WORCESTER POLYTECHNIC INSTITUTE

Investigating the Cleavage Specificity of
Mycobacterium smegmatis **RNase E** *in vitro*

A Major Qualifying Project
Authored by:

Joseph Dainis & Alexa Davis

Date Submitted:

May 17th, 2020

Submitted To:
Dr. Scarlet S. Shell, Ph.D. - Department of Biology & Biotechnology, advisor
Dr. Louis A. Roberts, Ph.D. - Department of Biology & Biotechnology, advisor
Dr. José M. Argüello, Ph.D. - Department of Chemistry & Biochemistry, co-advisor

This Major Qualifying Project is submitted to the Department of Biology & Biotechnology and the Department of Chemistry and Biochemistry at Worcester Polytechnic Institute for the partial fulfillment of the requirements for the Degrees of Bachelor of Science in Biology & Biotechnology and Bachelor of Science in Biochemistry.

Globally, there are an estimated 15 million active cases of tuberculosis, a bacterial disease caused by *Mycobacterium tuberculosis* that primarily affects the lungs. Upon infection, the bacterium is encapsulated within the granuloma lesion, where it is subject to many forms of stress. Therefore, the success of *M. tuberculosis* as a pathogen relies on tight regulation of gene expression related to growth, metabolism, and survival. One strategy the bacterium employs to achieve this is decreasing the rate of mRNA degradation, a process believed to be regulated by the activity of endoribonuclease E (RNase E), an essential enzyme in mycobacteria that cleaves mRNA and serves as the major catalytic and scaffolding element of a multi-protein complex known as the degradosome. Though thoroughly studied in *E. coli*, RNase E in *M. tuberculosis* remains relatively unexplored. To better understand the activity of mycobacterial RNase E we engineered, expressed, and purified several variants of the enzyme derived from *M. smegmatis*, a non-pathogenic species used as a model for *M. tuberculosis*. These enzymes were then used to investigate the cleavage specificity and putative autoregulation of the *rne* mRNA transcript. Due to similarities in observable RNA fragment sizes from the *in vitro* RNase E activity assay performed in this research and those predicted from cleavage sites previously mapped *in vivo*, there is evidence to suggest that the cleavage pattern seen for the *M. smegmatis rne* 5' UTR *in vivo* may be attributed to the activity of RNase E.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

INTRODUCTION

The World Health Organization (WHO) estimates that nearly one quarter of the global population is infected with *Mycobacterium tuberculosis*, the causative agent of tuberculosis (TB), a bacterial disease primarily affecting the lungs (WHO, 2019). While incidence of the disease is present on every continent, TB is most prevalent in developing countries, with nearly 70% of all cases worldwide occurring in regions of southern Africa and Southeast Asia (MacNeil et al., 2017). Individuals at highest risk for infection are those who are malnourished, HIV-positive, or otherwise immunologically compromised (WHO, 2017). Most infections are latent and asymptomatic; a person may live their entire life unknowingly a carrier of the disease. However, at any given time there are approximately 15 million active cases of TB, around 558,000 per year of which emerge as multidrug resistant (MDR) and unresponsive to front line TB drugs including isoniazid and rifampicin (Park, Satta, & Kon, 2019). Of MDR cases, around ten percent are expected to be extensively drug resistant (XDR), unresponsive to second line drugs as well (World Health Organization, 2019).

A key characteristic of TB disease is the formation of granulomas, nodular masses of inflamed tissue, within the host lungs. Upon infection, the bacterium is phagocytosed by alveolar macrophages, leading to a release of chemokines that recruit additional immune cells to the site as part of an inflammatory response (Ehlers & Schaible, 2012). This response then results in remodeling of the lung tissue that ultimately encapsulates the pathogen within the granuloma lesion. There, *M. tuberculosis* is subject to many forms of stress, including nutrient starvation, low pH, hypoxia, and if the individual is receiving TB treatment, antibiotic-related stress (Kiran et al., 2015).

The stressful nature of the granuloma microenvironment cues the pathogen to undergo changes in gene expression related to growth, metabolism, and survival. Thus, the long-term survival of *M. tuberculosis* within the lungs relies heavily on tight regulation of gene expression. Under physiologically normal circumstances, mRNA is degraded quickly, but in response to stressors, there is a global stabilization of the transcriptional profile (Rustad et al., 2013). Slowing of the rate of mRNA degradation (increasing the transcript half-life) favors conservation of energy in the metabolically unfavorable conditions of the granuloma, promoting survival. While there is not yet evidence of transcript-specific stabilization in mycobacteria, there is evidence in *E. coli* that some mRNAs are stabilized more than others under different growth conditions (Esquerré et. al, 2014). There are several plausible explanations for the mechanisms of global transcript stabilization, many related to the activity or abundance of ribonucleases (RNases); enzymes that facilitate the degradation of RNA into smaller fragments. Though this process is not yet fully understood in mycobacteria, it has been well demonstrated in *E. coli* that RNase E is a key enzyme involved in regulation of mRNA degradation.

