are thought to processively degrade RNAs in a 3'-5' direction. A key difference between *E. coli* and *B. subtilis* RNA degradation pathways is that *B. subtilis* harbors a 5'-3' exonuclease RNase J1 (Mathy *et al.,* 2007; Ghodge and Raushel, 2015; Jain, 2020). The concerted actions of various endo- and exoribonucleases results in the processing of long RNAs down until only short (< 5 nucleotides) RNA fragments remain. These RNA fragments are then degraded by a separate category of RNases that specialize in the processing of short RNAs.

In *E. coli* and other bacteria, the enzyme Oligoribonuclease (Orn) performs the final step of RNA degradation, in the processing of short RNA oligonucleotides. Orn was originally discovered in the 1960's and was shown to process short RNAs 2 – 5 nucleotides in length (Stevens and Niyogi, 1967; Datta and Niyogi, 1975; Datta and Niyogi, 1975; Ghosh and Deutscher, 1999; Mechold *et al.,* 2006). While Orn is essential for the viability of many organisms, it is unknown what the molecular basis for Orn's essentiality is. Previous genetic deletion experiments have shown the other exoribonucleases can be deleted without any severe

phenotypic effects. In fact, it is often the case the multiple general exoribonucleases must be deleted before any deleterious effects arise (Oussenko *et al.,* 2002). This essentiality makes Orn unique amongst other known 3'-5' exoribonucleases, which have been shown in some instances to be functionally redundant. Yet, a recent reassessment of Orn has revealed that this enzyme binds and cleaves dinucleotides with greater specificity than other short RNAs (Kim, *et al.,* 2019). Furthermore, X-ray crystal structures of Orn bound to various dinucleotide substrates have revealed that the substrate binding pocket seems remarkably restricted to dinucleotides. From this reassessment of Orn's activity, it seems as though Orn is most likely a 'diribonuclease', an enzyme that preferentially degrades diribonucleotides rather than long RNA substrates. This discovery implies that the degradation of short RNA molecules might occur through discrete steps and that diribonucleotides might represent a crucial bottleneck in RNA degradation.

The essentiality of Orn in *E. coli* has certainly aided in the discovery of proteins with 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 (Mechold *et al.,* 2007; Fang *et al.,* 2009; *Liu et al.,* 2012). 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 the discovery of $NrnA_{Bs}$ and $NrnB_{Bs}$, it has been assumed that these proteins are likely to behave in a redundant manner to cleave oligoribonucleotides. Orn is comprised of a DnaQ fold which contains a catalytic DEDD motif while $NrnA_{Bs}$ and $NrnB_{Bs}$ are part of the DHH superfamily of proteins (Aravind and Koonin, 1998). The DHH superfamily of proteins includes enzymes predicted to exhibit phosphodiesterase activity towards DNA or RNA molecules. $NrnA_{Bs}$ and

NrnB$_{Bs}$ are comprised of an N-terminal DHH (Pfam PF01368) domain connected to a C-terminal DHHA1 (DHH-associated domain 1, Pfam PF02272) domain. Despite sharing a similar domain architecture, NrnA$_{Bs}$ and NrnB$_{Bs}$ share only ~20% sequence identity. Additionally, the DHHA1 domain subfamily includes proteins of diverse functions, such as NrnA proteins, the c-di-AMP linearizing phosphodiesterase GdpP, alanyl-tRNA synthetase (AlaRS), and RecJ (Avarind and Koonin., 1998). RecJ is a 5'-3' exonuclease that processes single-stranded DNA (ssDNA) and is involved in homologous recombination and DNA repair. (Rao *et al*., 2010; Dianov *et al.,* 1994). Proteins comprised of DHH and DHHA1 domains do not display the same substrate specificity. Furthermore, the annotation of some subfamilies of the 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. For example, has 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 the 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 have led to several different opinions regarding what the enzymatic activity and substrate specificity is for NrnA. One possibility is that NrnA acts as a diribonuclease, functionally redundant with Orn. This speculation stems from the observation that overexpression of *B. subtilis nrnA* or *nrnB* in a *Pseudomonase aeruginosa* PA14 Δ*orn* strain resulted in the restoration of the wild-type PA14 phenotype, as compared to the PA14 Δ*orn* strain that shows a small colony phenotype (Orr *et al.,* 2018). Therefore, NrnA$_{Bs}$ and NrnB$_{Bs}$ either perform a similar function as Orn or exhibits 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' polarity, while others have reported 5'-3' polarity (Mechold *et al*., 2007; Wakamatsu *et al.,* 2011). Yet another study suggested that NrnA could degrade short RNAs with 3'-5' polarity but utilizes 5'-3' polarity to degrade longer RNA substrates (Schmier *et al.,* 2017). If NrnA degrades RNAs with a bidirectional mechanism, this would represent a rare instance where an RNase exhibits dual polarity on RNA substrates. Another analysis argues that NrnA can degrade ssDNA. Furthermore, some stand-alone DHH-DHHA1 domain containing proteins (e.g., *Staphylococcus aureus* Pde2, *Streptococcus pneumoniae* Pde2) were proposed to directly hydrolyze c-di-AMP to AMP in a two-step process via the intermediate AA, although there is not a uniform agreement on that either (Bowman *et al.,* 2016; Bai *et al*., 2013). Whether NrnA-like proteins directly degrade cyclic dinucleotide signaling molecules or possess an expanded range of RNA substrates has been further complicated by analyses of the *Mycobacterium tuberculosis* 'NrnA-like protein' that has been recently renamed to CnpB (or Rv2837c) (Postic *et al.,* 2012; He *et al*., 2016; Dey *et al.,* 2017). 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, other studies have argued that CnpB specifically hydrolyzed c-di-AMP and c-di-GMP to nucleotide monophosphates in addition to linearizing 2'3'-cGAMP (Postic *et al.,* 2012; He *et al*., 2016; Dey *et al.,* 2017). This cyclic nucleotide cleavage activity has also been reported for the *Mycobacterium smegmatis* 'NrnA-like' protein (Tang *et al.,* 2015).

To add further complexity to the intracellular function and substrate specificity of 'NrnA-like' proteins, other studies have linked NrnA-like proteins to sulfur metabolism through their influence on the molecule pAp. Initially, *B. subtilis* NrnA was identified in a pulldown-based

experiment in which *B. subtilis* cellular lysates were incubated with pAp tethered to agarose beads. The protein elution from the pAp based pulldown was run on a gel and bands were excised and subjected to LC-MS/MS (Mechold *et al.,* 2007). Of the candidate pAp binding proteins, $NrnA_{Bs}$ was identified along with GuaC which is a GMP reductase. 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. Moreover, a deletion of *B. subtilis nrnA* resulted in a slower growth rate in the absence of cysteine. However, the reduction in *B. subtilis* growth in for Δ*nrnA* was mild overall, hinting that there might be another undiscovered pAp phosphatase in *B. subtilis*.

There have certainly been a multitude of different claims regarding the substrate specificity and intracellular functions for *B. subtilis* NrnA, which has been further complicated by additional claims for other NrnA-like protein homologs. This could be attributed to the range of different experimental conditions and assay formats utilized to test the activities of the NrnA-like proteins. $NrnA_{Bs}$ and $NrnB_{Bs}$ both complement *orn* deletion mutants, therefore it is easiest to hypothesize that these proteins are all functionally redundant in the processing of diribonucleotides and possibly cellular roles (Orr *et al.,* 2018; Mechold *et al.,* 2007; Fang *et al.,* 2009). However, the aggregate data available for NrnA (and NrnB) paint a confusing picture of what exactly these proteins are doing inside of cells. Is NrnA a ‘diribonuclease’ akin to Orn? Or is it an RNase that specializes in the processing of ‘short’ or ‘long’ RNA substrates? Does NrnA process RNA and DNA substrates, as well as cyclic-di-nucleotides? And with what polarity does NrnA recognize and process substrates? Is this enzyme a unique bidirectional exoribonuclease or does it process RNAs in one direction as most RNases do? To address these questions, we have conducted a biochemical survey of multiple purified NrnA-like proteins and assayed them under

a set of common reaction conditions. This could lead to a clearer definition of NrnA's activity and could aid in the discovery of unique members of the DHH-DHHA1 protein family.

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 (Kim, *et al.,* 2019). We reasoned that this approach would provide the best evidence for $NrnA_{Bs}$'s substrate preferences, including the enzyme's directionality. To examine whether $NrnA_{Bs}$'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 $NrnA_{Bs}$ and $NrnB_{Bs}$ could affect the abundance of other short RNAs in vivo, we examined the cleavage pattern of a radiolabeled 10-merRNA in cellular extracts of wild-type, Δ*nrnA* and Δ*nrnB*. Finally, we also purified NrnA-like proteins from several species (*Streptococcus pyogenes* Pde2, *Enterococcus faecalis* NrnA, *Mycobacterium tuberculosis* CnpB, *Mycobacterium avium* CnpB, *Mycobacterium smegmatis* CnpB, and *Rhodococcus ruber* CnpB-like protein) and assayed them using the same reaction conditions as with $NrnA_{Bs}$. This, we reasoned, would show whether different NrnA proteins behave similarly when compared under the same assay conditions. Together, our aggregate data show that NrnA proteins exhibit broader substrate preferences as compared to Orn. While Orn might preferentially process dinucleotides, NrnA acts as a housekeeping enzyme for degradation of short RNAs between 2 and 4 nucleotides in length and processes them from their 5′ terminus.

3.3 Results

3.3.1 Δ*nrnA B. subtilis* accumulates cyclic-di-GMP

Previous analysis of Orn's biochemical activity has led us to reconsider the individual roles that $NrnA_{Bs}$ and $NrnB_{Bs}$ have *in vivo*. Orn's role as a diribonuclease (Kim *et al.,* 2019) has

implications in cyclic-di-GMP signaling, as the diribonucleotide 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* Δ*orn* mutants suggested that these enzymes are responsible for GG cleavage in *B. subtilis* in a manner that is analogous to Orn function (Orr *et al.,* 2018; Mechold *et al.,* 2007; Fang *et al.,* 2009). It has been shown previously that an increase in GG leads to a concomitant increase in c-di-GMP through feedback inhibition (Chan *et al.,* 2004; De *et al*., 2009; Yang *et al.*, 2011; Lacey *et al.,* 2010). Previously, a double Δ*nrnA*Δ*nrnB* mutant was assayed for c-di-GMP levels using a fluorescent c-di-GMP riboswitch reporter (Orr *et al.,* 2018; Weiss *et al.,* 2019; Weiss and Winkler 2021). This revealed that the levels of c-di-GMP, and consequentially 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, Δ*nrnA*, or Δ*nrnB* cellular backgrounds (Figure 3.1A). The c-di-GMP responsive riboswitch decreases downstream gene expression in response to an increase in c-di-GMP; therefore, resulting YFP expression is inversely coupled to c-di-GMP abundance in these strains. Assessment of a constitutive *yfp* control reporter showed mild to no differences in fluorescence between the wild-type strain and the single Δ*nrnA* or Δ*nrnB* mutants (Figure 3.1B, D). In agreement with our prior observations (Weiss *et al.,* 2019), 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

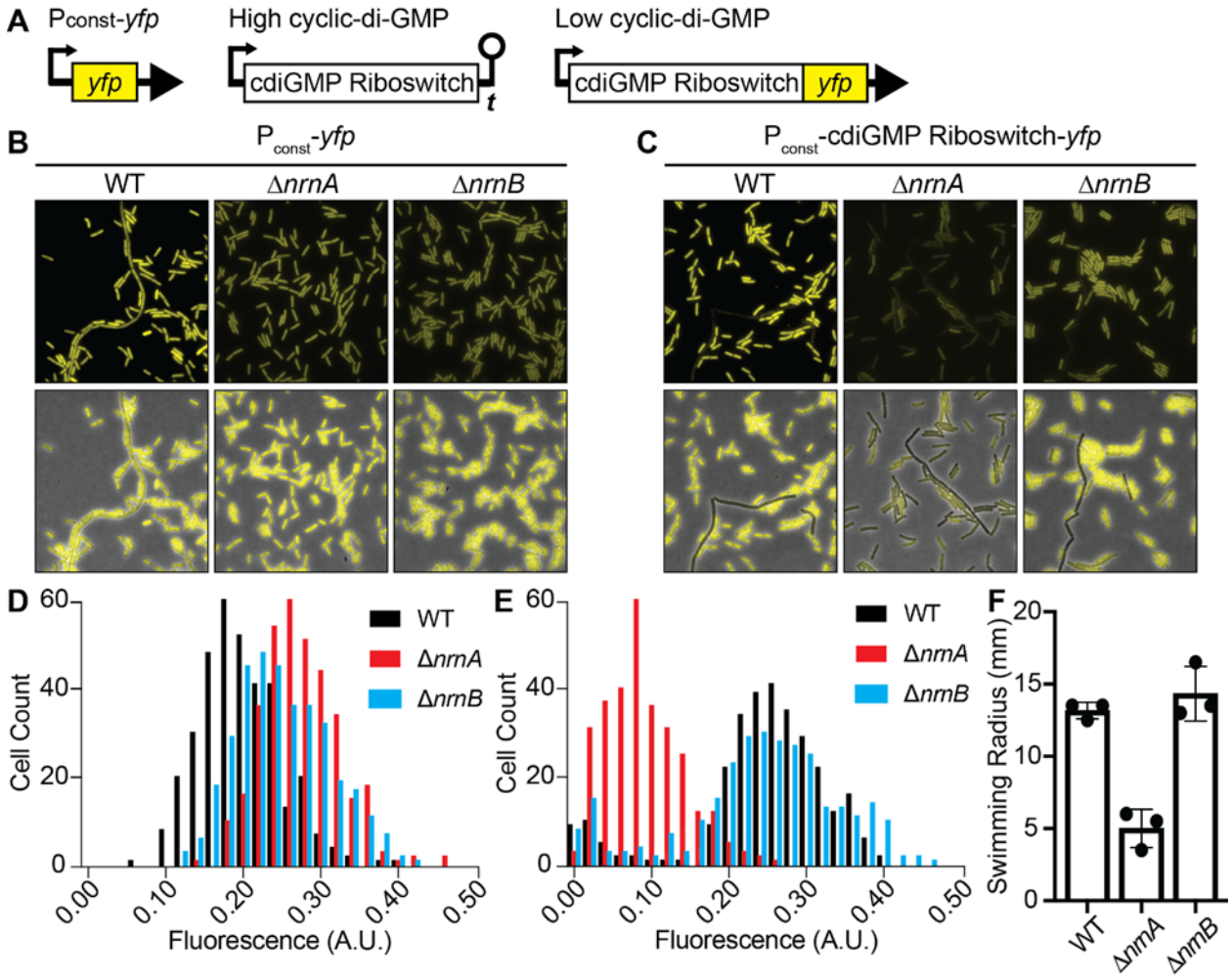

Figure 3.1. Δ*nrnA B. subtilis* accumulates cyclic-di-GMP (A) A schematic is shown for a Pconst-*yfp* reporter and a c-di-GMP-responsive riboswitch reporter, in the presence of high or low intracellular c-di-GMP levels. A similar riboswitch reporter was previously described (Orr et al., 2018). (B) Representative fluorescence microscopy images for cells containing the constitutively expressed YFP reporter Pconst-*yfp*. These strains consisted of wild-type (WT) *B. subtilis*, Δ*nrnA*, or Δ*nrnB*. (C) Representative microscopy images of the c-di-GMP-responsive riboswitch reporter expressed in WT, Δ*nrnA* or Δ*nrnB* strains of *B. subtilis*. The top row of pictures shows only fluorescence data while the bottom row shows the fluorescence data overlayed over phase contrast images. (D, E) Histograms of the quantification of average fluorescence intensity of individual cells for *B. subtilis* 168 WT, Δ*nrnA* or Δ*nrnB*, containing either Pconst-*yfp* (D) or the c-di-GMP responsive riboswitch reporter (E) (n = ~300). (F) Swimming motility analysis of WT, Δ*nrnA* or Δ*nrnB* strains of *B. subtilis* 168. Growth medium contained 0.2% agar and was analyzed after a 12-h incubation period at 30˚C. *Fluorescence microscopy performed by Dr. Cordelia Weiss, Winkler laboratory, University of Maryland.*

fluorescence) (Figure 3.1C, E). In the Δ*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 Δ*nrnB* strain, like wild-type (Weiss *et al.,* 2019) (Figure 3.1D). This suggests that if $NrnB_{Bs}$ does cleave GG during late exponential phase of growth (OD600 ~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 (Jenal, Reinders, and Lori, 2017). 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 (Chen *et al.,* 2012; Subramanian *et al*., 2017). *B. subtilis* uses flagella for swimming through liquid as well. Fortifying media with 0.2% agar allows the pores in the agar to be sufficiently large enough to discourage swarming over the surface, but rather, permit swimming (Subramanian *et al*., 2017). The Δ*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 3.1F). In contrast, the Δ*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 GG pools during exponential phase growth.

3.3.2 Purified $NrnA_{Bs}$ cleaves RNAs 2-4 nucleotides in length

To fully understand its length preferences, purified $NrnA_{Bs}$ was tested for cleavage of different RNA substrates. $NrnA_{Bs}$ was incubated with 5′-^{32}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$_{Bs}$ showed that it fully processed the dinucleotide AA into nucleoside monophosphates in 5 min (Figure 3.2A, B). Importantly, NrnA$_{Bs}$ also showed similar activity against a 3-mer RNA (AGG) within the same time frame. 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.

3.3.3 NrnA$_{Bs}$ does not cleave long RNA substrates *in vitro*

Using the same reaction conditions that were used to test short RNA substrates, we incubated NrnA$_{Bs}$ with 6-mer, 7-mer, 10-mer and 15-mer substrates. This revealed no detectable cleavage activity against the longer substrates (Figure 3.2B, C). Indeed, when NrnA$_{Bs}$ 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 (3.2D). These data mildly conflict with prior reports that NrnA$_{Bs}$ degrades longer nucleic acid substrates greater than four nucleotides in length (Mechold *et al.,* 2007; Wakamatsu *et al.,* 2011; Schmier *et al.,* 2017). Our data strongly suggest that NrnA$_{Bs}$ 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$_{Bs}$. Additionally, these data indicate that the processing of 2–4 mer RNAs by NrnA$_{Bs}$ is only

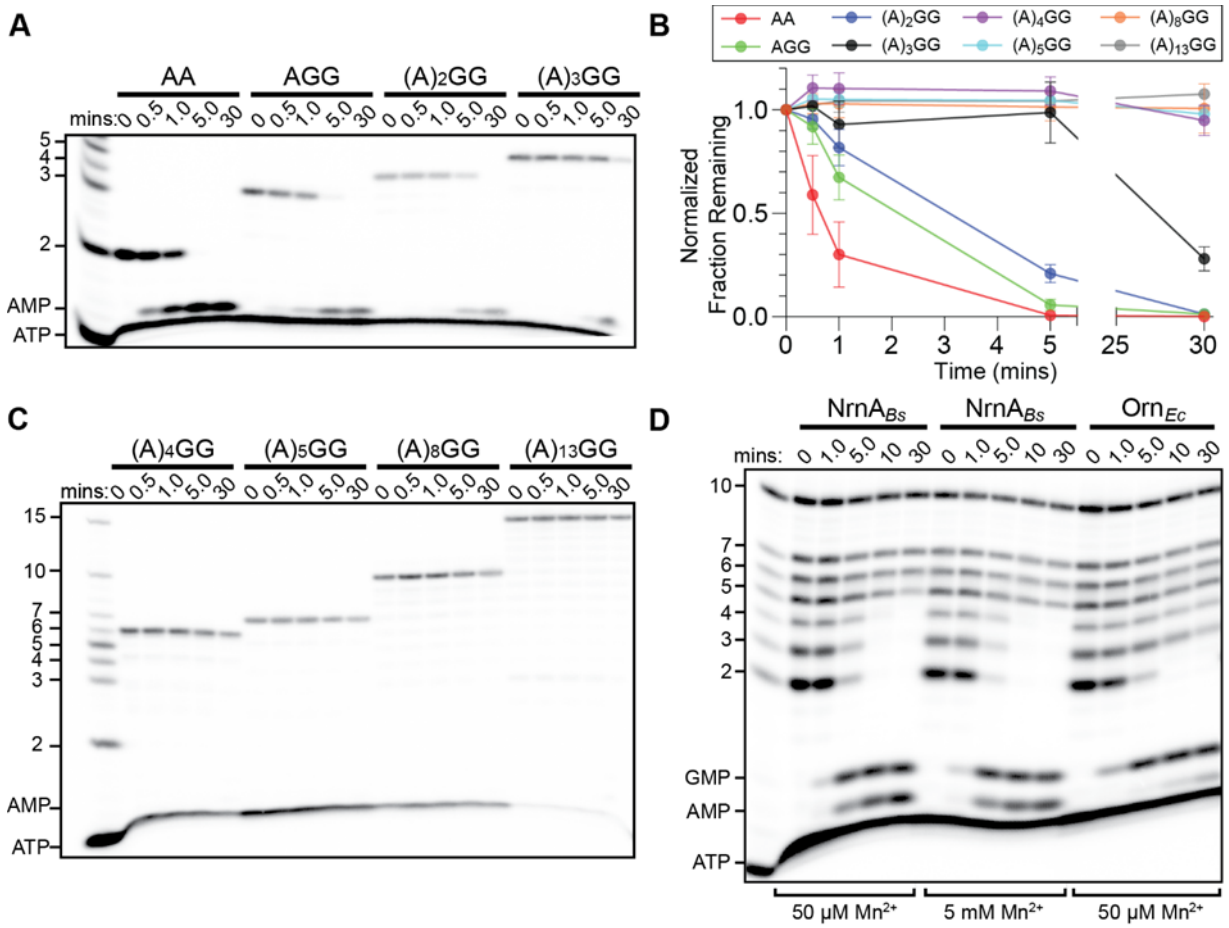

Figure 3.2. Purified NrnA$_{Bs}$ cleaves RNAs 2-4 mer RNAs. (A-C) 1 μM of RNA substrates containing trace amounts of ^{32}P-radiolabeled RNA were incubated with 100 nM of purified Nrn$_{Bs}$. The RNA substrates ranged from 2-7, 10 or 15 nucleotides in length. Aliquots were removed at various time points 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 ^{32}P-radiolabeled RNA substrates that ranged in size (including 2–7, and 10-mer RNAs) were simultaneously incubated with 20 nM of purified NrnA$_{Bs}$ or Orn$_{Ec}$. 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. *RNase activity assays in (A-C) were performed by Dr. Cordelia Weiss, Winkler laboratory, University of Maryland.*

modestly affected by increasing the concentration of Mn^{2+}. Interestingly, when Orn_{Ec} was simultaneously incubated with radiolabeled RNAs 2–7, and 10 nucleotides in length, dinucleotides were rapidly processed, and only minimal trinucleotide processing was seen over the longer time points (Figure 3.2D). In total, these data demonstrate that under reaction conditions where Orn_{Ec} displays diribonuclease activity, $NrnA_{Bs}$ instead displays a specific preference for short RNAs between 2–4 nucleotides in length.

3.3.4 $NrnA_{Bs}$ preferentially binds short RNAs

We next sought to correlate $NrnA_{Bs}$ cleavage preferences with binding affinities by using Differential Radial Capillary Action of Ligand Assay (DRaCALA) (Roelofs *et al*., 2011). Structural analyses of $NrnA_{Bs}$ and other NrnA-like proteins have demonstrated that the NrnA active site incorporates four aspartate residues ($NrnA_{Bs}$ D24, D26, D80, D156), and three histidine residues ($NrnA_{Bs}$ 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 (Wakamatsu *et al.,* 2011; Schmier *et al.,* 2017; Srivastav *et al.,* 2014), although magnesium can also support catalysis in vitro. Herein, we incubated $NrnA_{Bs}$ with 5′-radiolabeled AA in the presence of magnesium or manganese and assayed for protein-RNA complexes by DRaCALA (Figure 3.3A). 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 $NrnA_{Bs}$ that had been treated with the divalent cation chelator EDTA (Figure 3.3A). $NrnA_{Bs}$ is not catalytically active in the presence of excessive 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) (Drexler *et al.,* 2017). This mutant protein exhibited no detectable RNase cleavage activity, although it demonstrated binding activity to AA when incubated with divalent metals (Figure 3.3B). A survey of varying dinucleotide concentrations revealed that NrnA$_{Bs}$D80ND156N bound GG best, with an apparent equilibrium dissociation constant (K_d) of 210 nM, while other dinucleotides (GA, AA, UA and CA) exhibited modestly poorer binding affinities (Figure 3.3A, 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$_{Bs}$ in vitro (Figure 3.3C). This showed that the dinucleotides were all equally good substrates for NrnA$_{Bs}$, with only minor differences in their rates of cleavage. As it was previously reported that NrnA$_{Bs}$ could exhibit activity against single-stranded DNAs (Wakamatsu *et al.,* 2011), we also compared cleavage of a DNA dinucleotide alongside the RNA dinucleotides. This revealed that NrnA$_{Bs}$ is indeed fully active against a DNA dinucleotide substrate (Figure 3.3C). From these aggregate data, we conclude that NrnA$_{Bs}$ 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 bound of radiolabeled AA when co-incubated in the presence of unlabeled (10 or 100 μM) putative substrates (Figure 3.3D). 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

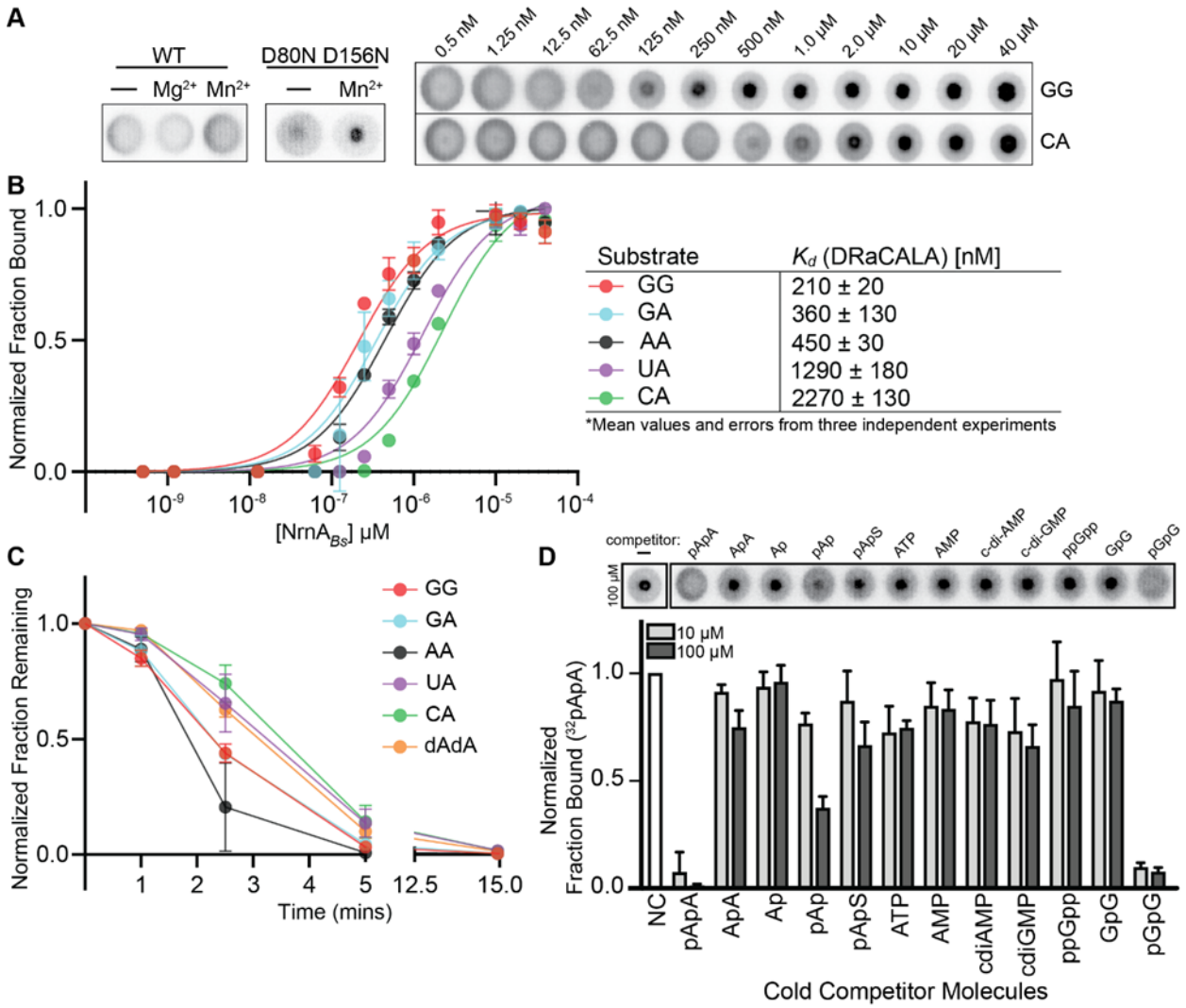

Figure 3.3. *B. subtilis* NrnA possesses a dinucleotide substrate length preference with a slight preference for 5′ purine residues. (A, B) DRaCALA, as described elsewhere (Roelofs *et al*., 2011), was used to measure binding of different radiolabeled dinucleotides to the catalytically inactive mutant NrnA$_{Bs}$ D80N D156N. (C) For cleavage assays, 1 µM of the indicated dinucleotide substrate was incubated with trace amounts of ^{32} 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$_{Bs}$ D80N D156N was bound to ~1 nM of ^{32}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$_{Bs}$ D80N D156N to ^{32}P-radiolabeled AA. (B-D) all data was plotted as the average and SD of three independent experiments. *Cold competition binding assays in (D) were performed by Dr. Cordelia Weiss, Winkler laboratory, University of Maryland.*

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 $NrnA_{Bs}$. pAp is a cytoplasmic carrier of sulfate when converted to pApS (Spiegelberg *et al.,* 1999); however, pApS was unable to compete for binding to $NrnA_{Bs}$. 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. Finally, the signaling molecule ppGpp did not compete for AA binding, despite previous reports that this molecule could compete with the membrane bound DHH-family member GdpP (Rao *et al.,* 2010; Bowman *et al.,* 2016). These data reveal that the $NrnA_{Bs}$ binding pocket is best suited for canonical dinucleotides and is unlikely to act directly on nucleotide signals.

3.3.4 The substrate preferences of $NrnA_{Ef}$ and $Pde2_{Sp}$ closely resemble $NrnA_{Bs}$

There have been many claims regarding the substrate preferences amongst the NrnA-like protein homologs. To extend our analysis, we assessed the short RNA degradation activities of purified NrnA-like proteins homologs encoded by *S. pyogenes* or *E. faecalis*. Both the *S. pyogenes* and *E. faecalis* NrnA-like proteins have been implicated in the processing of c-di-AMP, but neither have been assessed for cleavage of short RNAs (Fahmi *et al*., 2019; Kundra *et al.,* 2021). Yet, the sequence identity between these proteins and $NrnA_{Bs}$ are considerably high at 45% for *S. pyogenes* Pde2 and 47% for the *E. faecalis* NrnA. Therefore, we wanted to directly compare the substrate preferences of these enzymes ($Pde2_{Sp,}$ $NrnA_{Ef}$) against $NrnA_{Bs}$ using the same substrates and reaction conditions. For this analysis, we incubated purified $Pde2_{Sp}$ or $NrnA_{Ef}$ 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

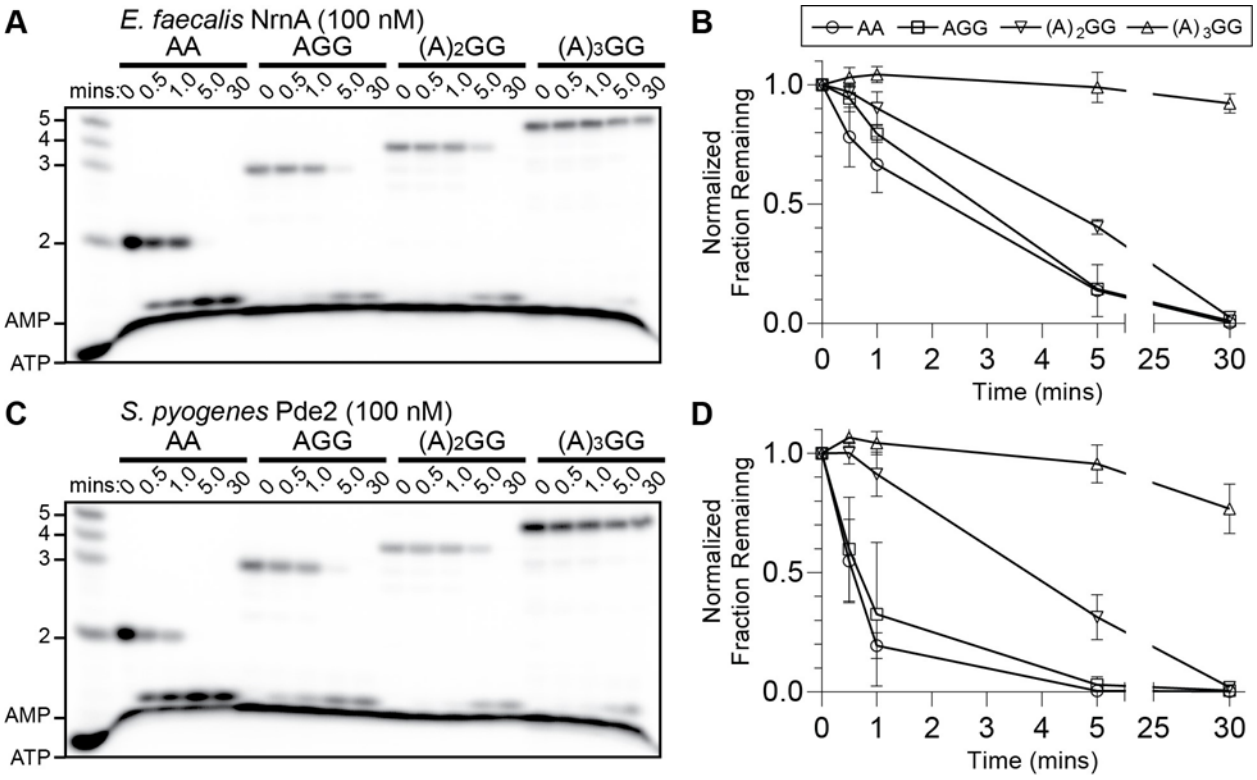

Figure 3.4. The substrate preferences of NrnA$_{Ef}$ and Pde2$_{Sp}$ closely resemble NrnA$_{Bs}$. (A, C) 100 nM of the respective protein was incubated with 1 μM of RNA substrate, 2–5 nucleotides in length, containing trace amounts of ^{32}P-radiolabeled substrates. Aliquots from reactions were removed at the indicated times and quenched in 150 mM EDTA and 4 M urea, and degradation products were resolved by denaturing PAGE. (B, D) The fraction of initial substrate was plotted over time as the average and SD of three independent experiments.

by 20% denaturing PAGE (Figure 3.4). From these data, it is apparent that NrnA$_{Ef}$, and Pde2$_{Sp}$ process short RNAs between 2–4 nucleotides in length, remarkably like NrnA$_{Bs}$. Moreover, both proteins fully processed the dinucleotide AA into nucleoside monophosphates within the same approximate time frame as NrnA$_{Bs}$ (Figure 3.4A–D). A trinucleotide RNA substrate was also fully processed by both proteins within a time frame that resembles dinucleotide substrate.

3.3.5 CnpB$_{Mt}$ exhibits limited RNase activity on 2-mers and 3-mers

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 (Postic *et al.,* 2012; He *et al*., 2016; Dey *et al*., 2017; Yang *et al*., 2014). To directly compare the activity of CnpB$_{Mt}$ to that of NrnA$_{Bs}$ we assessed the cleavage of 5'-radiolabeled RNAs 2-7, 10, and 15 nucleotides in length. Interestingly, while CnpB$_{Mt}$ 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 3.5A-D). Also, an assessment of longer substrates showed essentially no cleavage activity for CnpB$_{Mt}$ against them (Figure 3.5C-D). While there could slight differences in the percent active fractions of the difference purified protein preparations, these data indicate that the NrnA$_{Ef}$, and Pde2$_{Sp}$ proteins appear to exhibit substrate preferences that closely resemble NrnA$_{Bs}$, while CnpB$_{Mt}$ exhibits modest activity against linear RNAs.

3.3.6 NrnA$_{Bs}$ does not act on c-di-AMP

While our binding analysis of the NrnA$_{Bs}$ D80N D156N mutant suggested that the protein was not likely to bind c-di-NMPs, other DHH-DHHA1 domain-containing proteins have been observed to specifically hydrolyze c-di-AMP; therefore, we wanted to directly test for this

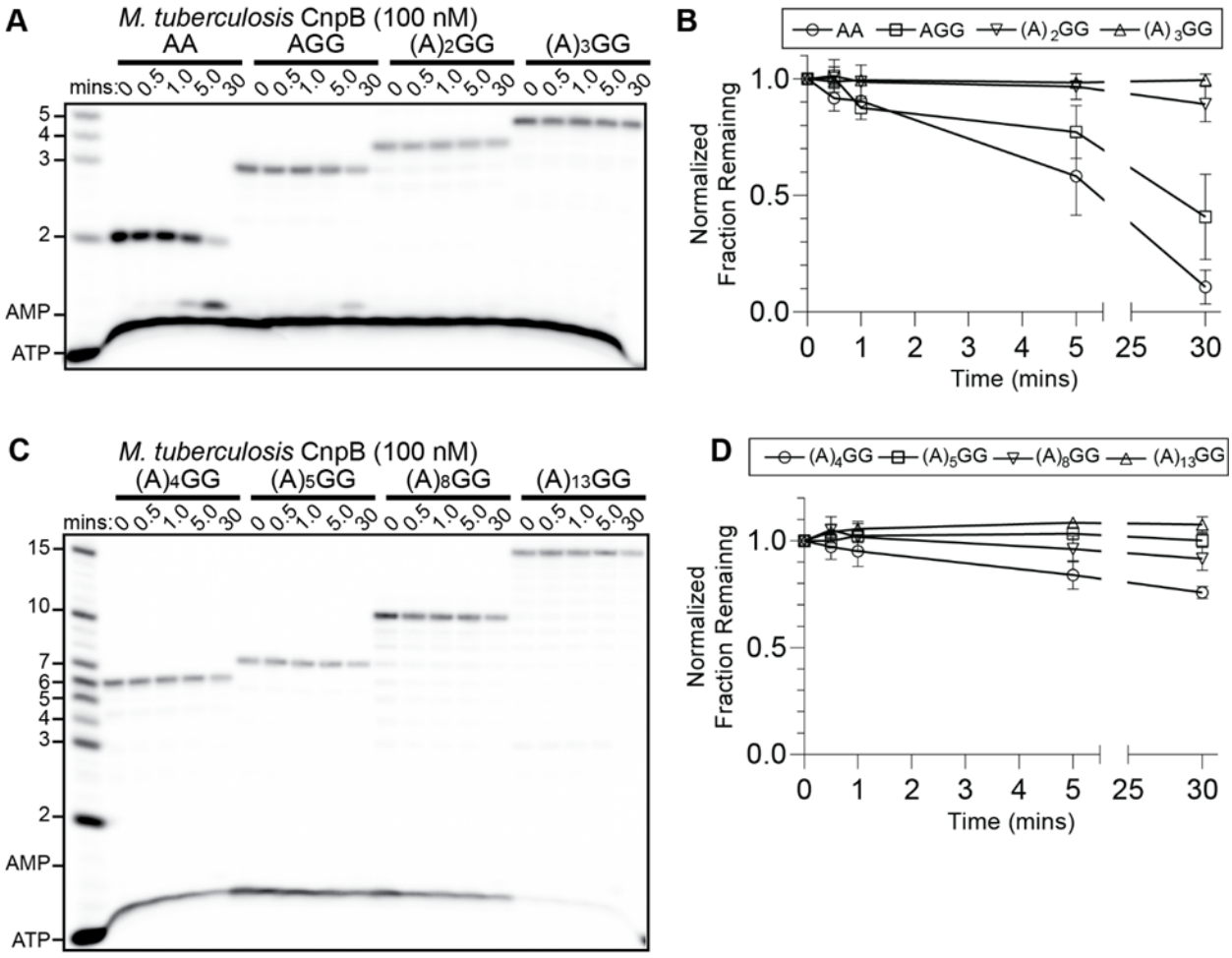

Figure 3.5. CnpB$_{Mt}$ exhibits limited RNase activity on 2-mers and 3-mers. (A) 100 nM of the purified CnpB$_{Mt}$ was incubated with 1 μM of RNA substrate, 2–5 nucleotides in length, or (C) 6, 7, 10 or 15 nucleotides in length containing trace amounts of ^{32}P-radiolabeled substrates. Aliquots from 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) The fraction of initial substrate was plotted over time as the average and SD of three independent experiments.

activity using the same reaction conditions as for cleavage of small linear RNA substrates. GdpP is one of the DHH-DHHA1 domain containing proteins that has been implicated in cleavage of c-di-AMP. Unlike NrnA, which only contains the DHH-DHHA1 domains, GdpP also includes an N-terminal transmembrane domain, a heme-binding PAS domain, and a degenerate GGDEF domain (Rao *et al.,* 2010). However, in contrast to the architectural complexity of GdpP, *M. tuberculosis* CnpB only contains the DHH-DHHA1 domains but has also been reported to exhibit c-di-NMP phosphodiesterase activity (He *et al.,* 2016; Dey *et al.,* 2017). Yet, this observation has been debated and, akin to $NrnA_{Bs}$, multiple other activities have been attributed to $CnpB_{Mt}$, such as cleavage of RNAs from 2–5 nucleotides in length, with an Orn-like preference for dinucleotides (Postic *et al.,* 2012). Furthermore, despite showing c-di-AMP hydrolysis activity in vitro, it was recently reported that expression of *cnpB in vivo* is entirely independent of c-di-AMP (Zhang *et al*., 2018). 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 (Bowman *et al.,* 2016, Bai *et al.,* 2013). To determine whether NrnA-like proteins had the ability to cleave c-di-AMP, $NrnA_{Bs}$, $CnpB_{Mt}$, $NrnA_{Ef}$ and $Pde2_{Sp}$ were purified and then tested for cleavage of ^{32}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 3.6A, C); no cleavage activity was detected for $NrnA_{Bs}$, $NrnA_{Ef}$ and $Pde2_{Sp}$. 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

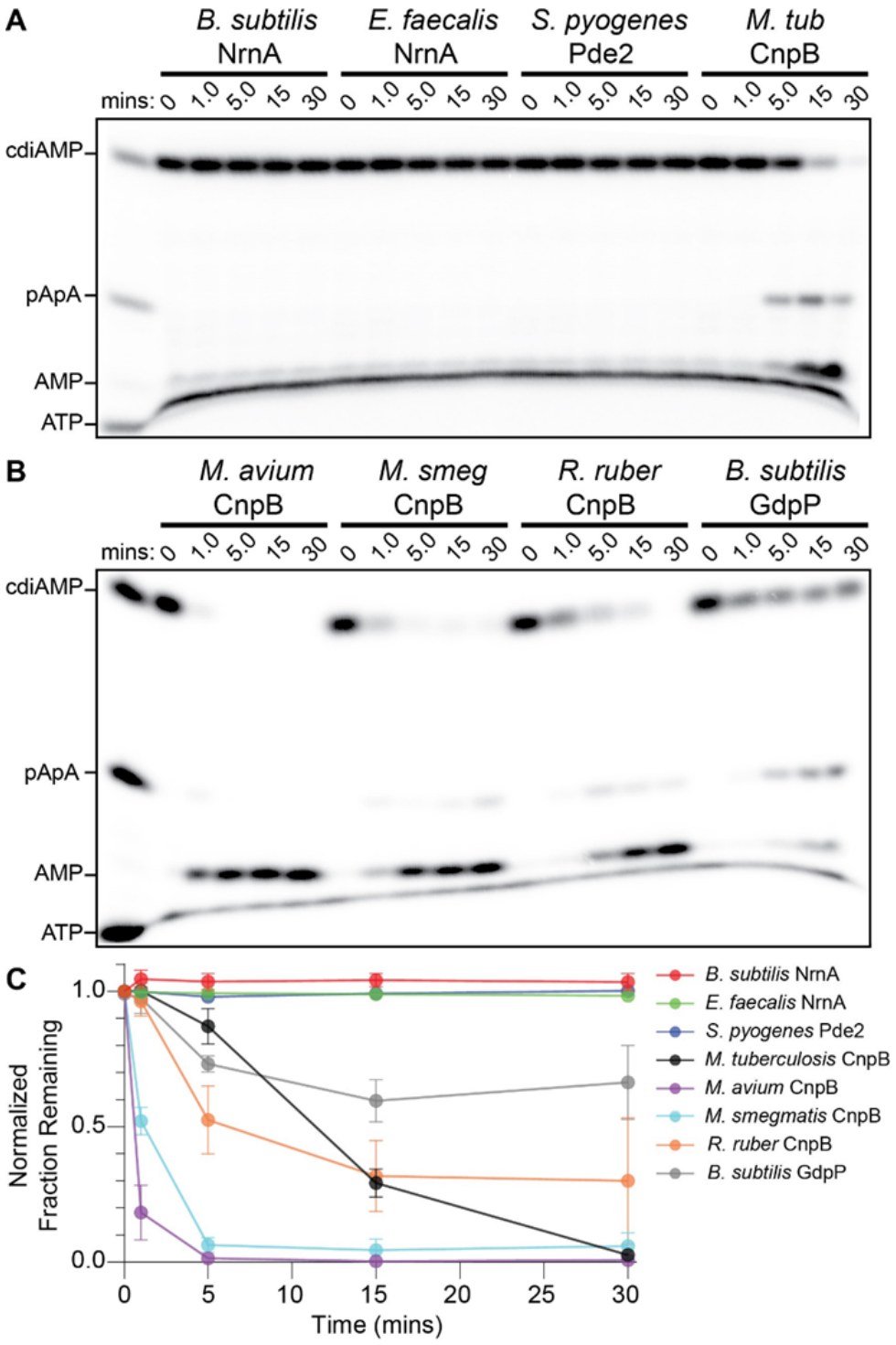

Figure 3.6. NrnA-like proteins differ in their capacities to process c-di-AMP. (A, B) 100 nM of the respective protein was assayed for c-di-AMP phosphodiesterase activity against 1 µM of substrate containing trace amounts of ^{32}P-labeled c-di-AMP. Aliquots were removed from the individual reactions 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.