RNase E is an endoribonuclease that functions in many bacteria as both the major catalytic and scaffolding unit of the degradosome, a large multi-enzyme complex that facilitates bulk mRNA degradation as well as pre-rRNA and pre-tRNA processing (Berstein et al., 2004). The scaffolding region of RNase E in *E. coli* is an intrinsically disordered domain that provides a platform upon which the core degradosome proteins and other accessory molecules that aid in degradation may assemble (Carpousis, 2002). The main components of the *E. coli* degradosome have been identified as PNPase, enolase, and Rh1B (Carpousis, 2007). The scaffolding region, through coordination of Zn^{2+}, also facilitates the assembly of the enzyme into a homotetrameric quaternary structure. (Callaghan et. al, 2005).

In *E. coli,* it has been found that RNase E will cleave single stranded RNA high in A+U content (McDowell et. al, 1994), and it is believed the enzyme is influenced more by this feature than a particular sequence or distance from the 5' end. A study mapping cleavage sites in *E. coli* reveals an approximate consensus sequence of G/A - N ↓ A/U - U - U, with N being any nucleotide (Chao et al., 2017), emphasizing the importance of the A+U content of the transcript in determining cleavage by RNase E. Substrate preference and mechanism of recognition by RNase E are also critical in determining the cleavage activity of the enzyme. A number of studies have found that in *E. coli,* RNase E will preferentially cleave 5'monophosphorylated transcripts over those that are 5' tri-phosphorylated (fully capped) and those that harbor a 5' hydroxyl group (Mackie, 1998; Richards & Belasco, 2015). It has been demonstrated both *in vitro* and *in vivo* that the removal of a pyrophosphate from the 5' end of tri-phosphorylated primary transcripts by RNA pyrophosphate hydrolase (RppH) creates the 5' monophosphate that initiates degradation by RNase E (Deana, A., Celesnik, H. & Belasco, J., 2008). 5' monophosphates are recognized by a sensing domain whose network of intermolecular forces increases affinity for the terminal phosphate of the RNA transcript and allows for its proper positioning within the catalytic domain for cleavage (Callaghan et. al, 2005). Despite this preference, however, it is shown that in *E. coli* RNase E there are multiple intrinsic mechanisms of substrate recognition that facilitate both 5' end dependent and independent cleavage. A proposed second method of substrate recognition, involving the bypass of the 5' sensing pocket and interaction with arginine-rich regions of the catalytic domain, may explain cleavage of RNA transcripts that lack a 5' monophosphate (Bandyra, K., Wandzik, J., & Luisi, B., 2018).

The mechanism of RNA cleavage carried out within the catalytic domain of RNase E has been found to be largely conserved in previously studied species like *E. coli*, *H. influenzae,* and *S. coelicolor* (Caruthers et al., 2006). Given the strong sequence homology within the catalytic domains of *E. coli* and *M. tuberculosis*, there is reason to believe the cleavage mechanism is conserved in mycobacteria as well. In *E. coli,* the catalytic activity of RNase E has been determined to take place within a DNase I-like subdomain of the catalytic region, where two aspartic acid residues facilitate cleavage (Callaghan et. al, 2005). To cleave an RNA substrate, the negatively

charged residues coordinate a divalent metal cation, often magnesium (Mg^{2+}). However, several *in vitro* studies have shown that manganese (Mn^{2+}) can act as an acceptable substitute for magnesium, resulting in comparable cleavage activity (Thompson, Zong, & Mackie, 2015). Though there generally seems to be a preference for magnesium in *in vitro* studies, it is not yet known if this preference exists *in vivo* (Mackie, 2013). Coordination of the divalent cation activates a hydroxyl group within the hydration shell that surrounds it, which can then act as a nucleophile in hydrolytic cleavage of the phosphate backbone of the RNA transcript (Callaghan et. al, 2005).

Though it is likely that the mechanism of cleavage is conserved, there is existing research to suggest that this may not be true for the domain structure, binding partners, and cleavage specificity of RNase E when comparing *E. coli* and mycobacteria. It is shown in previous studies that there is usually variability by species in the noncatalytic elements and binding partners of RNase E (Górna, Carpousis, & Luisi, 2012), all of which may affect the activity and cleavage specificity of the enzyme. With respect to mycobacterial RNase E, it differs from its counterpart in *E. coli* in that the catalytic domain, rather than lying next to a single C-terminal scaffold domain, is flanked by putative N- and C-terminal scaffolding domains (Kovacs et. al, 2005). In addition to a difference in domain structure, RNase E in mycobacteria has also been suggested to associate with different sets of accessory proteins relative to *E. coli*. These proteins are thought to be PNPase, RhIE, and RNase J in *M. tuberculosis* (Płociński et al., 2019), and were reported to be DIM1, EF-Tu, GroEL, MecA, relA, and rplB in *M. smegmatis* (Csanadi et al., 2009). Despite these differences in the scaffolding domain and associated proteins, the same preferential cleavage of 5' monophosphorylated transcripts and Zn^{2+} mediated tetramerization has also been observed in RNase E in *M. tuberculosis* as was previously described in *E. coli* (Zeller et. al, 2007).