(Figure 3.6B, C). As a positive control, we purified and analyzed the cytoplasmic portion of *B. subtilis* GdpP, which is known to hydrolyze c-di-AMP (Rao *et al.,* 2010). All CnpB-like proteins exhibited robust hydrolysis of c- di-AMP (Figure 3.6A-C). Together, these data suggest that $CnpB_{Mt}$ may broadly feature active site residues that make it functionally different from other NrnA-like proteins, which do not directly process c-di-AMP.

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

Multiple NrnA-like proteins have been described as pAp phosphatases which affect sulfur metabolism by altering the homeostasis of the molecule 3'-phosphoadenosine-5'-phosphosulfate (pApS) (Mechold *et al.,* 2007; Postic *et al*., 2012). This is likely due to the initial observation that $NrnA_{Bs}$ was pulldown using a pAp based pulldown assay (Mechold *et al.,* 2007). Additionally, subsequent cleavage analysis resolved by thin layer chromatography of purified $NrnA_{Bs}$ against pAp led researchers to conclude that $NrnA_{Bs}$ removes the 3' phosphate from pAp at a rate amenable to that of a bona-fide pAp phosphatase from *E. coli* referred to as CysQ. Interestingly, researchers reported that the addition of various concentrations of pAp to RNase cleavage assays only resulted in a modest effect on $NrnA_{Bs}$ in the cleavage of an RNA 5-mer. These data taken altogether suggest that $NrnA_{Bs}$ might possess a preference for short RNAs as compared to pAp. In our binding analysis, we found that pAp and pApS were only moderate and poor binding competitors for the $NrnA_{Bs}$ enzyme, respectively. To assess $NrnA_{Bs}$'s pAp cleavage activity more quantitatively, we wanted to directly compare the activity of various NrnA-like proteins alongside a positive control. Specifically, we analyzed the cleavage activity of purified $CysQ_{Ec}$, $NrnA_{Bs}$, $NrnA_{Ef}$, $Pde2_{Sp}$, and $CnpB_{Mt}$ against pAp in the presence of divalent cations. Each protein was incubated with 100 μM pAp and the release of inorganic phosphate was assessed using the malachite green detection assay. Under these conditions 5 nM of $CysQ_{Ec}$

Table 1 Concentration of free phosphate liberated by respective protein after incubation with 100 μM pAp detected by Malachite Green phosphate detection assay

	Concentration of phosphate liberated from pAp (mean ± SD) (μM)	
Protein	5 minutes	30 minutes
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

liberated 6.3 and 31.4 μM of phosphate in 5 and 30 minutes, respectively (Table 1). However, free phosphate was not detected at either the 5 minute or the 30 minute timepoint following the incubation of 5 nM of the other NrnA-like protein homologs ($NrnA_{Bs}$, $NrnA_{Ef}$, $Pde2_{Sp}$, and $CnpB_{Mt}$) with pAp. When we increased the concentration of purified $NrnA_{Bs}$ to 50 nM and 500 nM, we were able to detect free phosphate release. This observation might explain the previously published reports regarding the pAp phosphatase activity of NrnA-like proteins (Mechold *et al.,* 2007; Postic *et al.,* 2012). Furthermore, this analysis is consistent with a recent report (Gall *et al.,* 2021) that showed the NrnA-like protein homolog from *Listeria monocytogenes* exercise a strong cleavage preference for linear diribonucleotides compared to pAp. The lack of pAp phosphatase activity coupled with the inability of pAp or pApS to compete for diribonucleotide binding strongly suggests that NrnA-like proteins do not act on these molecules under physiological conditions.

3.3.8 $NrnA_{Bs}$ is required for processing of short RNAs *ex vivo*

Deletion of *nrnA* does not appear to affect growth in rich medium (Fang *et al.,* 2009); however, our *in vitro* data suggested that NrnA preferentially cleaves short RNAs. To investigate whether $NrnA_{Bs}$ might meaningfully affect these RNA substrates in vivo, we incubated a radiolabeled RNA with lysates extracted from *B. subtilis* cells grown to late-exponential phase (OD 600 ~1.0). Specifically, a 5′-radiolabeled 10-mer RNA was mixed with cell lysates that had Been extracted from wild-type, Δ*nrnA*, or Δ*nrnB* strains and the products of these reactions were then resolved by 20% urea-denaturing PAGE (Figure 3.7A). This revealed that the 10-mer was readily degraded to nucleoside monophosphates by wild-type lysates. It also showed that Δ*nrnB* lysates generated a degradation profile that was qualitatively identical to wild-type (Figure 3.7A). In contrast, there was specific accumulation of 2–4-mers when the 10-mer RNA was

incubated with Δ*nrnA* lysates, concurrent with reduced levels of nucleoside monophosphates (Figure 3.7A). This observation suggests that $NrnA_{Bs}$ 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 $nrnA_{Bs}$, $nrnB_{Bs}$, orn_{Vc} or $cysQ_{Ec}$ into a nonessential locus of the Δ*nrnA* strain and prepared lysates for RNA degradation (Figure 3.7B). An empty vector control was included in this analysis. Δ*nrnA* complemented with an empty vector resembled the Δ*nrnA* strain, showing again an accumulation of 2 – 4 mers. In contrast, bands corresponding to 2- to 4-mers decreased upon complementation with *nrnA*, resembling lysates from the wild-type strain. Although Δ*nrnB* lysates did not show any change in degradation of the 10-mer RNA, complementation of the Δ*nrnA* strain with *nrnB* surprisingly resulted in complete processing of the RNA substrate, including all intermediates. Consistent with its role in processing dinucleotides, the complementation with *orn* led to a marked decrease in 2-mers, alongside a moderate decrease in 3-mers. However, complementation of Δ*nrnA* with *cysQ* showed the same RNA accumulation as Δ*nrnA*, confirming that CysQ does not play a direct role in the processing of RNAs. Since addition of the 10-mer RNA led to accumulation of short RNAs in lysates from the Δ*nrnA* 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 (Orr *et al*., 2018; Bowman *et al*., 2016; Fahmi *et al.,* 2019). C-di-GMP abundance is decreased in *B. subtilis* motile cells (Weiss *et al*., 2019); 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 end-point assay (Figure 3.7C). This revealed that the Δ*nrnA* mutation

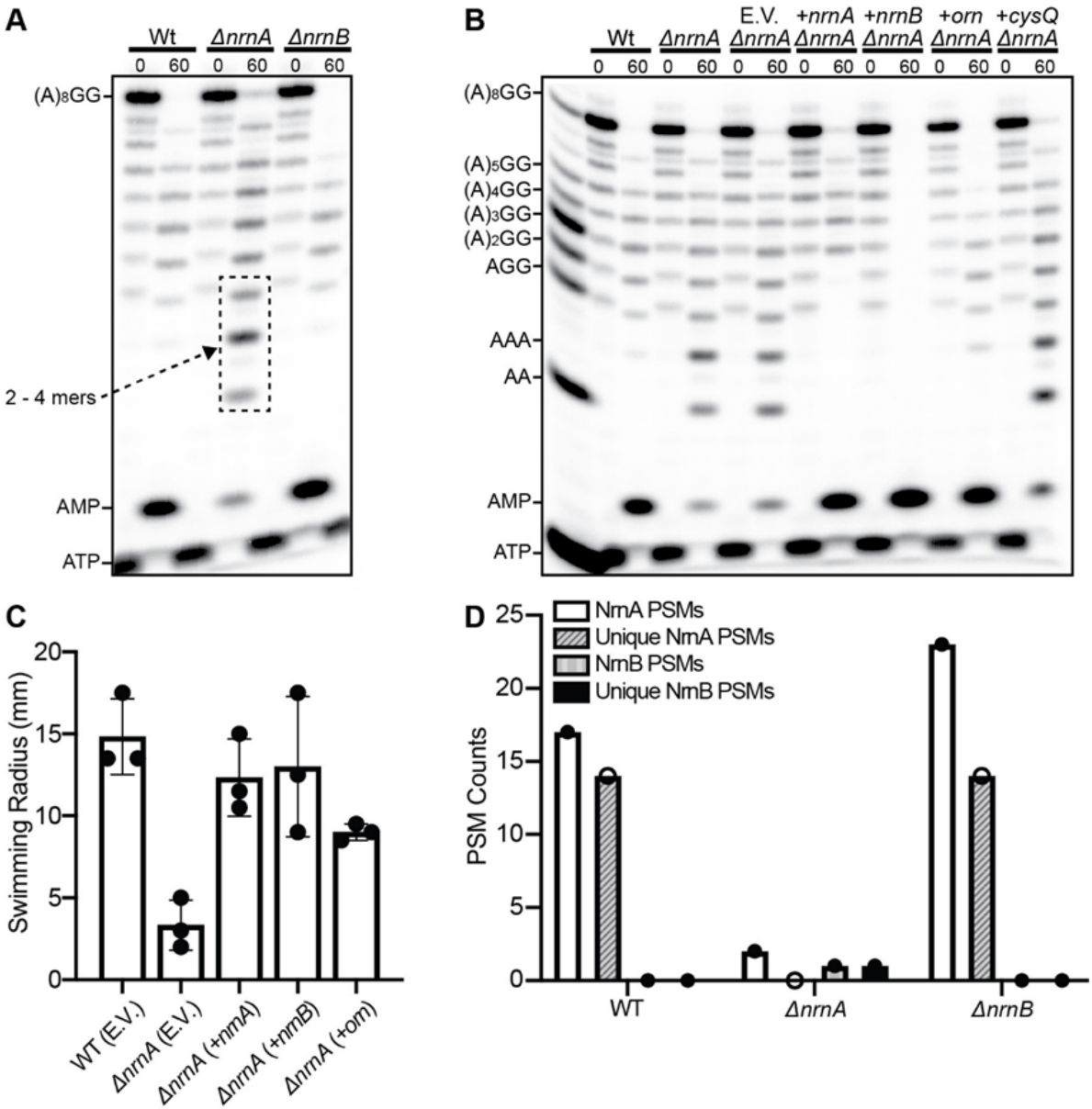

Figure 3.7. Cleavage of short RNAs by *B. subtilis* cellular lysates. (A) Whole cell lysates of WT, Δ*nrnA*, Δ*nrnB* or (B) Δ*nrnA* complementation strains containing IPTG-inducible expression of $nrnA_{Bs}$, $nrnB_{Bs}$, orn_{Vc}, or $cysQ_{Ec}$ were harvested during vegetative growth and incubated with 1 μM of a 5′-radiolabeled 10-mer RNA. Aliquots were removed at various time intervals. (A and B) Degradation products from cellular lysates were resolved by denaturing PAGE. (C) Swimming motility analysis of WT or Δ*nrnA* complementation strains that contained IPTG-inducible expression of $nrnA_{Bs}$, $nrnB_{Bs}$, or orn_{Vc}. Motility plates contained 0.2% agar and were analyzed after a 12-hour incubation period at 30˚C. E.V. indicates that the strain was transformed with an empty vector control. (D) Untargeted mass spectrometry analysis of wild-type, Δ*nrnA* or Δ*nrnB* strains yields high confidence peptide spectral matches attributed to unique peptides of the $NrnA_{Bs}$ proteins and a low confidence match to a unique $NrnB_{Bs}$ peptide. *Lysate based B. subtilis RNA degradation assays in (A) and (B) were conducted by Cordelia Weiss, Winkler laboratory, University of Maryland. Untargeted mass spectrometry of various B. subtilis strains in (D) were performed by Conor Jenkins, Winkler laboratory, University of Maryland.*

led to a significant decrease in swimming motility, suggesting that this strain exhibits elevated dinucleotides (both linear and cyclic). Correspondingly, the complementation with *nrnA* fully restored swimming motility. A swimming defect was not observed in the Δ*nrnB* mutant (Figure 3.7C), suggesting that $NrnA_{Bs}$ is primarily responsible for maintaining cellular GG pools during exponential phase growth; however, complementation of the Δ*nrnA* strain with *nrnB* 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 orn 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.

3.3.9 NrnA is expressed during exponential growth

To directly determine whether $NrnA_{Bs}$ or $NrnB_{Bs}$ are produced during exponential growth (OD600 = 0.8), an untargeted mass spectrometry proteomic analysis was performed on wild-type, Δ*nrnA* or Δ*nrnB* strains of *B. subtilis*. Peptide spectral matches (PSMs) associated with unique tryptic peptides from $NrnA_{Bs}$ versus the reference *B. subtilis* proteome were identified in the wild-type and Δ*nrnB* strains (Figure 3.7D) A single PSM identified as a unique $NrnB_{Bs}$ peptide was found in the Δ*nrnA* strain. This match was determined to be a low confidence match by manual observation of the PSM as the identifying peaks are commensurate with peaks from instrument noise, thus it can be surmised that $NrnB_{Bs}$ could not be detected under these growth conditions. Therefore, we conclude that $NrnA_{Bs}$ is produced during vegetative growth, whereas $NrnB_{Bs}$ is not. Perhaps the latter is produced under different cellular conditions.

3.3.10 NrnA$_{Bs}$ is widespread in Firmicutes

Our aggregate data have shown that, in our hands, purified NrnA$_{Bs}$ 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$_{Sp}$ and NrnA$_{Ef}$, 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$_{Mt}$ indicates that this protein likely 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$_{Bs}$, Pde2$_{Sp}$ and NrnA$_{Ef}$ 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$_{Bs}$ and other DHH family proteins includes several distinct sequence motifs, each including a conserved aspartate residue (Aravind and Koonin, 1998). 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$_{Bs}$) may be conserved in NrnA-like proteins and has also been suggested to participate in substrate recognition (Aravind and Koonin, 1998; Schmier *et al.,* 2017). 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 sub-classes of NrnA-like proteins may exhibit different substrate preferences and cellular purposes. Interestingly, a recent study demonstrated that a DHH-DHHA1 homolog from *Vibrio cholera* acts specifically on GG, at the exclusion of

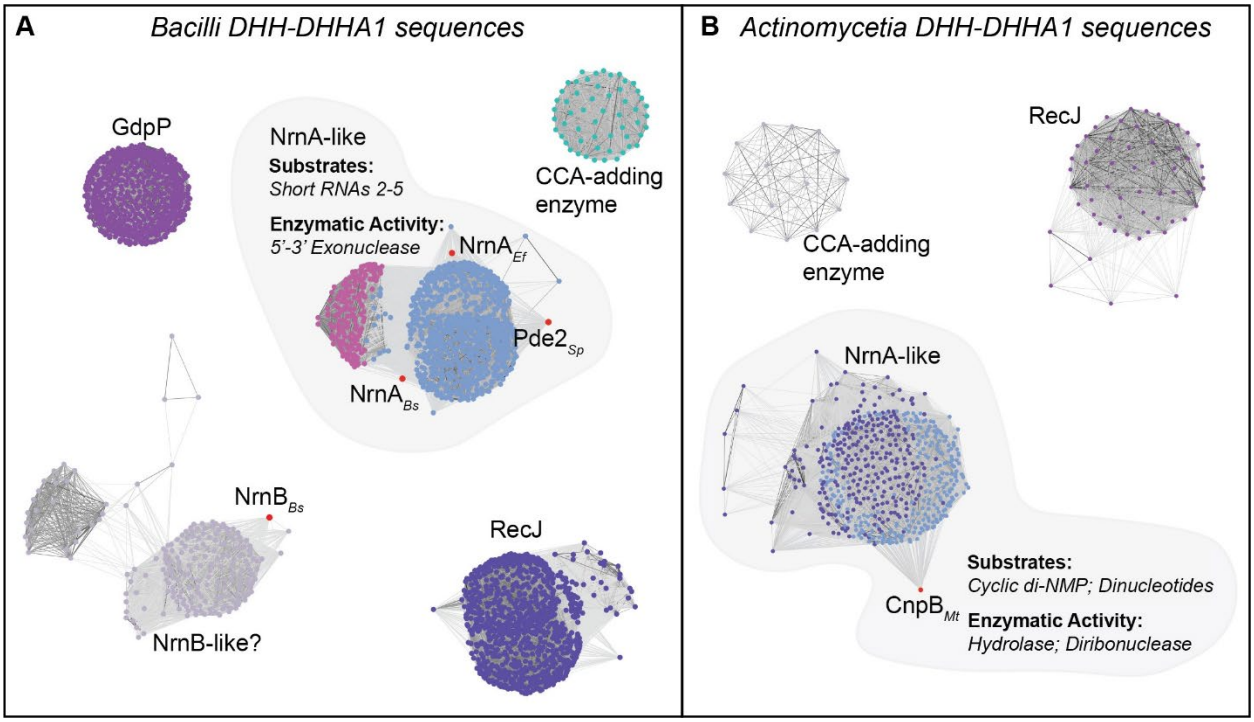

Figure 3.8. Sequence similarity network (SSNs) for DHH subfamily 1 proteins within selected taxonomical groups based on subsequences inclusive of only the DHH and DHHA1 domains. (A, B) Pairwise sequence similarities for DHH subfamily 1 proteins identified in two bacteria taxa (A) Bacilli, and (B) Actinomycetia are displayed in similarity networks, where nodes represent subsequences (concatenation of DHH and DHHA1 domain sequences) and edges are colored in grayscale to reflect the strength of the similarity, computed as $-\log_{10}$ (E-value). The E-value was thresholded at (A) $\leq 10^{-20}$ and (B) $\leq 10^{-15}$; these cutoffs were manually adjusted to reveal substructures within clusters that likely indicate functionally relevant subsequence diversity. Clauset-Newman-Moore greedy modularity maximization was applied to partition the network, finding clusters of sequences exhibiting high, in-group subsequence similarity; isolate nodes and clusters which consisted of fewer than four members were removed from the graph. Different clusters are denoted by color, and clusters are annotated based on selected representatives. *Sequence similarity network analysis was conducted by Chih Hao Wu, Winkler laboratory, University of Maryland*

other dinucleotides and other short RNA species (Heo *et al.,* 2022). 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 (Aravind and Koonin 1998; Rao *et al.,* 2010). For example, GdpP proteins, which specifically process c-di-AMP, utilize several domains alongside their DHH and DHHA1 domains. In contrast, the NrnA$_{Bs}$, NrnA$_{Ef}$, NrnA$_{Sp}$ and CnpB$_{Mt}$ proteins characterized herein do not have any recognizable domains other than the DHH and DHHA1 domains. As a preliminary test to determine if NrnA proteins could be identified by bioinformatics, we extracted the DHH and DHHA1 domain sequences of bacterial proteins from Bacilli and concatenated them together as a sub-sequence construct. Then, we performed a clustering analysis of these sequences with parameters that encouraged sequences to be closely clustered, even for pairs with large edit distances (Figure 3.8A-B). This revealed several distinct clusters, including different groups of proteins corresponding to RecJ, CCA-adding enzymes, and GdpP. This analysis also revealed a bimodal cluster, one portion of which contained the sequences for NrnA$_{Bs}$, NrnA$_{Ef}$, and NrnA$_{Sp}$. Given the similarities in their biochemical characteristics, as investigated herein, it is tempting to speculate that this overall collection of sequences corresponds to a group of NrnA proteins that closely resembles NrnA$_{Bs}$. Therefore, one would predict that these proteins process RNA substrates from 2- to 4-nucleotides in length, starting from their 5′ terminus. Notably, NrnB$_{Bs}$ was not found in the NrnA cluster of proteins. Instead, it could be identified within another cohesive group of related protein sequences, distinct from NrnA. From this, it is tempting to speculate that NrnB proteins may correspond to a specific class of DHH-DHHA1 proteins, functionally different

from NrnA. As little is known about the proteins in this cluster, NrnB warrants further biochemical and structural investigation.

A clustering analysis of Actinobacteria DHH-DHHA1 sequences revealed a complex arrangement of overlapping clusters that contained CnpB$_{Mt}$, CnpB$_{Ms}$, CnpB$_{Ma}$, and CnpB$_{Rr}$ (Figure 3.8B). From this, it is difficult to predict the overall diversity in cellular roles for these proteins, although it can be surmised that a large subset is likely to act specifically against c-di-NMP substrates. Without biochemically analyzing more proteins, the degree to which CnpB$_{Mt}$ is representative of these clusters of protein sequences is still unknown. Together, these data suggest that it may be possible to work toward prediction of NrnA proteins, at least from Bacilli, which share a core set of biochemical features. Yet these data also suggest that there are likely to be other groups of standalone DHH-DHHA1 proteins, including NrnB, that differ from NrnA in fundamental ways.

3.3.11 NrnA-like protein candidates from *Mycoplasma pneumonia* exhibit little to no activity on short RNAs *in vitro*

Other NrnA-like protein homologs with different reported biochemical activities have been identified to be encoded by the organism *M. pneumoniae*. Of these NrnA-like protein candidates, the NrnA-like protein homolog Mpn140 was originally reported to preferentially degrade short RNA and DNA oligonucleotides as well as complement an *E. coli cysQ* mutant strain (Postic *et al.*, 2012). More recently, the NrnA-like protein homolog Mpn549 was identified as a c-di-AMP phosphodiesterase (Blötz *et al.,* 2017). Given that these two proteins were both members of the DHH-DHHA1 protein family, we included them in our analysis to assess whether these proteins did indeed display differences in substrate preferences. However, these proteins showed reduced activity *in vitro* following purification. We observed only minimal

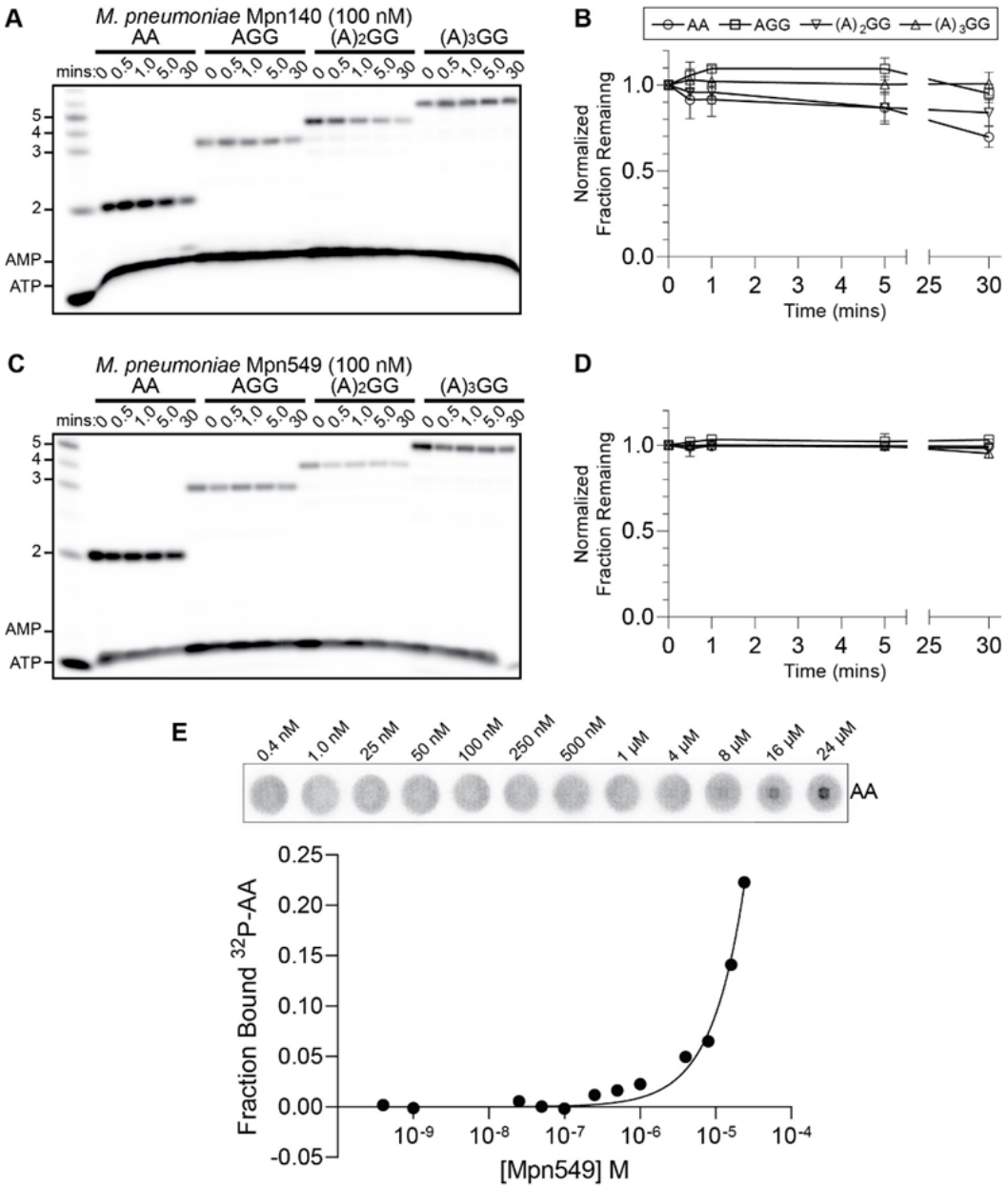

Figure 3.9. Purified Mpn140 and Mpn549 exhibit virtually no RNase activity *in vitro.* (A, C). 1 µM of RNA substrates 2-5 nucleotides in length containing trace amounts of ^{32}P-radiolabeled RNA were incubated with (A) 100 nM of purified Mpn140 or (C) Mpn549. Aliquots were removed at various time points 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 (B and D). (E) Single replicate DRaCALA, as described elsewhere (Roelofs *et al*., 2011), was used to measure binding of different radiolabeled diribonucleotide AA to Mpn549.

diribonucleotide degradation for Mpn140 (Figure 3.9A-B). Furthermore, we observed no short RNA degradation activity for Mpn549 (Figure 3.9B-C). However, we did observe that Mpn549 did exhibit diribonucleotide binding via DRaCALA, suggesting that this enzyme is not completely catalytically inactive (Figure 3.9E). It is possible that the differences in our reported activities deviate from the literature based on differences in assay conditions as well as in the protein purification methods utilized to obtain these enzymes.

3.4 Discussion

Our experiments on purified NrnA$_{Bs}$ 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$_{Bs}$ 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$_{Bs}$ resulted in accumulation of 2-mers,3-mers and 4-mers, while, in contrast, deletion of NrnB$_{Bs}$ did not result in accumulation of these short RNAs under normal vegetative growth conditions. And while deletion of NrnA$_{Bs}$ 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$_{Bs}$ is required for maintaining proper homeostasis of cyclic dinucleotide signaling molecules. We also purified NrnA-like proteins from several related species and tested their activities using the same reaction conditions as with NrnA$_{Bs}$. Pde2$_{Sp}$ and NrnA$_{Ef}$ exhibited a strong preference for short RNAs and closely resembled NrnA$_{Bs}$ overall. Moreover, our

bioinformatic analysis of DHH-DHHA1 proteins showed that these proteins may be closely related to NrnA$_{Bs}$. 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. ruber* exhibited 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$_{Bs}$, Pde2$_{Sp}$ and NrnA$_{Ef}$, 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 invitro under our assay conditions. Interestingly, a recent biochemical analysis of the *Lysteria monocytogenes* NrnA showed virtually no activity when assayed for pAp hydrolysis (Gall *et al.,* 2021). 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$_{Bs}$ 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$_{Bs}$ and NrnB$_{Bs}$ were previously assumed to be effectively synonymous in their biochemical activity and intracellular functions (Fang *et al.,* 2009). In fact, NrnA$_{Bs}$ and NrnB$_{Bs}$ were initially identified based on their ability to complement a conditional *E. coli orn* mutant strain (Mechold *et al.,* 2007; Fang *et al.,* 2009). 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$_{Bs}$ is not expressed during exponential growth, nor does it appear to process

RNAs during exponential growth, in contrast to NrnA$_{Bs}$. 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$_{Bs}$ is the housekeeping enzyme for degradation of short RNAs in B. subtilis, while NrnB$_{Bs}$ 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 *nrnA* does not present *B. subtilis* with a severe growth defect. 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 (Lormand *et al.,* 2021). 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 (Drexler *et al.,* 2017). 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 (Bowman *et al*., 2016).While our data agrees that a large subset of NrnA-like proteins is unlikely to hydrolyze c-di-AMP, our biochemical as-says 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 im-prove the functional annotation of different subclasses of DHH-DHHA1 protein sequences.

3.5 Materials and Methods

3.5.1 Bacterial Strains and Culture Conditions

E. coli strains were grown in 2xYT supplemented with 100 µg/mL carbenicillin, and *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* Δ*nrnA,* Δ*nrnB,* and Δ*nrnA*Δ*nrnB*, and the integration of fluorescent c-di-GMP riboswitch-*yfp* reporter constructs have been previously described (Orr *et al.,* 2018; Weiss *et al.,* 2019). To construct complementation strains, $nrnA_{Bs}$, $nrnB_{Bs}$, orn_{Vc}, and $cysQ_{Ec}$ were each PCR amplified from chromosomal DNA preparations. Sequences were subcloned via Gibson assembly (Gibson *et al.,* 2009) 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 (Jarmer *et al.,* 2014). To construct *E. coli* strains for overexpression of targeted proteins, $nrnA_{Bs}$, $nrnA_{Ef}$, $pde2_{Sp}$, $gdpP_{(82\text{-}659)Bs}$, and $cysQ_{Ec}$ were each PCR amplified from chromosomal DNA preparations. The sequences encoding $cnpB_{Mt}$, $cnpB_{Ma}$, $cnpB_{Ms}$, $cnpB_{Rr}$, $disA_{Bt}$, and $ppdk_{cs}$ were codon optimized for expression in *E. coli* and purchased from Integrated DNA Technologies. The different gene sequences were subcloned via Gibson assembly (Gibson *et al.,* 2009) into IPTG-inducible expression vector pAmr30, to yield an N-terminal 10xHis-SUMO tag that is cleavable by bdSENP1 protease (Fey and Görlich 2014). The $ppdk_{cs}$ and $gdpP_{(82\text{-}659)Bs}$ sequences were subcloned via Gibson assembly (Gibson *et al.,* 2009) into the IPTG-inducible expression vector pHis-parallel, to yield an N-terminal 6xHis tag. To create the point mutants of $nrnA_{Bs}$, mutations were generated using the Q5 Site-Directed Mutagenesis Kit (NEB). *E. coli* XL10-Gold® (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.

3.5.2 Fluorescence Microscopy and Quantification

Single colonies were used to inoculate liquid minimal salts glycerol glutamate (MSgg) medium (Branda *et al.,* 2001) and grown at 37°C with shaking overnight. The following morning, the cultures of each strain were inoculated 1:50 in fresh MSgg medium and cultured, shaking at 37°C until reaching an OD_{600} 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-C_2 electron-multiplying charge-coupled (EMCC) camera, enclosed within a temperature-controlled environmental chamber. Quantification was performed with Oufti and FIJI software (Paintdakhi *et al.,* 2016, Schindelin *et al.,* 2012).

3.5.3 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 100 µg/mL carbenicillin, 0.2 % glucose (w/v), and 3 mM $MgSO_4$. Cultures were grown shaking at 37˚C until reaching an OD_{600} ~ 0.4-0.8, at which point protein expression was induced with 1 mM IPTG. Cells expressing $NrnA_{Bs}$, $NrnA_{Ef}$, $CnpB_{Mt}$, $CnpB_{Ma}$, $CnpB_{Ms}$, $CnpB_{Rr}$, and $CysQ_{Ec}$ were grown for an additional 2 hours at 37˚C. Cells expressing $Pde2_{Sp}$, $GdpP_{(82\text{-}659)Bs}$, or $DisA_{Bt}$ were removed from the 37˚C incubator, induced with 1 mM IPTG, and grown at room temperature for 16 hours. Cells were harvested by

centrifugation and resuspended in 25 mM Tris-HCl pH 8.0, 300 mM NaCl, 0.25 mg/mL lysozyme, 0.01 mg/mL DNase I, and 1 mM PMSF. Cells were lysed by sonication and insoluble material was removed by centrifugation. Clarified soluble lysates were incubated with cOmplete™ His-Tag Purification Resin (Roche) for 1 hour. After incubation, the resin was washed with 10 column volumes of Wash Buffer 1 (25 mM Tris-HCl, 300 mM NaCl, 10 mM imidazole, pH 8), followed by 5 column volumes of Wash Buffer 2 (25 mM Tris-HCl, 300 mM NaCl, 25 mM imidazole, pH 8). The following modification was made for $GdpP_{(82\text{-}659)Bs}$, $CnpB_{Ma}$, $CnpB_{Ms}$, $CnpB_{Rr}$, and $NrnA(D80N\ D156N)_{Bs}$: all buffers also included 5% (v/v) glycerol. The following modification was made for $DisA_{Bt}$: wash buffers were comprised of 62.5 mM Tris-HCl, and 750 mM NaCl. All proteins were eluted with 5 column volumes of 25 mM Tris-HCl, 300 mM NaCl, 250 mM imidazole, pH 8. Eluates were then dialyzed overnight against 25 mM Tris-HCl, 300 mM NaCl, pH 8.0. $GdpP_{(82\text{-}659)Bs}$ was dialyzed into 25 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5% glycerol. $GdpP_{(82\text{-}659)Bs}$ was further purified using a Pierce™ Strong Anion Exchange Spin Column. The ion exchange column bound to $GdpP_{(82\text{-}659)Bs}$ was washed with 25 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5% glycerol, and $GdpP_{(82\text{-}659)Bs}$ was eluted with 25 mM Tris-HCl, pH 8.0, 500 mM NaCl, 5% glycerol. The 6XHis tag was not removed from $GdpP_{(82\text{-}659)Bs}$. $NrnA_{Bs}$, $NrnA_{Ef}$, $CnpB_{Mt}$, $CnpB_{Ma}$, $CnpB_{Ms}$, $CnpB_{Rr}$, $DisA_{Bt}$ and $CysQ_{Ec}$ were subjected to 10xHis-SUMO tag cleavage by 10xHis-bdSENP1 in the presence of 2 mM DTT and 2 mM $MgCl_2$ (30). Tag removal reactions were conducted on ice at 4°C for 4-12 hours. Reactions were then incubated with cOmplete™ His-Tag Purification Resin (Roche) for 45 minutes to separate 10xHis-bdSeENP1, free 10xHis-SUMO tag, and untagged proteins. Untagged proteins were recovered in the flow-through and dialyzed overnight against 25 mM

Tris-HCl pH 8.0, 300 mM NaCl. Untagged proteins and $GdpP_{(82\text{-}659)Bs}$ were aliquoted, flash frozen in liquid nitrogen, and stored at -80˚C.

3.5.4 Preparation of Whole Cell Lysates

Overnight cultures of *B. subtilis* strains were cultured shaking at 37˚C. The following morning, cells were diluted into 500 mL fresh LB and grown at 37°C with shaking to OD_{600} ~ 0.8. 250 μM IPTG was then added to the cultures and cells were grown for an additional 40 minutes. Cells were harvested by centrifugation and concentrated 10X 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.

3.5.5 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 1X 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 60 minutes, followed by heat inactivation of T4 PNK at 65°C for 20 minutes.

3.5.6 Synthesis of C-di-AMP

[^{32}P]c-di-AMP was synthesized from reactions comprising 0.5 μM [α-^{32}P]-ATP (Perkin Elmer), 0.5 μM unlabeled ATP, 50 mM Tris–HCl, 100 mM NaCl, 5 mM $MgCl_2$, and 1 μM of purified $DisA_{Bt}$. The reaction was incubated at 45°C for 5 h and then inactivated at 95°C for 10 min. The reaction was then centrifuged at 12 000 rpm in a 3 kDa MWCO NanoSep® centrifugal

device for 20 min to remove $DisA_{Bt}$. Conversion yield was determined by running an aliquot of the reaction on denaturing 20% PAGE and using a Cytiva Amersham Typhoon™ laser scanner platform. The intensity of radiolabeled c-di-AMP relative to the remaining ATP in the reaction was quantified using FIJI software (Schindelin *et al.,* 2012).

3.5.7 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 $NrnA_{Bs}$, $NrnA(D80N\ D156N)_{Bs}$, $NrnA_{Ef}$, $CnpB_{Mt}$, $CnpB_{Ms}$, $CnpB_{Ma}$, $CnpB_{Rr}$, $GdpP_{(82-659)Bs}$ or $Pde2_{Sp}$. These reactions were conducted in the presence of 25 mM Tris, pH 8.0, 300 mM NaCl, and 200 μM $MnCl_2$. 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. 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 $MnCl_2$, and 25 mM $MgCl_2$. 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 1X TBE and 4 M urea. The gels were imaged using Cytiva Amersham Typhoon™ laser scanner platform and analyzed for the appearance of truncated ^{32}P-labeled products. The intensity of the radiolabeled nucleotides was quantified using ImageJ software (Schindelin *et al.,* 2012).

3.5.8 Differential Radial Capillary Action of Ligand Assay (DRaCALA)

Apparent equilibrium binding reactions were performed by incubating trace amounts (~1 nM) 5'-[^{32}P]-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 $MnCl_2$) for 30 minutes at room temperature. For competition assays, 1 μM purified NrnA(D80N D156N)$_{Bs}$ was incubated with trace (~1 nM) 5'-[^{32}P]pApA in the presence of 10 μM 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 (Roelofs *et al.,* 2011; Patel *et al.,* 2014).

3.5.9 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 twelve hours. The diameter of the bacterial migration was measured with a ruler.

3.5.10 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 (Hatzios *et al.,* 2008). A standard curve was generated using 1 mM phosphate. All reactions were composed of 25 mM Tris-HCl pH 8.0, 300 mM NaCl, 100 μM pAp (Sigma), and either 200 μM of $MnCl_2$ (for NrnA-like proteins), or 1 mM $MgCl_2$ (for CysQ$_{Ec}$) and conducted at room temperature. 5 nM each protein was added to its respective reaction and allowed to incubate for the indicated reaction time of either 5 or 30

minutes. Reactions were then quenched with ¼ of the reaction volume of acidic malachite green dye solution and incubated for 30 minutes to allow for color development. Absorbance values at 620 nm (A_{620}) 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.

3.5.11 Collection of DHH-DHHA1 Protein Sequences and Subsetting 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 hmmsearch with a relaxed inclusion threshold, E-value ≤ 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 $\mathbf{S} = \{S_1, S_{2,\ldots,} S_n\}$. In **S** (n = 63390), 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 = 11068), and Actinomycetia (n = 1627)—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).

3.5.12 CNM Method-Based Approach to Detecting Sequence Clusters

Suppose a subset of sequences $T = \{n \in \mathbf{S} \mid n \text{ is encoded in a specific taxon (class)}\}$; 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 unaligned 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 neato 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.

3.5.13 Mass Spectrometry Proteomic Analysis

Wild type and Δ*nrnA* 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-C:Protein Promega Cat.# V5073). 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 (H_2O and 0.1% formic acid for buffer A and 95% acetonitrile with 5% H_2O 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 $2.04e^4$ 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.

3.5.14 Mass Spectrometry Data Analysis

Data analysis was performed with Metamorpheous Version 0.0.320 (Solntsev *et al.,* 2018). The database search was performed against the *B. subtilis* proteome from Uniprot (Bateman *et al.,* 2020) (Downloaded 5/11/2021). Search parameters included full Trypsin digest, 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 carbamidomethylation was added to cystines. The FlashLFQ algorithm was utilized for relative label free quantification (Millikin *et al.,* 2018). Q-values and false discovery rate (FDR) were calculated using Percolator 3.0 (MacCoss *et al.,* 2016). Results were filtered by requiring >0 unique peptides per protein, >1 PSMs per protein, and protein FDR ≤0.01.

Chapter 4: Development of a Purification Protocol for *Bacillus subtilis* NanoRNase B

4.1 Introduction

Two *B. subtilis* genes that were shown to complement a conditional *orn* mutant strain of *E. coli* are *ytqI* and *yngD* which were later renamed *nrnA* and *nrnB* with their translational products being referred to as NanoRNase A (nrnA) or NanoRNase B (nrnB). NrnA and NrnB are two proteins that belong to the DHH superfamily of proteins. Proteins in the DHH superfamily (Pfam PF01368) are comprised of a conserved N-terminal DHH domain and often include additional domains. A subset of proteins that contain an N-terminal DHH domain also include a C-terminal DHHA1 (DHH-associated domain 1, Pfam PF02272) domain. Previous analysis of *B. subtilis* NrnA and NrnB has suggested that these proteins are likely redundant in their biochemical and intracellular functions despite only sharing a very low ~20% sequence identity (Mechold *et al.,* 2007; Fang *et al.,* 2009). Interestingly, while there are crystallographic structures for *B. subtilis* NrnA and a host of different NrnA-like homologs, there is currently one crystallographic structure reported for a putative NrnB homolog encoded by *Helicobacter pylori* (Uemura *et al.,* 2013; He *et al.,* 2016; Drexler *et al.,* 2017; Schmier *et al.* 2017; Abendroth *et al.,* 2018). It remains unclear whether the *H. pylori* NrnB protein displays functional redundancy with the *B. subtilis* NrnB. Initial biochemical analysis of this *H. pylori* NrnB has lead researchers to believe that this proteins function best against cAMP (Choi *et al.,* 2013). Alphafold 2 structural predictions for *B. subtilis* NrnB indicates that there is potentially an additional small domain attached to the DHHA1 domain at the very C-terminus of the protein (Jumper *et al.,* 2021; Mirdita *et al.,* 2022). To elucidate the biochemical features and possibly determining the structure of *B. subtilis* NrnB, it is crucially important that we obtain purified

protein. While NrnA-like protein candidates can usually be purified using immobilized metal affinity chromatography (IMAC), we were unable to utilize this common method for obtaining purified $NrnB_{Bs}$. Herin, we described the development of a purification protocol for $NrnB_{Bs}$ utilizing primarily precipitation-based techniques.