For mycobacterial RNase E, there is inconsistent reporting with regards to cleavage specificity. An analysis of the activity and biochemical properties of RNase E in *M. tuberculosis* showed a U-rich dependency with favorable cleavage occurring downstream of a guanine nucleotide (Zeller et. al, 2007). However, a recent study mapping cleavage sites of the *M. smegmatis* genome found a +1 C (N↓C) cleavage pattern for 90% of the 3,344 high confidence cleavage sites that were analyzed (Martini et. al, 2019). Because of its key role in mRNA decay and maturation in many other bacterial species, there is reason to believe that RNase E is responsible for the observed pattern of cleavage. The same study found that transcriptional downregulation of RNase E in *M. smegmatis* leads to a global increase in mRNA half-life, emphasizing that the enzyme plays a large role in mRNA decay and stability.

Guided by previous research that addresses the domain structure and putative cleavage specificity of mycobacterial RNase E, this study aimed to further investigate the pattern of RNA cleavage and the kinetic parameters of substrate hydrolysis, divalent metal ion preference, and the extent to which the scaffolding domain plays a role in the cleavage activity of the enzyme. To

achieve this, we engineered the MSMEG_4626 coding sequence to produce three different variants of *M. smegmatis* RNase E harboring different length N-terminal truncations and a catalytically dead variant. Once expressed in *E. coli*, these constructs were then purified and concentrated for use in *in vitro* RNA cleavage assays. With these constructs, we aimed to learn more about the relationship between the scaffold domain and cleavage activity, which may identify regions of importance in the cleavage mechanism itself or the association of RNase E with binding partners that assemble into the mycobacterial degradosome. By studying the cleavage activity and specificity as well as metal ion coordination, we aimed to learn more about the biochemical properties of the enzyme.

METHODS

Designing RNase E Variants

Three RNase E variants (amino acids (aa) 331-824 full N-terminal truncation, 146-824aa catalytically active partial N-terminal truncation, and 146-824aa catalytically dead partial N-terminal truncation) were designed for this series of cleavage experiments. pSS348, carrying the *M. smegmatis rne* coding sequence with a Δ 1-145aa partial N-terminal deletion, Δ 825 - 1037aa full C-terminal deletion and an N-terminal addition of 6X polyhistidine tag (HIS), TEV protease cleavage site, and Gly linker sequences, was used as a template for creation of pSS420 (Δ 1-145aa catalytically active). pSS420 was then used as a template for creation of subsequent constructs, pSS421(Δ 1-145aa catalytically dead) and pSS459 (Δ 1-330aa catalytically active).

Due to putative differences in domain structure among bacteria, a sequence alignment was performed to determine the level of homology between the wild type RNase E amino acid sequences of *E. coli*, *M. tuberculosis*, *M. smegmatis*, and a closely related species *C. crescentus*. This alignment was performed using the Clustal Omega multiple sequence alignment tool available through EBI at https://www.ebi.ac.uk/Tools/msa/clustalo/.

The results of the sequence alignment (Fig.1) were then used to inform primer design for the 146-824aa catalytically dead variant to introduce the desired point mutations to the *M. smegmatis rne* coding sequence. Benchling informatics software was used to map and design primers, and the New England Biolabs Primer Analysis Tool was used to determine primer melting temperatures for PCR. The pET42 plasmid, which harbors an inducible T7 promoter system, was used in these experiments. All primers used in this research and their purposes are summarized in Table 1. All plasmids, their purposes and antibiotic markers are summarized in Table 2. All constructs were sequenced to confirm the success of point mutations and truncations. Primers and sequencing reactions were ordered through Eton Bioscience Inc.

Table 1 - Primers. All primers used during cloning, their sequences, and descriptions.

Primer Number	Sequence	Purpose
SSS1841	GCACCGAGGCCATGACCGTCGTCCGCGTCAACACCGGCAAG	Forward primer to amplify RNase E to construct D694R and D737R mutations in *M. smegmatis rne* catalytic domain
SSS1842	GATTCCAGGACCATGTCGATGAAGCGGATGACGACGATGCC	Reverse primer to amplify RNase E to construct D694R and D737R mutations in *M. smegmatis rne* catalytic domain
SSS1843	CGACATGGTCCTGGAATCCAACCGCG	Forward primer to HiFi RNase E D694R and D737R mutations into pET42

		backbone.
SSS1844	CGGTCATGGCCTCGGTGCGG	Reverse primer to HiFi RNase E D694R and D737R mutations into pET42 backbone
SSS1995	TTAAGAAGGAGATATACAATGCACCACC ACCACCACCACGATTACAAG	Forward primer to create 146-824aa full N terminal deletion of *M. smegmatis rne* coding sequence
SSS1996	TGCTCGAGTGCGGCCGCACTACGAGTC GGACTTGCGCCC	Reverse primer to create 146-824aa full N terminal deletion of *M. smegmatis rne* coding sequence
SSS1997	GACAAGAGCGACGACTCCGAGATC	Forward primer to create 331-824aa full N terminal deletion of *M. smegmatis rne* coding sequence
SSS1998	CGTCGCTCTTGTCGCCGCCGCCGCCCT GGAA	Reverse primer to create 331-824aa full N terminal deletion of *M. smegmatis rne* coding sequence

Table 2 - Plasmids. All plasmids used in this study, their descriptions, and antibiotic selection markers.

Plasmid Name	Description	Antibiotic
pSS348	pJEB402 nat reverse plasmid with Δ1-145aa and Δ825 - 1037aa deletions of RNase E	Nourseothricin
pET42	Plasmid harboring inducible T7 promoter	Kanamycin
pSS420	Partially N-terminal truncated (146-824aa) catalytically active RNase E inserted into pET42	Kanamycin
pSS421	Partially N-terminal truncated (146-824aa) catalytically dead RNase E inserted into pET42	Kanamycin
pSS459	Full N-terminal truncated (331-824aa) catalytically active Rnase E inserted into pET42	Kanamycin

High Fidelity Polymerase Chain Reaction (PCR)

All PCR was performed following the New England Biolabs recommended protocol, using Q5 High Fidelity DNA polymerase. Q5 polymerase was chosen over Taq polymerase for the purpose of avoiding introduction of unintended mutations in PCR products. All PCR reactions began with an initial denaturation for 2 minutes at 98°C. This was followed by thirty cycles of

denaturation (95℃ for 30 seconds), annealing (60℃ for 30 seconds), and extension (72℃ for 30s/kb)] and a final extension for 2 minutes at 72℃. Six sets of primers were used to create the three constructs harboring modified *rne* coding sequences (Table 1).

DpnI Digestion

DpnI digestion was performed to remove the remaining methylated plasmid template leaving only amplified unmethylated PCR products intact. All PCR reactions were treated with 1μL of DpnI and incubated at 37℃ for 60 minutes. The Dpn1 enzyme was then heat inactivated at 80℃ for 20 minutes. All PCR products were stored at -20℃.

Gel Electrophoresis and Extraction

Agarose gel electrophoresis was performed to confirm the success of each PCR reaction. All PCR samples were mixed with 6X DNA sample buffer, loaded onto a 1% agarose-TBE gel, and run for 1 hour at 90V. Bands of appropriate size were excised and purified following the instruction of the Macherey-Nagel gel clean-up kit. All DNA was eluted in 30μL of nuclease free water. Eluent concentration and purity were measured using a Nanodrop spectrophotometer.

NdeI/HinDIII Restriction Digest and Fragment Assembly

Before insertion of the 146-824aa *rne* sequence amplified from pSS348 into the pET42 cloning vector to create pSS420, the pET42 plasmid was first linearized by an NdeI and HinDIII restriction digest. Assembly was performed following the New England Biolabs recommended NEBuilder® HiFi DNA Assembly protocol. Two fragment assemblies of plasmid and insert were used to create pSS420 and pSS421. For these assemblies, 2.5ng (1.4μL) of insert was combined with 50ng (1.7μL) of plasmid, 1.9μL of ddH2O, and 5μL of NEBuilder® HiFi Assembly Master Mix (REF #E2621S), then the reaction was incubated at 50℃ for 60 minutes. pSS459 was generated by a single fragment assembly using the same reaction conditions as mentioned above.

DH5α and BL21(DE3) pLysS E. coli Transformation

DH5α *E. coli* competent cells (NEB 5α, New England BioLabs REF #C2987I) were used in cloning the RNase E variants. All competent cells were selected in kanamycin (Kan) supplemented medium (30μg/mL). 2 uL of assembly reaction were added directly to 50μL of cells and mixed by gentle rocking. The cells were incubated on ice for 15 minutes, heat shocked for 45 seconds at 42℃, and then placed on ice for 2 minutes. 500μL of SOC recovery medium was then added and the cells were incubated at 37℃ on a shaker for 1 hour. Cells were pelleted with a short spin and resuspended in a final volume of 300μL of SOC medium. The 300μL resuspension was

spread on two plates: one plated with 50μL of cells and one with 250μL of cells. Plates were incubated overnight at 37°C.