4.2 Results

4.2.1 Purification of polyhistidine variants of $NrnB_{Bs}$

My initial attempts to purify $NrnB_{Bs}$ were complicated by my inability to purify $NrnB_{Bs}$ by immobilization metal chelate affinity chromatography (IMAC). In the initial biochemical analyses of $NrnB_{Bs}$ researchers attempted to purify recombinant C-terminally his-tagged NrnB from *E. coli* but had issues obtaining purified preparations due to complications with protein solubility (Fang *et al.,* 2009). To obtain purified preparation of $NrnB_{Bs}$ these researchers resuspended the insoluble cell debris overexpressing $NrnB_{Bs}$ into water. Using this method, it was reported that using this method improved the purity of the enzyme to an estimated ~95%. For this purification attempt, I grew *E. coli* T7 express cells that harbored a plasmid that encoded the $NrnB_{Bs}$ gene which was to be expressed with an N-terminal 6x polyhistidine sequence. These cells were grown shaking at 37°C to mid-log phase and induced by the addition of 1 mM IPTG. After the addition of IPTG, the cells were grown shaking at room temperature for 16 hours. These cells were harvested by centrifugation and lysed via sonication. After sonication, the cell lysate was centrifuged at 12,000 rpm for 15 minutes. These expression and lysis conditions were used for all of the different tag variations described in this section unless indicated otherwise. The insoluble cell debris was resuspended in water, or 25 mM Tris-HCl, pH 8.0. From this resuspension an aliquot from the sample resuspended in water or Tris-HCl was resolved by SDS-

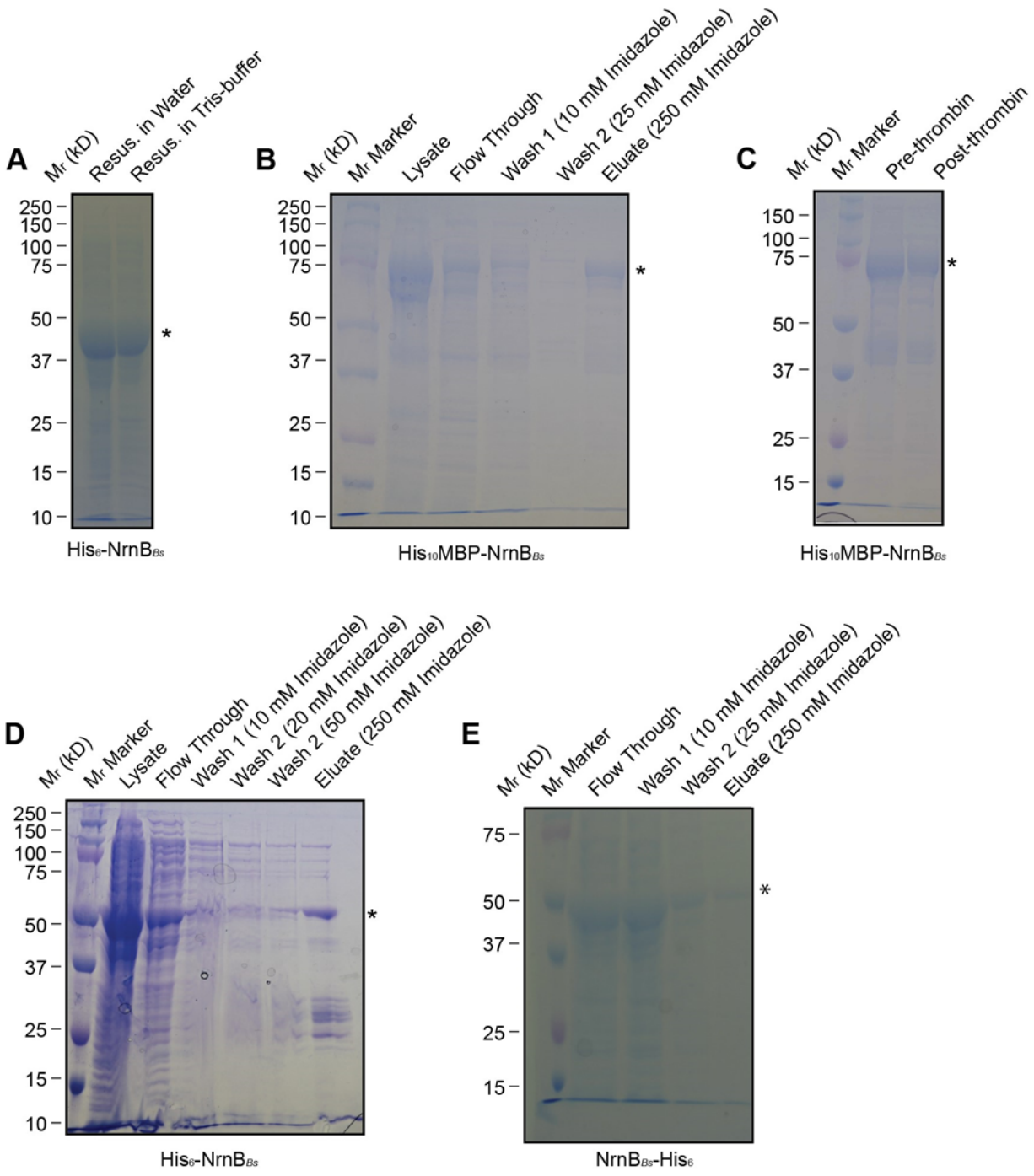

Figure 4.1. Purification attempts for polyhistidine and solubility tag variants of NrnB*Bs*. SDS-PAGE analysis of aliquots taken from T7 expression cells induced overnight with 1 mM IPTG in which (A) the insoluble pellet 6xHis-NrnB$_{Bs}$ was resuspended in water or Tris-HCl pH 8.0 buffers or (B) 10xHis-MBP-NrnB$_{Bs}$ was purified by IMAC and (C) cleavage with thrombin was attempted or (D) 6xHis-NrnB$_{Bs}$ was purified by IMAC or (D) NrnB$_{Bs}$-6xHis was purified by IMAC. The black asterisk corresponds to the molecular weight that would be anticipated for the various NrnB constructs.

PAGE (Figure 4.1A). In the analysis of these samples in an SDS-PAGE gel stained with Coomassie brilliant blue, I would not estimate this procedure to result in ~95% purity (Figure 4.1A). Due to the impurities contained in this preparation, we decided to concurrently attempt multiple different purification procedures.

Our initial preliminary biochemical analysis of NrnB$_{Bs}$ was conducted with protein purified by IMAC with an N-terminal 10xhis-tagged maltose binding protein (MBP) solubility tag. For this purification attempt, *E. coli* T7 express cells that harbored a plasmid that encoded the NrnB$_{Bs}$ gene which was to be expressed with an N-terminal 10x polyhistidine sequence in frame with MBP and NrnB$_{Bs}$. These cells were grown shaking at 37°C to mid-log phase and induced by the addition of 1 mM IPTG. After the addition of IPTG, the cells were grown shaking at room temperature for 16 hours. These cells were harvested by centrifugation, resuspended in buffer 1 (25 mM Tris-HCl, pH 8.0, 300 mM NaCl, 5% glycerol) and lysed via sonication. After sonication, the cell lysate was centrifuged twice at 12,000 rpm for 15 minutes. The soluble portion of this fraction was incubated with Roche nickel resin for 1 hour. The nickel resin was washed with 10 column volumes (CV) of buffer 1 containing 10 mM imidazole Then, the nickel resin was washed with 5 CV of buffer 1 containing 25 mM imidazole. Finally, the sample was eluted with buffer 1 containing 250 mM imidazole. Analysis of the IMAC elution by Coomassie strained SDS-PAGE shows partially purified 10xhis-MBP-NrnB$_{Bs}$ (Figure 4.1B). Since it is possible that the large MBP tag could impact NrnB$_{Bs}$'s activity, we attempted to take advantage of the thrombin cleavage site between the his-MBP tag and NrnB$_{Bs}$. The tag removal was conducted in accordance with the Sigma-Aldrich Thrombin CleanCleave Kit™ in which 10xhis-MBP-NrnB$_{Bs}$ was incubated with the protease overnight. However, analysis of the pre and post

thrombin treated 10xhis-MBP-NrnB$_{Bs}$ samples by SDS-PAGE indicated that none of the MBP tag was removed (Figure 4.1C).

Due to the issues surrounding tag removal for the 10xhis-MBP-NrnB$_{Bs}$ construct, we decided to try to purify NrnB$_{Bs}$ with a small 6xhis-N-terminal tag that we would not attempt to remove. For our initial attempt to purify this construct by IMAC, we utilized the same expression and buffer and similar column washing conditions as we did for the 10xhis-MBP-NrnB$_{Bs}$ purification outlined above. However, we changed the concentrations of imidazole in the wash buffers to 10 mM imidazole, 20 mM imidazole, and 50 mM imidazole. In the analysis of the wash and elution fractions by way of SDS-PAGE, we observed very low yields of 6xhis-NrnB$_{Bs}$ that were heavily contaminated with proteins (Figure 4.1D). Additionally, we observed that the 6xhis-NrnB$_{Bs}$ were heavily contaminated with nucleic acids, as seen by a spectrophotometric analysis in which theA260/280 ratio was > 1.0. Seeking a quick alternative to the N-terminal 6xhis tag, we attempted IMAC for an NrnB$_{Bs}$ construct containing a C-terminal-6xhis tag. However, analysis by SDS-PAGE showed us substantial contamination as well as nucleic acid contamination once again (Figure 4.1E).

Concomitantly with the previous purification attempts, we also tried to purify NrnB$_{Bs}$ with a 10xhis N-terminal sequence with a SUMO solubility tag. This method was previously utilized for the purification of many NrnA-like proteins. However, we still observed very low protein yields for NrnB$_{Bs}$ with this method. Utilizing the same expression, lysis, insoluble debris removal, and IMAC protocol outlined above, one way that were able to make a slight improvement to the purification was by increasing the salt concentration in buffer 1 from 300 mM NaCl to 750 mM NaCl to decrease non-specific binding of nucleic acids and contaminating proteins to NrnB$_{Bs}$. Examination of the fraction by SDS-PAGE for this purification attempt

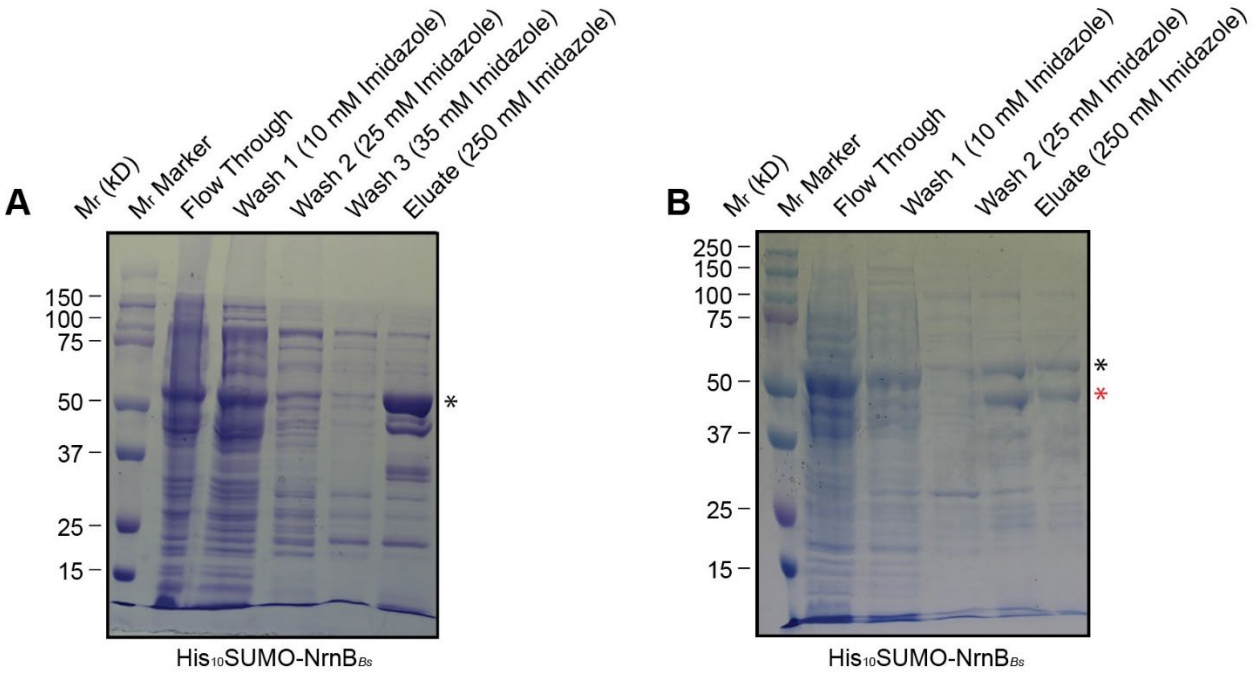

Figure 4.2. Purification attempts for 10xHis-SUMO-NrnB$_{Bs}$ under high salt. SDS-PAGE of aliquots taken during various purification attempts for 10xHis-SUMO-NrnB$_{Bs}$. SDS-PAGE of fractions from T7 expression cells induced overnight with 1 mM IPTG to express 10xHis-SUMO-NrnB$_{Bs}$ in which the cells were purified by IMAC. Aliquots taken from the purification of 10xHis-SUMO-NrnB$_{Bs}$ in which the purification was conducted (A) in the presence of buffer comprised of 25 mM Tris-HCl, pH 8.0, 750 mM NaCl, 5% glycerol or (B) in the presence of buffer comprised of 25 mM Tris-HCl, pH 8.0, 1 M NaCl, 5% glycerol. The black asterisk indicates the molecular weight corresponding to 10xHis-SUMO-NrnB$_{Bs}$. The red asterisk indicates the molecular weight of a potential proteolytic degradation product of 10xHis-SUMO-NrnB$_{Bs}$.

revealed a slight qualitative increase in the protein yield in the elution (4.2A). However, contamination by nucleic acids and other proteins continued. We tried to conduct the purification in the presence of very high salt (1 M NaCl), and we did not observe any improvements in purity or yield for 10xhis-SUMO-NrnB$_{Bs}$ (Figure 4.2B). In fact, a second band can be observed at a molecular weight corresponding to the tag-less form of 10xhis-SUMO-NrnB$_{Bs}$ can be observed (Figure 4.2B red asterisk). With this, we ultimately decided that perhaps IMAC is not a viable first step for NrnB$_{Bs}$ purification, and that should utilize some other purification methods. At this point we speculated that since NrnB$_{Bs}$ exhibits nucleic acid binding, that we should attempt to remove nucleic acids immediately to aid in the solubility of NrnB$_{Bs}$.

4.2.2 PEI Precipitation improves the purity of NrnB$_{Bs}$ while removing nucleic acids

In the attempts to purify NrnB$_{Bs}$ by IMAC, we frequently observed that the A260/280 was always > 1.0 indicating that nucleic acids were in high abundance in the preparations of NrnB$_{Bs}$. This observation led us to speculate that perhaps NrnB$_{Bs}$ was binding nucleic acids in solution and that this was altering NrnB$_{Bs}$'s solubility. With this in mind, we decided to add the polymer polyethyleneimine (PEI) to the initial steps of our purification protocol. PEI is a positively charged insoluble polymer in neutral buffer conditions. Analogous to ion exchange chromatography, PEI will bind to negatively charged proteins, but this is highly dependent on the salt concentration of the solution. Since NrnB$_{Bs}$ is predicted to be a very acidic protein with an estimated pI of ~5.5 we first tested to see if NrnB$_{Bs}$ would bind to PEI and by extension precipitate out of solution. In this analysis, we added increasing concentrations of PEI (0.1 to 0.5%) to cellular lysates prepared as mentioned above to assess whether NrnB$_{Bs}$ would precipitate (Figure 4.3A). We observed in a cellular lysate induced to overexpress NrnB$_{Bs}$ that there was a large band corresponding to NrnB$_{Bs}$. When we added increasing amounts of PEI to

this culture and spun pelleted the PEI, we took an aliquot of the supernatant and analyzed this by SDS-PAGE. We observed that while NrnB$_{Bs}$ seems to be present in the soluble portion of the cellular lysate, it is absent upon the addition of 0.1% PEI indicating that NrnB$_{Bs}$ is in fact precipitating with the PEI. Next, we tested to see at what salt concentration NrnB$_{Bs}$ would elute from PEI. We took NrnB$_{Bs}$ that had been precipitated by 0.1% PEI and added increasing concentrations of NaCl and found that NrnB$_{Bs}$ elutes at 800 mM NaCl (Figure 4.3B). Now knowing that NrnB$_{Bs}$ will not precipitate in PEI under high salt conditions >800 mM NaCl we tested two possible methods for utilize PEI in a purification protocol. For method 1, NrnB$_{Bs}$ would never be pelleted in the PEI polymer as the cellular lysate was generated using buffer 1 comprised of 1 M NaCl. These cells were lysed via sonication and then PEI was added to a final concentration of 0.1% prior to centrifugation to remove insoluble debris. The insoluble debris and PEI was removed by two rounds of centrifugation at 12,000 rpm. Then, an aliquot from the soluble portion of this sample was analyzed via SDS-PAGE (Figure 4.3C Method 1). For method 2, the cells overexpressing NrnB$_{Bs}$ were lysed in buffer 1 containing 300 mM NaCl. The cell debris was removed by two rounds of centrifugation at 12,000 rpm. Next, 0.1% PEI was added to soluble cellular lysate. This sample was centrifuged at 10,000 rpm for 10 minutes to pellet NrnB$_{Bs}$ bound PEI. Next, the soluble portion of this sample was disposed of, and the NrnB$_{Bs}$ containing PEI was resuspended in buffer 1 containing 500 mM NaCl. This sample was centrifuged again at 10,000 rpm and the soluble portion was disposed of. Finally, the PEI was resuspended in buffer 1 containing 1 M NaCl. At this salt concentration NrnB$_{Bs}$ will no longer be bound to the insoluble PEI. We left NrnB$_{Bs}$ in the 1 M NaCl buffer for 15 minutes to elute NrnB$_{Bs}$ or allowed the NrnB$_{Bs}$ to elute overnight. Next, the PEI was removed by another round of centrifugation while NrnB$_{Bs}$ remained in the soluble fraction and transferred into another conical

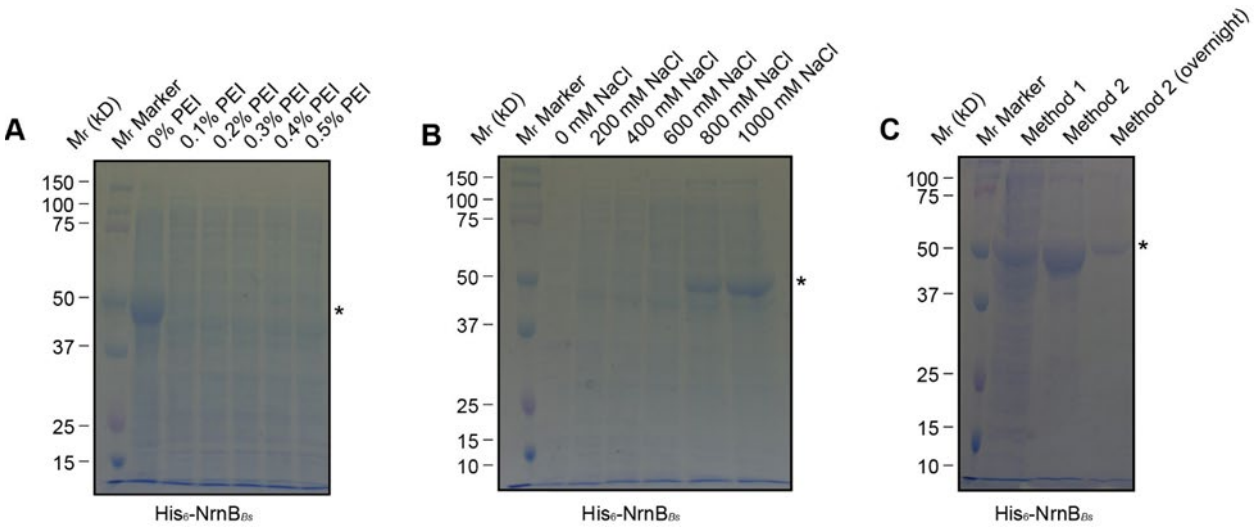

Figure 4.3. PEI Precipitation improves the purity of NrnB$_{Bs}$. SDS-PAGE of aliquots of the soluble portion of cellular lysates overexpressing *nrnB* that have been treated with (A) increasing concentrations of polyethyleneimine (PEI) and. (B) SDS-PAGE of aliquots taken from cellular lysates overexpressing *nrnB* that have been first treated with 0.1% PEI and centrifuged at 10,000 rpm to pellet the PEI with the initial soluble portion being removed, prior to eluting proteins from the PEI using 25 mM Tris-HCl, pH 8.0, with increasing concentrations of salt. (C) SDS-PAGE of aliquots taken from cellular lysates overexpressing *nrnB* in which method 1 – cell lysis in buffer containing 1 M NaCl, method 2 – pelleting the target acid protein in the positively charged PEI and washing the PEI pellet with the buffer containing the highest concentration of salt that will not elute the target protein prior to eluting the target protein by incubating for 15 minutes with buffer containing 1 M NaCl and method – 2 with the elution being overnight.

tube to remove the insoluble PEI pellet. An aliquot from these samples was analyzed via SDS-PAGE (Figure 4.3C Method 2). In the analysis of this SDS-PAGE gel we concluded that utilizing method 2 with a 15-minute elution time was the best protocol since we were able to remove many contaminating proteins and nucleic acids while maximizing out protein yield.

4.2.3 Ammonium sulfate precipitation improves the purity of NrnB$_{Bs}$

One complication associated with using PEI precipitation is that it is possible that PEI could persist in the protein sample interfering with future biochemical tests. To improve the purity of NrnB$_{Bs}$ and remove all of the PEI, we turned to ammonium sulfate precipitation. For this we reasoned that we could include another high-speed centrifugation step while possibly removing some contaminating proteins to improve the purity of our preparation. For this analysis we added different amounts of solid ammonium sulfate to five different test tubes containing NrnB$_{Bs}$ lysate preparations which were described above to reach 20 – 60% saturation. The ammonium sulfate was dissolved, and the lysates were incubated on ice shaking for 30 minutes to allow for precipitation. Next, these fractions were centrifuged at 10,000 x *g* for 10 minutes. The supernatant from the centrifuged tubes were transferred to a new tube. Then, to the new tubes containing different concentrations of ammonium sulfate (20 – 60% saturation) more solid ammonium sulfate was added to increase the percent saturation of each sample by 10%. The ammonium sulfate concentrations were 30, 40, 50, 60, & 70% saturation respectively. These fractions were incubated again on ice shaking for 30 minutes to allow for precipitation. Following precipitation, these samples were centrifuged again at 10,000 x *g* for 10 minutes. After centrifugation, the supernatant from these test tubes was discarded and the protein pellet remaining was resuspended in buffer 1. A small aliquot from the resuspension of the protein pellet from each of the different test tubes was analyzed by SDS-PAGE (Figure 4.4A). From the

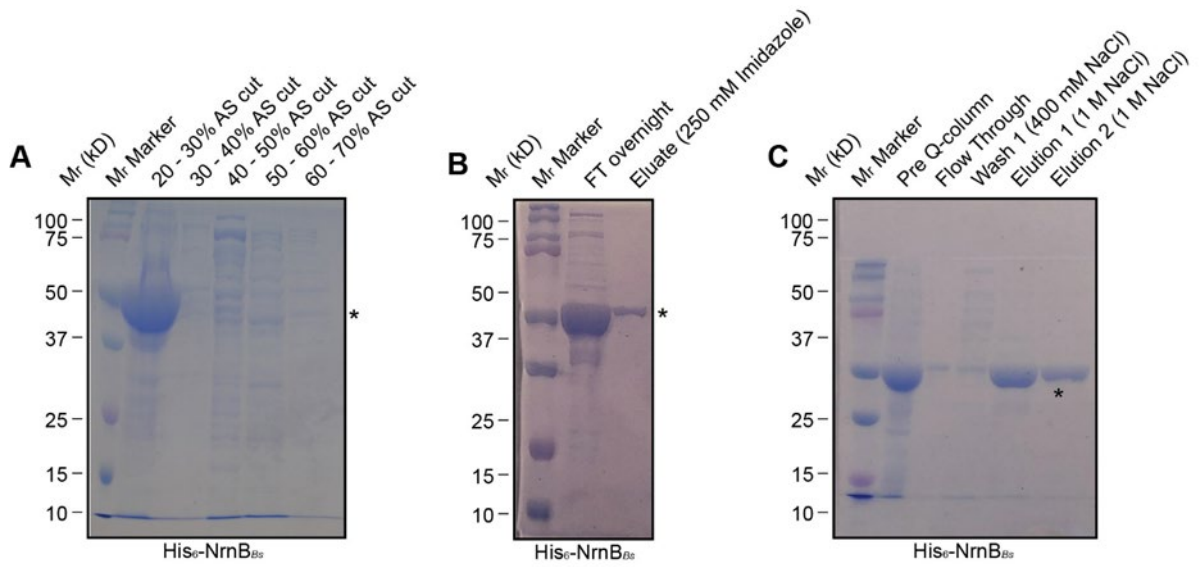

Figure 4.4. Ammonium sulfate precipitation and anion exchange chromatography improves the purity of NrnB$_{Bs}$. (A) SDS-PAGE in which aliquots were taken from cellular lysates overexpressing *nrnB* that had been initially treated with the first ammonium sulfate in which the first indicated amount of ammonium sulfate was added prior to centrifugation and the addition of 10% more ammonium sulfate. (B) SDS-PAGE where 6xhis-NrnB$_{Bs}$ was treated with PEI and then AS and tested for nickel binding by IMAC. (C) SDS-PAGE of NrnB$_{Bs}$ that was previously treated with PEI and AS further purified by exchange chromatography.

analysis of the different ammonium sulfate treated lysates, we found that treating the lysate with 20% ammonium sulfate, and then adding 10% more ammonium sulfate resulted in complete precipitation of NrnB$_{Bs}$. From this analysis, we observe that many protein contaminants are removed when the 30% ammonium sulfate supernatant is discarded because most proteins present in the extract require >30% ammonium sulfate to precipitate.

We had speculated that the nucleic acids present in the initial attempts to purify NrnB$_{Bs}$ by IMAC could have hindered the binding of NrnB$_{Bs}$ to the nickel resin. With this in mind, we tested whether NrnB$_{Bs}$ would bind the nickel resin post treatment with PEI and ammonium sulfate. For this analysis we generate cellular lysates in buffer 1 as described above. We clarified the lysate of insoluble debris via two rounds of centrifugation at 12,000 rpm and then PEI was added to the cellular lysate to a final concentration of 0.1%. Next, we utilized method 2 described in (4.2.2). Then, we further purified NrnB$_{Bs}$ by using a 20-30% ammonium sulfate cut. We resuspended the ammonium sulfate pellet which contained mostly NrnB$_{Bs}$ and resuspended this sample in a buffer 1 containing 20 mM imidazole. We incubated this NrnB$_{Bs}$ sample with Roche nickel resin overnight to allow the protein to bind the nickel resin. Next, we collected the flow-through and eluted protein from the nickel resin with buffer 1 containing 250 mM imidazole. We took a small aliquot of the flow-through and elution and analyzed their contents by SDS-PAGE (Figure 4.4B) We found that ~80% of NrnB$_{Bs}$ did not bind to nickel resin (Figure 4.4B). With this, we speculated that perhaps the histidine tag could be sequestered in the NrnB$_{Bs}$ tertiary structure, so we did not pursue IMAC further, and instead attempted strong-ion exchange by way of a Q-column.

4.2.4 Anion exchange chromatography improves the purity of NrnB$_{Bs}$

While we have observed significant improvements in the purity of our NrnB$_{Bs}$ preparation using PEI and ammonium sulfate precipitation, we decided to pursue additional steps for improving the purity of NrnB$_{Bs}$. For this, we decided to take advantage of NrnB$_{Bs}$'s negative charge and employ the use of anion-exchange chromatography using a Q-column. For this, we dialyzed NrnB$_{Bs}$ that had been purified with PEI and ammonium sulfate treatment described in previous sections, into buffer containing 25 mM Tris-HCl, pH 8.0, 5% glycerol. Interestingly, we were unsuccessful in our attempts to further purify NrnB$_{Bs}$ using a HiTrap Q HP column by way of an ÄKTA Pure FPLC, as the protein never bound the column and simply passed into the flow-through upon loading. Since it stands as a possibility that NrnB$_{Bs}$ could have been interacting with the matrix of the HiTrap Q HP column, we decided to try a Pierce™ Strong Anion Exchange spin column. The Pierce ion exchange columns are comprised of a highly porous membrane-based adsorption matrix structure in which the pores are larger than 3000 nm. We found that NrnB$_{Bs}$ bound very strongly to the Pierce anion exchange column, and also required a buffer containing 600 mM NaCl to elute the protein from the matrix. With this knowledge, we included a 400 mM NaCl wash step to aid in improving the purification of NrnB$_{Bs}$. NrnB$_{Bs}$ was eluted from the Pierce anion exchange column with buffer composed of 25 mM Tris-HCl, pH 8.0, 1 M NaCl, 5% glycerol. From SDS-PAGE analysis of the ion-exchange fractions, we estimated that NrnB$_{Bs}$ is roughly 90% pure at this stage of purification (Figure 4.4C).

4.2.5 Size exclusion chromatography did not dramatically improve the purity of NrnB$_{Bs}$

As a possible final polishing step in the purification of NrnB$_{Bs}$ we decided to try to improve the purity of this protein by way of gel filtration chromatography. We loaded 500 µl of ~50 µM NrnB$_{Bs}$ onto a Superdex 200 10/300 GL size exclusion column on a ÄKTA Pure FPLC

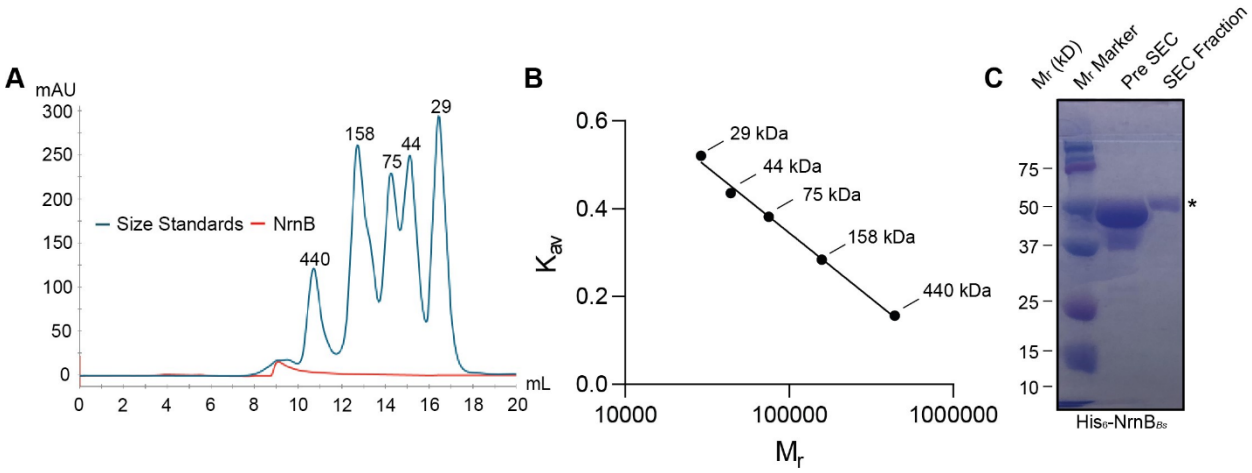

Figure 4.5. Size exclusion chromatography of NrnB$_{Bs}$. (A) Size exclusion trace for NrnB$_{Bs}$ in red elution compared to protein size standards in blue run on a Superdex 200 10/300 GL using a 0.3 mL/min flow rate in 25 mM Tris-HCl, pH 8.0, 150 mM NaCl. (B) A graph showing the partitioning coefficients of the protein standards as compared to their reported molecular weights. (C) SDS-PAGE of NrnB$_{Bs}$ post treatment with PEI, AS, and ion exchange chromatography pre and post gel filtration.

at a flow rate of 0.3 mL/min in 25 mM Tris-HCl, pH 8.0, 150 mM NaCl. As seen by the red elution trace for $NrnB_{Bs}$, this protein eluted at the approximate void volume of the column which is 8.25 mL (Figure 4.5A). We decided to compare the elution profile of $NrnB_{Bs}$ to Cytiva molecular weight standards represented by the blue elution trace (Figure 4.4A). Then, we graphed the partitioning coefficients of the known molecular weight size standards against their known molecular weights to ensure that our column was properly functioning (Figure 4.4B). We were unable to determine an estimated molecular weight for $NrnB_{Bs}$ because the protein eluted at a mass much greater than our largest protein size standard. This led us to believe that $NrnB_{Bs}$ could be interacting with the Superdex resin and that we might have to utilize another method to elucidate the oligomeric state of $NrnB_{Bs}$. Also, when we analyzed the fraction of $NrnB_{Bs}$ pre and post SEC by SDS-PAGE, we found that there was no significant increase in the purity of this preparation. Therefore, we have not pursued further purification measures for $NrnB_{Bs}$.

4.3 Discussion

Obtaining purified enzymes is of the greatest importance when conducting biochemical assays. For example, impurities in partially purified protein fractions can influence the capacity to accurately determine the concentration of the sample as the contaminating proteins will contribute to the A_{280}. If the concentration of an enzyme is not reflective of the concentration of the target protein, this can influence the capacity of a scientist to determine binding constants, or accurately elucidate the kinetics parameters of an enzyme. Herein, we utilized a combination of different approaches to optimize the expression and the purification protocol for $NrnB_{Bs}$. Through the use of PEI precipitation, ammonium sulfate precipitation, and ion-exchange chromatography, we were ultimately able to obtain purified fractions of $NrnB_{Bs}$.

4.4 Materials and Methods

4.4.1 $NrnB_{Bs}$ overproduction and purification method

Plasmids containing recombinant protein sequences were transformed into T7 Express *E. coli*. Cells were grown in 2xYT that had been supplemented with 0.2% glucose (w/v) and 3 mM $MgSO_4$ at 37˚C until an OD_{600} of 0.4-0.6 was reached. At this point the cells expressing $NrnB_{Bs}$ were induced with 1 mM IPTG and removed from the 37˚C incubator to overexpress proteins at room temperature overnight (~16 hours). Cells were harvested the next day by centrifugation at 5,000 rpm for 20 minutes. Cells containing *B. subtilis* NrnB were resuspended in 25 mM Tris-HCl pH 8, 300 mM NaCl, 5% glycerol (v/v) (1 gram of cell pellet to 10 mL lysis buffer). Lysozyme was added prior to cell lysis at a final concentration of 0.25 mg/mL and PMSF was added to reach a concentration of 1 mM. Cells were lysed via sonication. The cell lysates were clarified by two rounds of centrifugation at 12,000 rpm at 4˚C for 15 minutes. We then subjected the soluble fraction to polyethyleneimine (PEI) precipitation followed by ammonium sulfate precipitation (Burgess, 2009). The volume of the clarified lysate was determined and 10% (v/v) polyethyleneimine (PEI) was added to the clarified lysate to reach a final concentration of 0.1%. The protein was then pelleted in the PEI at 5,000 rpm for 5 minutes. The supernatant was discarded and the PEI pellet containing the protein was washed in an equal volume of 25 mM Tris-HCl pH 8, 500 mM NaCl, 5% glycerol resuspension buffer. The sample was centrifuged again for 5 minutes at 5000 rpm to pellet the PEI. The supernatant was discarded. The PEI pellet was resuspended in an equal volume of 25 mM Tris-HCl pH 8, 1 M NaCl, 5% glycerol to elute $NrnB_{Bs}$ from the PEI. At this point the sample was left for 15 minutes to elute the protein. The PEI was pelleted again by centrifugation at 10,000 rpm for 10 minutes. The supernatant containing $NrnB_{Bs}$ was collected, and the PEI pellet was discarded. To the supernatant, solid

ammonium sulfate was added to reach a 20% saturation at 0°C. The sample was rocked gently for 30 minutes on ice. After shaking, the sample was centrifuged at 10,000 rpm for 10 minutes. The supernatant was collected, and more solid ammonium sulfate was added to reach 30% saturation. Samples were shaken on ice for 30 minutes at which point it was assumed that $NrnB_{Bs}$ had precipitated. The sample was centrifuged at 10,000 rpm for 10 minutes and the pellet was resuspended in 25 mM Tris-HCl pH 8, 300 mM NaCl, 5% glycerol. Next, the samples were dialyzed overnight into 25 mM Tris-HCl pH 8, 5% glycerol. The dialyzed samples were loaded into a Pierce™ Strong Anion Exchange Spin Column. In accordance with the manufacturer's protocol, the ion exchange spin column was centrifuged at 2,000 x *g* for 5 minutes for loading, washing, and eluting steps. The ion exchange spin column was washed with 25 mM Tris-HCl pH 8, 400 mM NaCl, 5% glycerol. Finally, $NrnB_{Bs}$ was eluted from the column with 25 mM Tris-HCl pH 8, 1 M NaCl, 5% glycerol. Samples of purified $NrnB_{Bs}$ were dialyzed into 25 mM Tris-HCl pH 8, 300 mM NaCl, 5% glycerol overnight. Purified protein samples were individually aliquoted, and flash frozen in liquid nitrogen.

Chapter 5: *Bacillus subtilis* NrnB is expressed during sporulation and acts as a unique 3’-5’ exonuclease

5.1 Copyright Notice

5.2 Introduction

Degradation of intracellular RNA is an essential process involving multiple classes of ribonucleases (RNases) (Hui et al., 2014, Bechhoffer & Duetscher, 2019, Trinquier et al., 2020). Some RNases act as indiscriminate, non-specific, and redundant enzymes whose primary purpose is to recycle ribonucleotides for further rounds of transcription. Other RNases can exhibit specialized substrate specificity to play key roles in gene expression. Together, these assemblies of RNases vary between bacteria. RNases are classified by the mechanisms that they employ to degrade or mature transcripts. Endoribonucleases cleave RNA transcripts internally, rendering them more accessible to cleavage by exoribonucleases. Exoribonucleases sequentially remove ribonucleoside monophosphates from the ends of transcripts. Exoribonucleases are further classified by the polarity in which they recognize and degrade RNA substrates (e.g., 5’-3’ versus 3’-5’) in addition to their propensity to make a single (distributive) or multiple (processive) cleavages per binding event. Dozens of RNases have been discovered that, among

other functions, degrade mRNAs, and process pre-tRNAs and pre-rRNAs; however, these RNases do not act on short RNA molecules (Hui et al., 2014, Bechhofer & Murray 2019, Trinquier et al., 2020). For instance, exoribonucleases such as PNPase, RNase R, RNase II, and RNase J1 do not readily recognize and degrade RNA molecules between 2-5 nucleotides in length (Dorléans et al., 2011, Cheng et al., 2002, Amblar et al., 2007). Previous reports of the phosphorolytic 3'-5' exonuclease PNPase indicated that long RNAs are preferred substrates for this enzyme as compared to very short RNAs ~2-6 nucleotides in length (Singer 1958, Chou et al., 1970, Casinhas et al., 2018, Unciuleac et al., 2021) Additionally, RNase R and RNase II are 3'-5' exonucleases that have been reported to degrade long RNA substrates processively and release short RNA oligonucleotides as they reach the 5' terminus of their substrates (Cheng et al., 2002). RNase J1 is a processive 5'-3' exonuclease that was shown to be capable of acting on short RNAs but that possesses a potentially strong preference for RNA molecules ≥ 10 nucleotides in length (Dorleans et al., 2011). Therefore, short RNAs, from 2-5 nucleotides in length, represent a discrete class of RNA substrates during RNA degradation. To specifically process these short RNAs, all cells are likely to harbor at least one specialized RNase that is dedicated to this purpose (Lee et al., 2022).

Escherichia coli Oligoribonuclease (Orn) was the first enzyme discovered that specifically degrades short RNAs 2-5 nucleotides in length (Niyogi et al., 1975, Datta et al., 1975). The depletion of *orn* is either lethal or, at minimum, severely deleterious for the bacteria in which it has been examined (Ghosh et al., 1999, Orr et al., 2015, Cohen et al., 2015). However, *orn* is not found in all bacterial genomes. Using a conditional *orn* depletion strain of *E. coli*, investigators previously discovered several genes that could complement the growth defect caused by depletion of Orn: NanoRNase A (NrnA), NanoRNase B (NrnB), and NanoRNase C

(NrnC) (Mechold et al., 2007, Fang et al., 2009, Lui et al., 2012). Orn and NrnC proteins display structural similarities between their nucleotide binding pockets (Lormand et al., 2021) and are composed of DnaQ-like folds with DEDD active site motifs. In contrast, NrnA and NrnB proteins are members of the protein superfamily that features DHH (Pfam: PF01368) and DHHA1 (Pfam: PF02272) domains. This superfamily includes, among other subfamilies, proteins corresponding to RecJ, GdpP, and certain aminoacyl-tRNA synthases; however, unlike the other protein subfamilies, NrnA and NrnB are standalone DHH-DHHA1 proteins, lacking any other recognizable protein domains. In general, Orn and NrnA are more broadly distributed across most taxonomic groups relative to NrnC and NrnB (Lormand et al., 2021). NrnC is largely restricted to the alphaproteobacterial and cyanobacteria phyla of organisms, while the distribution of NrnB is less clear and may be more narrowly found. Based on their distributions, it appears that all organisms are likely to encode for at least one Orn, NrnA, NrnB or NrnC homolog. Yet, some bacteria also encode for multiple classes of nanoRNase genes, suggesting that the additional genes may encode for either redundant functions or for proteins that act as specialized paralogs. Indeed, not all NrnA homologs act as general exoribonucleases of short, linear RNAs, such as Orn, NrnA, NrnB and NrnC. For example, *Mycobacterium tuberculosis* CnpB is an NrnA-like protein that is known to preferentially process cyclic di-adenosine monophosphate (c-di-AMP) (He et al., 2016, Dey et al., 2017, Weiss et al., 2022). Also, *Vibrio cholera* PggH is an NrnA-like protein that specifically processes pGpG dinucleotides, at the exclusion of other RNAs (Heo et al., 2021). Therefore, the full range of the intracellular functions of Orn, NrnA, NrnB and NrnC proteins has not been assessed and their actual substrate preferences cannot be predicted by their current gene annotation.