From each plate, 3-4 individually defined colonies were picked and grown in 10mL LB + Kan for 24 hours at 37°C and mini-prepped following the Plasmid DNA purification protocol found in the manual for the Macherey-Nagel Nucleospin kit (REF #740609.50). All DNA was eluted in 30μL of nuclease free water and assessed for concentration and purity using a Nanodrop spectrophotometer. Sequences were confirmed and then these plasmids were used to transform the BL21(DE3) pLysS *E. coli* (Invitrogen™, REF #C606010) expression strain. This strain contains an integrated DE3 lysogen, constitutively expressing the T7 polymerase, and a pLysS plasmid which confers chloramphenicol resistance and constitutively expresses low levels of the T7 lysozyme, allowing for minimal uninduced expression of the target gene.

Transformation and plating were performed using the same protocol as previously described. 3-4 individually defined colonies were picked and grown in 10mL liquid culture (LB + Kan + chloramphenicol (Cm) (25μg/mL)) overnight shaking at 37°C, then stored at 4°C until expression and purification could be completed.

Expression of RNase E Variants

The procedure for expression of RNase E variants was adapted from the New England Biolabs protocol for protein expression using BL21(DE3) cells. For each construct, sequence confirmed colonies were picked and used to initiate starter liquid cultures. For expression and harvest of RNase E, a liquid culture was inoculated by adding 20μL BL21 *E. coli* starter culture to 15mL of LB + Kan + Cm. Cultures were incubated overnight shaking at 37°C.

On the day of expression, 10mL of starter culture was added to 1 liter of LB + Kan + Cm. This culture was then incubated shaking at 37°C until an OD600 of ~0.5 was attained. 25mL of uninduced culture was pelleted at 3,000 x g for 10 minutes at 4°C and saved for analysis. To induce expression of RNase E, 500μL of 0.8M IPTG was added to a working concentration of 400μM and incubated shaking at 28°C for four hours. 25mL of induced culture was pelleted and saved for analysis. The remaining culture volume was pelleted, weighed, and stored at 4°C until purification.

Cell Lysis and Clarification of Extract for Purification

For each expression, a total pellet weight of approximately 5 - 8g was resuspended in 5mL of 1X IMAC buffer (Supplemental Material S2) containing 10mM imidazole and pooled into a 50mL conical tube. The cell resuspension was then lysed using a BioSpec Tissue-Tearor in 10 cycles (15 seconds at maximum speed, then a 60 second incubation on ice). The homogenized lysate was then aliquoted into 1.5mL microcentrifuge tubes and centrifuged at maximum speed for 15 minutes at 4°C to pellet cell debris. The supernatant was pipetted from each microcentrifuge

tube and pooled, with a final volume of approximately 12 - 15mL. 500μL of this crude extract was saved for analysis.

Immobilized Metal Affinity Chromatography (IMAC)

IMAC was used to separate each N-terminal 6x His-tagged RNase E variant from unwanted proteins. For convenience, two IMAC columns were packed and run in parallel to facilitate purification of the total volume of crude extract. 2mL total of His-Pur Ni-NTA resin (Thermo Scientific, REF #88223) was mixed with the clarified protein extract (12 - 15mL) and incubated on ice for 30 minutes to allow binding of the 6X HIS tag to the Ni^{2+} of the resin. The resin/extract mixture was split equally between two polypropylene columns, packed to a height of approximately 5cm, and flow through (approximately 5 - 6mL each) was collected upon removal of the end cap. The following series of buffers were applied to each column and collected: 2mL of 10 mM imidazole to wash, 4mL 150 mM imidazole to elute, and 2mL 500 mM imidazole to strip. A total of 8mL of eluent was collected, with 1mL saved for analysis. All buffers also contained final concentrations of 20mM Tris base, 150mM NaCl, 5% v/v glycerol, and 0.01% v/v IGEPAL. All step-by-step IMAC buffer recipes can be found in Supplemental Material (S1-S3; S5).

The RNase E eluent was concentrated from 7mL to approximately 400μL using Microcon PL-30 (30,000 NMWL) protein concentrators (Millipore Sigma, REF #MRCF0R030). The concentrators were spun in 15 minute intervals at maximum speed (14,500 rpm) at 4°C until the desired volume of 400μL was reached (2 - 3 hours). 100μL of concentrated elute were saved for analysis.

Size Exclusion Chromatography (SEC)

The concentrated eluent was run through SEC to further refine RNase E variants. The 0.75 inch diameter size exclusion column was packed with 38 mL of Sephacryl S-200 High Resolution resin (GE Healthcare, REF #17058401) to a height of approximately 17 inches. 300μL of concentrated elute was applied to the top of the column and ran through the column at 20μL/min (slowest speed). The flow rate was regulated using a Masterflex C/L pump, and a fraction collector was used to collect 90 - 100 300μL fractions per run. All fractions were transferred to 1.5mL microcentrifuge tubes and stored at 4°C for analysis and concentration. The SEC buffer contained final concentrations of 2mM Tris base, 150mM NaCl, 5% v/v glycerol, 0.01% v/v IGEPAL, and 1mM EDTA. All step-by-step SEC instructions and buffer recipes can be found in Supplemental Material (S4, S5).