Although NrnA and NrnB are both members of the DHH-DHHA1 protein superfamily, it has been unclear whether they serve redundant functions, or what features may or may not differentiate NrnA-like and NrnB-like proteins. Indeed, they are oftentimes mentioned in the literature as NrnA/NrnB, suggesting a potential redundancy in cellular function. Recently, we investigated whether $NrnA_{Bs}$ and $NrnB_{Bs}$ were individually important for cleavage of short RNAs in *B. subtilis* during vegetative growth conditions (Weiss et al., 2022). This revealed that *nrnA* was specifically required for degradation of short RNAs but that *nrnB* was dispensable under the growth conditions tested. Furthermore, a proteomics analysis revealed the presence of $NrnA_{Bs}$, but not $NrnB_{Bs}$ during exponential growth in rich medium. Taken together, these data suggested that $NrnA_{Bs}$ is likely to act as the primary enzyme for degrading short RNAs during vegetative growth conditions. But if NrnA is the 'housekeeping' enzyme for degradation of short RNAs, when is NrnB expressed? And is NrnB biochemically like NrnA or does it display RNA substrate preferences that are significantly different? Is it appropriate to lump NrnA and NrnB together as if they comprise a single functional entity? In this study, we sought to answer these questions.

5.3 Results

5.3.1 *B. subtilis nrnB* is expressed in the forespore during endospore development

In a recent analysis of *B. subtilis* NrnA (Weiss *et al.,* 2022) we found that this protein primarily functions as the housekeeping RNase for degrading short RNAs, from 2–4 nucleotides in length, during exponential growth conditions. A prior investigation of *nrnA* and *nrnB* transcript abundance, as reported by a publicly available suite of transcriptomic datasets (Nicolas *et al*., 2012), revealed that *nrnA* is expressed under a broad range of growth conditions but that *nrnB* may exhibit an increase in transcript abundance during endospore formation. Furthermore,

our proteomic analysis of wild-type *B. subtilis* cells also suggested that $NrnB_{Bs}$ is likely to function outside of exponential-phase growth conditions (Weiss *et al.,* 2022). Inspired by these two observations, we manually searched for a putative promoter located in the intergenic space between the *yngJIHGFE* and *nrnB* operons. This revealed a short sequence segment upstream of *nrnB* that was consistent with the reported consensus patterns for promoters recognized by the sigma factors SigF and SigG (Figure 5.1A) (Wang *et al.,* 2006). Notably, SigF and SigG are sigma factors that primarily control the expression of proteins within the forespore during later stages of endospore development; therefore, we hypothesized that *nrnB* is expressed within the forespore.

To examine *nrnB* expression, we constructed *B. subtilis* strains in which the *nrnB* gene was translationally fused to open reading frames that encode either Yellow Fluorescent Protein (YFP) or Superfolder GFP (sfGFP). Importantly, the complete intergenic region upstream of *nrnB* was also included in these reporter fusions; therefore, any detectable expression of the *nrnB-yfp* fusions would result from transcription determinants located in the *yngJIHGFE-nrnB* intergenic region. These genetic reporters (P_{nrnB}-*nrnB-yfp*, P_{nrnB}-*nrnB-sfgfp*, respectively) were then integrated into a non-essential locus (Figure 5.1B). For this set of experiments, we included a control strain harboring a synthetic promoter ('P_{const}') that resulted in constitutive transcription of *yfp*. We also included a genetic reporter featuring a known forespore-specific promoter (P_{sspB}) transcriptionally fused to *yfp*. The P_{const}-*yfp* strain was included to ensure cells were actively expressing housekeeping genes, while the P_{sspB}-*yfp* strain served to indicate forespore specific gene expression. When we imaged cells containing the P_{nrnB}-*nrnB-sfgfp* translational fusion during vegetative growth, we did not observe any fluorescent signal, suggesting that *nrnB* was not being expressed during this phase of growth (Figure 5.2A). Since we did not see any

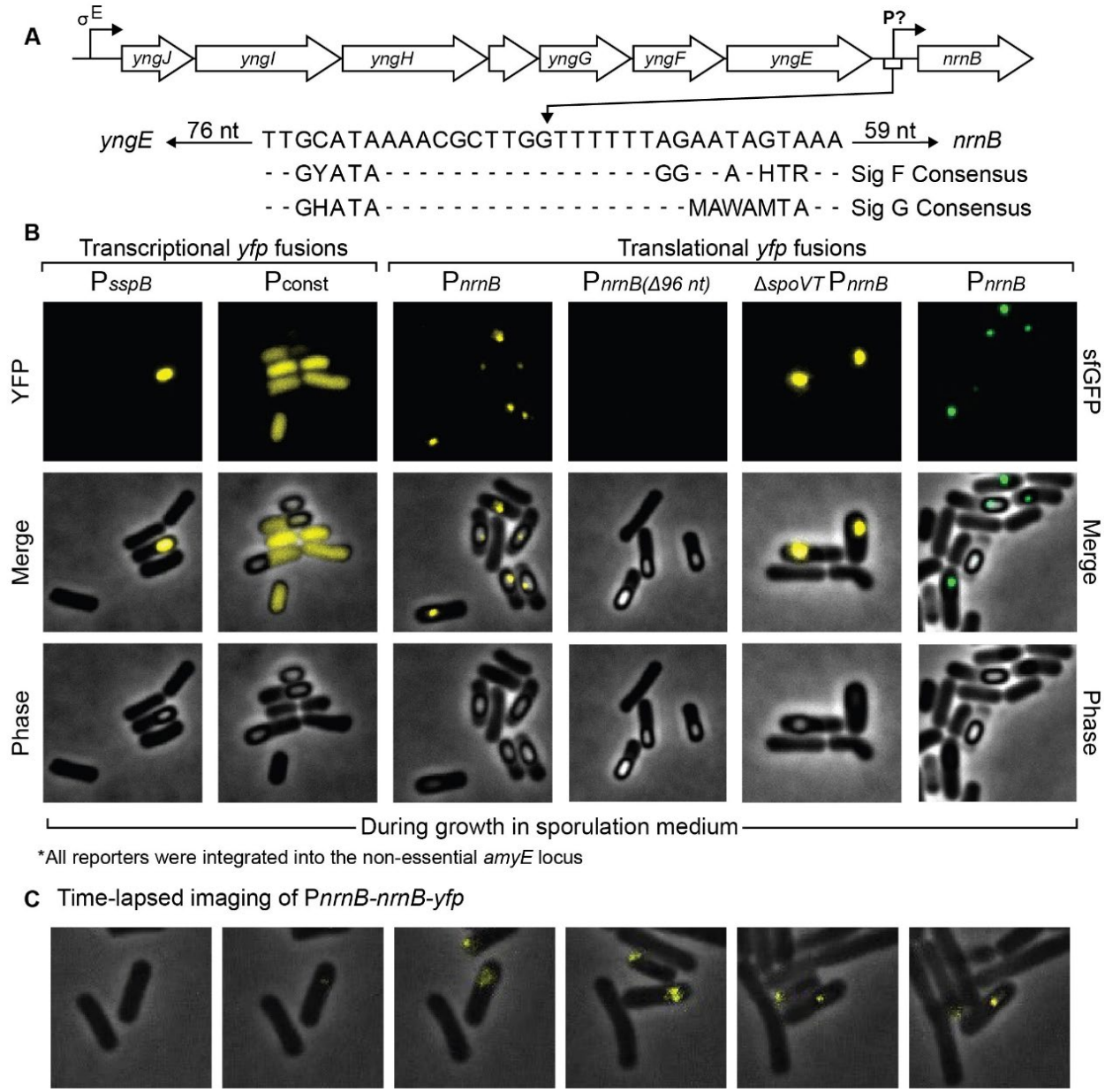

Figure 5.1. *Bacillus subtilis nrnB* is expressed during sporulation within the forespore. (A) The gene arrangements of the *yngJIHGFE* operon and *nrnB* are shown. The putative Sigma F/G-dependent promoter located within the *yngJIHGFE-nrnB* intergenic region is shown (H is A or C or T; M is A or C; Y is C or T; W is A or T; R is A or G). The Sigma F/G consensus patterns have been published previously (Wang *et al.,* 2006). (B) Representative microscopy images of *B. subtilis* 168 that was subcultured into Difco sporulation media and incubated, shaking at 37°C. The *B. subtilis* strains harboring P_{sspB}-*yfp* and P_{const}-*yfp* were imaged using a shorter exposure time than the P_{nrnB}-*nrnB-yfp,* $P_{nrnB(\Delta 96\ nt)}$-*nrnB-yfp*, P_{nrnB}-*nrnB-yfp* in a Δ*spoVT* strain background, and P_{nrnB}-*nrnB-sfgfp* reporter fusions. (C) Representative images from an overnight time-lapsed microscopy experiment showing sporulation of *B. subtilis* cells that contained the P_{nrnB}-*nrnB-yfp* reporter fusion.

appreciable signal for the *nrnB* translational fusions during vegetative growth, we turned our attention to sporulation conditions. For this analysis, strains of *B. subtilis* harboring the reporter fusions for P_{const}-*yfp*, P_{sspB}-*yfp,* P_{nrnB} -*yfp* (*i.e.,* the putative *nrnB* promoter region transcriptionally fused to the *yfp* gene), or P_{nrnB}-*nrnB*-*yfp* were subcultured into DIFCO sporulation media and incubated, shaking, at 37°C. Sporulation is known to be triggered by these conditions. Cells harboring P_{const}-*yfp* displayed fluorescence signal across the entirety of the cells, consistent with a constitutive promoter that is governed by a housekeeping sigma factor (Figure 5.1B). The P_{sspB}-*yfp* strains showed fluorescence intensity spread across the forespore, confirming expression of the P_{sspB} promoter during sporulation (Figure 5.1B). Cells containing P_{nrnB}-*yfp* generated a modest signal in the phase-bright forespore (Figure 5.2B). Similarly, cells that harbored the P_{nrnB}--*nrnB*-*yfp* translational fusion displayed a fluorescent signal localized within the developing forespore. However, the fluorescent signal for the *nrnB*-*yfp* translational fusion appeared to be contained within a few foci, unlike the transcriptional fusion P_{nrnB}-*yfp*, which showed fluorescent signal throughout the forespore. To ensure that there was not a cryptic promoter sequence contained within the *nrnB* coding sequence that affected expression of *yfp*, a 96-nucleotide portion of the *yngE*-*nrnB* intergenic space, including half of the putative SigF/G promoter, was removed from the *yfp* reporter construct (Figure 5.1A). Correspondingly cells, harboring $P_{nrnB(\Delta 96\ nt)}$-*nrnB*-*yfp* were subcultured into DIFCO sporulation media to induce sporulation; however, we observed no appreciable fluorescence signal for this strain (Figure 5.1B). Taken together, these data strongly suggested that *nrnB* is activated by a forespore-specific promoter and that the NrnB protein is largely contained within the developing forespore.

Transcriptomic datasets have been previously published for deletion mutants of *B. subtilis* sporulation sigma factors (Arrieta-Ortiz *et al.,* 2015). An assessment of these data

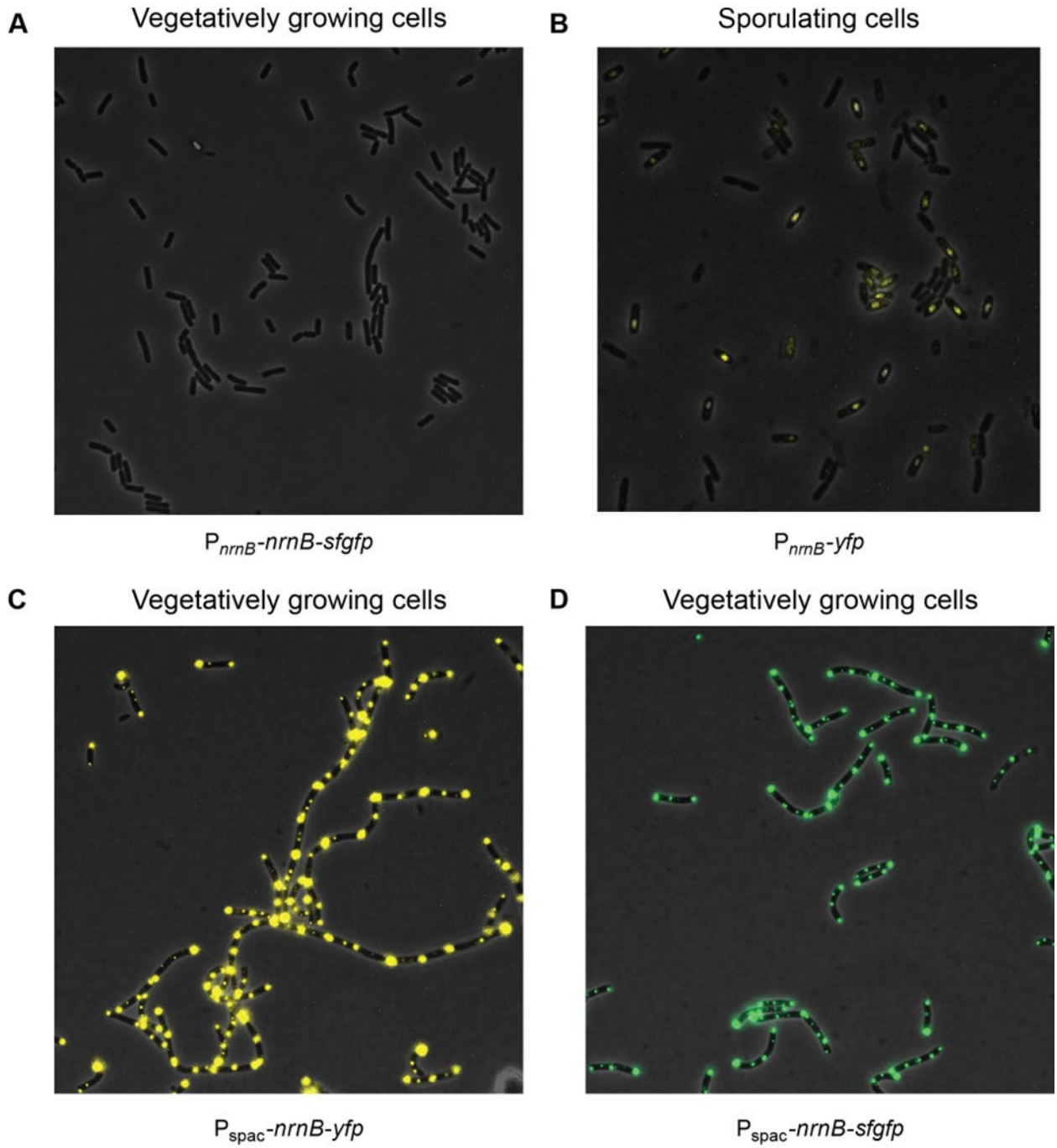

Figure 5.2. NrnB exhibits sporulation-specific expression. (A) *Bacillus subtilis* harboring a translational fusion for *nrnB-sfgfp,* imaged at an OD_{600} = 1. (B) Sporulating *B. subtilis* harboring a transcriptional fusion of P*nrnB-yfp*. (C & D) Representative images are shown for cells harboring IPTG-inducible copies of *nrnB* translationally fused to *yfp* or *sfgfp*. These cells were cultured continuously in the presence of 250 μM IPTG and imaged at an OD_{600} = 0.8.

suggested that *nrnB* is likely to be activated by SigG and potentially repressed by the SigG-activated transcription factor SpoVT, which is required for the development of the late stages of forespore development (Wang *et al.,* 2006; Bagyan *et al.,*1996). SpoVT has been shown to act as either a repressor or activator for many known SigG-dependent genes. To further investigate the forespore-specific regulation of *nrnB,* we constructed a Δ*spoVT* strain containing the P_{nrnB}-*nrnB-yfp* reporter integrated into the genome (Figure 5.1B). Analysis of this strain revealed that the YFP signal intensity for P_{nrnB}-*nrnB-yfp* was increased in the Δ*spoVT* strain as compared to the wild-type background. These data confirm that *nrnB* is expressed within the SigG and SpoVT regulons.

To identify whether *nrnB* was expressed exclusively within the forespore throughout spore development, we conducted time-lapsed fluorescence microscopy of cells containing P_{nrnB}-*nrnB-yfp* as they progressed through endospore formation (Figure 5.1C). We observed that a sporulating *B. subtilis* cell initially contained a diffuse YFP signal that rapidly coalesced to form a subcellular focal point of YFP intensity. Additionally, we noticed that when spores harboring P_{nrnB}-*nrnB-yfp* were induced to germinate with excess L-alanine, the fluorescent focal point persisted inside the mature spore until spore cortex lysis (Figure 5.3). However, to address whether this subcellular localization was somehow an artifact of the YFP do- main, we translationally fused *nrnB* to super folder GFP (*sfgfp;* (Pédelacq *et al.,* 2006)) and integrated it into the non-essential locus. Like the YFP reporter fusion, mature spores harboring the P_{nrnB}-*nrnB-sfgfp* translational fusion displayed evidence of subcellular localization (Figure 5.1B). From these observations, we speculate that the subcellular localization of fluorescently tagged $NrnB_{Bs}$ might be an intrinsic property of $NrnB_{Bs}$.

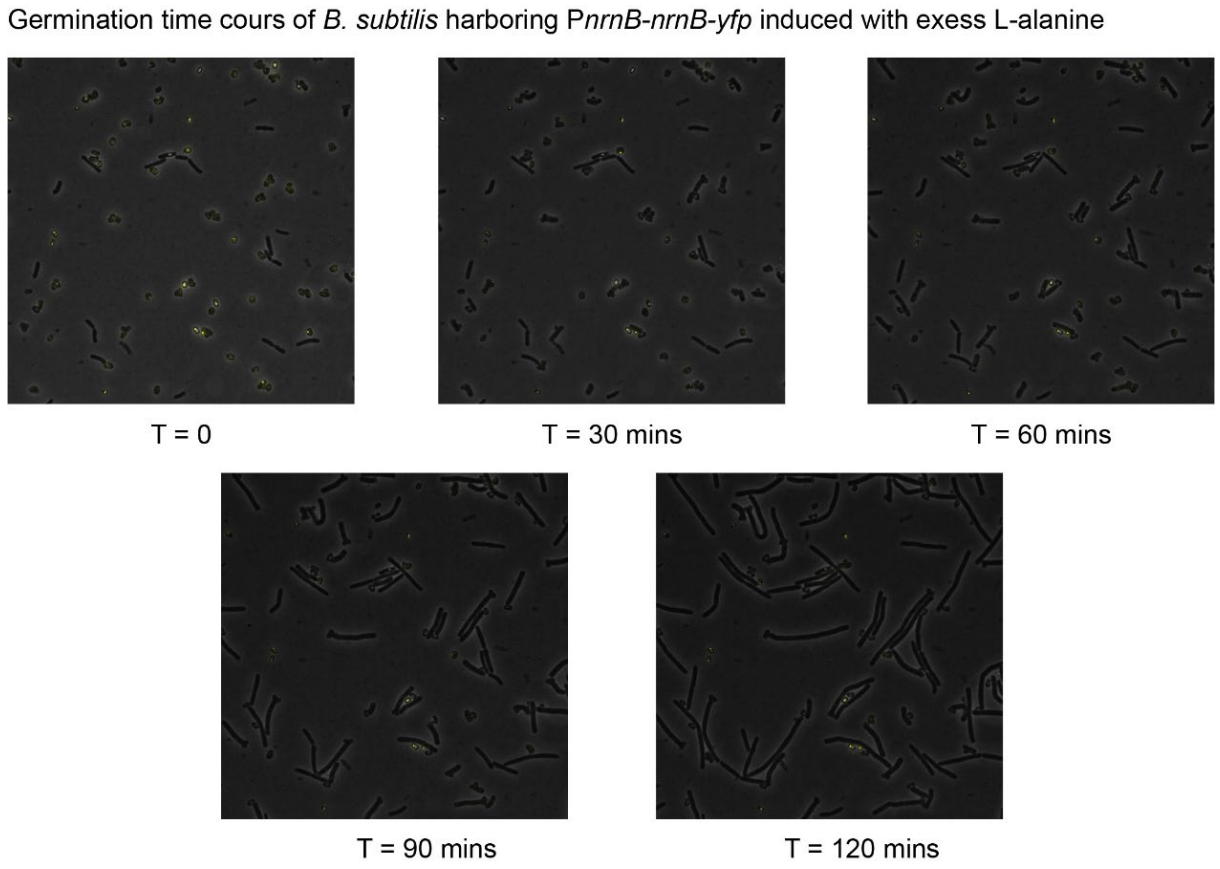

Figure 5.3. Imaging a translational *nrnB-yfp* fusion during germination. Representative images are shown for germinating spores harboring P_{nrnB}-*nrnB-yfp nrnB* that were induced to germinate by the addition of 10 mM L-alanine. Images were taken every 30-minutes for 2 hours.

Table 2. A moist heat killing assay reveals a moderate defect in sporulation efficiency in *nrnA* or *nrnB* deletion strains

Strain	CFU/mL ($x10^6$)	Spores/mL ($x10^6$)	Spores/CFU	Average Sporulation Efficiency (with SEM)
WT 168	790	750	0.95	0.94 ± 0.14
	940	670	0.71	
	450	530	1.18	
Δ*nrnA*	740	60	0.08	0.04 ± 0.02
	690	6.7	0.01	
	890	19.8	0.02	
Δ*nrnB*	790	530	0.67	0.56 ± 0.08
	1090	550	0.50	
	860	450	0.52	
Δ*nrnA* Δ*nrnB*	570	1.13	0.002	0.0047 ± 0.0015
	710	3.2	0.005	
	570	3.8	0.007	

To investigate whether depletion of *nrnB* resulted in impaired *B. subtilis* sporulation, we conducted sporulation efficiency tests on wild-type, Δ*nrnA,* Δ*nrnB,* and Δ*nrnA*Δ*nrnB* strains of *B. subtilis* (Table 2). Wild-type *B. subtilis* displayed an average sporulation efficiency of 0.94 ± 0.14. Cells lacking *nrnB* resulted in a modest ~50% decrease in sporulation, where the average sporulation efficiency was 0.56 ± 0.08. Depletion of *nrnA* resulted in a more severe sporulation defect, displaying an average sporulation efficiency of 0.04 ± 0.02, which corroborates previous reports (Meeske *et al.,* 2016). When both *nrnA* and *nrnB* are deleted the sporulation efficiency was severely reduced to 0.0047 ± 0.0015. These data suggest that degradation of short RNAs is likely to be an important process for generating heat resistant spores.

5.3.2 *B. subtilis* NrnB is active against RNA substrates of varying lengths

Prior analyses suggested that NrnA-like enzymes preferentially degrade short RNAs 2–4 nucleotides in length (Fang *et al.,* 2009). In contrast, Orn and NrnC exhibit a preference for dinucleotides (Kim *et al.,* 2019; Lormand *et al.,* 2021). Herein, we sought to determine whether $NrnB_{Bs}$ possessed substrate length restrictions similar to those reported for NrnA-like, NrnC or Orn proteins. The protein was purified to homogeneity, as analyzed by SDS-PAGE and by size exclusion chromatography as seen in chapter 4. We then incubated purified $NrnB_{Bs}$ at an enzyme to substrate ratio of 1:20 with $5'$-^{32}-P-radiolabeled RNAs of varying lengths. Reactions were quenched at the indicated time points and the degradation products were analyzed by 20% denaturing gel electrophoresis. Unexpectedly, $NrnB_{Bs}$ was active against all the RNA substrates used in this assay. $NrnB_{Bs}$ processed short RNAs 2–5 nucleotides in length by 15 min (Figure 5.4A, B), while substrates six nucleotides and longer took 30 min for the initial substrate to deplete (Figure 5.4B, C). We also noted that RNA intermediates were visible during processing of short RNAs 2–5 nucleotides in length, while degradation intermediates appeared to be in

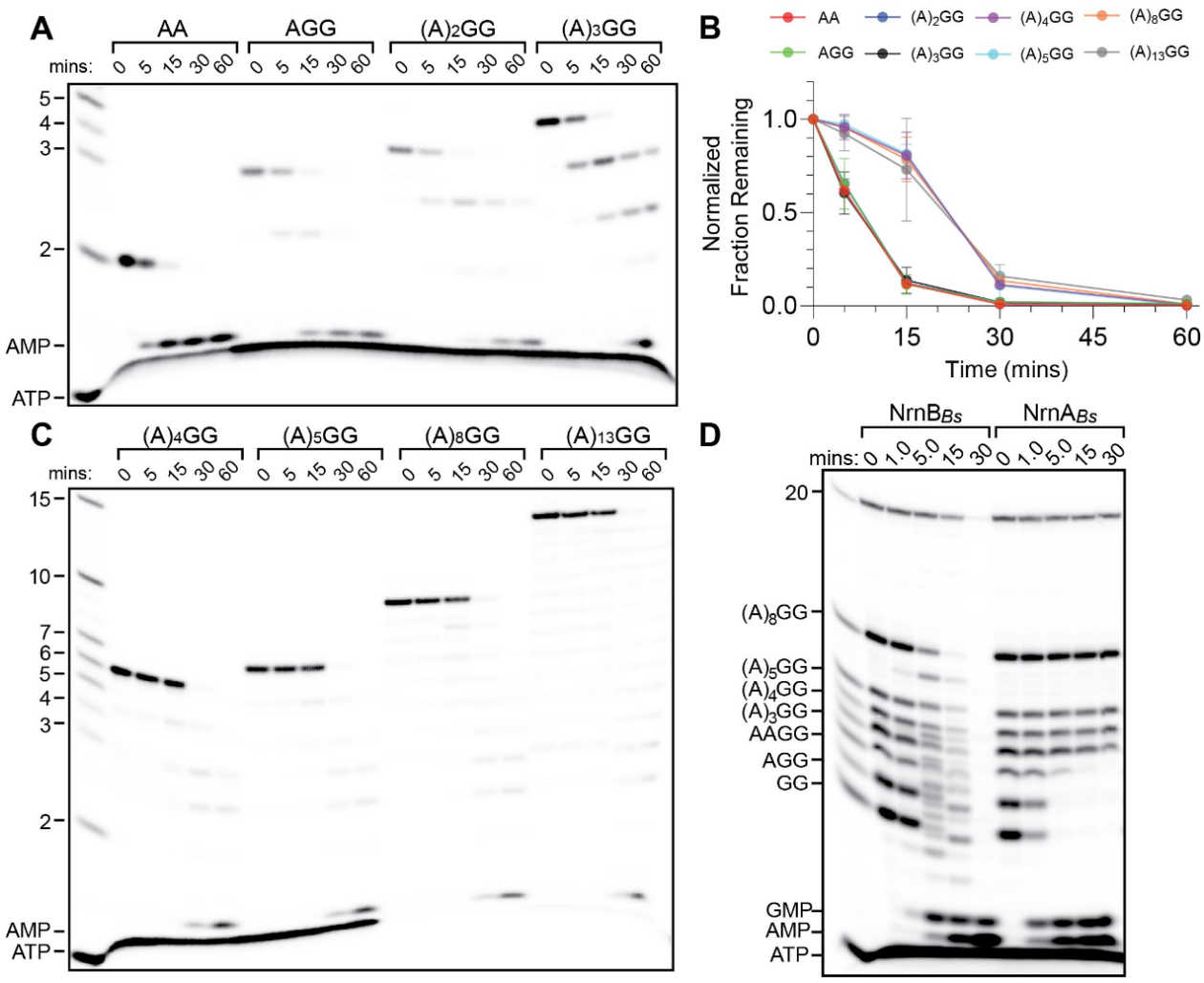

Figure 5.4. *B. subtilis* NrnB hydrolyzes RNAs of varying lengths. (A, C, and D) Native RNA molecules 2-7, 10, 15, and 20 nucleotides in length were incubated at a final concentration of 1 µM with 50 nM of purified NrnB. The reactions also contained a trace amount of 5'-^{32}P-radiolabeled RNA. Samples were removed at time intervals and analyzed by urea-denaturing 20% PAGE. (B) Quantification of the normalized radioactive intensity of the initial substrate depletion over time plotted as the average and SD of 3 independent experiments in (A and C). (D) Trace amounts of 5' ^{32}P-radiolabeled RNA molecules that were 2-7, 10, and 20 nucleotides in length were mixed and simultaneously incubated with 100 nM of purified $NrnB_{Bs}$ or $NrnA_{Bs}$. Aliquots were removed from reactions and quenched in 150 mM EDTA and 4 M urea. Degradation products were resolved by 20% denaturing PAGE.

lower abundance when NrnB$_{Bs}$ was incubated with RNA substrates greater than six nucleotides in length. We also tested the capacity of NrnB$_{Bs}$ to degrade DNA substrates 5, 10 and 15 nucleotides in length (Figure 5.5A, B). This revealed that NrnB$_{Bs}$ can hydrolyze the DNA 5-mer at approximately the same rate as an RNA substrate. However, the longer DNA substrates were degraded at a much slower rate as compared to RNA substrates. While these data suggest that the 2$^{'}$ hydroxyl is not essential for NrnB catalyzed reactions, NrnB$_{Bs}$ may prefer long RNA substrates to DNA.

Our data strongly suggests that NrnB$_{Bs}$ does not possess the same RNA substrate length restrictions as seen for NrnA$_{Bs}$. It was previously suggested that NrnB$_{Bs}$ might possess a greater propensity than NrnA$_{\boldsymbol{Bs}}$ to hydrolyze long RNA substrates, but the general activity of NrnA$_{Bs}$ and NrnB$_{Bs}$ against a 24-mer RNA was reported to be minimal (Mechold et al., 2007, Fang et al., 2009). Also, the activities of NrnA$_{Bs}$ and NrnB$_{Bs}$ have not been directly compared under the exact same assay conditions using nonmodified RNA substrates. Therefore, to directly address whether there are key differences in substrate length accommodation, we simultaneously incubated 5'-radiolabeled RNAs of different lengths (2-7, 10, and 20 nucleotides in length) with purified NrnA$_{Bs}$ or NrnB$_{Bs}$ under the same reaction conditions (Figure 5.4D). When NrnA$_{Bs}$ was simultaneously presented with RNAs 2-7, 10, and 20 nucleotides in length, only short RNAs (2-4 nucleotides) were processed, while longer RNAs remained unperturbed. In contrast, NrnB$_{Bs}$ processed all of the RNA substrates, short and long, within the same incubation period. In total, these data demonstrate that NrnA$_{Bs}$ is restricted to RNAs 2-4 nucleotides in length while NrnB$_{Bs}$ does not display substrate length restrictions. Therefore, we speculate that long RNA processing may constitute a biologically relevant function for NrnB$_{Bs}$, and that substrate length

accommodation may provide a diagnostic criterion in the annotation of NrnA-like and NrnB-like proteins.

5.3.3 NrnB$_{Bs}$ does not display activity against additional nucleic acid substrates

Interestingly, other members of the same DHH-DHHA1 protein family have been shown to display robust activity against c-di-AMP as seen in chapter 3 (Tang *et al.,* 2015; He *et al.,* 2016; Dey *et al.,* 2017; Weiss *et al.,* 2022). Given the capacity of some DHH-DHHA1 protein family members, such as *M. tuberculosis* CnpB, to hydrolyze c-di-AMP, we tested whether this was a relevant substrate for NrnB$_{Bs}$ (Figure 5.5C, D). For this analysis we incubated purified NrnB$_{Bs}$ or CnpB$_{Mt}$ at an enzyme to substrate ratio of 1:20 with ^{32}P-radiolabeled c-di-AMP. As expected, CnpB$_{Mt}$ hydrolyzed the c-di-AMP to completion in 30 minutes (Figure 5.5C, D). However, NrnB *Bs* did not exhibit any appreciable activity against c-di-AMP under the same reaction conditions. This suggests that there is functional diversity amongst the proteins contained within the DHH-DHHA1 protein family.

Additionally, since other DHH-DHHA1 protein family members have been proposed to utilize CysQ-like phosphatase activity against adenosine 3', 5'-diphosphate (pAp), we directly compared the activity of purified CysQ$_{Ec}$ to that of NrnB$_{Bs}$. 5 nM of purified CysQ$_{Ec}$ or NrnB$_{Bs}$ was incubated with 100 μM pAp and phosphatase activity was measured by the absorbance change that occurs upon complexing of Malachite Green with free phosphate. While significant free phosphate was liberated in the reaction contain ng CysQ, no phosphate signal was observed for NrnB$_{Bs}$ (Figure 5.5E).

A previous analysis suggested that an NrnB-like protein encoded by *Helicobacter pylori* was capable of cAMP phosphodiesterase activity (Choi *et al.,* 2014). Therefore, to address whether NrnB$_{Bs}$ also possessed cAMP phosphodiesterase activity, we incubated 5 nM of purified

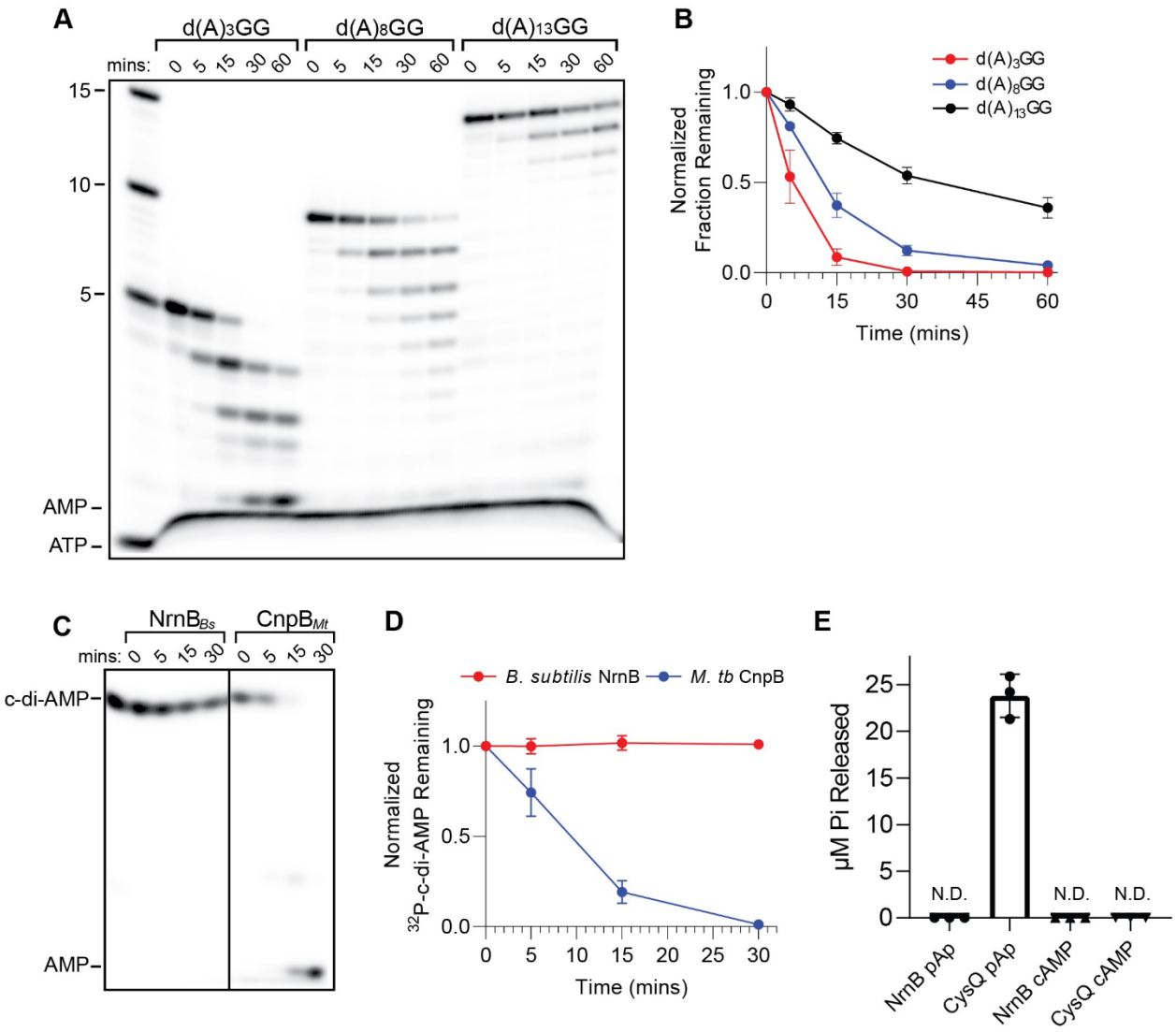

Figure 5.5. NrnB$_{Bs}$ does not display activity against additional nucleic acid substrates. (A) 1 µM DNA molecules 5, 10, and 15 nucleotides in length (containing a trace amount of 5' ^{32}P-radiolabeled DNA) were subjected to cleavage by 50 nM of purified NrnB$_{Bs}$. (B) Quantification of the normalized radioactive intensity of the initial substrate depletion over time plotted as the average and SD of 3 independent experiments. (C) 1 µM c-di-AMP containing trace amounts of 5' ^{32}P-c-di-AMP was subjected to cleavage by 50 nM of NrnB$_{Bs}$ or CnpB$_{Mt}$. (D) Quantification of the normalized radioactive intensity of the initial substrate depletion over time plotted as the average and SD of 3 independent experiments. (A & C) Aliquots were removed from reactions and quenched in 150 mM EDTA and 4 M urea. Degradation products were resolved by 20% denaturing PAGE. (E) Phosphate release was determined by the Malachite green assay after subjecting 100 µM pAp or cAMP to degradation by 5 nM of NrnB$_{Bs}$ or CysQ$_{Ec}$. For cAMP phosphodiesterase activity, 10 U of alkaline phosphatase was added to the reactions to liberate 5' phosphate if the 3'-5' phosphate linkage had been broken. N.D. indicates that the phosphate release was below our limits of detection.

NrnB$_{Bs}$ with 100 μM of cAMP (Figure 5.5E). 10 U of alkaline phosphatase were added to these reactions to couple cAMP phosphodiesterase activity directly to the release of phosphate from AMP, which could then be monitored using the Malachite Green assay, as described above. After subjecting cAMP to hydrolysis by NrnB$_{Bs}$ for 30 min, we observed no change in free phosphate. From these aggregate data, we conclude that NrnB$_{Bs}$ is a unique NanoRNase that degrades linear RNA molecules of varying lengths, which contrasts with the activities reported by other NanoRNases such as NrnA$_{Bs}$.

5.3.4 *B. subtilis* NrnB degrades RNAs from the 3’ terminus

The processing of short RNAs 2-5 nucleotides in length by NrnB$_{Bs}$ (Figure 5.4A) shows clear accumulation of intermediate degradation products. However, degradation intermediates are generally absent or in low abundance in reactions where longer are subjected to degradation by NrnB$_{Bs}$ (Figure 5.4C). We recently found that NrnA employs a 5’-3’ exonucleolytic strategy to degrade short RNAs (Weiss *et al.,* 2022). Thus, it is possible that NrnB$_{Bs}$ utilizes an indiscriminate distributive mechanism against short RNAs but degrades long RNAs with a 5’-3’ polarity akin to NrnA$_{Bs}$. Nevertheless, the simplest explanation is that NrnB$_{Bs}$ hydrolyzes RNAs with 3’-5’ polarity, opposite that of NrnA. To test for directionality, we utilized a fluorescence-based approach. Specifically, we determined whether NrnB$_{Bs}$ would release a terminal 2-aminopurine residue from either the 5’- or 3’-terminus for a 4-mer RNA substrate that also contained an internal phosphorothioate linkage (*i.e.,* 5’-(2AP)A*ps*GG or 5’-AA*ps*G(2AP)). When 2AP is liberated from even short RNA molecules, there is a resulting increase in fluorescence output. Previously we found that the fluorescence intensity of 5’-(2AP)A*ps*GG or 5’-AA*ps*G(2AP) remains constant in control reactions where 2AP is not liberated, yet incubation of these RNAs with NrnA$_{Bs}$ resulted in an increase in fluorescence for only the 5’-(2AP)A*ps*GG

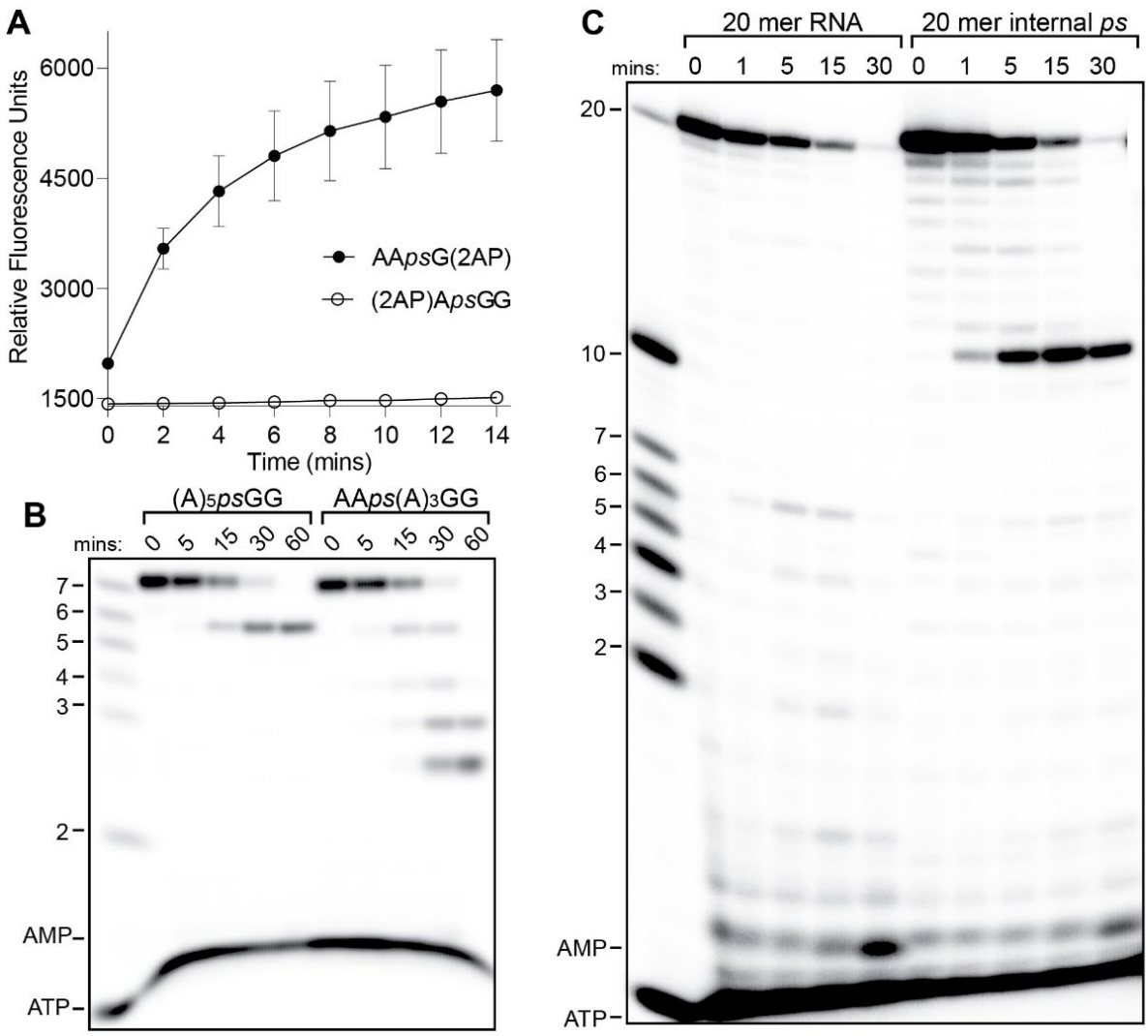

Figure 5.6. *B. subtilis* NrnB degrades RNAs from the 3' terminus. (A) A 10 μM solution of RNA molecules 4, 7, or 20 nucleotides in length was incubated with 100 nM NrnB. These RNA substrates also contained an internal non-hydrolyzable phosphorothioate linkage (backbone placement of the linkage represented by '*ps*') and a 2-aminopurine (nucleotide replacement represented by '2AP') located at the terminus. Intrinsic fluorescence of 2AP increases upon its release from an RNA polymer; therefore, the fluorescence emission changes shown herein represent release of 2AP from either the 5' or 3' terminus. Data is plotted as the average and SD of 3 independent experiments. (B) 1 μM of 7-mer RNA substrate (with trace amounts of 5' -^{32}P-radiolabeled RNA) was incubated with 100 nM purified NrnB. These RNA substrates also contained an internal phosphorothioate linkage (indicated by '*ps*'). Aliquots were removed at varying time intervals and resolved by urea-denaturing 20% PAGE. (C) Trace amounts of 5'-^{32}P-radiolabeled 20-mer RNA was incubated with 100 nM of purified NrnB. NrnB was also incubated with an equivalent 20-mer RNA substrate that contained an internal phosphorothioate linkage (at the tenth residue). The products of these reactions were resolved by urea-denaturing 20% PAGE.

substrate (Weiss et al., 2022). Herein, we incubated NrnB$_{Bs}$ with 5'-(2AP)A*ps*GG or 5'-AA*ps*G(2AP) and only observed an increase in fluorescence with the 5'-AA*ps*G(2AP) substrate (Figure 5.6A). These data indicate that NrnB$_{Bs}$ can remove the 3' residue of a short RNA substrate and cannot hydrolyze through a phosphorothioate linkage. To ensure that there were no mechanistic differences in the degradation of short RNAs and longer RNAs, we correspondingly assessed the cleavage of two different 7-mer RNAs containing a non-hydrolyzable phosphorothioate linkage. When the 7-mer substrate 5'-^{32}P-AAAAApsGG was incubated with NrnB$_{Bs}$, we observed accumulation of a degradation product corresponding to 5'-^{32}P-AAAAApsG (*i.e.,* removal of a single 3' guanosine monosphosphate residue; Figure 5.6B). In reactions where we subjected 5'-^{32}P-AApsAAAGG to cleavage by NrnB$_{Bs}$, we visualized degradation products that corresponded to 5'-^{32}P-AApsAAAG, 5'-^{32}P-AApsAAA, 5'-^{32}P-AApsAA, and 5'-^{32}P-AApsA (Figure 5.6B). Additionally, to ensure that NrnB$_{\boldsymbol{Bs}}$ was acting exclusively as a 3'-5' exonuclease, we compared the degradation of a native 5' labeled 20-mer RNA to a 20-mer RNA containing an internal phosphorothioate linkage at the tenth residue (Figure 5.6C). In agreement with the data in (Figure 5.4D) NrnB$_{Bs}$ processed the native 20-mer to almost exclusively monoribonucleotides in 30 minutes. Interestingly, NrnB$_{Bs}$ rapidly processed the 20-mer containing the internal phosphorothioate to the length that corresponded to the enzyme reaching the phosphorothioate linkage and stalling. Taken together, these data suggest that NrnB$_{Bs}$ degrades RNAs of different lengths using a 3'-5' exonucleolytic mechanism.