Analysis of Fractions

Saved fractions from SEC were analyzed to determine in which fractions RNase E variants eluted, and therefore which fractions to pool and concentrate for use in RNA cleavage assays. Denaturing protein gel electrophoresis was used to analyze the relative purity, abundance, and molecular weight of the RNase E variants. Samples were prepared for the SDS-PAGE gels (10% resolving) (BioRad, REF #4561035) by combining 20μL of fraction with 5μL of 5X SDS-PAGE Protein Loading Buffer and heating for 5 minutes at 90°C. After running for 60 minutes at 150V, gels were stained with GelCode™ Blue Stain Reagent (REF #24592) following the Thermo Fisher Scientific microwave procedure to visualize the protein. Gels were destained rocking in deionized H$_2$O and imaged using a BioRad ChemiDoc XRS+ Gel Documentation System.

Total protein assays were done to determine protein concentrations of purified fractions. Fractions were analyzed in a two-fold dilution series and compared to a 0.5mg/mL BSA standard. To each well, 200μL of Dilute Coomassie (Bradford) Protein Assay Reagent (Thermo Fisher Scientific, REF #23200) was added, and the plate was read at 595 nm.

In Vitro Transcription

The native *M. smegmatis rne* 5' UTR + 45nts *rne* coding sequence (See Supplemental Material S6) used in the RNase E activity assay was generated through *in vitro* transcription using the New England BioLabs HiScribe™ T7 High Yield RNA Synthesis Kit (REF #E2040S). A master mix containing nuclease free water, 2μL of 10X reaction buffer, ATP, CTP, GTP, and UTP concentrations of 10μM, 1μg of *M. smegmatis rne* template DNA, and 2μL of T7 RNA polymerase was assembled, mixed thoroughly, and incubated overnight at 37°C. The reaction was diluted by adding 70μL of RNase-free water and 10μL of DNase I buffer, then treated with 2μL of DNase I and incubated at 37°C for 15 min. to remove the remaining template.

In vitro transcribed RNA was purified using a Zymo Research Clean & Concentrator-25 RNA purification kit (REF #R1018) according to the manufacturer's instructions.

RNase E Activity Assay

For this assay, 360ng of substrate was used per reaction, combined with 1.25X reaction buffer to a total volume of 8μL. The mRNA substrate was first heated to 65°C for three minutes to denature secondary structure, then at 37°C for one minute. At T= 0min, 2μL of purified enzyme (~80ng) was added, and the reactions were incubated at 37°C for 60 minutes. To stop the reaction, 1μL of a 0.1M EDTA solution was added, followed by 11μL of 2X Loading Buffer II (Thermo Fisher Scientific, REF #AM8546G). Reactions were incubated at 65°C for three minutes immediately prior to analysis to denature any residual secondary structure. A step-by-step protocol for this assay can be found in Supplemental Material S7.

Visualization and Analysis of mRNA Degradation Products

To visualize fragment size and abundance of cleavage products, $5\mu L$ from each reaction were loaded into a BioRad Mini- PROTEAN Urea-TBE 10% denaturing nucleic acid gel (REF #4566036) and run for one hour at 100V. The gel was stained in 1X SYBR Gold Nucleic Acid Gel (Thermo Fisher, REF #S11494) stain for 15 minutes and then imaged using a BioRad ChemiDoc XRS+ Gel Documentation System.

RESULTS

Designing A Catalytically Dead RNase E Variant

Design of the 146-824aa catalytically dead variant (pSS421) was inspired by an experiment performed in *E. coli,* where it was shown that site-directed mutagenesis to introduce D → R changes of catalytic residues abolishes cleavage activity by impeding metal ion coordination (Bandyra et. al, 2018). To determine the level of homology between the catalytic domains of RNase E from *E. coli* and *M. smegmatis*, a Clustal omega alignment of the full *rne* amino acid sequence was performed (Fig. 1). The sequence alignment revealed that the aspartic acid residues involved in coordination of divalent metal cations in the active site of *E. coli* RNase E (D303 and D346) are conserved in *M. tuberculosis, M. smegmatis,* and a closely related species *C. crescentus*.