5.3.5 Elevated expression of *nrnB* in *B. subtilis* cellular lysates leads to faster processing of long RNA substrates

To address whether general exonucleolytic activity of NrnB$_{Bs}$ is a physiologically relevant trait, we incubated 1 μM of phosphorylated 22-mer RNA containing trace amounts of 5'

radiolabeled RNA with total cell lysates extracted from late-exponential phase cells (OD_{600} ≈ 1.0). For these experiments, we used cell lysates that were extracted from wild-type, Δ*nrnB*, or a Δ*nrnB* strain that harbored an ectopic, IPTG-inducible *nrnB*. The degradation products were resolved by 20% denaturing PAGE (Figure 5.7A). Quantification of the initial substrates revealed that the processing of the 22-mer RNA fit to a single exponential decay ($R^2 > 0.95$). Given that the initial substrate depletion fit to a single exponential decay for all the reactions, we decided to compare the half-life of the initial substrate amongst the cellular lysates from the different strains (Figure 5.7B). The half-life of the 22-mer in the presence of the wild-type lysate is 21.8 ± 0.8 minutes, which also resembles the half-life of the 22-mer incubated with the Δ*nrnB* lysate (21.0 ± 0.8 minutes). These data agree with our expression data which strongly suggests that *nrnB* is not expressed or functionally active during vegetative growth conditions. However, the half-life of the 22-mer is significantly decreased (5.4 ± 1.0 minutes) when *nrnB* is forcibly induced by an IPTG-responsive promoter. Interestingly, there are fewer intermediate degradation products when expression of *nrnB* is induced. One potential explanation for the lack of intermediate degradation products is that $NrnB_{Bs}$ may use a processive exonucleolytic mechanism on long RNA substrates akin to the mechanisms reported for RNase R or PNPase.

5.3.6 $NrnB_{Bs}$ degrades a 44-mer RNA substrate at a similar rate to that of *E. coli* Rnr

In our analyses of $NrnB_{Bs}$, we found that it displayed robust activity against RNA molecules from 2 to 22 nucleotides in length, using both *in vitro* and *ex vivo* assay conditions. To address whether the capacity of $NrnB_{Bs}$ to degrade longer RNA substrates was similar to other general exonucleases, we decided to directly compare the hydrolysis of a native, unstructured 44-mer by $NrnB_{Bs}$ or RNase R_{Ec}. We incubated purified $NrnB_{Bs}$ or RNase R_{Ec} at an enzyme to substrate ratio of 1:10 with ^{32}P-radiolabeled 44-mer (Figure 5.8A). Under these reaction

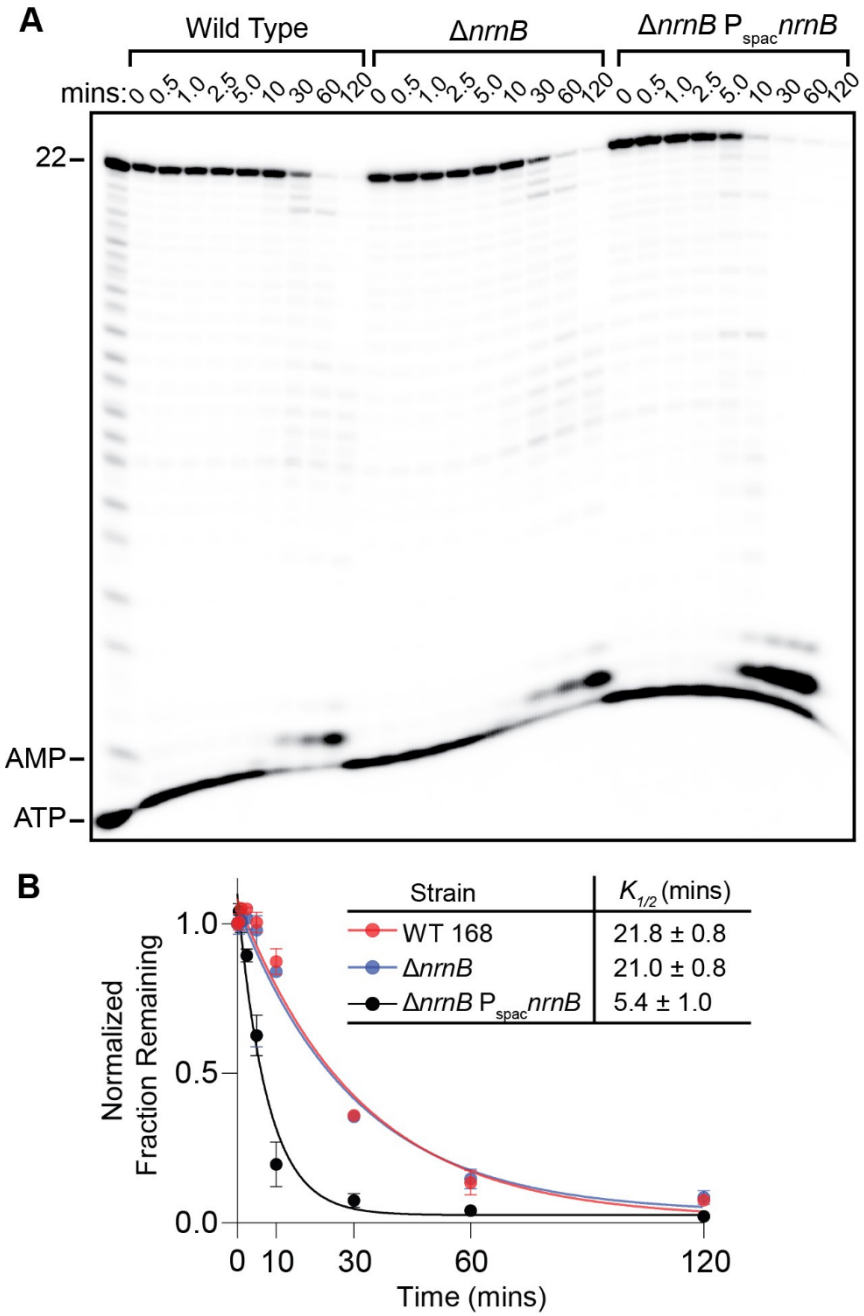

Figure 5.7. Elevated expression of *nrnB* in *B. subtilis* cellular lysates leads to faster processing of long RNA substrates. (A) Cellular lysates from wild-type, Δ*nrnB*, or Δ*nrnB amyE::nrnB (*i.e., containing an IPTG-inducible *nrnB* integrated into the non-essential *amyE* gene), were harvested during vegetative growth (OD_{600} = 1.0) and incubated with 1 mM of a native 22-mer RNA and trace amounts of the corresponding 5'-^{32}P-radiolabeled RNA. The reaction buffer was comprised of 25 mM Tris-HCl, pH 8.0, 300 mM NaCl, 25 mM $MgCl_2$, and 200 µM $MnCl_2$. Degradation products were quenched in 4 M urea and 150 mM EDTA and resolved by 20% urea-denaturing PAGE. (B) The normalized intensity of the initial substrate depletion over time is shown plotted as an average with SD of 3 independent experiments. The fit to a single exponential decay is also shown.

conditions, the 44-mer is degraded nearly to completion by 60 minutes for both enzymes (Figure 5.8B, C, Figure 5.9 A, B**)**. Interestingly, the 44-mer is degraded primarily to monoribonucleotides when incubated with NrnB$_{Bs}$ (Figure 5.8B). In contrast, the primary degradation products for RNase R$_{Ec}$ are short RNAs between 2-4 nucleotides in length. To ensure that the 44-mer degradation is a specific activity of NrnB$_{Bs}$ and not a contaminating protein in the purification preparation, we assessed the activity of an active site mutant NrnB$_{BsD61ND145N}$. We purified and assessed the activity of the NrnB$_{BsD61ND145N}$ protein against the 44-mer substrate and found that this enzyme was completely inactive (Figure 5.9 C, D**)**. This data indicates that the cleavage of the 44-mer is indeed a specific activity of NrnB$_{Bs}$ and not a contaminating protein. Together, these data suggest that NrnB$_{Bs}$ is capable of degrading long RNA substrates at a rate like that of purified RNase R$_{Ec}$, while also displaying proficient activity against short RNAs.

5.3.7 Identification of a promoter mutation that elevates *nrnB* expression during vegetative growth

Currently, *B. subtilis* is known to harbour four known exoribonucleases that use a 3'-5' polarity: PNPase, RNase R, RNase PH, and YhaM (Bechhofer and Deutscher 2019). Although *B. subtilis* utilized multiple 3'-5' exoribonucleases, PNPase has been shown to be the primary mRNA degrading 3'-5' exoribonuclease (Liu *et al.,* 2014; Chhabra *et al.,* 2022). However, due to the partial redundancy in function exhibited by general exoribonucleases in *B. subtilis,* multiple exoribonuclease genes must be deleted prior to observing any major growth phenotypes (Bechhofer and Deutscher 2019). Interestingly, a *B. subtilis* strain mutant for the genes encoding RNase R, RNase PH, and YhaM only exhibits a modest growth phenotype (Oussenko *et al.,* 2005). In a recent analysis, a *B. subtilis* strain mutant for the genes encoding RNase R, RNase PH, YhaM, and the endoribonuclease YloC was obtained (Ingle *et al.,* 2022). This quadruple

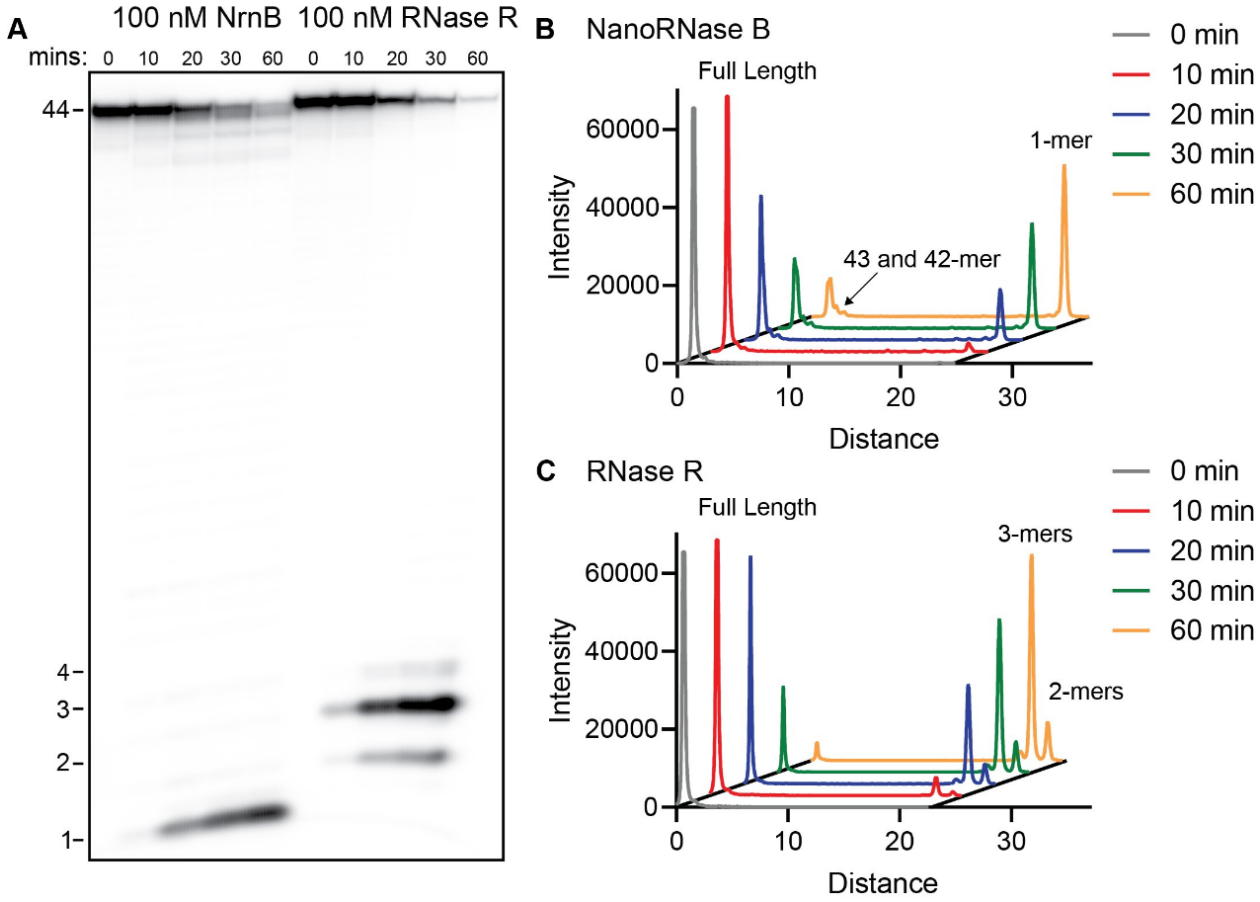

Figure 5.8. NrnB displays general exonucleolytic activity akin to *E. coli* RNase R. 1 µM 44mer RNAs (containing a trace amount of ^{32}P radiolabeled RNA) was incubated with 100 nM of purified NrnB or *E. coli* RNase R. Aliquots were removed from reactions and quenched in 150 mM EDTA and 4 M urea and resolved by 20% denaturing PAGE. (B, C) Line trace quantification of the indicated time point degradation products in (A).

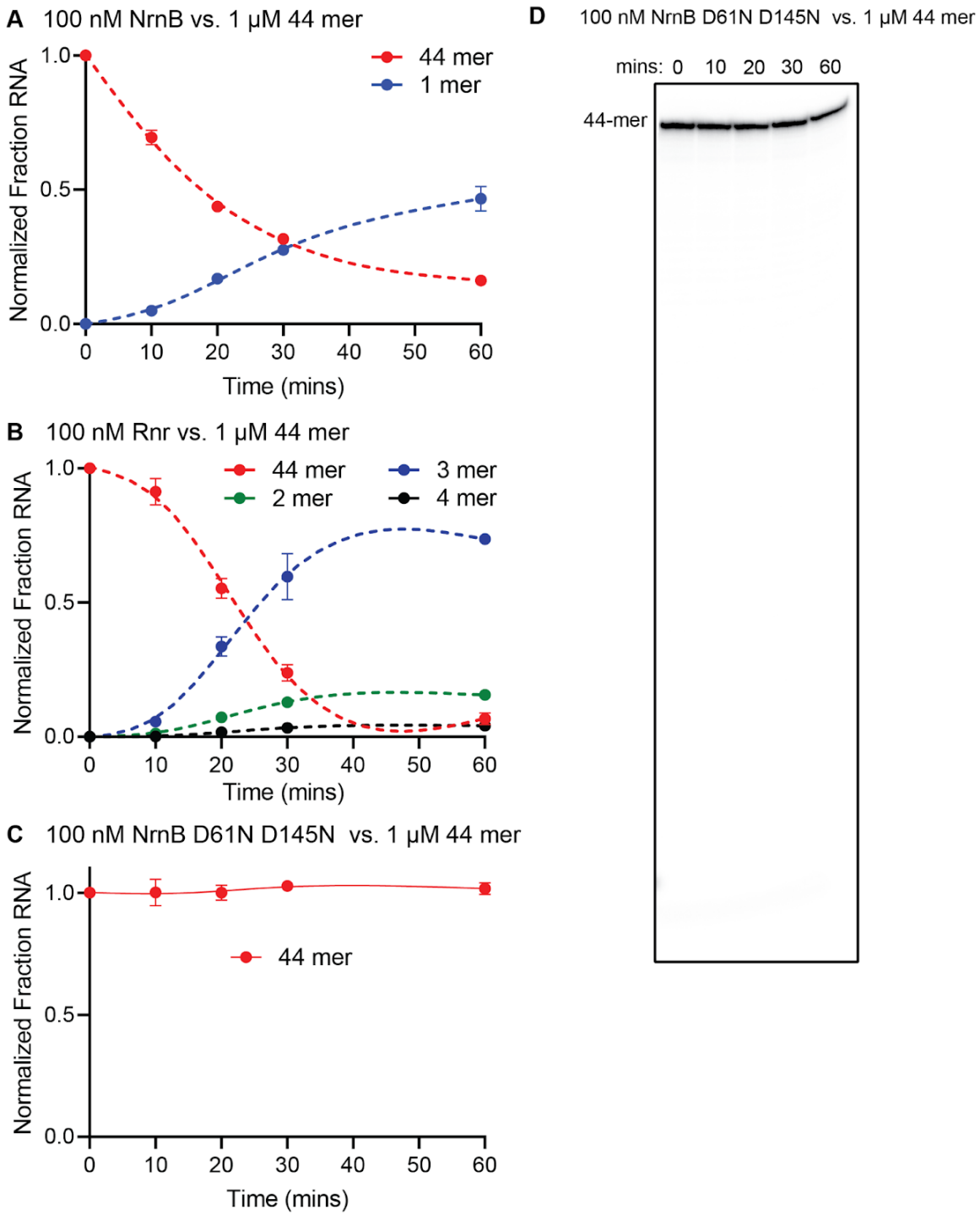

Figure 5.9. A purified inactive NrnB$_{Bs}$ mutant is not active against a 44-mer RNA. 1 μM RNA molecules 44 nucleotides in length were incubated with a trace amount of ^{32}P-radiolabeled RNA and 100 nM of purified NrnB, an inactivated mutant NrnB$_{D61ND145}$, or *E. coli* RNase R. Aliquots were removed from reactions and quenched in 150 mM EDTA and 4 M urea and resolved by 20% denaturing PAGE. (A, B, C) Quantification of the normalized radioactive intensity of the initial substrate depletion and product accumulation over time plotted as the average and SD of 2 or 3 independent experiments. (D) Gel image showing NrnB$_{D61ND145}$ does not exhibit cleavage activity against the 44-mer RNA.

RNase mutant background was utilized to generate a quintuple RNase mutant strain by knocking out the gene encoding PNPase. This strain grew very well as compared to that of a quadruple RNase R, RNase PH, YhaM, and PNPase deficient strain that had been previously constructed. This restoration in the growth phenotype exhibited by the quadruple RNase mutant strain led us to believe that suppressor mutations may have arisen in this strain. Genomic sequencing of the quintuple RNase mutant strain revealed a point mutation in the *nrnB* promoter region (Figure 5.10A), which was also observed in genomic sequencing of the quadruple mutant strain. From this observation, we hypothesized that the point mutation might be a suppressor mutation that triggered vegetative expression of *nrnB*. This T → G change created a TG dinucleotide at the −14, -15 positions of the *nrnB* promoter sequence has been noted as a conserved element 'extended -10 region' in Sig A dependent promoters (Helmann, 1995; Keilty *et al.,* 1987). In the case of extended −10 promoters, the −35 region is not essential for promoter function, but the TG dinucleotide at −14, −15 is critical (Camacho and Salas, 1999). As such, we hypothesized that the poorly conserved −35 region and non-optimal spacing between putative −35 and −10 regions in this case (Figure 5.10A) were not relevant, and that *nrnB* might be expressed in the quintuple RNase mutant strain during vegetative growth. To test this, the T → G mutation was introduced into the promoter region of the *nrnB-yfp* translational fusion; microscopy images of this strain showed that, in fact, *nrnB* was expressed during vegetative growth (Figure 5.10B). As a positive control for these experiments, we also made constructs where an IPTG-inducible promoter (P_{spac}) controlled expression of *nrnB-yfp* or *nrnB-sfgfp* reporters. Brief exposure of these cells to IPTG resulted in a 'speckling' pattern of fluorescent signals across the full length of the cells (Figure 5.10B). However, when cells were cultured in the prolonged presence of 250

μM IPTG throughout exponential phase growth, we observed strong fluorescence localization at the poles of the cells

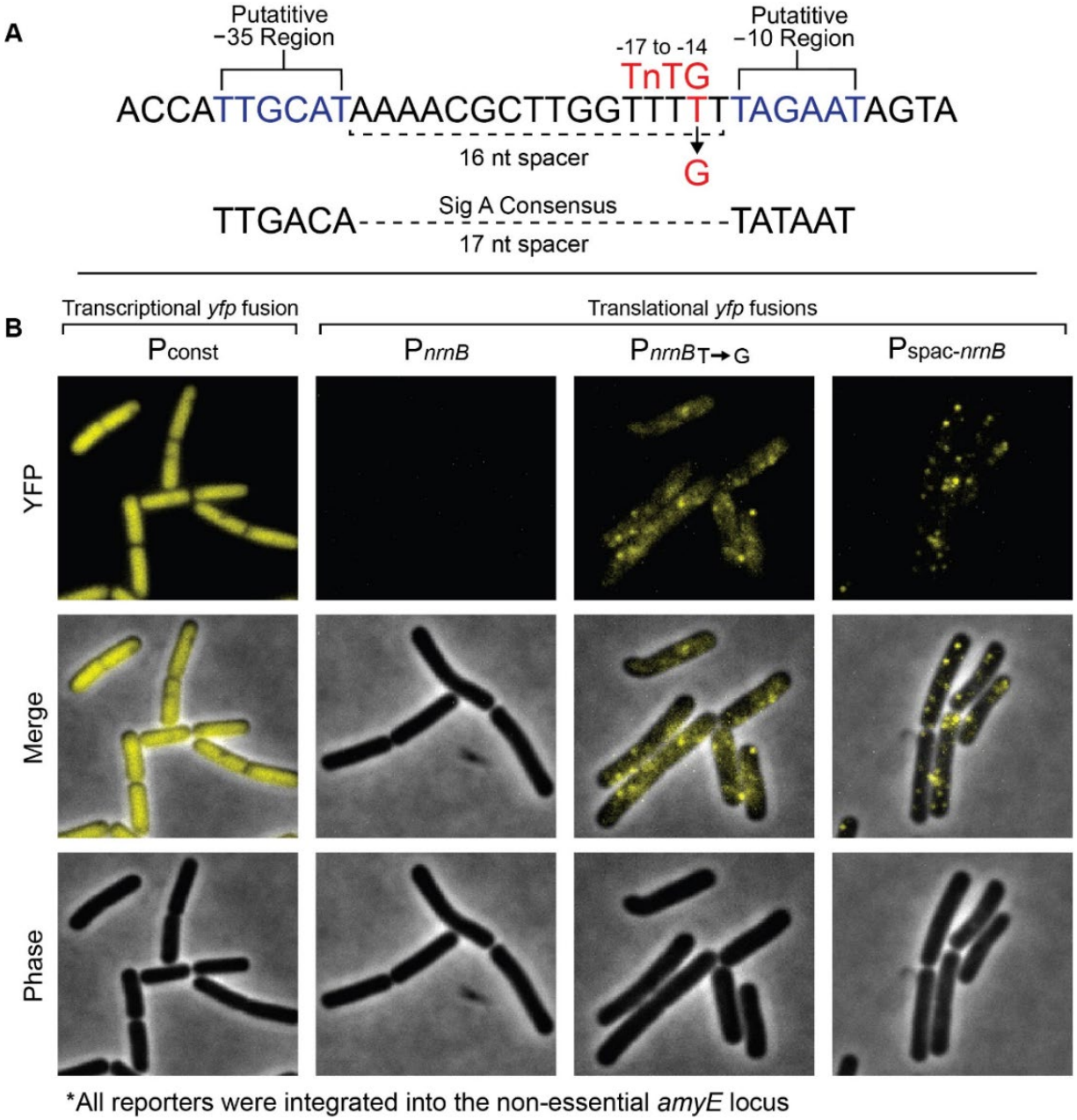

Figure 5.10. A suppressor point mutation within the intergenic region upstream of *B. subtilis nrnB* resulted in increased NrnB expression during exponential growth conditions. (A) When the intergenic region upstream of *nrnB* was fused with the *yfp* gene it did not result in detectable fluorescence during exponential growth; instead, fluorescence signal was observed within the forespore during endospore formation (**Figure 5.1**). This suggested that the DNA sequence upstream of *nrnB* does not feature a housekeeping promoter but instead contains a sporulation-specific promoter. Yet, a manual inspection of this region identified a putative promoter sequence for the housekeeping sigma factor, Sigma A, although the spacer distance between the -35 and -10 boxes is less than consensus at 16 nucleotides. A strain of *B. subtilis* lacking five exoribonuclease genes (Δ*rnr* Δ*pnpA* Δ*rph* Δ*yhaM* Δ*yloC*) was previously shown to exhibit a moderate growth defect. However, a few suppressor mutants were obtained as result of an enrichment for improved growth for the RNase-deficient strain. One of these mutant strains contained a T➔G mutation upstream of *nrnB*, resulting in the formation of an extended -10 region that conforms to the TnTG consensus sequence found in many strong sigma A promoter regions (Helmann *et al.,* 1995). (B) This point mutation was then added to the *nrnB-yfp* reporter fusion, alongside a *yfp* reporter containing the native sequence. Representative microscopy images are shown for *B. subtilis* 168 cultured in LB media, shaking at 37°C, to a final OD_{600} = 1.0. While a shorter exposure time was used to image the *B. subtilis* strain harboring constitutive *yfp* (P_{const}-*yfp*), a longer, but consistent, exposure time was used for imaging the P_{nrnB}-*nrnB-yfp*, $P_{nrnB\ T\rightarrow G}$-*nrnB-yfp,* and P_{spac}-*nrnB-yfp* strains. Cells containing P_{spac}-*nrnB-yfp* were plated on an agarose pad containing 250 μM IPTG to induce expression of the P_{spac} promoter. *These data were informed by a suppressor screen conducted by Dr. Shakti Ingle, Bechhofer laboratory, Department of Pharmacological Sciences, Icahn School of Medicine at Mount Sinai.*

(Figure 5.2C, D). From this we speculate that overexpression of *nrnB-yfp* or *nrnB-sfgfp* will result in protein aggregation. Taken together, these data demonstrate that the native *nrnB* gene is expressed during sporulation from within the forespore, where it may be localized into distinct foci, although the basis of subcellular localization is not known. Furthermore, our experiments suggest that expression of *B. subtilis* $NrnB_{Bs}$ may alleviate the burden in mRNA processing that occurs when the full complement of 3'-5' exonucleolytic RNases is not present.

5.3.8 A potential structural feature that may contribute to the longer substrate length permitted by *B. subtilis* NrnB

Previous reports have suggested that $NrnA_{Bs}$ and $NrnB_{Bs}$ are effectively synonymous in their biochemical activity and possibly their intracellular functions (Mechold *et al*., 2007, Fang *et al.,* 2009). In a previous study we conducted a comprehensive analysis of NrnA's substrate specificity and polarity and found that NrnA-like proteins degrade short RNAs 2-4 nucleotides in length specifically from the 5' terminus (Weiss *et al*., 2022). Yet, in this investigation we find that $NrnB_{Bs}$ utilizes 3'-5' polarity and is not restricted by substrate length. Given these mechanistic disparities, and the current lack of structural information for $NrnB_{Bs}$, we utilized AlphaFold 2 to generate a putative model for $NrnB_{Bs}$ to compare with crystallographic data of $NrnA_{Bs}$ PDB ID 5J21 (Schmier et al., 2017, Jumper *et al*., 2021, Mirdita *et al*., 2022). This analysis predicted that the general domain structures of $NrnA_{Bs}$ and $NrnB_{Bs}$ are similar in that both proteins are comprised of a two-lobed structure containing a DHH domain (dark blue), a linker (grey), and a DHHA1 domain (red) (Figure 5.11A). However, unlike NrnA, $NrnB_{Bs}$ is predicted to possess a long linker region (grey) and a unique C-terminal extension (cyan). We speculated that the large linker might allow for greater conformational dynamics for $NrnB_{Bs}$, perhaps enabling longer RNA substrate accommodation. Although the oligomeric state is

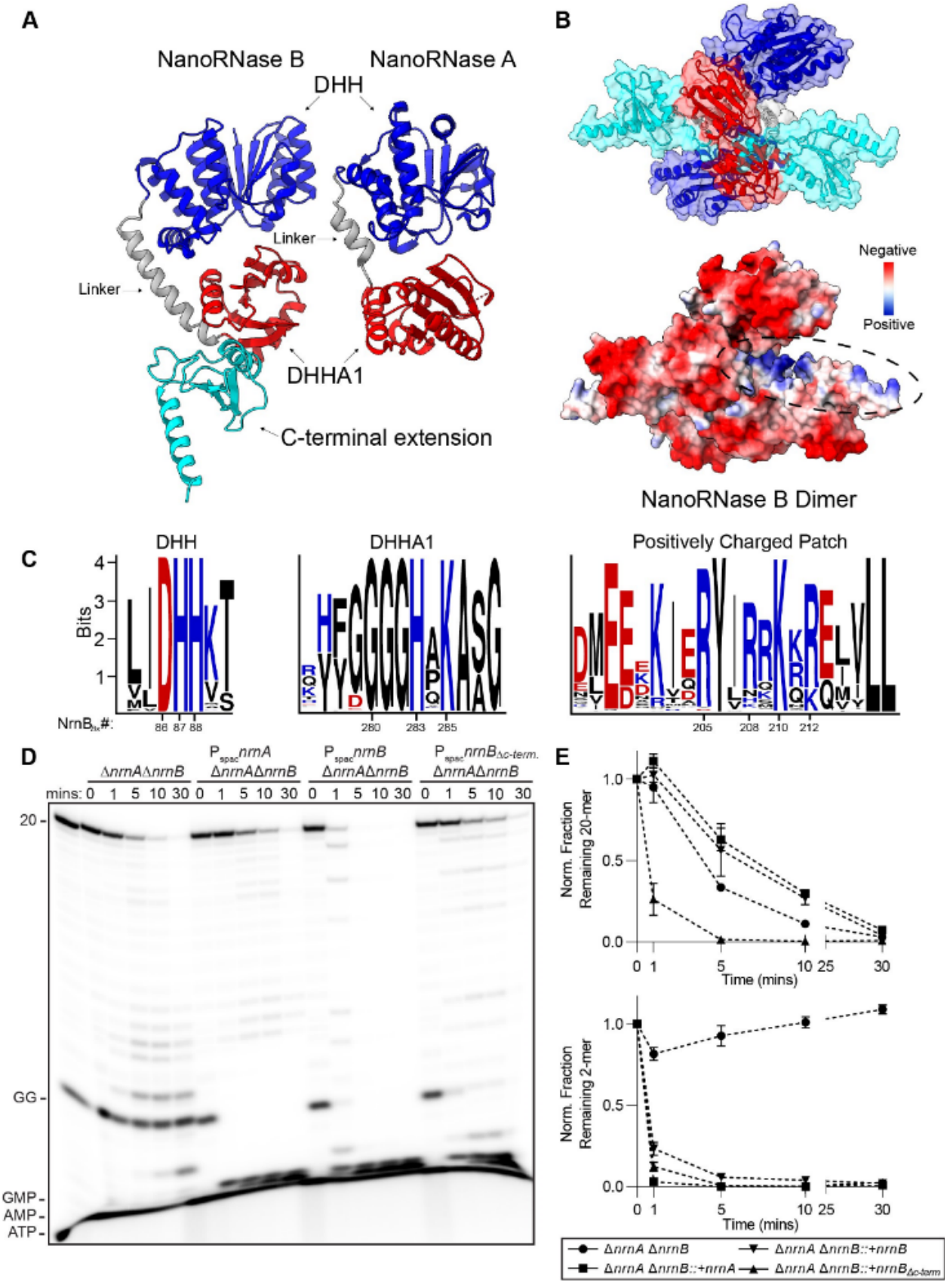
A
NanoRNase B
NanoRNase A
DHH
Linker
Linker
DHHA1
C-terminal extension
B
Negative
Positive
NanoRNase B Dimer
C
DHH
DHHA1
Positively Charged Patch
Bits
NrnB$_{Bs}$#:
D
ΔnrnAΔnrnB
P$_{spac}$nrnA ΔnrnAΔnrnB
P$_{spac}$nrnB ΔnrnAΔnrnB
P$_{spac}$nrnB$_{Δc-term.}$ ΔnrnAΔnrnB
mins: 0 1 5 10 30
20
GG
GMP
AMP
ATP
E
Norm. Fraction Remaining 20-mer
Norm. Fraction Remaining 2-mer
Time (mins)
ΔnrnA ΔnrnB
ΔnrnA ΔnrnB::+nrnB
ΔnrnA ΔnrnB::+nrnA
ΔnrnA ΔnrnB::+nrnB$_{Δc-term}$

Figure 5.11. A potential structural feature that may contribute to the longer substrate length permitted by *B. subtilis* NrnB. (A) Ribbon structures are shown for the AlphaFold2 structural prediction of NrnB alongside the previously published crystallographic structure of NrnA PDB 5J21 (Jumper et al., 2021, Mirdita et al., 2022, Schmier et al., 2015). The DHH domain is indicated in blue. The DHHA1 domain is indicated in red and a short C-terminal extension that is present in *B. subtilis* NrnB is shown in cyan. Images were generated using ChimeraX 1.4 (Pettersen *et al.,* 2021). (B) AlphaFold2 was also used to predict the structure of dimeric NrnB. An electrostatic surface potential map of the latter revealed a strong electropositive patch that overlaps with the NrnB C-terminal extension. (C) We prepared a multiple sequence alignment of 430 NrnB sequences found in the phylum *Firmicutes*, as collected from Uniprot. These protein sequences ranged from 375-450 amino acids in length and did not contain any additional detectable domains (other than DHH and DHHA1). This sequence alignment was used to generate sequence logos using WebLogo (Crooks *et al.,* 2004). This analysis suggested that the conserved residues for the DHH domain (DHH) and the DHHA1 domain (GGGH) are conserved, as is a highly conserved positive patch (RRKR) that overlaps with the C-terminal extension. The sequence numbering in these images is based on the residue positions found in the *B. subtilis* NrnB sequence. (D) Cellular lysates for Δ*nrnA*Δ*nrnB*, or Δ*nrnA*Δ*nrnB* that had been complemented with an IPTG inducible copy of *nrnA, nrnB,* or $nrnB_{\Delta c\text{-}term}$ were harvested during vegetative growth (OD_{600} = 1.0). A mixture of trace amounts of 5'-^{32}P-radiolabeled 20-mer and 2-mer RNA were added to the lysates in in a buffer comprised of 25 mM Tris-HCl, pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, and 50 µM $MnCl_2$. Aliquots were quenched in 4 M urea and 150 mM EDTA and resolved by 20% urea-denaturing PAGE. (E) The normalized radioactive intensity of the initial substrate depletion over time is shown plotted as the average and SD of 3 independent experiments.

Previously for NrnB$_{Bs}$, we instructed Alpha Fold to generate a dimeric model for NrnB$_{Bs}$ based on previous reported claim that other members of the DHH-DHHA1 protein family form dimers (Drexler *et al.,* 2017). In the dimeric structure, the C-terminal extension of NrnB$_{Bs}$ appeared to provide a reasonable scaffolding region for guiding RNA substrates into the putative active site region (Figure 5.11B).

Given the potential role of the C-terminal extension in ligand binding, we generated an surface electrostatic potential model of this assembly to look for regions of charged residues that could be implicated in ligand binding. This revealed a highly basic region consisting mostly of arginine and lysine residues that extended out of the potential active site and across the C-terminal extension. To identify whether the positive patch leading into the predicted active sight of NrnB$_{Bs}$ is conserved amongst other NrnB sequences, we generated a consensus sequence using 430 predicted NrnB sequences from the phylum *Firmicutes* that were constrained to a sequence length of 375-450 amino acid residues (Figure 5.11C) (Crooks *et al.,* 2004). From this sequence analysis we identified the two totally conserved hallmark motifs for the DHH domain (DHH) and the DHHA1 domain (GGGH), giving us confidence that these are NrnB proteins, a subset of DHH-DHHA1 family phosphodiesterase. Interestingly, we also identified that the positive patch seen in the NrnB$_{Bs}$ Alpha Fold structure is mostly conserved to generate a consensus pattern (RYIRRKXR), corresponding with residue positions 205-212.

To further investigate the functional significance of the C-terminal extension found in NrnB$_{Bs}$, we generated total cell lysates from various mutant strains of *B. subtilis* including Δ*nrnA*Δ*nrnB* or complementation strains that harbored ectopic IPTG-inducible copies of *nrnA*, *nrnB*, or *nrnB*$_{\Delta c\text{-}term.}$ which were harvested from late-exponential phase. The *nrnB*$_{\Delta c\text{-}term.}$ mutant was constructed by depleting residues 307-391 which correspond with the C-terminal extension

while preserving the entire DHH and DHHA1 domains. The cell lysates were simultaneously incubated with trace amounts (40 nM) of 5' radiolabeled 2-mer and 20-mer and the degradation products were resolved by 20% denaturing PAGE (Figure 5.11D**)**. Quantification of the initial substrate depletion showed that the Δ*nrnA*Δ*nrnB* lysate degraded the 20-mer to near completion by 10-minutes while the 2-mer went virtually unprocessed (Figure 5.11D, E). We speculate that the slight increase in 2-mer abundance at 30 minutes could have arisen due to the accumulation of terminal di- or trinucleotides that were liberated in the processing of the 20-mer, possibly by RNase R in or PNPase that could be present in the cellular extract. In agreement with our prior investigations of $NrnA_{Bs}$, when we forcibly induced expression of *nrnA* in the Δ*nrnA*Δ*nrnB* background we observed that the 2-mer was rapidly degraded by 1-minute, while the 20-mer was processed to near completion by 10 minutes at a rate amenable to that of the Δ*nrnA*Δ*nrnB* lysate. In agreement with our data in (Figure 5.7A) we found that when we forcibly overexpressed *nrnB* in the Δ*nrnA*Δ*nrnB* background the 2-mer and 20-mer were both degraded to near completion by the 1-minute timepoint. However, when we overexpressed $nrnB_{\Delta c\text{-}term.}$, the 2-mer was processed at a similar rate as the other strains, yet the 20-mer substrate was degraded more slowly than the lysate containing full-length *nrnB*. These data suggest that the C-terminal extension of $NrnB_{Bs}$ aids in the recognition of long RNA substrates but is not required in the degradation of short RNAs. Although further structure and function studies are needed, we speculate from these data that one might be able to discern NrnB-like proteins from NrnA-like proteins based on their linker length, the presence of conserved sequence motifs, and the presence of an additional C-terminal domain. This data in total indicates that $NrnB_{Bs}$ functions as a unique exoribonuclease that can recognize short and long RNA substrates.

5.4 Discussion

In a previous publication, we conducted a pairwise sequence clustering analysis of members of the DHH-DHHA1 protein family contained in the Firmicutes or Actinobacterial phyla (Weiss *et al.,* 2022). For this analysis, we used only the DHH and the DHHA1 domains, excluding any other subdomains or sequences found in other members of this protein family. This analysis revealed distinct clusters of DHH-DHHA1 domain-containing proteins within the phylum *Firmicutes*. These clusters separated into classes of proteins that featured distinct functional specialties. For example, GdpP is a DHH-DHHA1 protein family member that also contains a PAS domain as well as a degenerate GGDEF domain. We found that based on just the DHH-DHHA1 domain, GdpP proteins clustered tightly together. Similarly, a pairwise sequence clustering analysis of the DHH-DHHA1 proteins contained within the phylum Actinobacteria revealed that the proteins all formed one distinct cluster (Weiss *et al.,* 2022). Biochemical analysis of several of these Actinobacteria DHH-DHHA1 proteins revealed that they all possessed specific c-di-AMP hydrolysis activity, suggesting a common function in c-di-AMP homeostasis. However, protein sequences corresponding to NrnA and NrnB did not cluster together, even though they are both standalone DHH-DHHA1 proteins that had been previously considered to be potentially redundant NanoRNases. Instead, the Firmicutes NrnB proteins formed a cluster separate from Firmicutes NrnA proteins. This led us to hypothesize that, in contrast to expectations, NrnA and NrnB may exhibit functionally distinct properties within their DHH and DHHA1 domains, and these proteins may even perform different functions in cells.

Together, our data revealed striking differences between *B. subtilis* NrnB and NrnA. First, we found that their expression patterns differed dramatically. While $NrnA_{Bs}$ is expressed during most cellular conditions, $NrnB_{Bs}$ is specifically expressed within the forespore during

endospore formation. Second, we found that NrnA$_{Bs}$ and NrnB$_{Bs}$ showed important differences in their enzymatic activities. We purified NrnB$_{Bs}$ and NrnA$_{Bs}$ and directly compared their enzymatic preferences in vitro. This revealed that, despite their similar domain structures, NrnA$_{Bs}$ and NrnB$_{Bs}$ act at different termini of their RNA substrates. It was previously shown that NrnA$_{Bs}$ acts at the 5' terminus of short RNA substrates, acting as a 5'-3' exoribonuclease (Weiss *et al.,* 2022). However, in the current study, we found that NrnB$_{Bs}$ is fundamentally different in that it processes its RNA substrates from their 3' terminus, thereby acting as a 3'-5' exoribonuclease. Moreover, while NrnA$_{Bs}$ acts only on short RNAs, from 2-4 nucleotides in length, NrnB$_{Bs}$ exhibited a surprisingly broad range of RNA substrates. Purified NrnB$_{Bs}$ protein acted against short (<5 nucleotides) and long RNA substrates, fully processing them to nucleoside monophosphates. Furthermore, our data revealed that purified NrnB$_{Bs}$ degrades long RNA substrates at a rate resembling *E. coli* RNase R. Moreover, we found a single nucleotide mutation contained within the *nrnB* promoter region that led to expression of *nrnB* during vegetative growth for a *B. subtilis* strain depleted for many exoribonuclease genes. This suggests that NrnB$_{Bs}$ plays an expanded role in RNA turnover that could help alleviate the loss of multiple other exoribonucleases. These data were further bolstered by *ex vivo* experimentation, where we examined the cleavage profiles of short and long RNA substrates that had been simultaneously added to cellular extracts of *nrnA nrnB* double mutant strains. These experiments revealed that NrnA$_{Bs}$ was capable of rapidly processing dinucleotides but exhibited no activity against a 20 mer RNA. In contrast, addition of NrnB$_{Bs}$ resulted in rapid processing of both dinucleotides and the 20 mer RNA. Moreover, when a short C-terminal extension was deleted from the NrnB$_{Bs}$ sequence, the mutant protein retained the ability to process a dinucleotide but exhibited a significant decrease in activity against the 20 mer, suggesting that the C-terminal extension is

likely to play an important role in the protein's substrate specificity. From our data together, we conclude that NrnB$_{Bs}$ unexpectedly functions as a general 3'-5' exoribonuclease *in vitro* and *in vivo*. Our data show that NrnB$_{Bs}$ is not only capable of processing long RNAs as a general 3'-5 exoribonuclease but can also perform 'nanoRNase' activity by processing short RNAs. This suggests that some bacteria may be able to bypass the need for a dedicated NanoRNase enzyme by utilizing an NrnB RNase. Overall, we hypothesize that NrnB proteins differ significantly from NrnA and are likely to have evolved for distinct cellular purposes.

NrnA and NrnB have both been previously reported to act as nanoRNases, enzymes that are specifically active against short RNAs, given that they can complement the growth defect caused by depletion of Orn (Mechold *et al.,* 2007; Fang *et al.,* 2009). Indeed, the name given to NrnA and NrnB derives from this 'nanoRNase' complementation assay. However, while Orn and NrnC appear to act as dinucleases, and the NrnA acts as a genuine nanoRNase, the discovery that NrnB cleaves longer RNAs calls into question whether it is truly a professional nanoRNase, or if instead it plays a different, and perhaps more expanded, role in RNA degradation. Based on the observation that NrnB acts on longer substrates *in vitro,* one would predict that either: (1) NrnB cleaves a broad range of substrates *in vitro*, but still acts as a nanoRNase *in vivo*, or (2) NrnB acts as a general exonuclease *in vivo*, uniquely capable of processing longer RNAs down to nucleoside monophosphates. The latter would differ from what has been reported for most bacterial 3'-5' exoribonucleases, such as PNPase and RNase R, which efficiently process longer RNAs but are unable to process the final oligonucleotides from 2-4 nucleotides in length, thereby requiring dedicated nanoRNases for resolving a bottleneck in degradation of short RNAs (Lormand *et al.,* 2021). Yet our evidence supports this hypothesis, that NrnB acts as both a general exoribonuclease and a nanoRNase *in vivo*. In *ex vivo* experiments, where radiolabeled

RNA substrates were added to cellular lysates from different bacterial strains, the expression of NrnB correlated with the complete processing of longer RNAs, including fully processing them to nucleoside monophosphates. This contrasted when RNase R was expressed within lysates, which resulted in processing of longer RNAs but also accumulation of 2-mers and 3-mers. From this, we can conclude that NrnB was indeed able to process short RNAs within cellular-like conditions. We also show that the C-terminal extension is a functionally relevant substructure that aids in the recognition and degradation of long RNA substrates, but not in the processing of diribonucleotides. However, there are still many outstanding questions. For example, why is NrnB specifically expressed within the forespore and mature spore? Does the fact that NrnB appeared to compensate for the loss of other RNases in the multiple-RNase mutant strain suggest that NrnB has a role in RNA turnover in the forespore? What role does it play beyond the housekeeping RNases? Similarly, what are the structural determinants that switch termini selectivity and govern substrate preferences? The answer to the latter question is critical for improving the currently murky annotation of NrnA and NrnB-like proteins.

While there have been multiple biochemical analyses of proteins that functionally resemble *B. subtilis* NrnA or *M. tuberculosis* CnpB, NrnB candidates have been significantly understudied. Given the functional diversity of the DHH-DHHA1 protein family, it remains a possibility that many of the putative NrnB-like proteins could be misannotated. However, it is also possible that there may be subclasses of NrnB-like proteins that display differences in their enzymatic capabilities. For example, while many NrnB homologs are found within the phylum *Firmicutes*, there are also proteins annotated as NrnB found within the *Epsilonproteobacteria*. There is one report that suggests that the *Helicobacter pylori* NrnB possesses cAMP phosphodiesterase activity (Choi *et al.,* 2013). Therefore, it remains to be seen whether the NrnB

homologs found within the *Epsilonproteobacteria* Class (or found in any other phyla) are cAMP phosphodiesterases or if they possess RNase activity amenable to *B. subtilis* NrnB. Until other NrnB-like proteins are individually characterized, the data presented in this study are likely to be directly relevant only to *Firmicutes* NrnB-like proteins. In total, current literature indicates that proteins containing standalone DHH-DHHA1 domains can thus far be subclassified based on the substrates that they act upon, including: (1) degradation of short RNAs 2-4 nucleotides in length, (2) specific processing of the dinucleotide pGpG, and (3) specific processing of c-di-AMP (Tang *et al.,* 2015; He *et al.,* 2016; Dey *et al.,* 2017; Weiss *et al.,* 2022; Heo *et al.,* 2022). The current study now adds a new function, that of general 3'-5' exoribonucleolytic activity against both short and long RNAs, which for some bacteria may allow them to bypass the need for an otherwise dedicated nanoRNase enzyme.

5.5 Materials and Methods

5.5.1 Bacterial strains and culture conditions

E. coli strains were grown in LB containing (as needed) 100 µg/mL carbenicillin at 37°C in a shaking incubator. *B. subtilis* strains were grown in LB or 2xYT in the presence of 5 µg/mL chloramphenicol, 100 µg/mL spectinomycin, 1 µg/mL erythromycin and 25 µg/mL lincomycin, or 5 µg/mL kanamycin, as needed. A forespore-specific *yfp* reporter strain was constructed by PCR-amplifying the region corresponding to the *sspB* promoter region and sub-cloning this sequence followed by *yfp* into vector pDG1662, which integrates into the *amyE* gene. Additionally, we created P_{const}-*yfp* reporter strain by sub-cloning a constitutively active promoter sequence upstream of the *yfp* gene into the vector pDG1662, for integration into *amyE.* To create a translational reporter fusion to *nrnB*, we PCR-amplified the region corresponding to the *nrnB*

gene, including the intergenic space between *nrnB* and the upstream *yngJIHGFE* operon, and subcloned *nrnB* so that its C-terminus would be in-frame with either *yfp* or *sfgfp* (encoding superfolder GFP). These sequences were sub-cloned via Gibson assembly into plasmid pDG1662. To construct an inducible *nrnB-yfp* construct, we amplified the *nrnB-yfp* fusion from plasmid pTMM66 and subcloned this sequence via Gibson assembly into the *amyE* integration vector pDR111, which harbors an IPTG-inducible promoter upstream of the target gene. The protocol for making markerless deletion strains of *B. subtilis* Δ*nrnA,* Δ*nrnB,* and Δ*nrnA* Δ*nrnB* was also previously described (Orr *et al.,* 2018). To construct complementation strains of $nrnA_{Bs}$, $nrnB_{Bs}$, $nrnB_{Bs\Delta c\text{-}term.}$, or rnr_{Bs} these sequences were each PCR-amplified from chromosomal DNA preparations. These sequences were then subcloned via Gibson assembly (Gibson *et al.,* 2009) into the plasmid pDR111. All transformations of *B. subtilis* were performed using a previously described protocol (Jamer *et al.,* 2002).

To construct *E. coli* strains for the overexpression and purification of proteins used in this study, $nrnB_{Bs}$ was PCR-amplified from a genomic DNA preparation of *B. subtilis* 168. The $nrnB_{Bs}$ coding sequence was subcloned via Gibson assembly into the IPTG-inducible expression vector for overexpression and purification of a protein containing an N-terminal hexahistidine tag or into a vector containing no tag. The construction of bacterial strains harboring $nrnA_{Bs}$, $cysQ_{Ec}$, $cnpB_{Mt}$ and $disA_{Bt}$, as well as the protocol for purifying the respective protein products of these gene sequences was previously reported (Weiss *et al.,* 2022). *E. coli* XL10-Gold (Agilent) was initially transformed with all plasmids, and the 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.

5.5.2 Fluorescence microscopy and quantification

Single colonies of *B. subtilis* were used to inoculate DIFCO sporulation media and grown at 37˚C with vigorous aeration. Cells undergoing sporulation were washed with 1X phosphate-buffered saline three times, to aid in the removal of cell debris. The cells were then diluted and spotted on 1.5% agarose pads containing 1X phosphate-buffered saline. For an overnight time-lapsed series of *B. subtilis* undergoing sporulation, we spotted the cells that were growing in DIFCO sporulation media onto an agarose pad containing 1.5% agarose made with 1X resuspension buffer (Nicholson and Setlow, 1990). Cell images were taken at room temperature using a Zeiss Axio-Observer Z1 inverted fluorescence equipped with a Rolera EM-C2 electron-multiplying charge-coupled (EMCC) camera.

5.5.3 Sporulation efficiency assay

Measurements of *B. subtilis* sporulation efficiency was conducted as previously published (Mastny *et al*., 2013). *B. subtilis* strains were induced to sporulate by nutrient exhaustion after 24 hr in DIFCO sporulation medium. Sporulation efficiency was determined by serially diluting the sporulated cultures in T-Base medium and dividing the serially diluted cultures into two fractions. One fraction was immediately spread-plated onto rich- medium agar plates, while the second fraction was heated in a water bath at 85°C for 20 minutes prior to plating on rich medium agar plates. After overnight incubation at 37°C, colonies from the heat-treated and non-heat-treated fractions were counted and used to calculate the percentage of sporulation efficiency.

5.5.4 Protein overproduction and purification

See chapter 4.4 for the $NrnB_{Bs}$ purification protocol.

5.5.5 Preparation of whole cell lysates

Cultures of *B. subtilis* strains were grown overnight shaking at 37˚C. The following morning, cells were diluted into fresh LB and grown at 37˚C with shaking to OD_{600} ~ 1.0. IPTG was added to the cultures to a final concentration of 250 μM and cells were cultured for an additional 45 minutes. Cells were harvested by centrifugation and concentrated 10X in 25 mM Tris-HCl, 100 mM NaCl, pH 8.0. Prior to cell lysis, PMSF was added to the resuspension to reach a final concentration of 1 mM. Cells were lysed by sonication, and the lysates were aliquoted and stored at -80˚C. The lysates were diluted from ~10X to 1X in the reactions for testing the RNase activity of the lysates.

5.5.6 Oligoribonucleotide labeling

Synthetic RNAs (2-7-mers) were purchased from TriLink Biotechnologies or Sigma-Aldrich. Synthetic RNAs (10-15-mers) were purchased from Integrated DNA Technologies. The synthetic RNA 22-mer and 44-mer were purchased from Dharmacon. RNAs were subjected to 5' radioactive end labeling or non-radioactive phosphorylation by T4 Polynucleotide Kinase (NEB). Each RNA was phosphorylated using equimolar concentrations of [γ-^{32}P]-ATP or ATP, T4 PNK, and 1X T4 PNK Reaction Buffer. Reactions comprising a final concentration of either 0.5 μM 5'-[^{32}P]-radiolabeled RNA or 2.0 μM phosphorylated RNA were incubated at 37°C for 60 minutes, followed by heat inactivation of T4 PNK at 65°C for 20 minutes.

5.5.7 Synthesis of c-di-AMP

See chapter 3.5.6 for the protocol for the c-di-AMP generation protocol.

5.5.8 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 50 nM of purified $NrnB_{Bs}$ or $CnpB_{Mt}$. For the reactions in (Figure 5.8), 1 μM of phosphorylated 44-mer containing trace amounts of radiolabeled RNA was subjected to cleavage by 100 nM of $NrnB_{Bs}$ or RNase R_{Ec}. Purified RNase R_{Ec} was purchased from BIOSEARCH^{TM} TECHNOLOGIES. These reactions were conducted in the presence of 25 mM Tris, pH 8.0, 300 mM NaCl, and 5 mM $MnCl_2$ for $NrnB_{Bs}$ or 25 mM Tris, pH 8.0, 300 mM NaCl, and 5 mM $MgCl_2$ for RNase R_{Ec}. At the appropriate times, aliquots of the reaction were removed and quenched in the presence of 4 M urea and 150 mM EDTA on ice. For reactions with *B. subtilis* whole cell lysates in (Figure 5.7), 1.0 μM of phosphorylated 22-mer including 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 $MnCl_2$, and 25 mM $MgCl_2$. The lysate reactions in (Figure 5.11) were conducted in the presence of 25 mM Tris-HCl, pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, and 50 μM $MnCl_2$, against trace amounts of 5' -radiolabeled 20-mer and the diribonucleotide pGpG. At the indicated time points, aliquots of the reaction were removed and quenched in the presence of 4M urea and 150 mM EDTA on ice. All cleavage reactions were separated on denaturing 20% PAGE containing 1X TBE and 4 M urea. The gels were imaged using Cytiva Amersham Typhoon^{TM} laser scanner platform. The intensity of the radiolabeled nucleotides was quantified using FIJI software (Schindelin et al., 2012).

5.5.9 2-aminopurine hydrolysis assay

To determine the polarity of $NrnB_{Bs}$ we utilized a fluorescence-based assay as described previously (Zhou *et al.,* 2017). This assay is based on the differential fluorescence output of the nucleotide analog 2-aminopurine. 2-Aminopurine exhibits reduced fluorescence output when

base-stacked with other nucleobases in a ssRNA molecule, as compared to free 2-aminopurine, which exhibits increased fluorescence. We monitored the fluorescence output of 2-aminopurine generated from the phosphodiester hydrolysis of a synthetic 4-mer RNA substrate that was purchased from GE Healthcare Dharmacon. For this analysis, 100 nM $NrnB_{Bs}$ was incubated with 10 μM of the RNAs containing a 2-aminopurine and a specific internal phosphorothioate modification (as described in Results section). These reactions also contained 25 mM Tris-HCl, pH 8.0, 300 mM NaCl, and 200 μM $MnCl_2$. Reactions were conducted in a black 384-well plate using a Spectramax M5 plate reader and fluorescence was measured every two minutes using an excitation wavelength of 310 nm and an emission wavelength of 375 nm.

5.5.10 pAp and cAMP phosphodiesterase assays

$NrnB_{Bs}$ was 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 (Hatzio *et al.,* 2008). A standard curve was generated using 1 mM phosphate. All reactions were composed of 25 mM Tris-HCl pH 8.0, 300 mM NaCl, 100 μM pAp (Sigma), and either 5 mM of $MnCl_2$ (for $NrnB_{Bs}$), or 1 mM $MgCl_2$ (for $CysQ_{Ec}$), and conducted at room temperature. For experiments testing cAMP phosphodiesterase activity, we modified the assay to include 100 μM cAMP and 10 U of alkaline phosphatase (NEB). For each reaction 5 nM each protein was added to its respective reaction and allowed to incubate for the indicated reaction time of 30 minutes. Reactions were then quenched with ¼ volume of acidic malachite green dye solution and incubated for 30 minutes to allow for color development. Absorbance values at 620 nm (A_{620}) 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.

Chapter 6: Elucidating the biochemical features of NanoRNase B protein homologs

6.1 Introduction

RNA degradation is a crucially important intracellular process that occurs in all living organisms. NanoRNAs are short RNA molecules that are ~2-6 nucleotides in length are generated through the processing of long RNA polymers (Hui *et al.,* 2014). Interestingly, a breadth of current and historical analyses of the RNases that participate in degrading long RNA polymers such as mRNA do not efficiently recognize and degrade nanoRNAs (Lee *et al.,* 2022). In fact, cells harbor specialized enzymes referred collectively to as nanoRNases to degrade these terminal breakdown products. Also, a very recent phylogenetic analysis of the nanoRNase protein families has provided evidence that all organisms harbor at least one of these specialized enzymes (Figure 1.4) (Lormand *et al.,* 2021). Furthermore, genetic experiments in many different bacterial organisms have shown that nanoRNase encoding genes are frequently encoded by essential genes or that deletion of a nanoRNase encoding gene will result in deleterious effects (Lee *et al.,* 2022). NanoRNA accumulation has been shown to interfere with second messenger signaling pathways, transcriptional priming, and feedback inhibition of general exoribonucleases (Lee *et al.,* 2022). Altogether, the turnover of nanoRNAs represents a critical point in both RNA degradation and second messenger signaling.

Two broadly distributed nanoRNases found in most taxonomic groups of bacterial organisms are the proteins NanoRNase A (NrnA) and NanoRNase B (NrnB) (Lormand *et al.,* 2022). NrnA and NrnB proteins are very similar proteins in that they are members of the DHH protein family (Aravind and Koonin, 1998) and are comprised of two lobed structures that harbors a DHH domain connected by a linker to a DHHA1 (DHH associated domain 1).NrnA

and NrnB were first discovered in *B. subtilis,* and the initial biochemical analyses suggested that these enzymes were essentially redundant in their enzymatic activities and intracellular functions (Mechold *et al.,* 2007; Fang *et al.,* 2009). Nevertheless, our recent investigations of $NrnA_{Bs}$ and $NrnB_{Bs}$ have indicated that there were previously unrecognized differences in substrate accommodation and recognition amongst the DHH-DHHA1 protein family. We found that $NrnA_{Bs}$ specifically degrades RNAs 2-4 nucleotides in length by using a 5'-3' exonucleolytic mechanism during vegetative growth in *B. subtilis* (Weiss *et al.,* 2022). Contrasting with $NrnA_{Bs}$, we identified that $NrnB_{Bs}$ does not exhibit profound substrate length discrimination and this protein degrades RNAs of varying lengths utilizing a 3'-5' exonucleolytic mechanisms during spore development in *B. subtilis* (Myers *et al.,* 2023). Furthermore, we found that the C-terminal extension is a feature contained within $NrnB_{Bs}$ that contributes to the capacity of this protein to degrade long RNAs (Myers *et al.,* 2023). While the C-terminal extension is found in many of the Firmicutes NrnB homologs, it is conspicuously absent from a subset of putative protein homologs. Moreover, putative NrnB proteins found in the Epsilonproteobacterial phyla do not seem to possess the C-terminal extension found in $NrnB_{Bs}$. These observations have sparked many questions guiding our analysis such as 1. Does the C-terminal extension represent a global diagnostic criterion for identifying NrnB proteins? 2. Will NrnB proteins that do not contain C-terminal extensions be restricted to processing short RNAs? 3. Are the putative NrnB proteins found in the phylum Epsilonproteobacteria functionally redundant with *B. subtilis* NrnB? These are the questions that we have sought to answer in this biochemical survey of putative NrnB protein homologs.

6.2.1 Alphafold structural predications of putative NrnB homologs

Previously we found that the C-terminal extension found in *B. subtilis* NrnB likely plays a key role in long RNA substrate accommodation (Myers *et al.,* 2023). With this information, we specifically searched to identify whether other putative NrnB-like proteins exhibited different amino acid sequence lengths. To our surprise, we found that NrnB-like proteins from the Bacilli class of organisms are 380-410 amino acids in length, while there exists a subset of NrnB-like proteins that are ~350 amino acids in length. Additionally, there are putative NrnB-like proteins that are encoded by the Epsilonproteobacterial class of organisms. These proteins generally are ~350 amino acids in length as well. We hypothesized that these different sequence lengths could indicate the presence or absence of the C-terminal extension found in the *B. subtilis* NrnB. The Alphafold structural prediction for the *B. licheniformis* NrnB is unsurprisingly similar to that of the *B. subtilis* NrnB; likely due to these proteins sharing ~60% sequence identity (Figure 6.1A). However, the structural prediction for the NrnB protein encoded by *B. thuringiensis* shows a protein that seems to possess all of the same structural features as the *B. subtilis* and *B. licheniformis* NrnB's but is missing the C-terminal extension. To our surprise the structural prediction for the *Helicobacter pylori* NrnB was dramatically different to the other Firmicutes NrnB proteins. The key difference between the *H. pylori* NrnB and the *B. licheniformis* or *B. thuringiensis* NrnB proteins is that composition of the linker region is markedly different (Figure 6.1A). Instead of being one contiguous alpha-helix as seen in the *B. licheniformis* or *B. thuringiensis* proteins, the helix is broken by disordered regions and extends inward into the putative ligand binding site between the DHH and DHHA1 domains. While there are differences in the overall protein structures amongst these different NrnB-like homologs, an overlay of the

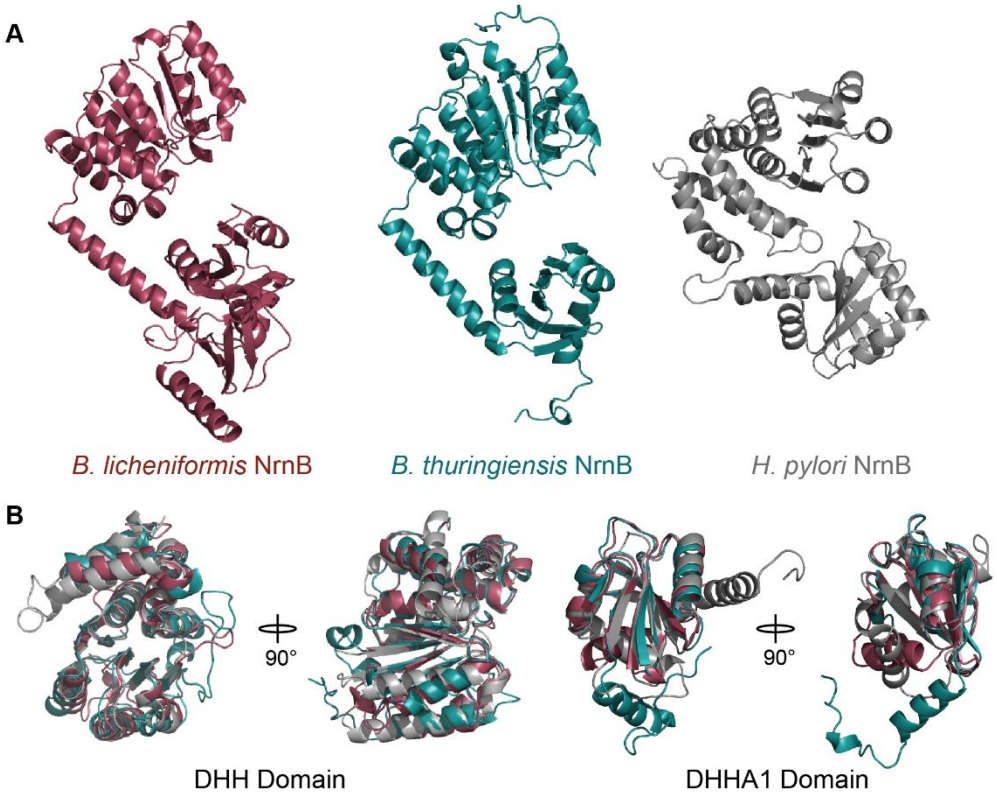

Figure 6.1. Alphafold 2 structural predictions of putative NrnB homologs. (A) Monomeric Alphafold 2 structural predictions for putative NrnB-like protein homologs that exhibit different structural elements using PyMOL (Jumper *et al.,* 2021; Mirdita *et al.,*2022; Schrödinger and DeLano, 2020). (B) An overlay of the DHH and the DHHA1 domain indicates that the predicted global folds are similar amongst these protein homologs.

DHH or the DHHA1 domains are strikingly similar in their global folds (Figure 6.1B) Given the predicted differences in these putative NrnB homologs, we felt that these were perfect candidates to biochemically analyze for RNase activity.

6.2.2 NanoRNase B proteins from *B. licheniformis, B. thuringiensis,* and *H. pylori* degrade short and long RNA substrates.

In our prior analyses we found that there are key differences in substrate preferences amongst the DHH-DHHA1 protein family. Previously, we found that *B. subtilis* NrnA and NrnB exhibit differences in substrate length preferences in *ex vivo* and *in vitro* assays. $NrnA_{Bs}$ exclusively degrades short RNAs 2-4 nucleotides in length, while $NrnB_{Bs}$ is unique in that this enzyme exhibits dual functions as a nanoRNase that can degrade short RNAs in addition to general exoribonuclease activity against long RNA substrates 20-mer and greater. Also, by utilizing Alphafold 2 structural prediction modeling, we identified that there are two potential structural features contained within $NrnB_{Bs}$ that could confer this protein with nano- and general exoribonuclease activity. One such feature that we speculate enhances the general exoribonuclease activity of $NrnB_{Bs}$ is the long linker region connecting the DHH and DHHA1 domains relative to crystallographic structures of $NrnA_{Bs}$. The C-terminal extension of $NrnB_{Bs}$ is a small structural element that aids in the recognition and degradation of longer RNA substrates. Based on this knowledge, we conducted a small biochemical survey in which we interrogated the capacity of purified NrnB-like proteins homologs that were predicted to have differences in their linker regions or C-termini to degrade RNAs (Figure 6.1). We chose to purify $NrnB_{Bl}$ based on the presence of a long linker region as well as the presence of the C-terminal extension. Interestingly, upon looking for sequences for NrnB-like protein homologs, we identified that the C-terminal extension is absent in a subset of NrnB-like protein sequences and Alphafold 2

structure predictions. Of these proteins, we purified the plasmid encoded NrnB-like protein candidate from *B. thuringiensis,* which is predicted to harbor a long linker region, but does not have a C-terminal extension. Furthermore, we identified that there are NrnB-like proteins contained in the phylum Epsilonproteobacteria that are predicted to have a very different linker-region as well as no C-terminal extension. While NrnB-like proteins from various Firmicutes organisms are predicted to harbor a long alpha helical linker connecting the DHH and DHHA1 domains, the linker predicted for the *H. pylori* NrnB protein is of an alpha helical bundle. With Alphafold 2 structural predictions for these proteins, we hypothesized that the *B. licheniformis* NrnB would exhibit RNase activity consistent with that which was previously reported for *B. subtilis* NrnB. Additionally, we thought that the *B. thuringiensis* and *H. pylori* NrnB proteins would exhibit an NrnA-like short RNA length preference.

For this analysis we conducted a time-course RNase assay in which we simultaneously incubated trace amounts of 5'-^{32}P-radiolabeled RNAs 2-7, 10, and 20 nucleotides in length with 100 nM of purified Nrn_{Bl}, $NrnB_{Bt}$, $NrnB_{Hp}$, and RNase R_{Ec}. Overall, the substrate degradation profiles for the various NrnB-like protein homologs look very similar, as all the RNA substrates are mostly degraded to monoribonucleotides by 30 minutes (Figure 6.2). However, one key difference is that while $NrnB_{Bt}$ degrades RNAs 2-7, and 10 nucleotides in length at a similar rate to that of $NrnB_{Bl}$ and $NrnB_{Hp}$, the degradation profile of the 20-mer is sufficiently different between these three proteins. The $NrnB_{Bl}$ and $NrnB_{Hp}$ proteins degrade the 20-mer RNA to near completion by the 10-minute timepoint, while the $NrnB_{Bt}$ protein displays clear degradation product intermediates remaining at the 10-minute and 30-minute timepoint. This differential activity could indicate that the $NrnB_{Bt}$ protein acts in a more distributive fashion against longer RNA substrates than that of the $NrnB_{Bl}$ and $NrnB_{Hp}$ proteins. Yet another small difference in

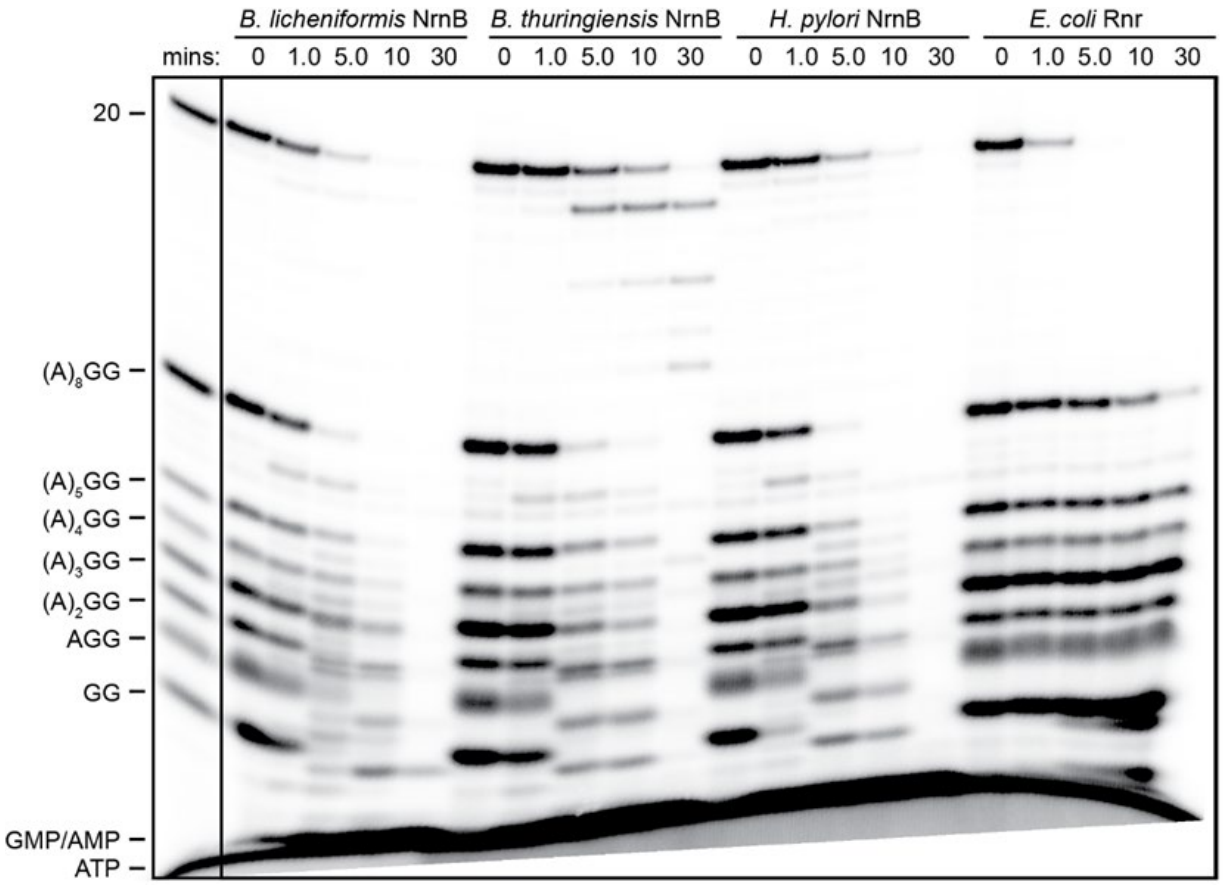

Figure 6.2. NanoRNase B proteins from *B. licheniformis, B. thuringiensis,* and *H .pylori* degrade short and long RNA substrates. Trace amounts of 5' ^{32}P-radiolabeled RNA molecules that were 2-7, 10, and 20 nucleotides in length were mixed and simultaneously incubated with 100 nM of purified $NrnB_{Bl}$, $NrnB_{Bt}$, $NrnB_{Hp}$, or Rnr_{Ec}. Aliquots were removed from reactions and quenched in 150 mM EDTA and 4 M urea. Degradation products were resolved by 20% denaturing PAGE.

protein activities as seen by this analysis is that NrnB$_{Hp}$ processes the diribonucleotide GG to near completion by 1 minute. This may indicate that the NrnB$_{Hp}$ might exhibit a diribonucleotide substrate preference. Another striking result observed in this analysis is that the bona-fide general 3'-5' exoribonuclease RNase R$_{Ec}$ degraded the long 20-mer substrate to near completion by 5 minutes, and the 10-mer substrate after 30 minutes. The other short RNA substrates 2-7 nucleotides in length were not appreciably acted upon, corroborating previous analyses that strongly suggest that general exoribonuclease exhibit long RNA substrate preferences.

6.2.3 *H. pylori* NrnB degrades dinucleotides at low protein concentrations

Since our previous analysis in (Figure 6.2) only utilized ~40 nM of radiolabeled RNA substrates with the protein concentration being at 100 nM or >2-fold excess of substrate, we decided to take a closer look at the protein substrate ratios required for long and short RNA processing by NrnB$_{Hp}$. In this analysis, we observed that 5 nM NrnB$_{Hp}$ processed 1 μM of the diribonucleotide pGpG containing trace amounts of 5'-^{32}P-radiolabeled RNA by 30 minutes (Figure 6.3A, C). When the concentration of NrnB$_{Hp}$ was increased 10-fold to 50 nM, pGpG was completely processed by the 20-minute time point. Additionally, another 10-fold increase in NrnB$_{Hp}$ lead to even faster pGpG processing as the dinucleotide was processed to near completion by the 2.5-minute timepoint. When we assessed the degradation of the 20-mer by NrnB$_{Hp}$ we observed virtually no processing at 5 nM of protein by the 30-minute timepoint (Figure 6.3B, C). Since NrnB$_{Hp}$ could efficiently process a diribonucleotide at this protein to substrate ratio (Figure 6.3A) we speculate that NrnB$_{Hp}$ exhibits a greater specificity for diribonucleotides as compared to longer RNAs. When we increased the concentration of NrnB$_{Hp}$ to 50 nM we observed moderate 30% processing of the 20-mer substrate by the 30-minute time point (Figure 6.3B, C). However, when we incubated 500 nM of NrnB$_{Hp}$ with the 20-mer, we

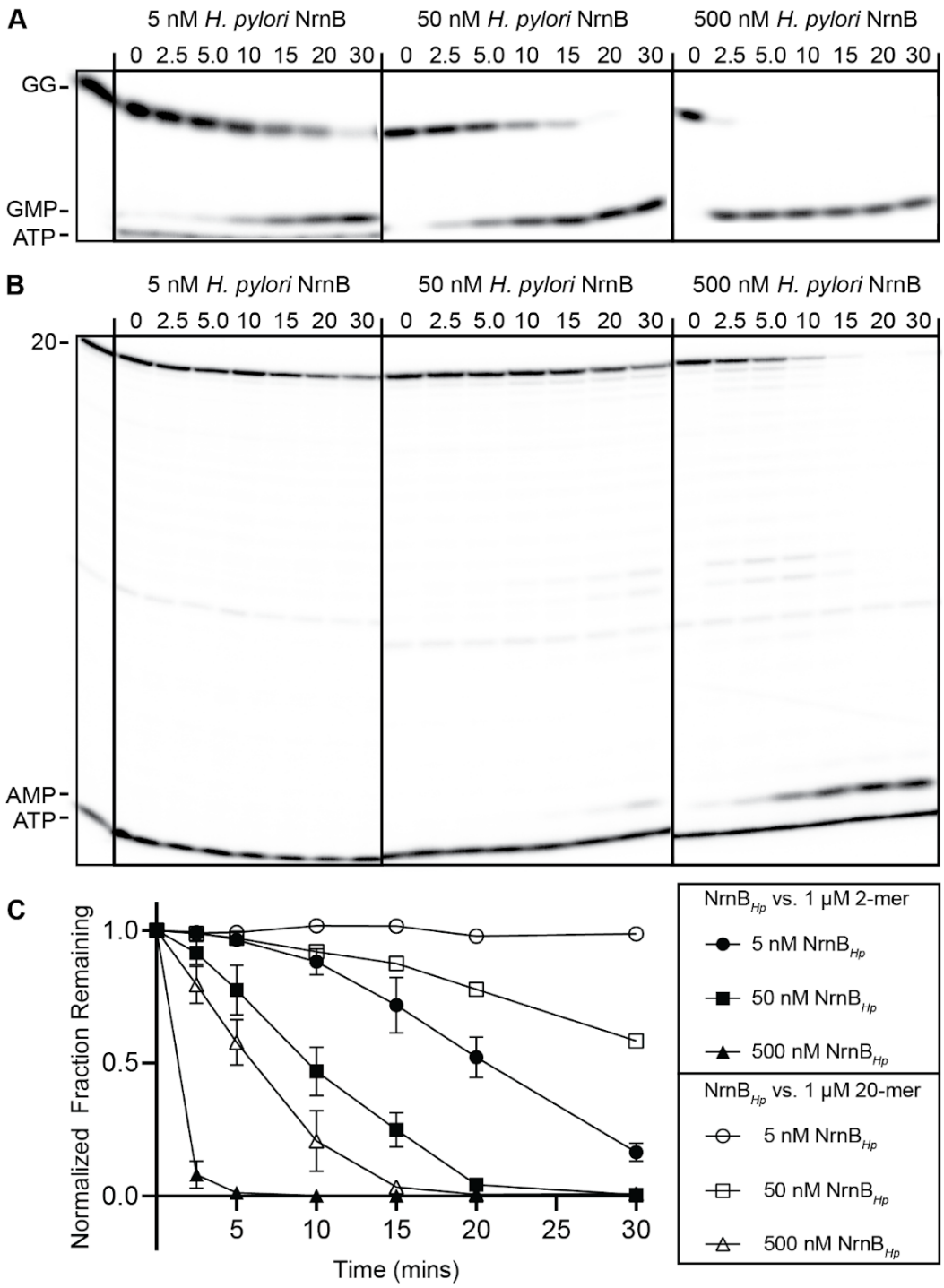

Figure 6.3. *H. pylori* NrnB degrades dinucleotides at low protein concentrations. (A - B). 1 μM of RNA molecules 2 or 20 nucleotides in length containing a trace amount of 5'-^{32}P-radiolabeled RNA were incubated with purified NrnB$_{Hp}$ at enzyme concentrations ranging from 5 nM to 500 nM. Samples were removed at time intervals and analyzed by urea-denaturing 20% PAGE. (C) Quantification of the normalized radioactive intensity of the initial substrate depletion over time plotted as the average and SD of 3 independent experiments in (A and B). Aliquots were removed from reactions and quenched in 150 mM EDTA and 4 M urea. Degradation products were resolved by 20% denaturing PAGE

observed that the 20-mer was processed to near completion by 15 minutes. Taken altogether, we speculate that NrnB$_{Hp}$ might possess a short RNA substrate length preference.

6.2.4 *H. pylori* NrnB preferentially degrades short RNAs 2-4 nucleotides in length

Building off our previous analysis, we wanted to further interrogate the substrate length preferences of the NrnB-like protein encoded by *H. pylori.* For this set of experiments, we individually incubated RNAs of varying lengths with purified NrnB$_{Hp}$. In this analysis we subjected 1 μM of RNAs 2-7, 10, or 15 nucleotides in length containing trace amounts of 5'-^{32}P-radiolabled tracer RNA to cleavage by 50 nM NrnB$_{Hp}$. We observed by denaturing PAGE that the 2-mer and 3-mer substrate were processed to mononucleotides by the 15-minute timepoint (Figure 6.4A, B). We also found that NrnB$_{Hp}$ degraded the 4-mer RNA to near completion ~80% by 15 minutes. The 5-mer substrate was only degraded to ~50% completion by 60 minutes with visible intermediates corresponding to RNAs 2-4 mer in length present in the gel. There was only a very small amount of activity of NrnB$_{Hp}$ against the RNAs 6, 7, 10, and 15 nucleotides in length. In fact, only ~ 20% of the initial substrate corresponding to 6, 7, 10, or 15 nucleotides was processed after 60 minutes (Figure 6.4A, B). Taken together these data strongly suggest that NrnB$_{Hp}$ exhibits an NrnA-like preference for short RNAs 2-4 nucleotides in length, while displaying the capacity to act as a general exoribonuclease akin to NrnB$_{Bs}$.

6.2.5 *B. thuringiensis* NrnB preferentially degrades RNAs 2-4 nucleotides in length

In order to further understand the substrate preferences of the NrnB-like protein homologs, we decided to take a deeper look at the capacity of NrnB$_{Bt}$ to degrade RNA substrates ranging in different lengths. For this set of experiments, we individually incubated RNAs of varying lengths with purified NrnB$_{Bt}$. In this analysis we subjected 1 μM of RNAs 2-7, 10, or 15

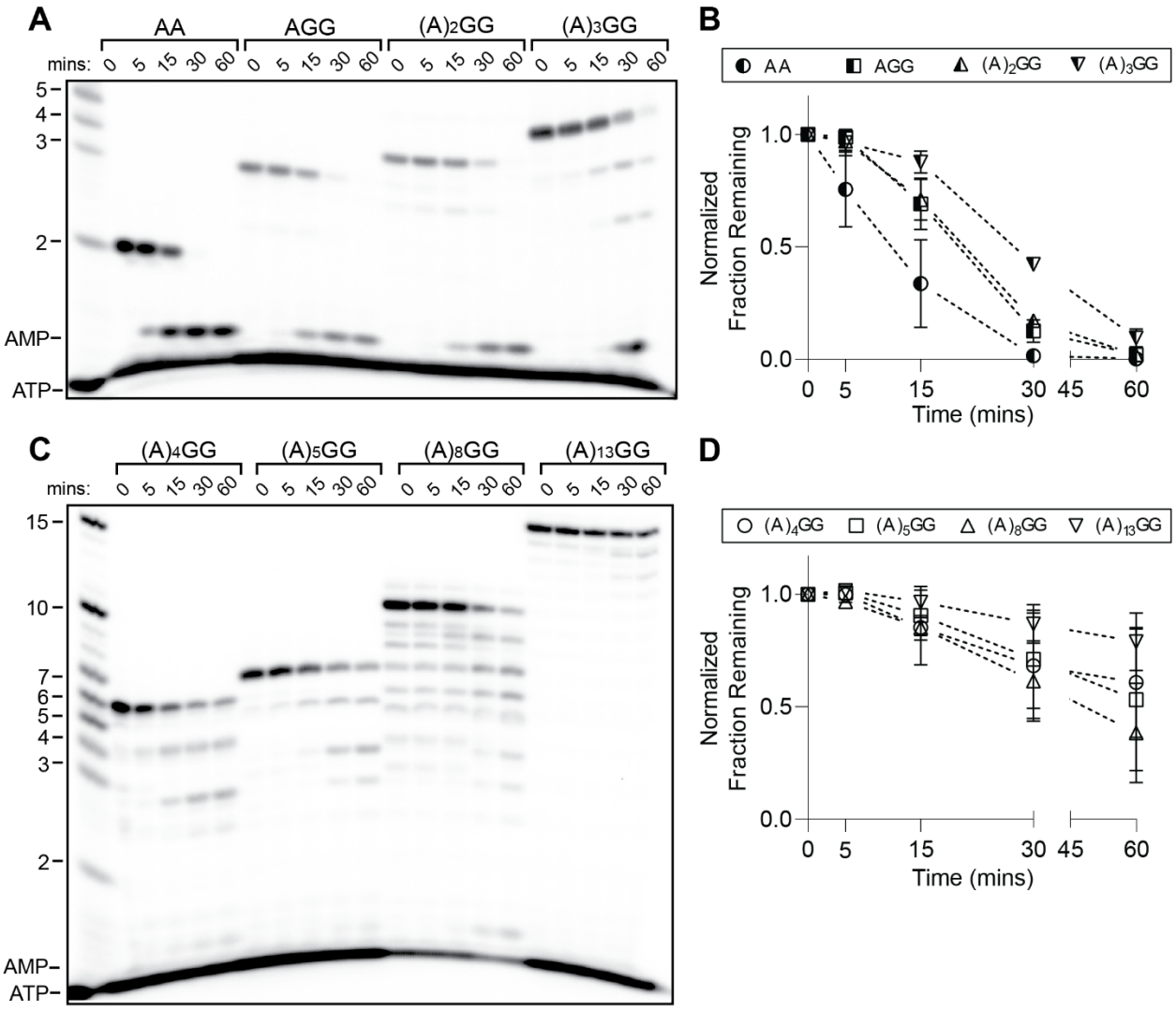

Figure 6.4. *H. pylori* NrnB preferentially hydrolyzes short RNAs 2-4 nucleotides in length. (A and C). RNA molecules 2-7, 10, and 15 nucleotides in length were incubated at a final concentration of 1 µM with 50 nM of purified $NrnB_{Hp}$. The reactions also contained a trace amount of 5'-^{32}P-radiolabeled RNA. Samples were removed at time intervals and analyzed by urea-denaturing 20% PAGE. (B and D) Quantification of the normalized radioactive intensity of the initial substrate depletion over time plotted as the average and SD of 3 independent experiments in (A and C). Aliquots were removed from reactions and quenched in 150 mM EDTA and 4 M urea. Degradation products were resolved by 20% denaturing PAGE

nucleotides in length containing trace amounts of 5'-^{32}P-radiolabled tracer RNA to cleavage by 50 nM NrnB$_{Bt}$. We observed by denaturing PAGE that the 2-mer, 3-mer, and 4-mer substrates were processed to mononucleotides by the 15-minute time point. Interestingly, the 5-mer substrate was only partially degraded by the 60-minute time point, in which there are still multiple cleavage product intermediates visible. We also found that NrnB$_{Bt}$ processed the 6-mer, 7-mer, and 10-mer RNA to near completion ~80% in 60 minutes. There was only a very moderate amount of 15-mer processing by NrnB$_{Bt}$ against the 15-mer RNA. This analysis indicates that NrnB$_{Bt}$ exhibits an NrnA-like preference for short RNAs 2-4 nucleotides in length, while displaying a slightly greater capacity to act on other short RNA substrates 6-10 nucleotides in length. This contrasts with previous reports for the *in vitro* analysis of NrnB$_{Bs}$. In prior studies NrnB$_{Bs}$ degraded longer RNAs 6-15 nucleotides in length to near completion by the 30-minute time point using the same assay conditions. This is further *in vitro* evidence that the lack of a C-terminal extension might reduce the proclivity of NrnB-like proteins to degrade long RNA substrates.