Figure 1 - Clustal Omega Alignment of *rne* Amino Acid Sequence. The amino acid sequence of the full length *rne* from *E. coli, M. tuberculosis, M. smegmatis*, and *C. crescentus* were aligned using an EMBL-EBI Clustal Omega Multiple Sequence Alignment tool.
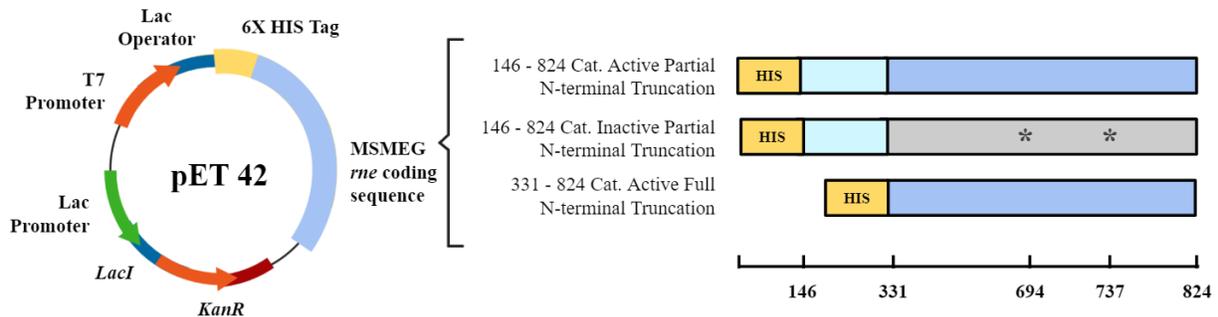
Figure 2 - pET42 Plasmid Construct and RNase E Variant Domain Structures. All RNase E variants were cloned into a pET42 plasmid system (left), which utilizes the T7 promoter system, to induce expression of the His-tagged constructs. Three RNase E variants were designed and engineered using this system (right), including a 146-824aa partial N-terminal truncated catalytically active, 331-824aa full N-terminal truncated catalytically active, and a 146-824aa partial N-terminal truncated catalytically dead.

With this information, a primer set was designed (Table 1) for site directed mutagenesis of the codons corresponding to the catalytic aspartic acid residues within the *M. smegmatis* RNase E active site, mimicking a D → R mutagenesis experiment performed in *E. coli* (Bandyra et. al, 2018). Additional primer sets were designed to truncate the N-terminus of *M. smegmatis* RNase

E, generating a construct containing either 146 - 824aa (partial truncation), or 331-824aa (full deletion) (Fig. 2).

To confirm the success of site-directed mutagenesis and N-terminal truncations, each plasmid was mini-prepped and sequenced. These sequences were aligned to the plasmid map used to design each construct in Benchling. An example of such an alignment, 146 - 824aa catalytically dead, is shown in Figure 3.
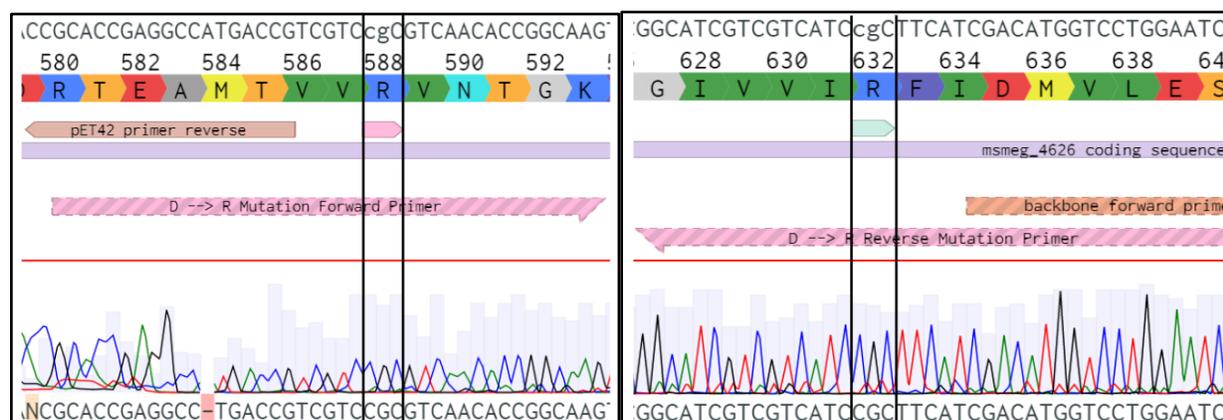
Figure 3 - Sequence Alignment for 146 - 824aa Catalytically Dead RNase E. Sequencing results were compared to plasmid designs using the consensus alignment tool in Benchling.

IMAC Purification of RNase E Variants

Once modified *rne* cloning was completed, all RNase E variants were expressed in BL21 (DE3) pLysS *E. coli* and purified first by immobilized metal affinity chromatography (IMAC), and then by size exclusion chromatography (SEC).