6.2.6 *B. subtilis* cellular lysates expressing *nrnB* protein homologs result in enhanced long and short RNA processing in a Δ*nrnA*Δ*nrnB B. subtilis* background

In an effort to improve our understanding of all of the relevant biochemical features of NrnB-like proteins we decided to test various NrnB candidate proteins using whole cell lysate RNase cleavage assays. Specifically, we generated whole cell lysates of Δ*nrnA*Δ*nrnB B. subtilis* or Δ*nrnA*Δ*nrnB B. subtilis* complementation strains that encoded ectopic IPTG-inducible copies of *nrnB*$_{Bm}$, *nrnB*$_{Bm\ \Delta\text{C-term.}}$, *nrnB*$_{Hp}$, or *nrnB*$_{Cj}$. The *nrnB*$_{Bm\ \Delta\text{C-term.}}$ mutant construct was created by deleting the residues (306-395) corresponding to the small C-terminal extension subdomain predicted by Alphafold 2. These cellular lysates were simultaneously incubated with trace

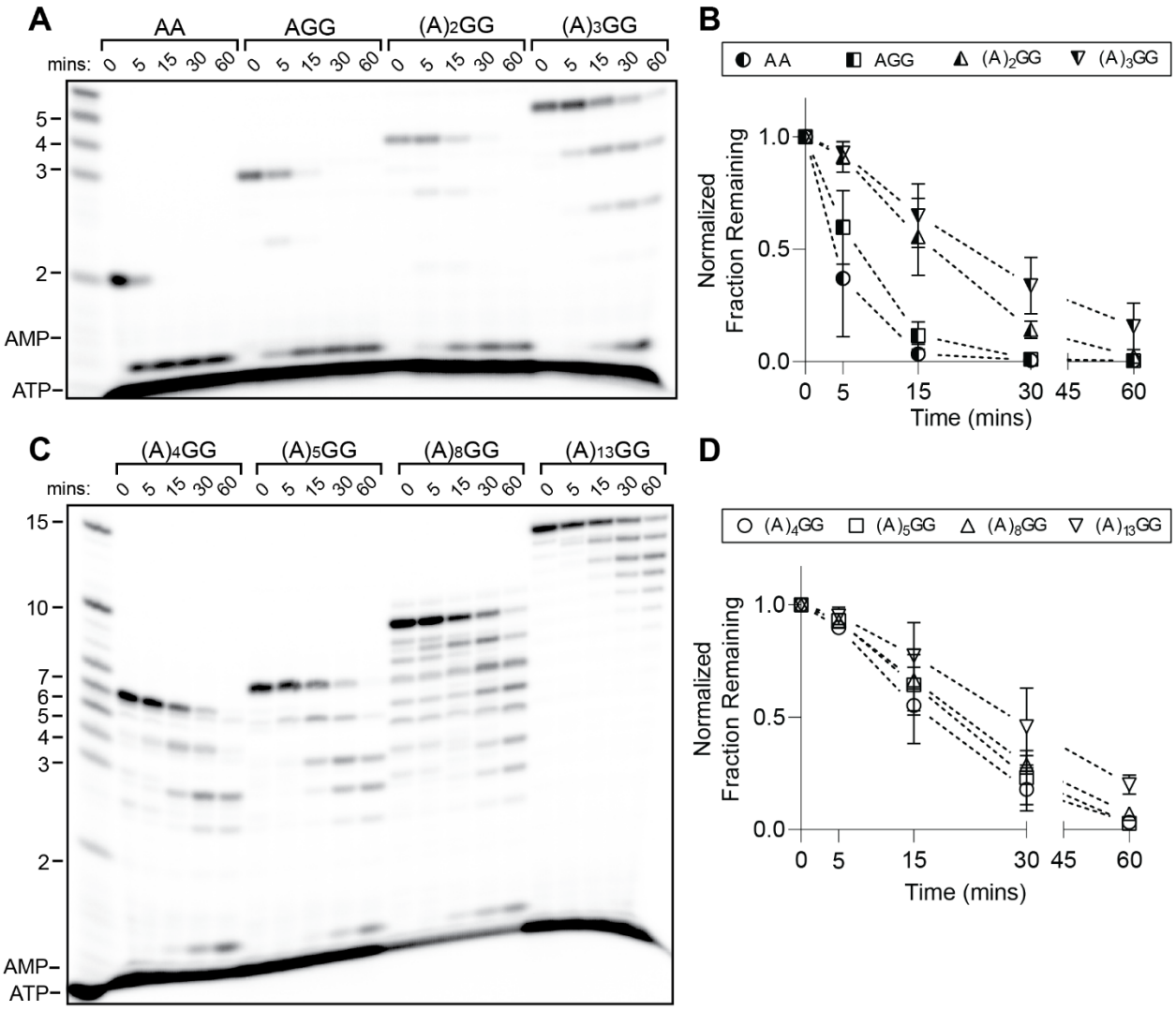

Figure 6.5. *B. thuringiensis* NrnB preferentially hydrolyzes short RNAs 2-4 nucleotides in length. A and C). RNA molecules 2-7, 10, and 15 nucleotides in length were incubated at a final concentration of 1 μM with 50 nM of purified $NrnB_{Bt}$. The reactions also contained a trace amount of 5'-^{32}P-radiolabeled RNA. Samples were removed at time intervals and analyzed by urea-denaturing 20% PAGE. (B and D) Quantification of the normalized radioactive intensity of the initial substrate depletion over time plotted as the average and SD of 3 independent experiments in (A and C).

amounts (~40 nM) of the 5'-radiolabeled diribonucleotide pGpG and 20-mer. The degradation products from these lysate experiments were resolved by 20% denaturing PAGE. In agreement with our previous analysis, we found that the diribonucleotide pGpG, went virtually uncleaved when incubated with Δ*nrnA*Δ*nrnB B. subtilis* lysate (Figure 6.6A). Additionally, the 20-mer was processed by the Δ*nrnA*Δ*nrnB B. subtilis* lysate to near completion in 15-minutes (Figure 6.6A-B). When we forcibly overexpressed $nrnB_{Bm}$, $nrnB_{Bm\ \Delta\text{C-term.}}$, $nrnB_{Hp}$, or $nrnB_{Cj}$ the diribonucleotide was processed to near completion by the 1-minute timepoint (Figure 6.6A). Interestingly, the $nrnB_{Hp}$ and $nrnB_{Cj}$ expression lysates resulted in very rapid processing of the 20-mer as the RNA was >90% degraded by the 5-minute time point. Also, when the 20-mer was subjected to cleavage by the lysates that had expressed $nrnB_{Bm}$, this substrate was rapidly processed to near completion by the 5-minute time point (Figure 6.6A-B). Corroborating our previous analysis, when we subjected the 2-mer and 20-mer to degradation by the cellular lysate for $nrnB_{Bm\ \Delta\text{C-term.}}$ we observed rapid dinucleotide processing similar to that of the wild-type $nrnB_{Bm}$. However, the processing of the 20-mer by the $nrnB_{Bm\ \Delta\text{C-term.}}$ containing lysate remained a rate similar to the Δ*nrnA*Δ*nrnB B. subtilis* control strain. These data further suggest the functional significance of the C-terminal extension in the processing of long RNA substrates. Altogether, these data suggest that NrnB-like proteins function as nanoRNases that process dinucleotides and long RNAs *in vivo.*

6.3 Discussion

The degradation of nanoRNA molecules is often a crucial process in bacterial organisms. *B. subtilis* NrnA and NrnB are two members of the DHH-DHHA1 protein family that were previously thought to exhibit synonymous enzymatic function in the degradation of nanoRNA.

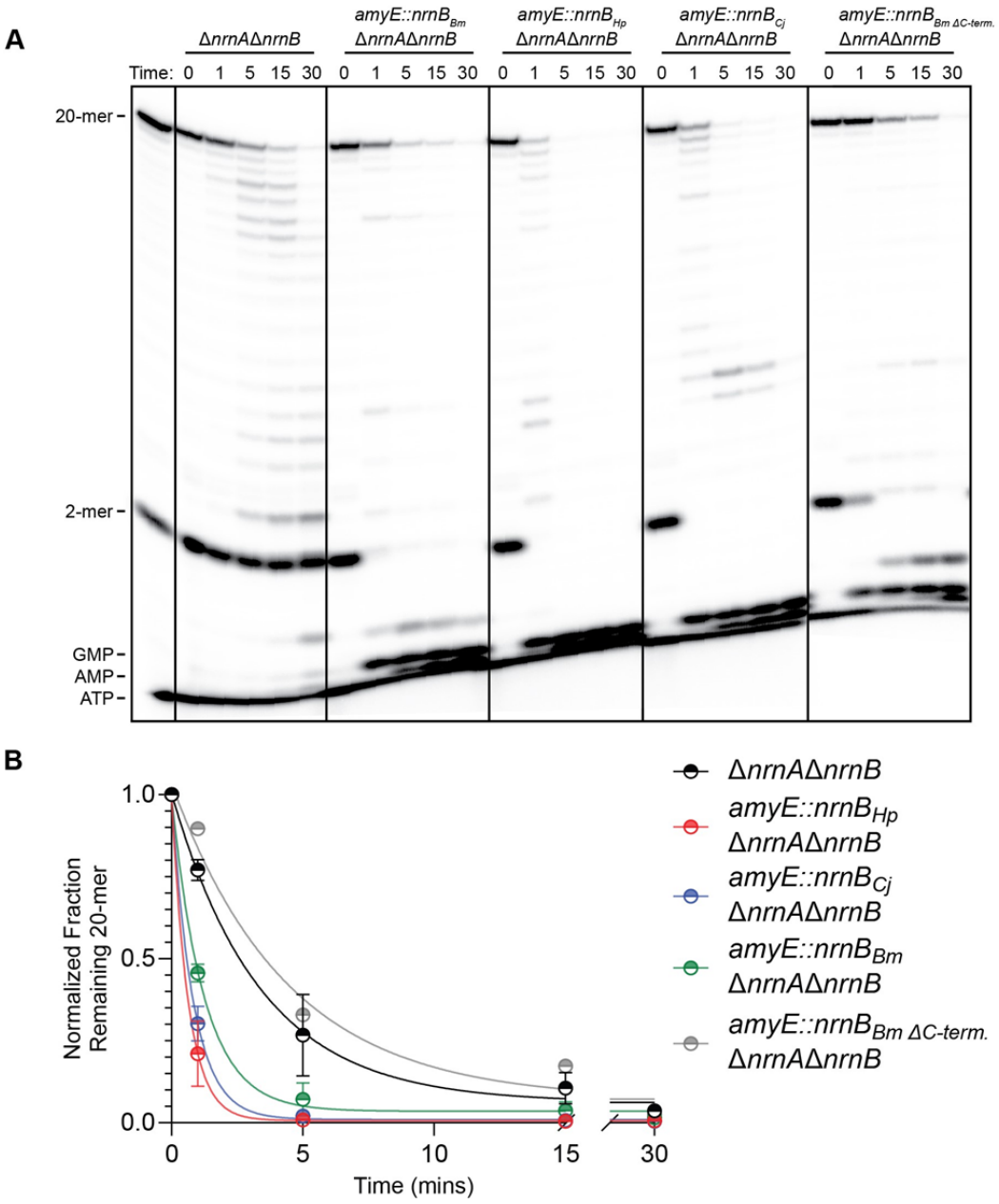

Figure 6.6. *B. subtilis* cellular lysates expressing *nrnB* protein homologs result in enhanced long and short RNA processing in a Δ*nrnA*Δ*nrnB B. subtilis* background (A) Cellular lysates for Δ*nrnA*Δ*nrnB*, or Δ*nrnA*Δ*nrnB* that had been complemented with an IPTG inducible copy of $nrnB_{Bm}$, $nrnB_{Hp}$, $nrnB_{Cj}$, or $nrnB_{Bm\Delta c\text{-}term}$ were harvested during vegetative growth (OD_{600} = 1.0). A mixture of trace amounts of 5'-^{32}P-radiolabeled 20-mer and 2-mer RNA were added to the lysates in in a buffer comprised of 25 mM Tris-HCl, pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, and 50 µM $MnCl_2$. Aliquots were quenched in 4 M urea and 150 mM EDTA and resolved by 20% urea-denaturing PAGE. (B) The normalized radioactive intensity of the initial substrate depletion over time is shown plotted as the average and SD of 2 independent experiments for Δ*nrnA*Δ*nrnB*, $nrnB_{Bm}$, $nrnB_{Hp}$, $nrnB_{Cj}$, *and* 1 independent experiment for $nrnB_{Bm\Delta c\text{-}term}$.

However, our recent analyses of NrnA$_{Bs}$ and NrnB$_{Bs}$ has provided evidence that these two proteins display significant differences in their substrate recognition, accommodation, and intracellular functions. We have found the NrnA$_{Bs}$ preferentially degrades RNAs 2-4 nucleotides in length using a 5'-3' exonucleolytic mechanism, while NrnB$_{Bs}$ degrades RNAs of varying lengths ranging from very short diribonucleotides to long 44-mer substrates with a 3'-5' exonucleolytic mechanism (Weiss *et al.,* 2022; Myers *et al.,* 2023). Adding layers of complexity to the DHH-DHHA1 protein family, we have also provided evidence corroborating previous reports that there is a subfamily of NrnA-like proteins found in the Actinobacterial class of organisms that process c-di-AMP directly and do not process long RNAs (Weiss *et al.,* 2022; He *et al.,* 2016). Furthermore, there is currently a report suggesting that *Vibrio cholera* harbors a DHH-DHHA1 protein that specifically processes the diribonucleotide pGpG (Heo *et al.,* 2022). With the mounting evidence that there are many different subfamilies of DHH-DHHA1 proteins, we conducted a small biochemical survey to ascertain whether NrnB protein homologs were uniform in their activities.

In this study, we investigated the *in vitro* RNase activities of putative NrnB-like protein homologs from *B. licheniformis, B. thuringiensis,* and *H. pylori*. Furthermore, we utilized cellular lysate-based experiments to understand the possibly *in vivo* functions of these NrnB-like proteins. In assessing the *in vitro* activity of purified NrnB$_{Bl,}$ NrnB$_{Bt,}$ and NrnB$_{Hp}$ we found that these enzymes all exhibit RNase activity against RNAs 2-7, 10, 15 and 20 nucleotides in length. We find that NrnB$_{Bl}$ and NrnB$_{Hp}$ efficiently degrade the long RNA 20-mer efficiently, while NrnB$_{Bt}$ does not show robust activity against the RNA 20-mer. We speculate that NrnB$_{Bt}$ cannot readily accommodate and process long RNA substrates due to the lack of a C-terminal extension that likely aids in the accommodation of long RNAs found in NrnB$_{Bs}$ and NrnB$_{Bl}$. Furthermore,

ex vivo lysate experiments indicate that long and short RNA substrates are likely *in vivo* substrates for NrnB$_{Hp}$, NrnB$_{Cj}$, and NrnB$_{Bm}$. Altogether, our aggregate data provides evidence that there are functional dissimilarities amongst the NrnB protein family.

6.4 Materials and methods

6.4.1 Bacterial strains and culture conditions

E. coli strains were grown in LB containing (as needed) 100 µg/mL carbenicillin at 37°C in a shaking incubator. *B. subtilis* strains were grown in LB or 2xYT in the presence of 5 µg/mL chloramphenicol, or 100 µg/mL spectinomycin. The method utilized to make deletion strains of *B. subtilis* Δ*nrnA,* Δ*nrnB,* and Δ*nrnA* Δ*nrnB* is described previously (Orr *et al.,* 2018). To build *E. coli* overexpression strains for protein purification, the genes encoding *nrnB$_{Bl}$*, *nrnB$_{Bt}$*, and *nrnB$_{Hp}$* were purchased from IDT as codon optimized gene fragments and were subcloned into their respective expression vector by Gibson assembly (Gibson *et al.,* 2009). The gene encoding *nrnB$_{Bl}$* was cloned in-frame into an expression vector so that this protein would be expressed without any affinity tags. The gene encoding *nrnB$_{Bt}$* was cloned in-frame into the plasmid pHisll to be expressed with a C-terminal 6xHis tag. The gene for *nrnB$_{Hp}$* was cloned in-frame to be expressed with an N-terminal 10xHis-SUMO tag. The complementation strains of *nrnB$_{Bm}$*, *nrnB$_{Hp}$*, or *nrnB$_{Cj}$* were sequences purchased from IDT as codon optimized gene fragments. These sequences were then subcloned via Gibson assembly (Gibson *et al.,* 2009) into the plasmid pDR111 which harbors an IPTG controllable promoter region upstream of the target gene and regions of homology to the non-essential *amyE* locus for ectopic integration. All transformations of *B. subtilis* were performed using a previously described protocol (Jamer *et al.,* 2002). *E. coli* XL10-Gold (Agilent) was initially transformed with all plasmids, and the 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 target proteins.

6.4.2 Expression and purification of proteins

T7 Expression strains of *E. coli* harboring expression vectors were grown overnight, shaking, at 37°C. The next morning, these cells were subcultures into fresh 2xYT containing 100 µg/mL carbenicillin. These cultures were grown shaking at 37°C until reaching mid-log phase OD_{600} of ~0.6. At this point, the cultures were removed from the 37°C incubator, induced with IPTG, and grown overnight at room temperature for another 16-18 hours. Cells expressing $nrnB_{Bl}$ or $nrnB_{Hp}$ were induced by adding IPTG to a final concentration of 1 mM, while the cells expressing $nrnB_{Bt}$ were induced with IPTG to a final concentration of 400 µM. The cells were pelleted by centrifugation at 5,000 rpm for 20 minutes. The purification of $NrnB_{Bl}$ was completed using the method described in chapter 4.4. The purification of $NrnB_{Hp}$ was conducted using the method for purifying $NrnA_{Bs}$ in chapter 3.5.3 except that 5% glycerol was added to all of the buffers, and the 10xHis-SUMO tag was removed by an overnight cleavage on ice by the enzyme bdSENP1. The cell pellet for $NrnB_{Bt}$ was resuspended in 25 mM Tris-HCl, pH 8.0, 300 mM NaCl, 5% glycerol at 10 mL of buffer per 1 gram of cell pellet. PMSF was added to the resuspension to a final concentration of 1 mM. The cells were lysed via sonication. The insoluble material was removed by two 15-minute rounds of centrifugation at 12,000 rpm. 10% polyethyleneimine (PEI) was added to the clarified lysate to reach a final concentration of 0.5%. NrnB was then pelleted in the PEI at 5000 rpm for 5 minutes. The supernatant was discarded and the PEI pellet containing the protein was resuspended in an equal volume of 25 mM Tris-HCl pH 8, 500 mM NaCl, 5% glycerol resuspension buffer to elute $NrnB_{Bt}$. At this point the sample was left for 15 minutes to elute the protein. The PEI was pelleted again by centrifugation at 10,000

rpm for 10 minutes. The supernatant containing NrnB was collected, and the PEI pellet was discarded. Solid ammonium sulfate was added to the sample to reach 70% saturation using the 0°C convention. The sample was rocked gently for 30 minutes on ice. After shaking, the sample of $NrnB_{Bt}$ was centrifuged at 10,000 x *g* for 10 minutes. At this point $NrnB_{Bt}$ precipitated out of solution. The protein pellet was resuspended in 25 mM Tris-HCl pH 8, 300 mM NaCl, 5% glycerol. $NrnB_{Bt}$ was incubated with Roche™ nickel resin for 1 hour shaking gently on ice. Next, the nickel resin was washed with 10 column volumes of 25 mM Tris-HCl, pH 8, 300 mM NaCl, 10 mM imidazole, 5 % glycerol. Then, the resin was washed with 5 column volumes of 25 mM Tris-HCl, pH 8, 300 mM NaCl, 25 mM Imidazole, 5% glycerol. Finally, $NrnB_{Bt}$ was eluted from the resin with 5 column volumes of 25 mM Tris-HCl, pH 8, 300 mM NaCl, 250 mM imidazole, 5% glycerol. $NrnB_{Bt}$ was dialyzed to remove imidazole, and then aliquots were flash frozen in liquid nitrogen and stored at -80°C.

6.4.3 Oligoribonucleotide labeling

These reactions were conducted using the method reported in chapter 5.5.6

6.4.4 Oligoribonucleotide cleavage reactions

RNA cleavage assays were conducted utilizing the method in chapter 5.5.8.

6.4.5 Preparation of whole cell lysates

Whole cell lysates were generated using the method reported in chapter 5.5.5.

Chapter 7: Outlooks and perspectives

7.1 A reservoir of nanoRNAs

Short RNA degradation by nanoRNases is a vitally important intracellular process that is crucial for survival and virulence of many organisms. In fact, a recent phylogenetic analysis has shown that effectively all organisms encode for at least one nanoRNase whether it is Orn, NrnA, NrnB, or NrnC (Figure 1.4) (Lormand *et al.,* 2021). Many pathogens including *Campylobacter jejuni, Vibrio cholera, Pseudomonas aeruginosa, Streptococcus pneumoniae, Mycobacterium tuberculosis, Bartonella henselae* and *Escherichia coli* will not survive the depletion of their respective nanoRNase, or they will exhibit major physiological defects (Dey *et al.,* 2017; Yang *et al.,* 2014; Bai *et al.,* 2013; Fahmi *et al.,* 2019; Mandal *et al.,* 2017; Lui *et al.,* 2012; Ghosh and Deutscher, 1999; Orr *et al.,* 2015). It is largely assumed that defects that arise in the nanoRNase deficient strains occur due to the accumulation of short RNAs byproducts that can arise through the processing of long mRNAs by cellular exoribonucleases (Bechhofer and Deutscher, 2019; Trinquier *et al.,* 2020; Hui *et al.,* 2014). However, one often overlooked process that could be the primary source of metabolically burdensome short RNAs is during transcription initiation. During the process of transcription initiation short RNAs 2-15 nucleotides (i.e., abortive transcripts) in length are generated prior to the RNA polymerase switching from initiation to elongation (Carpousis and Gralla, 1980; Carpousis *et al.,* 1982; Hsu *et al.,* 2003; Godman *et al.,* 2009). While it is generally unknown what the distribution of abortive initiation products are for bacterial polymerases, the analysis of phage polymerases and has suggested that diribonucleotides might be the most prominent short RNA synthesized (Carpousis and Gralla, 1980; Carpousis *et al.,* 1982; Hsu *et al.,* 2003).

Previous analysis of nanoRNases has suggested that diribonucleotides are likely the preferred substrate amongst of the different nanoRNase subfamilies, however, it remains unknown whether short RNAs bearing 5' triphosphosphates are suitable substrates for any of these enzymes (Lee *et al.,* 2022). The capacity of different nanoRNases to recognize and degrade triphosphorylated short RNA substrates could differ amongst the different subfamilies of nanoRNases due to differences in substrate recognition and accommodation. Specifically, it is easy to predict that NrnB will process triphosphorylated RNAs due to its ability to recognize and degrade RNAs using a 3'-5' mechanism, and that NrnB recognizes RNA substrates of varying lengths (Myers *et al.,* 2023). However, we predict that the 5'-3' mechanism utilized by NrnA will preclude this enzyme from degrading triphosphorylated RNAs (Weiss *et al.,* 2022). This is due to previous analyses where another 5'-3' exonuclease RNase J1/J2 was shown to display minimal activity on triphosphorylated RNAs (Mathy *et al.,* 2007). Additionally, we speculate from the analysis of crystallographic structures of Orn and NrnC bound to monophosphorylated diribonucleotides that these enzymes might harbor active sites that might be too selective to accommodate the additional negative charge of a triphosphorylated dinucleotides (Kim *et al.,* 2019; Lormand *et al.,* 2021). Altogether, we predict that some organisms might harbor yet another enzyme that might be required to participate in the turnover of triphosphorylated RNAs.

RppH is a nudix family pyrophosphohydrolase that is broadly distributed amongst most taxonomic groups of bacteria (Celesnik *et al.,* 2007; Deana *et al.,* 2008; Piton *et al.,* 2013). The primary function of RppH is to catalyze the removal of pyrophosphate from the 5' ends of mRNA transcripts rendering these RNAs highly susceptible to degradation by 5'-3' exonucleolytic degradation by RNase J in *B. subtilis,* or endonucleolytic cleavage by the endonuclease RNase E found in *E. coli* (Celesnik *et al.,* 2007; Richards *et al.,* 2011). Moreover,

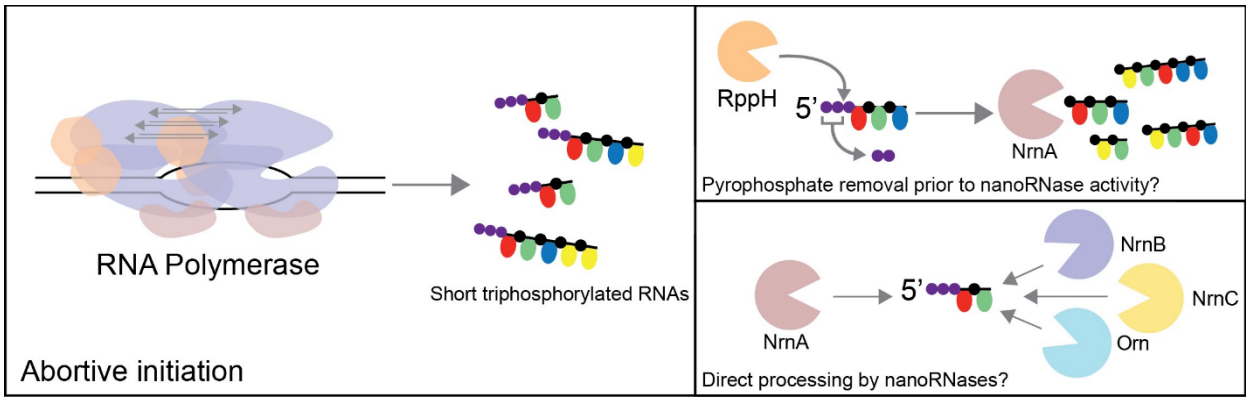

Figure 7.1. Abortive initiation product degradation schematic. RNA polymerase generates short RNA oligonucleotides that have a triphosphorylated 5' end. It remains unknown whether the degradation of abortive initiation products occurs through a two enzyme mechanism where RppH removed the pyrophosphate group of abortive initiation products prior to nanoRNase processing. Also, it remains unknown whether nanoRNases can directly process triphosphorylated RNAs.

RppH has also been shown to function in the removal of other 5' capping molecules such as dinucleoside tetraphosphates (Np4), NAD, and FAD which can alter the half-life of mRNA transcripts (Levenson-Palmer *et al.,* 2022). Previous studies have shown that RppH can remove pyrophosphate from short RNA oligoribonucleotides, but it is unknown whether this enzyme exhibits stark substrate length preferences analogous to the nanoRNases. However, since RppH has been shown to act on short RNA oligonucleotides, it is tempting to speculate that the degradation of abortive initiation transcripts could proceed by way of a two-enzyme mechanism in which RppH removes pyrophosphate from abortive initiation products prior to the degradation of short RNAs. Whether or not nanoRNases and RppH function together to recycle abortive initiation products represents a gap in the literature (Figure 7.1). This gap in knowledge is likely due to the difficulty in generating short 5'-triphosphorlyated RNAs. The reason why these observations have not been heavily interrogated is because it is generally difficult to obtain the RNAs with 5' triphosphates.

7.2 A potential method for the rapid synthesis of short oligoribonucleotides

Utilizing unmodified and highly purified substrates is a crucial step for the analysis of any enzyme. In order to analyze RNases or RppH, frequently one must purchase chemically synthesized RNAs that are purified by HPLC. Synthetic RNA oligonucleotides are very expensive compared to the longer RNAs that can simply be transcribed at high quantities in the T7 transcription kits. Herein we proposed a potential method for exploiting the catalytic mechanism of the *glmS* ribozyme to generate short triphosphorylated RNA substrates to mimic abortive initiation products.

In prior analyses of the *glmS* ribozyme it was determined that this RNA has minimal 5' end sequence requirements for self-cleavage. The only requirement is that there must be a purine

residue next to the cleavage site of the ribozyme (Winkler *et al.,* 2004). With this, it seems entirely plausible to to *in vitro* transcribe the *glmS* ribozyme from a DNA template using T7 RNA polymerase so that the cleavage products from the reaction can range from a purine mono ribonucleotide to a length of our choosing (i.e., di-, tri-, tetraribonucleotides) (Figure 7.2). After transcription of the full-length RNA, the unincorporated NTP's can be removed by a Zymo-RNA clean-up kit. Next, the *glmS* ribozyme can be activated by the addition of glucosamine-6-phosphate and the short cleavage product can be generated. One complication that remains is that the cleavage by the *glmS* ribozyme generates a 5' fragment that harbors a 2'3'-cyclic phosphate that would have to be removed. However, T4 polynucleotide kinase has been previously shown to exhibit 3' phosphatase activity that can relieve 2'3'-cyclic phosphates and generate a canonical 3' hydroxyl (Das and Shuman, 2013). Another possibility that could be particularly beneficial to researchers would be to include α-^{32}P-GTP to initiate the transcription reaction. This would allow researchers to generate short radiolabeled RNAs in which they could directly analyze using phosphorimaging. Altogether, this could provide a quick method for the generation of RNA substrates that are frequently not analyzed. Recently Jena Biosciences has included pppApG as a commercially available RNA dinucleotide. This RNA could be utilized in a myriad of different ways to directly and indirectly assess the activities of nanoRNases an RppH to process triphosphorylated short RNAs. For example, researchers could utilize the pppApG substrate and monitor the cleavage by coupling the RNase cleavage to a reaction that utilizes ATP as a substrate such as luciferase or lactate dehydrogenase. Another possibility would be to monitor the degradation products of pppApG by mass spectroscopy or by chromatographically separating the degradation product by HPLC or FPLC. In all, it is crucially important for researchers investigating nanoRNase specificity to include various short RNA substrates with the

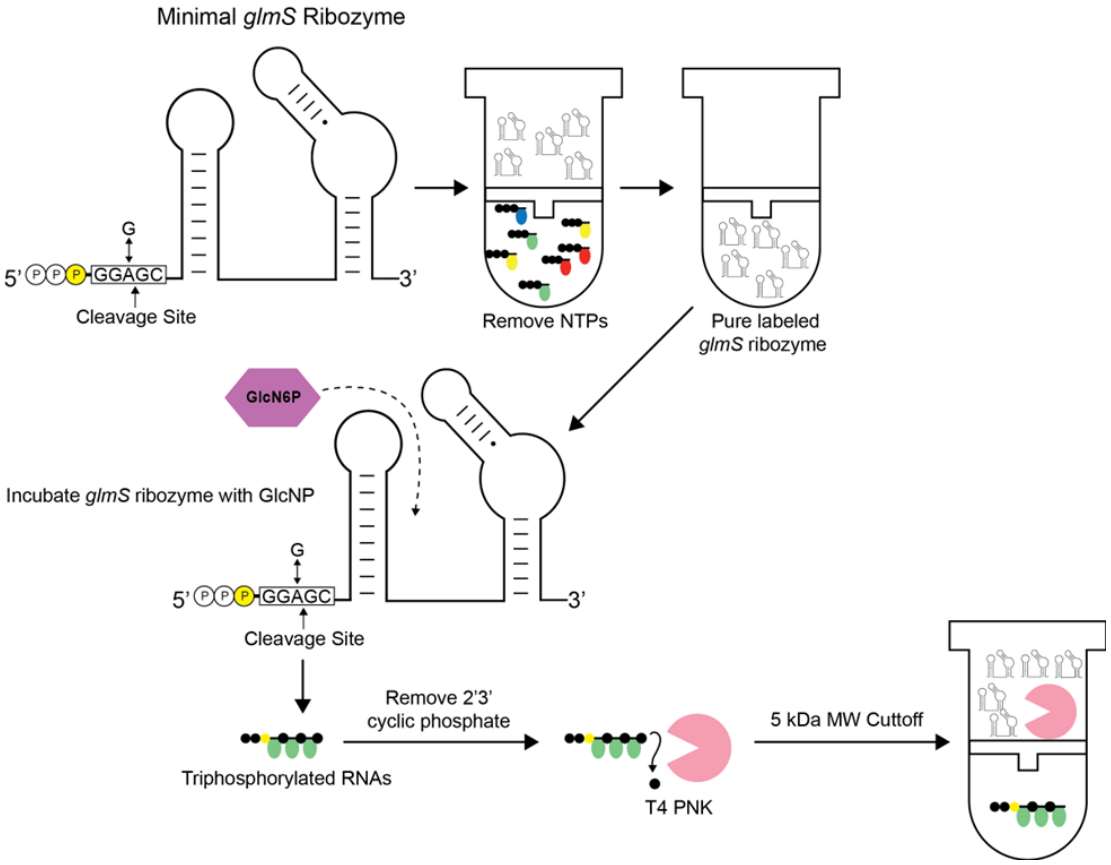

Figure 7.2. A theoretical method for the generation of short triphosphorylated RNAs. First RNAs encoding the minimal *glmS* ribozyme sequence is transcribed utilizing an *in vitro* transcription kit. Then, the *glmS* ribozyme is separated by remaining nucleotide triphosphates and RNA polymerase by the use of an RNA clean up kit. Glucosamine-6-phosphate is incubated with the *glmS* RNA to induce ribozyme self-cleavage. The short triphosphorylated RNA is incubated with T4 polynucleotide kinase to remove the 2'3' cyclic phosphate. The triphosphorylated short RNA is separated from the *glmS* RNA through the use of a 5 kDa MW cutoff filter.

different phosphorylation states to better understand the physiological functions of these enzymes.

7.3 The search for undiscovered nanoRNases

The known nanoRNases NrnA, NrnB, NrnC, and Orn have been shown by phylogenetic analyses to be broadly distributed amongst virtually all organisms including Archaea and Eukaryota (Lormand *et al.,* 2021). Interestingly, genetic experiments have revealed that in some cases nanoRNase encoding genes are essential for growth in some organisms such as *E. coli*, *V. cholera*, *C. jejuni*, and many others (Mandal *et al.,* 2017; Lui *et al.,* 2012; Ghosh and Deutscher, 1999; Orr *et al.,* 2015). However, there are also many instances where deletion of an organism's nanoRNase encoding gene is completely dispensable, resulting in no major cellular viability phenotypes as in the case of *B. subtilis*. These genetic experiments could indicate that short RNA oligonucleotides are toxic to some organisms and not others. Another possibility is that these organisms encode for other unknown or previously unrecognized nanoRNases.

In *B. subtilis* we have previously observed that the single and double deletions of *nrnA* and *nrnB* do not affect cell growth in rich media. Additionally, we found that these two proteins are likely serving different functions in that *nrnA* is expressed during vegetative growth while *nrnB* is expressed within the developing forespore. To investigate whether there are other nanoRNases present within *B. subtilis,* we decided to test the capacity of *B. subtilis* cellular lysates to degrade 5'- ^{32}P-radiolabeled lshort RNAs 2-5 nucleotides in length. For this analysis we simultaneously subjected trace amounts of RNAs corresponding to the sequences (GG, AGG, AAGG, and AAAGG) as well as a 20-mer to observe background general exoribonucleolytic activity (Figure 7.3). We observed that wild type cellular extracts quickly processed long and

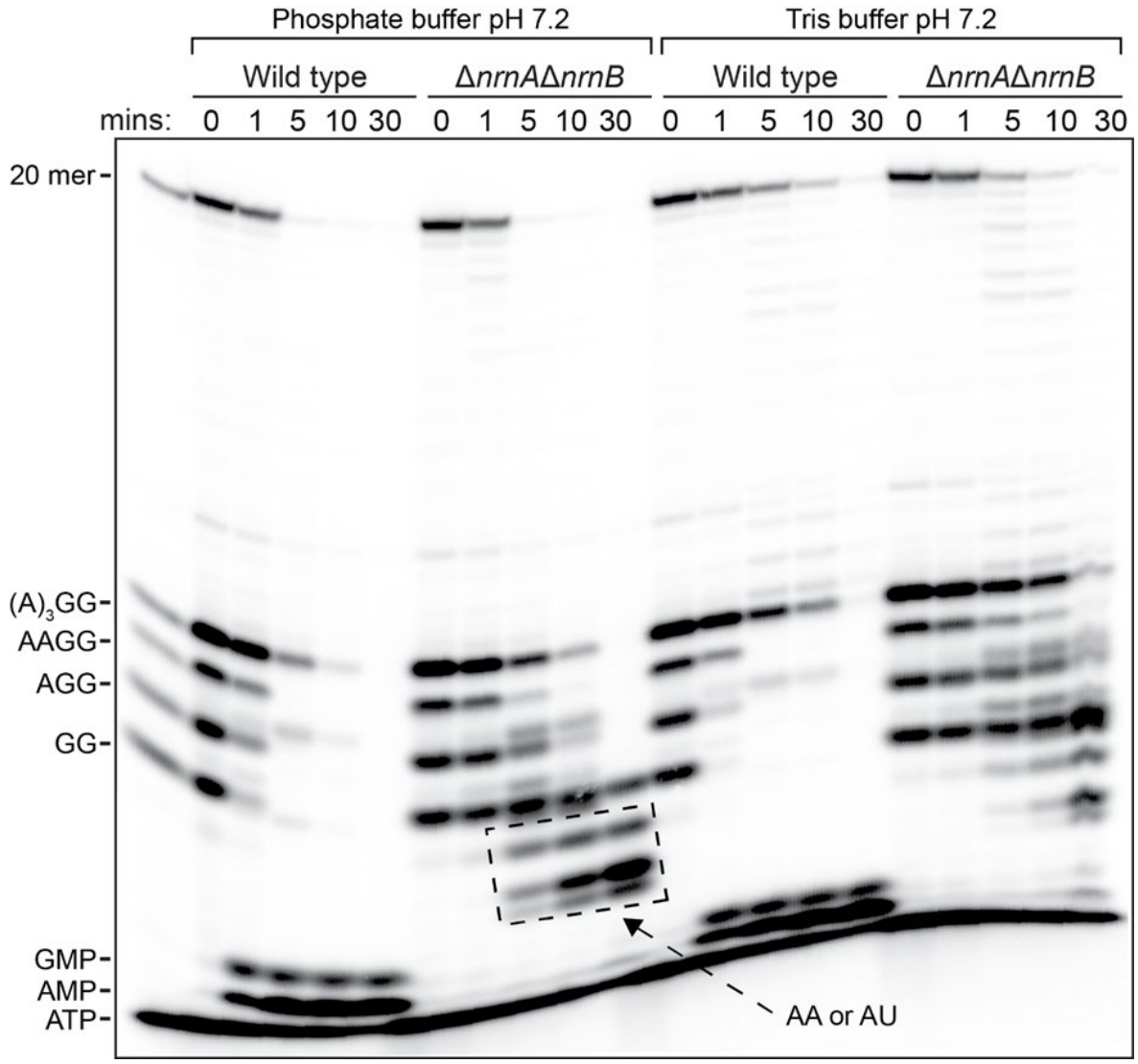

Figure 7.3. Cellular extracts of Δ*nrnA*Δ*nrnB B. subtilis* display robust nanoRNase activity against short RNAs 3 to 5 nucleotides in length. Cellular lysates of wild-type or Δ*nrnA*Δ*nrnB* were simultaneously incubated with trace amounts of 5'-^{32}P-RNAs 2-5 and 20 nucleotides in length in a reaction containing phosphate buffer (25 mM NaH_2PO_4, pH 7.2, 100 mM NaCl, 25 mM $MgCl_2$, 500 μM $MnCl_2$) or tris buffer (25 mM Tris-HCl, pH 7.2, 100 mM NaCl, 25 mM $MgCl_2$, 500 μM $MnCl_2$). Reactions were quenched in 4 M urea and 150 mM EDTA prior to resolving degradation products by 20% denaturing PAGE.

short RNAs to near completion by the 5 minute time point in the phosphate buffer. When we -assessed both long and short RNA cleavage by the Δ*nrnA*Δ*nrnB* cellular lysate in phosphate buffer and observed that the processing of the 5-mer and 20-mer were virtually identical to that of the wild type extract. To our surprise, the RNA diribonucleotide went unprocessed by the Δ*nrnA*Δ*nrnB* cellular lysate in phosphate buffer. Yet another interesting finding is that there is 3-mer and 4-mer processing in reactions containing the Δ*nrnA*Δ*nrnB* cellular lysate strongly suggesting that there are other nanoRNases present in *B. subtilis*. To speculate whether the 3-mer and 4-mer processing activity is due to a hydrolytic of phosphorolytic enzyme we also tested the wild type and Δ*nrnA*Δ*nrnB* cellular lysates activity in tris buffer. When we compared the RNase activity of the wild type cellular lysate in tris buffer, we found that the qualitative rate of degradation for the short RNAs 2-5 nucleotides in length remained virtually the same to that of the lysate reactions in the phosphate buffer. However, we noticed that the processing of the 20-mer was greatly reduced in assays conducted in tris buffer. The reduction in rate could indicate that the enzyme degrading the 20-mer is phosphorolytic, while the processing of the short RNAs could be accomplished by a hydrolytic RNase that does not require inorganic phosphate for its reaction mechanism. We still observe virtually no processing of the 2-mer in tris buffer, but we do see a reduction in the rate of processing of the 5-mer, 4-mer, and 3-mer which could suggest that the enzyme responsible for the processing of short RNAs utilizes a phosphorolytic mechanism.

Currently, there are only two known phosphorolytic enzymes that have been identified in *B. subtilis*. These enzymes include the major general 3'-5' exoribonuclease PNPase and the tRNA processing enzyme RNase PH. Based on this current state of knowledge we hypothesize that the nanoRNase activity that we are observing against 3-mers and 4-mer in the Δ*nrnA*Δ*nrnB*

cellular lysates is due to PNPase or RNase PH. Historical studies of PNPase have suggested that this enzyme can degrade short RNA substrates down to diribonucleotides (Singer, 1958; Chou *et al.,* 1970; Chou *et al,* 1975). More recent studies of a PNPase homolog from *Mycobacterium smegmatis* have shown that this enzyme can efficiently degrade long RNAs down to 3-mers and then process the 3-mer at a slower rate (Unciuleac *et al.,* 2021). RNase PH has been shown to recognize the 3' overhang of tRNAs and process 3 to 4 nucleotides from this RNA to keep it from being charged by an amino acid (Bechhofer and Murray 2019). Perhaps RNase PH can act on short RNA substrates that are amenable to the 3' overhangs observed on mature tRNAs. We do not think that the processing of the short RNAs is due to activity of RNase J1/J2. Since RNase J1/J2 has been shown to exhibit 5'-3' exoribonuclease activity we would anticipate to observe 5'-^{32}P monoribonucleotides accumulating as degradation products in the Δ*nrnA*Δ*nrnB* lysates, however, we do not see any evidence of this activity. However, prior analyses of RNase J1 have suggested that this enzyme exhibits a slight proclivity to degrade nanoRNases (Fang *et al.,* 2009).

Although we speculate that degradation of the 3-mer and the 4-mer might be due to either PNPase or RNase PH, there are many possible candidate RNases that could be responsible for this activity. Yham is a 3'-5' exoribonuclease that was previously shown to degrade both short oligonucleotides with a preference for processing DNA over RNA. A recent transcriptomic analysis to analyze the processing sites of the 3'-5' exoribonucleases found in *S. pyogenes* strongly suggests that YhaM participates in the the removal of a few nucleotides (3 nucleotides on average) from the 3' ends of transcripts produced by transcription termination or endonuclease cleavage (Lécrivain *et al.*, 2018). This data could hint that YhaM can

process short RNA substrates *in vivo*. Furthermore, researchers exploited the aggregation phenotype of the Δ*orn Pseudomonas aeruginosa* to screen for functional Orn homologs (Orr *et al.,* 2015). In this analysis when YhaM was forcibly overexpressed in the Δ*orn P. aeruginosa* there was a modest reduction in cellular aggregation, suggesting that perhaps YhaM exhibits some diribonucleotide cleavage *in vivo*. Another interesting find was that overexpression of RNase J1 seemed to moderately decrease the Δ*orn* aggregation phenotype (Orr *et al.,* 2018). There also remains multiple genes encoding putative metallophosphodiesterases in the *B. subtilis* genome that have not been biochemically characterized, such as *yutE, ysnB, ykoQ*, *yunD*, or *yybD*.

7.5 Concluding remarks

While there are many outstanding questions in the field of bacterial RNA degradation and c-di-NMP signaling, my research has improved our understanding of RNA degradation in *B. subtilis*. My research strongly suggests that the enzymes NrnA and NrnB serve different intracellular functions in *B. subtilis*. NrnA functions as a housekeeping nanoRNase during vegetative growth, and specifically processes short RNAs 2-4 nucleotides in length by way of a 5'-3' exonucleolytic mechanism (Figure 7.4). Contrastingly, I have shown that NrnB is a previously unrecognized general 3'-5' exoribonuclease akin to PNPase and RNase R that is expressed during sporulation.

The initial genetic and biochemical experiments surrounding Orn have paved the way for significant breakthroughs in the field of RNA decay. Also, investigations into the metabolism of second messenger signaling have indicated that there is overlap between diribonucleotide processing and c-di-NMP levels. Yet another breakthrough in the short RNA degradation field has been in the transcriptomic analysis in which researchers have asked whether short RNAs can

alter transcription initiation *in vivo* (Goldman *et al.,* 2011). While nanoRNAs may play a myriad of roles in cellular metabolism, the next great breakthrough in understanding these enzymes might be in the direct investigation of how nanoRNAs generated through abortive initiation are processed. This question has been complicated largely due to the complications in generating di- and triphosphorylated short RNAs to include in the biochemical analysis of nanoRNases. Several questions remain in this research area such as 1. Is the processing of abortive initiation products accomplished by selecting nanoRNases? 2. Does RppH and other pyrophosphorohydrolases participate in the degradation of abortive transcript? 3. Is abortive initiation transcript accumulation deleterious to cells? 4. Are there specific nanoRNases that process abortive initiation products? Certainly, there are many experiments that must be done to improve the collective understanding of abortive initiation product degradation.

There are certainly previously unrecognized nanoRNases or novel nanoRNases present in the *B. subtilis* genome, and the next great task will be identifying these proteins. This task of identifying nanoRNases could be carried out by a series of biochemical protein fractionations while monitoring the specific activity of the 3-mer and 4-mer cleavage. Historically, coupling this method to mass spectrometry has been extremely successful in determining the identity of proteins with specific activities. Another approach that could aid in the identification of this protein might be through forcibly overexpressing all of the known exoribonucleases in this strain background. It remains unclear how many RNases exhibit nanoRNase and general exoribonuclease activity akin to $NrnB_{Bs}$. However, the forcible expression of known RNases in Δ*nrnA*Δ*nrnB B. subtilis* cellular extracts incubated with RNAs 2-5 nucleotides in length could provide a quick method to provide clues as to whether this dual activity is exhibited by other RNases.

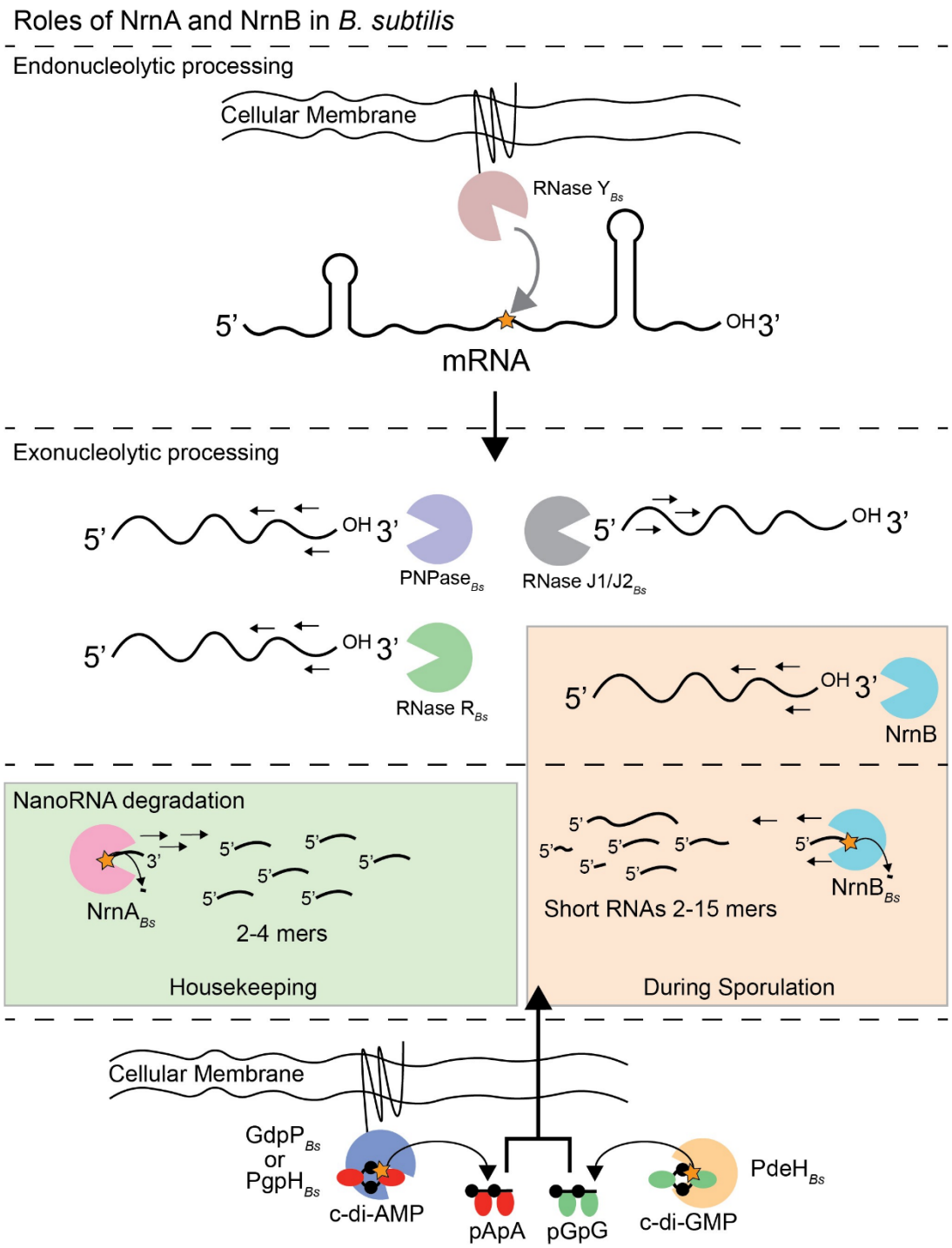

Figure 7.4. Cellular functions of NrnA and NrnB in the context of *B. subtilis* RNA degradation and c-di-NMP signaling. RNA processing in *B. subtilis* generally begins with an internal processing event by RNase Y. The RNA fragments generated by this destabilization even can be processed by the 3'-5' exoribonucleases PNPase, RNase R, NrnB, and the 5'-3' exoribonuclease RNase J1/J2. Diribonucleotides can be generated through the processing of c-di-NMP molecules by PdeH, GdpP, or PgpH. These linearized diribonucleotides contribute to cellular nanoRNA pools. NrnA degrades short RNAs 2-4 nucleotides in length with a 5'-3' polarity during vegetative growth while NrnB degrades short RNAs ranging from diribonucleotides to longer RNAs during sporulation.

Bibliography

1. Jones,A.C. and Neely,R.K. (2015) 2-aminopurine as a fluorescent probe of DNA conformation and the DNA–enzyme interface. *Q. Rev. Biophys.*, **48**, 244–279.

2. Nilsen,T.W. (2014) 3′-End Labeling of RNA with [5′- 32 P]Cytidine 3′,5′-Bis(Phosphate) and T4 RNA Ligase 1. *Cold Spring Harb. Protoc.*, **2014**, pdb.prot080713.

3. Carpousis,A.J., Stefano,J.E. and Gralla,J.D. (1982) 5′ Nucleotide heterogeneity and altered initiation of transcription at mutant lac promoters. *J Mol Biol*, **157**, 619–633.

4. Mathy,N., Bénard,L., Pellegrini,O., Daou,R., Wen,T. and Condon,C. (2007) 5′-to-3′ Exoribonuclease Activity in Bacteria: Role of RNase J1 in rRNA Maturation and 5′ Stability of mRNA. *Cell*, **129**, 681–692.

5. Chao,Y. and Vogel,J. (2016) A 3′ UTR-Derived Small RNA Provides the Regulatory Noncoding Arm of the Inner Membrane Stress Response. *Mol. Cell*, **61**, 352–363.

6. Bagyan,I., Hobot,J. and Cutting,S. (1996) A compartmentalized regulator of developmental gene expression in Bacillus subtilis. *J. Bacteriol.*, **178**, 4500–4507.

7. Roembke,B.T., Zhou,J., Zheng,Y., Sayre,D., Lizardo,A., Bernard,L. and Sintim,H.O. (2014) A cyclic dinucleotide containing 2-aminopurine is a general fluorescent sensor for c-di-GMP and 3′,3′-cGAMP. *Mol Biosyst*, **10**, 1568–1575.

8. Py,B., Higgins,C.F., Krisch,H.M. and Carpousis,A.J. (1996) A DEAD-box RNA helicase in the Escherichia coli RNA degradosome. *Nature*, **381**, 169–172.

9. 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 as the terminal step of RNA degradation. *Elife*, **8**, e46313.

10. Levenson-Palmer,R., Luciano,D.J., Vasilyev,N., Nuthanakanti,A., Serganov,A. and Belasco,J.G. (2022) A distinct RNA recognition mechanism governs Np4 decapping by RppH. *Proc National Acad Sci*, **119**, e2117318119.

11. 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.

12. Niyogi,S.K. and Datta,A.K. (1975) A novel oligoribonuclease of Escherichia coli. I. Isolation and properties. *J Biological Chem*, **250**, 7307–12.

14. Datta,A.K. and Niyogi,K. (1975) A novel oligoribonuclease of Escherichia coli. II. Mechanism of action. *J Biological Chem*, **250**, 7313–9.

15. Luciano,D.J., Vasilyev,N., Richards,J., Serganov,A. and Belasco,J.G. (2017) A Novel RNA Phosphorylation State Enables 5′ End-Dependent Degradation in Escherichia coli. *Mol Cell*, **67**, 44-54.e6.

16. Heo,K., Lee,J.-W., Jang,Y., Kwon,S., Lee,J., Seok,C., Ha,N.-C. and Seok,Y.-J. (2022) A pGpG-specific phosphodiesterase regulates cyclic di-GMP signaling in Vibrio cholerae. *J Biol Chem*, **298**, 101626.

17. 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**, JB.00300-18.

18. Arrieta-Ortiz,M.L., Hafemeister,C., Bate,A.R., Chu,T., Greenfield,A., Shuster,B., Barry,S.N., Gallitto,M., Liu,B., Kacmarczyk,T., *et al.* (2015) An experimentally supported model of the Bacillus subtilis global transcriptional regulatory network. *Mol Syst Biol*, **11**, 839.

19. Huynh,T.N., Luo,S., Pensinger,D., Sauer,J.-D., Tong,L. and Woodward,J.J. (2015) An HD-domain phosphodiesterase mediates cooperative hydrolysis of c-di-AMP to affect bacterial growth and virulence. *Proc National Acad Sci*, **112**, E747–E756.

20. Richards,J., Liu,Q., Pellegrini,O., Celesnik,H., Yao,S., Bechhofer,D.H., Condon,C. and Belasco,J.G. (2011) An RNA Pyrophosphohydrolase Triggers 5′-Exonucleolytic Degradation of mRNA in Bacillus subtilis. *Mol Cell*, **43**, 940–949.

21. Figaro,S., Durand,S., Gilet,L., Cayet,N., Sachse,M. and Condon,C. (2013) Bacillus subtilis Mutants with Knockouts of the Genes Encoding Ribonucleases RNase Y and RNase J1 Are Viable, with Major Defects in Cell Morphology, Sporulation, and Competence. *J Bacteriol*, **195**, 2340–2348.

22. Myers,T.M., Ingle,S., Weiss,C.A., Sondermann,H., Lee,V.T., Bechhofer,D.H. and Winkler,W.C. (2023) Bacillus subtilis NrnB is expressed during sporulation and acts as a unique 3′-5′ exonuclease. *Nucleic Acids Res.*, **51**, 9804–9820.

23. Oussenko,I.A., Sanchez,R. and Bechhofer,D.H. (2002) Bacillus subtilis YhaM, a Member of a New Family of 3′-to-5′ Exonucleases in Gram-Positive Bacteria. *J Bacteriol*, **184**, 6250–6259.

24. Bechhofer,D.H. and Deutscher,M.P. (2019) Bacterial ribonucleases and their roles in RNA metabolism. *Crit Rev Biochem Mol*, **54**, 242–300.

25. Choi,H.-P., Juarez,S., Ciordia,S., Fernandez,M., Bargiela,R., Albar,J.P., Mazumdar,V., Anton,B.P., Kasif,S., Ferrer,M., *et al.* (2013) Biochemical Characterization of Hypothetical Proteins from Helicobacter pylori. *Plos One*, **8**, e66605.

26. Kundra,S., Lam,L.N., Kajfasz,J.K., Casella,L.G., Andersen,M.J., Abranches,J., Flores-Mireles,A.L. and Lemos,J.A. (2021) c-di-AMP Is Essential for the Virulence of Enterococcus faecalis. *Infect. Immun.*, **89**, e00365-21.

27. Stelitano,V., Giardina,G., Paiardini,A., Castiglione,N., Cutruzzolà,F. and Rinaldo,S. (2013) C-di-GMP Hydrolysis by Pseudomonas aeruginosa HD-GYP Phosphodiesterases: Analysis of the Reaction Mechanism and Novel Roles for pGpG. *PLoS ONE*, **8**, e74920.

28. Gomelsky,M. (2011) cAMP, c-di-GMP, c-di-AMP and now cGMP: bacteria use them all! *Mol. Microbiol.*, **79**, 562–565.

29. Burgess,R.R. (2009) Chapter 20 Protein Precipitation Techniques. *Methods Enzymol*, **463**, 331–342.

30. Patel,D.K., Gebbie,M.P. and Lee,V.T. (2014) Chapter Twenty-One Assessing RNA Interactions with Proteins by DRaCALA. *Methods Enzymol*, **549**, 489–512.

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

32. Spiegelberg,B.D., Xiong,J.-P., Smith,J.J., Gu,R.F. and York,J.D. (1999) Cloning and Characterization of a Mammalian Lithium-sensitive Bisphosphate 3′-Nucleotidase Inhibited by Inositol 1,4-Bisphosphate*. *J Biol Chem*, **274**, 13619–13628.

33. Mirdita,M., Schütze,K., Moriwaki,Y., Heo,L., Ovchinnikov,S. and Steinegger,M. (2022) ColabFold: making protein folding accessible to all. *Nat Methods*, **19**, 679–682.

34. Helmann,J.D. (1995) Compilation and analysus of Bacillus Subtilis σ A -dependent promoter sequences: evidence for extended contact between RNA polymerse and upstream promoter DNA. *Nucleic Acids Res*, **23**, 2351–2360.

35. Nicolas,P., Mäder,U., Dervyn,E., Rochat,T., Leduc,A., Pigeonneau,N., Bidnenko,E., Marchadier,E., Hoebeke,M., Aymerich,S., *et al.* (2012) Condition-Dependent Transcriptome Reveals High-Level Regulatory Architecture in *Bacillus subtilis*. *Science*, **335**, 1103–1106.

36. Keilty,S. and Rosenberg,M. (1987) Constitutive function of a positively regulated promoter reveals new sequences essential for activity. *J Biol Chem*, **262**, 6389–6395.

37. Winkler,W.C., Nahvi,A., Roth,A., Collins,J.A. and Breaker,R.R. (2004) Control of gene expression by a natural metabolite-responsive ribozyme. *Nature*, **428**, 281–286.

38. Carpousis,A.J., Houwe,G.V., Ehretsmann,C. and Krisch,H.M. (1994) Copurification of E. coli RNAase E and PNPase: Evidence for a specific association between two enzymes important in RNA processing and degradation. *Cell*, **76**, 889–900.

39. Uemura,Y., Nakagawa,N., Wakamatsu,T., Kim,K., Montelione,G.T., Hunt,J.F., Kuramitsu,S. and Masui,R. (2013) Crystal structure of the ligand-binding form of nanoRNase from Bacteroides fragilis, a member of the DHH/DHHA1 phosphoesterase family of proteins. *Febs Lett*, **587**, 2669–2674.

40. Stülke,J. and Krüger,L. (2020) Cyclic di-AMP Signaling in Bacteria. *Annu Rev Microbiol*, **74**, 1–21.

41. Corrigan,R.M. and Gründling,A. (2013) Cyclic di-AMP: another second messenger enters the fray. *Nat Rev Microbiol*, **11**, 513–524.

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

43. Römling,U., Galperin,M.Y. and Gomelsky,M. (2013) Cyclic di-GMP: the First 25 Years of a Universal Bacterial Second Messenger. *Microbiol Mol Biol R*, **77**, 1–52.

44. Carpousis,A.J. and Gralla,J.D. (1980) Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation in vitro at the lac UV5 promoter. *Biochemistry-us*, **19**, 3245–3253.

45. Khemici,V., Prados,J., Linder,P. and Redder,P. (2015) Decay-Initiating Endoribonucleolytic Cleavage by RNase Y Is Kept under Tight Control via Sequence Preference and Sub-cellular Localisation. *PLoS Genet.*, **11**, e1005577.

46. Fang,M., Zeisberg,W.-M., 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.

47. De,N., Navarro,M.V.A.S., Raghavan,R.V. and Sondermann,H. (2009) Determinants for the Activation and Autoinhibition of the Diguanylate Cyclase Response Regulator WspR. *J. Mol. Biol.*, **393**, 619–633.

48. Roelofs,K.G., Wang,J., Sintim,H.O. and Lee,V.T. (2011) Differential radial capillary action of ligand assay for high-throughput detection of protein-metabolite interactions. *Proc National Acad Sci*, **108**, 15528–15533.

49. Ingle,S., Chhabra,S., Chen,J., Lazarus,M.B., Luo,X. and Bechhofer,D.H. (2022) Discovery and initial characterization of YloC, a novel endoribonuclease in Bacillus subtilis. *Rna*, **28**, 227–238.

50. Pédelacq,J.-D., Cabantous,S., Tran,T., Terwilliger,T.C. and Waldo,G.S. (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol*, **24**, 79–88.

51. Solntsev,S.K., Shortreed,M.R., Frey,B.L. and Smith,L.M. (2018) Enhanced Global Post-translational Modification Discovery with MetaMorpheus. *J Proteome Res*, **17**, 1844–1851.

52. Gibson,D.G., Young,L., Chuang,R.-Y., Venter,J.C., Hutchison,C.A. and Smith,H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods*, **6**, 343–345.

53. Kobayashi,K., Ehrlich,S.D., Albertini,A., Amati,G., Andersen,K.K., Arnaud,M., Asai,K., Ashikaga,S., Aymerich,S., Bessieres,P., *et al.* (2003) Essential Bacillus subtilis genes. *Proc. Natl. Acad. Sci.*, **100**, 4678–4683.

54. Lacey,M.M., Partridge,J.D. and Green,J. (2010) Escherichia coli K-12 YfgF is an anaerobic cyclic di-GMP phosphodiesterase with roles in cell surface remodelling and the oxidative stress response. *Microbiology+*, **156**, 2873–2886.

55. 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.

56. The,M., MacCoss,M.J., Noble,W.S. and Käll,L. (2016) Fast and Accurate Protein False Discovery Rates on Large-Scale Proteomics Data Sets with Percolator 3.0. *J Am Soc Mass Spectr*, **27**, 1719–1727.

57. Schindelin,J., Arganda-Carreras,I., Frise,E., Kaynig,V., Longair,M., Pietzsch,T., Preibisch,S., Rueden,C., Saalfeld,S., Schmid,B., *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods*, **9**, 676–682.

58. Somsen,O.J.G., Hoek, van A. and Amerongen, van H. (2005) Fluorescence quenching of 2-aminopurine in dinucleotides. *Chem. Phys. Lett.*, **402**, 61–65.

59. Zhou,J., Zheng,Y., Roembke,B.T., Robinson,S.M., Opoku-Temeng,C., Sayre,D.A. and Sintim,H.O. (2017) Fluorescent analogs of cyclic and linear dinucleotides as phosphodiesterase and oligoribonuclease activity probes. *Rsc Adv*, **7**, 5421–5426.

60. Branda,S.S., González-Pastor,J.E., Ben-Yehuda,S., Losick,R. and Kolter,R. (2001) Fruiting body formation by Bacillus subtilis. *Proc National Acad Sci*, **98**, 11621–11626.

61. Tang,Q., Luo,Y., Zheng,C., Yin,K., Ali,M.K., Li,X. and He,J. (2015) Functional Analysis of a c-di-AMP-specific Phosphodiesterase MsPDE fromMycobacterium smegmatis. *Int J Biol Sci*, **11**, 813–824.

62. Meeske,A.J., Rodrigues,C.D.A., Brady,J., Lim,H.C., Bernhardt,T.G. and Rudner,D.Z. (2016) High-Throughput Genetic Screens Identify a Large and Diverse Collection of New Sporulation Genes in Bacillus subtilis. *Plos Biol*, **14**, e1002341.

63. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Žídek,A., Potapenko,A., *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583–589.

64. Stevens,A. and Niyogi,S.K. (1967) Hydrolysis of oligoribonucleotides by an enzyme fraction from Escherichia coli. *Biochem Bioph Res Co*, **29**, 550–555.

65. Sharma,S., Yang,J., Doamekpor,S.K., Grudizen-Nogalska,E., Tong,L. and Kiledjian,M. (2022) Identification of a novel deFADding activity in human, yeast and bacterial 5′ to 3′ exoribonucleases. *Nucleic Acids Res.*, **50**, 8807–8817.

66. 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.

67. Blötz,C., Treffon,K., Kaever,V., Schwede,F., Hammer,E. and Stülke,J. (2017) Identification of the Components Involved in Cyclic Di-AMP Signaling in Mycoplasma pneumoniae. *Front Microbiol*, **08**, 1328.

68. Hsu,L.M., Vo,N.V., Kane,C.M. and Chamberlin,M.J. (2003) In Vitro Studies of Transcript Initiation by Escherichia coli RNA Polymerase. 1. RNA Chain Initiation, Abortive Initiation, and Promoter Escape at Three Bacteriophage Promoters †. *Biochemistry-us*, **42**, 3777–3786.

69. Lécrivain,A.-L., Rhun,A.L., Renault,T.T., Ahmed-Begrich,R., Hahnke,K. and Charpentier,E. (2018) In vivo 3′-to-5′ exoribonuclease targetomes of Streptococcus pyogenes. *Proc. Natl. Acad. Sci.*, **115**, 11814–11819.

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

71. Laalami,S., Zig,L. and Putzer,H. (2014) Initiation of mRNA decay in bacteria. *Cell. Mol. Life Sci.*, **71**, 1799–1828.

72. Celesnik,H., Deana,A. and Belasco,J.G. (2007) Initiation of RNA Decay in Escherichia coli by 5′ Pyrophosphate Removal. *Mol Cell*, **27**, 79–90.

73. Chou,J.Y., Singer,M.F. and McPhie,P. (1975) Kinetic studies on the phosphorolysis of polynucleotides by polynucleotide phosphorylase. *J Biol Chem*, **250**, 508–514.

74. DeLoughery,A., Lalanne,J.-B., Losick,R. and Li,G.-W. (2018) Maturation of polycistronic mRNAs by the endoribonuclease RNase Y and its associated Y-complex in Bacillus subtilis. *Proc National Acad Sci*, **115**, 201803283.

75. Hui,M.P., Foley,P.L. and Belasco,J.G. (2014) Messenger RNA Degradation in Bacterial Cells. *Annu Rev Genet*, **48**, 1–23.

76. Dorléans,A., Li de la Sierra-Gallay,I., Piton,J., Zig,L., Gilet,L., Putzer,H. and Condon,C. (2011) Molecular Basis for the Recognition and Cleavage of RNA by the Bifunctional 5′–3′ Exo/Endoribonuclease RNase J. *Structure*, **19**, 1252–1261.

77. Subramanian,S., Gao,X., Dann,C.E. and Kearns,D.B. (2017) MotI (DgrA) acts as a molecular clutch on the flagellar stator protein MotA in Bacillus subtilis. *Proc National Acad Sci*, **114**, 13537–13542.

78. Lee,V.T., Sondermann,H. and Winkler,W.C. (2022) Nano-RNases: oligo- or dinucleases? *Fems Microbiol Rev*, **46**, fuac038.

79. Goldman,S.R., Sharp,J.S., Vvedenskaya,I.O., Livny,J., Dove,S.L. and Nickels,B.E. (2011) NanoRNAs Prime Transcription Initiation In Vivo. *Mol Cell*, **42**, 817–825.

80. Bowman,L., Zeden,M.S., Schuster,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.

81. Commichau,F.M., Rothe,F.M., Herzberg,C., Wagner,E., Hellwig,D., Lehnik-Habrink,M., Hammer,E., Völker,U. and Stülke,J. (2009) Novel Activities of Glycolytic Enzymes in Bacillus subtilis INTERACTIONS WITH ESSENTIAL PROTEINS INVOLVED IN mRNA PROCESSING*. *Mol. Cell. Proteom.*, **8**, 1350–1360.

82. Gall,A.R., Hsueh,B.Y., Siletti,C., Waters,C.M. and Huynh,T.N. (2021) NrnA Is a Linear Dinucleotide Phosphodiesterase with Limited Function in Cyclic Dinucleotide Metabolism in Listeria monocytogenes. *J Bacteriol*, **204**, e00206-21.

83. Yuan,Z., Gao,F., Yin,K. and Gu,L. (2019) NrnC, an RNase D-Like Protein From Agrobacterium, Is a Novel Octameric Nuclease That Specifically Degrades dsDNA but Leaves dsRNA Intact. *Front. Microbiol.*, **09**, 3230.

84. 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.

85. Ghosh,S. and Deutscher,M.P. (1999) Oligoribonuclease is an essential component of the mRNA decay pathway. *Proc National Acad Sci*, **96**, 4372–4377.

86. Yu,D. and Deutscher,M.P. (1995) Oligoribonuclease is distinct from the other known exoribonucleases of Escherichia coli. *J. Bacteriol.*, **177**, 4137–4139.

87. Paintdakhi,A., Parry,B., Campos,M., Irnov,I., Elf,J., Surovtsev,I. and Jacobs-Wagner,C. (2016) Oufti: an integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. *Mol Microbiol*, **99**, 767–777.

88. Singer,M.F. (1958) Phosphorolysis of oligoribonucleotides by polynucleotide phosphorylase. *J Biological Chem*, **232**, 211–28.

89. Donovan,W.P. and Kushner,S.R. (1986) Polynucleotide phosphorylase and ribonuclease II are required for cell viability and mRNA turnover in Escherichia coli K-12. *Proc. Natl. Acad. Sci.*, **83**, 120–124.

90. Adams,P.P. and Storz,G. (2020) Prevalence of small base-pairing RNAs derived from diverse genomic loci. *Biochim. Biophys. Acta (BBA) - Gene Regul. Mech.*, **1863**, 194524.

91. Cummings,J.A., Vetting,M., Ghodge,S.V., Xu,C., Hillerich,B., Seidel,R.D., Almo,S.C. and Raushel,F.M. (2014) Prospecting for Unannotated Enzymes: Discovery of a 3′,5′-Nucleotide Bisphosphate Phosphatase within the Amidohydrolase Superfamily. *Biochemistry*, **53**, 591–600.

92. Cheng,Z.-F. and Deutscher,M.P. (2002) Purification and Characterization of the Escherichia coli Exoribonuclease RNase R. *J Biol Chem*, **277**, 21624–21629.

93. Tang,J. and Breaker,R.R. (1997) Rational design of allosteric ribozymes. *Chem Biol*, **4**, 453–459.

94. Ross,P., Weinhouse,H., Aloni,Y., Michaeli,D., Weinberger-Ohana,P., Mayer,R., Braun,S., Vroom,E. de, Marel,G.A. van der, Boom,J.H. van, *et al.* (1987) Regulation of cellulose synthesis in Acetobacter xylinum by cyclic diguanylic acid. *Nature*, **325**, 279–281.

95. Trinquier,A., Durand,S., Braun,F. and Condon,C. (2020) Regulation of RNA processing and degradation in bacteria. *Biochimica Et Biophysica Acta Bba - Gene Regul Mech*, **1863**, 194505.

96. 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.

97. Even,S., Pellegrini,O., Zig,L., Labas,V., Vinh,J., Bréchemmier-Baey,D. and Putzer,H. (2005) Ribonucleases J1 and J2: two novel endoribonucleases in B.subtilis with functional homology to E.coli RNase E. *Nucleic Acids Res*, **33**, 2141–2152.

98. Weiss,C.A. and Winkler,W.C. (2021) RNA Scaffolds, Methods and Protocols. *Methods Mol Biology*, **2323**, 153–170.

99. Jain,C. (2020) RNase AM, a 5′ to 3′ exonuclease, matures the 5′ end of all three ribosomal RNAs in E. coli. *Nucleic Acids Res*, **48**, 5616–5623.

100. Cohen,S.N. and McDowall,K.J. (1997) RNase E: still a wonderfully mysterious enzyme. *Mol Microbiol*, **23**, 1099–1106.

101. 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)nucleotide Degradation. *J Biol Chem*, **286**, 2807–2816.

102. Hatzios,S.K., Iavarone,A.T. and Bertozzi,C.R. (2008) Rv2131c from Mycobacterium tuberculosis is a CysQ 3'-phosphoadenosine-5'-phosphatase. *Biochemistry-us*, **47**, 5823–31.

103. Weiss,C.A., Hoberg,J.A., Liu,K., Tu,B.P. and Winkler,W.C. (2019) Single-Cell Microscopy Reveals That Levels of Cyclic di-GMP Vary among Bacillus subtilis Subpopulations. *J Bacteriol*, **201**.

104. Foley,P.L., Hsieh,P., Luciano,D.J. and Belasco,J.G. (2015) Specificity and Evolutionary Conservation of the Escherichia coli RNA Pyrophosphohydrolase RppH*. *J Biol Chem*, **290**, 9478–9486.

105. Hsieh,P., Richards,J., Liu,Q. and Belasco,J.G. (2013) Specificity of RppH-dependent RNA degradation in Bacillus subtilis. *Proc National Acad Sci*, **110**, 8864–8869.

106. Wang,F., He,Q., Su,K., Wei,T., Xu,S. and Gu,L. (2018) Structural and biochemical characterization of the catalytic domains of GdpP reveals a unified hydrolysis mechanism for the DHH/DHHA1 phosphodiesterase. *Biochem J*, **475**, 191–205.

107. 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.

108. Drexler,D.J., Müller,M., Rojas-Cordova,C.A., Bandera,A.M. and Witte,G. (2017) Structural and Biophysical Analysis of the Soluble DHH/DHHA1-Type Phosphodiesterase TM1595 from Thermotoga maritima. *Structure*, **25**, 1887-1897.e4.

109. Cheng,K., Xu,H., Chen,X., Wang,L., Tian,B., Zhao,Y. and Hua,Y. (2016) Structural basis for DNA 5′-end resection by RecJ. *eLife*, **5**, e14294.

110. Schmier,B.J., Nelersa,C.M. and Malhotra,A. (2017) Structural Basis for the Bidirectional Activity of Bacillus nanoRNase NrnA. *Sci Rep-uk*, **7**, 11085.

111. Chan,C., Paul,R., Samoray,D., Amiot,N.C., Giese,B., Jenal,U. and Schirmer,T. (2004) Structural basis of activity and allosteric control of diguanylate cyclase. *Proc. Natl. Acad. Sci.*, **101**, 17084–17089.

112. Witte,G., Hartung,S., Büttner,K. and Hopfner,K.-P. (2008) Structural Biochemistry of a Bacterial Checkpoint Protein Reveals Diadenylate Cyclase Activity Regulated by DNA Recombination Intermediates. *Mol Cell*, **30**, 167–178.

113. 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.

114. Unciuleac,M.-C., Ghosh,S., Cruz,M.J. de la, Goldgur,Y. and Shuman,S. (2021) Structure and mechanism of Mycobacterium smegmatis polynucleotide phosphorylase. *Rna*, **27**, 959–969.

115. Babitzke,P. and Kushner,S.R. (1991) The Ams (altered mRNA stability) protein and ribonuclease E are encoded by the same structural gene of Escherichia coli. *Proc. Natl. Acad. Sci.*, **88**, 1–5.

116. Deana,A., Celesnik,H. and Belasco,J.G. (2008) The bacterial enzyme RppH triggers messenger RNA degradation by 5′ pyrophosphate removal. *Nature*, **451**, 355–358.

117. Tamayo,R., Tischler,A.D. and Camilli,A. (2005) The EAL Domain Protein VieA Is a Cyclic Diguanylate Phosphodiesterase. *J Biol Chem*, **280**, 33324–33330.

118. Chou,J.Y. and Singer,M.F. (1970) The effect of chain length on the phosphorolysis of oligonucleotides by polynucleotide phosphorylase. *J Biological Chem*, **245**, 1005–11.

119. Christen,B., Abeliuk,E., Collier,J.M., Kalogeraki,V.S., Passarelli,B., Coller,J.A., Fero,M.J., McAdams,H.H. and Shapiro,L. (2011) The essential genome of a bacterium. *Mol. Syst. Biol.*, **7**, 528–528.

120. Wang,S.T., Setlow,B., Conlon,E.M., Lyon,J.L., Imamura,D., Sato,T., Setlow,P., Losick,R. and Eichenberger,P. (2006) The Forespore Line of Gene Expression in Bacillus subtilis. *J. Mol. Biol.*, **358**, 16–37.

121. Houseley,J. and Tollervey,D. (2009) The Many Pathways of RNA Degradation. *Cell*, **136**, 763–776.

122. Lehnik-Habrink,M., Pförtner,H., Rempeters,L., Pietack,N., Herzberg,C. and Stülke,J. (2010) The RNA degradosome in Bacillus subtilis: identification of CshA as the major RNA helicase in the multiprotein complex. *Mol. Microbiol.*, **77**, 958–971.

123. Amblar,M., Barbas,A., Gomez-Puertas,P. and Arraiano,C.M. (2007) The role of the S1 domain in exoribonucleolytic activity: Substrate specificity and multimerization. *Rna*, **13**, 317–327.

124. Fahmi,T., Faozia,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, SpeB Expression, and Virulence in Streptococcus pyogenes. *Infect Immun*, **87**.

125. Yang,C.-Y., Chin,K.-H., Chuah,M.L.-C., Liang,Z.-X., Wang,A.H.-J. and Chou,S.-H. (2011) The structure and inhibition of a GGDEF diguanylate cyclase complexed with (c-di-GMP)2 at the active site. *Acta Crystallogr. Sect. D: Biol. Crystallogr.*, **67**, 997–1008.

126. Cheng,Z.-F., Zuo,Y., Li,Z., Rudd,K.E. and Deutscher,M.P. (1998) The vacB Gene Required for Virulence inShigella flexneri and Escherichia coli Encodes the Exoribonuclease RNase R*. *J. Biol. Chem.*, **273**, 14077–14080.

127. McMurry,L.M. and Levy,S.B. (1987) Tn5 insertion in the polynucleotide phosphorylase (pnp) gene in Escherichia coli increases susceptibility to antibiotics. *J. Bacteriol.*, **169**, 1321–1324.

128. Jarmer,H., Berka,R., Knudsen,S. and Saxild,H.H. (2002) Transcriptome analysis documents induced competence of Bacillus subtilis during nitrogen limiting conditions. *Fems Microbiol Lett*, **206**, 197–200.

129. Sternon,J.-F., Godessart,P., Freitas,R.G. de, Henst,M.V. der, Poncin,K., Francis,N., Willemart,K., Christen,M., Christen,B., Letesson,J.-J., *et al.* (2018) Transposon Sequencing of Brucella abortus Uncovers Essential Genes for Growth In Vitro and Inside Macrophages. *Infect. Immun.*, **86**.

130. Bai,Y., Yang,J., Eisele,L.E., Underwood,A.J., Koestler,B.J., Waters,C.M., Metzger,D.W. and Bai,G. (2013) Two DHH Subfamily 1 Proteins in Streptococcus pneumoniae Possess Cyclic Di-AMP Phosphodiesterase Activity and Affect Bacterial Growth and Virulence. *J Bacteriol*, **195**, 5123–5132.

131. Manikandan,K., Sabareesh,V., Singh,N., Saigal,K., Mechold,U. and Sinha,K.M. (2014) Two-Step Synthesis and Hydrolysis of Cyclic di-AMP in Mycobacterium tuberculosis. *Plos One*, **9**, e86096.

132. Millikin,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.

133. Srivastav,R., Kumar,D., Grover,A., Singh,A., Manjasetty,B.A., Sharma,R. and Taneja,B. (2014) Unique subunit packing in mycobacterial nanoRNase leads to alternate substrate recognitions in DHH phosphodiesterases. *Nucleic Acids Res*, **42**, 7894–7910.

134. Crooks,G.E., Hon,G., Chandonia,J.-M. and Brenner,S.E. (2004) WebLogo: A Sequence Logo Generator. *Genome Res*, **14**, 1188–1190.

135. Mechold,U., Fang,G., Ngo,S., Ogryzko,V. and Danchin,A. (2007) YtqI from Bacillus subtilis has both oligoribonuclease and pAp-phosphatase activity. *Nucleic Acids Res*, **35**, 4552–4561.

136. Rao,F., See,R.Y., Zhang,D., Toh,D.C., Ji,Q. and Liang,Z.-X. (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.

137. The UniProt Consortium, Bateman,A., Martin,M.-J., Orchard,S., Magrane,M., Agivetova,R., 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.

138. Crystal structure of an uncharacterized protein from Helicobacter pylori G27. Abendroth, J., Horanyi, P.S., Lorimer, D.D., Edwards, T.E. (2018) Protein Data Bank.

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

140. Bagyan,I., Hobot,J. and Cutting,S. (1996) A compartmentalized regulator of developmental gene expression in Bacillus subtilis. *J. Bacteriol*., **178**, 4500–4507.

141. Nicholson,W.L. and Setlow,P. (1990) Sporulation, germination, and outgrowth. In: Harwood,C.R. and Cutting,S.M. (eds.) *Molecular Biological Methods for Bacillus.* John Wiley and Sons, Sussex, England, pp. 391–450.

142. Pettersen,E.F., Goddard,T.D., Huang,C.C., Meng,E.C., Couch,G.S., Croll,T.I., Morris,J.H. and Ferrin,T.E. (2021) UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.,* **30**, 70–82.

143. Weiss,C.A., Myers,T.M., Wu,C.H., Jenkins,C., Sondermann,H., Lee,V.T. and Winkler,W.C. (2022) NrnA is a 5′-3′ exonuclease that processes short RNA substrates in vivo and in vitro. *Nucleic Acids Res*, **50**, 12369–12388.

144. Liu,B., Deikus,G., Bree,A., Durand,S., Kearns,D.B. and Bechhofer,D.H. (2014) B. subtilis mRNA processing and effect of PNPase deletion. *Mol Microbiol*, **94**, 41–55.

145. Mastny,M., Heuck,A., Kurzbauer,R., Heiduk,A., Boisguerin,P., Volkmer,R., Ehrmann,M., Rodrigues,C.D.A., Rudner,D.Z. and Clausen,T. (2013) CtpB Assembles a Gated Protease Tunnel Regulating Cell-Cell Signaling during Spore Formation in Bacillus subtilis. *Cell*, **155**, 647–658.

146. Cohen,D., Mechold,U., Nevenzal,H., Yarmiyhu,Y., Randall,T.E., Bay,D.C., Rich,J.D., Parsek,M.R., Kaever,V., Harrison,J.J., et al. (2015) Oligoribonuclease is a central feature of cyclic diguanylate signaling in Pseudomonas aeruginosa. *Proc National Acad Sci*, **112**, 11359–11364.

147. Casinhas,J., Matos,R.G., Haddad,N. and Arraiano,C.M. (2018) Biochemical characterization of Campylobacter jejuni PNPase, an exoribonuclease important for bacterial pathogenicity. *Biochimie*, **147**, 70–79.

148. 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.

149. Yang,J., Bai,Y., Zhang,Y., Gabrielle,V.D., Jin,L. and Bai,G. (2014) Deletion of the cyclic di-AMP phosphodiesterase gene (cnpB) in Mycobacterium tuberculosis leads to reduced virulence in a mouse model of infection. *Mol Microbiol*, **93**, 65–79.

150. Camacho,A. and Salas,M. (1999) Effect of mutations in the "extended −10" motif of three Bacillus subtilis σA-RNA polymerase-dependent promoters11Edited by I. B. Holland. *J Mol Biol,* **286**, 683-693

151. Mandal,R.K., Jiang,T. and Kwon,Y.M. (2017) Essential genome of Campylobacter jejuni. *Bmc Genomics,* **18**, 616