All RNase E variants were tagged with a 6X polyhistidine (HIS) tag so they could be purified with specificity from unwanted cell debris and proteins by IMAC. After binding of RNase E variants to the nickel-NTA resin, increasing concentrations of imidazole (a competitive inhibitor of the HIS tag) are used to first wash out unwanted proteins, then elute RNase E. A pilot test using a series of buffers increasing in imidazole concentration was performed to determine the appropriate concentration for optimal RNase E elution. It was determined that a 20mM imidazole concentration is sufficient to wash most of the unwanted proteins that remained in the IMAC column after packing, while the majority of bound RNase E protein elutes at a concentration range of 50-150mM imidazole (Fig. 4). Informed by this pilot experiment, buffers with 20mM and 150mM imidazole were used to wash the column and elute RNase E, respectively.

Five fractions were collected during each IMAC purification of RNase E variants: crude extract (CE), flow through (FT), wash (W), elution (E), and strip (S). CE is expected to contain both RNase E and a high concentration of unwanted *E. coli* proteins of a wide range of molecular
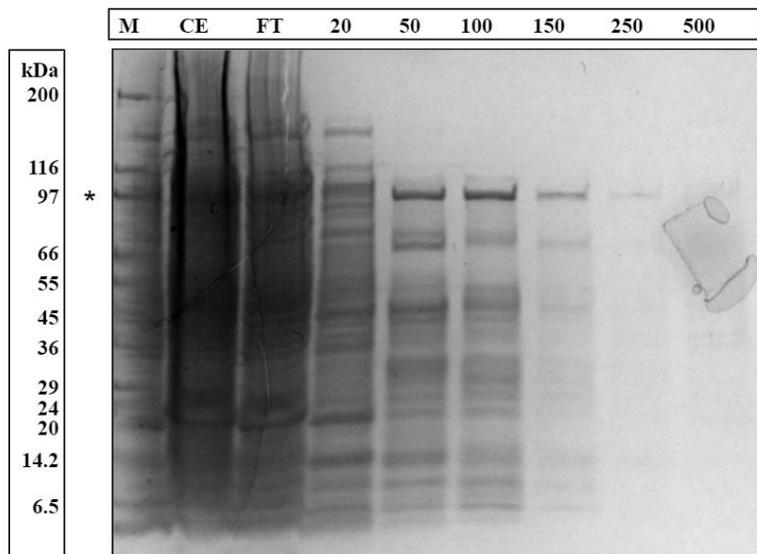
Figure 4 - SDS-PAGE Gel Pilot Test IMAC Fractions With Increasing Imidazole Concentrations. Molecular weight standard (M), crude extract (CE), flowthrough (FT), and elution fractions from increasing imidazole concentrations were loaded and run on an SDS-Page denaturing gel. The position of the presumed RNase E band in is denoted on the left by an asterisk (*).
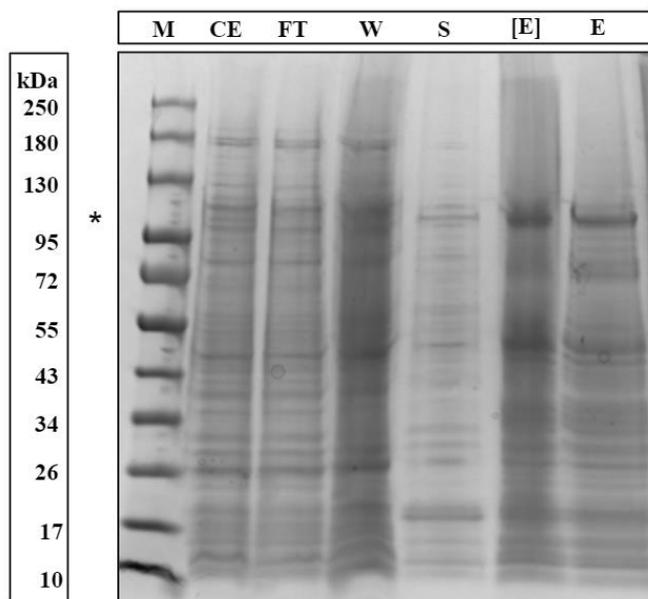
Figure 5 - SDS-PAGE Gel 146 - 824aa Catalytically Active **Partial Truncated RNase E IMAC Fractions.** Molecular weight standard (M), crude extract (CE), flowthrough (FT), wash (W), strip (S), elute (E), and concentrated elute ([E]) were loaded and run on an SDS-PAGE denaturing gel. The position of the assumed RNase E band is denoted on the left by an asterisk (*).

weights. FT is expected to be visually similar to CE when separated on a gel, with the exception of the absence of RNase E, which is now bound to the IMAC column. Unwanted proteins are removed in W, which is anticipated to contain a large number of *E. coli* proteins of varying molecular weights. Elution fractions are expected to be more refined, with fewer unwanted proteins and a defined band of presumptive RNase E protein. The final fraction, strip, where all remaining proteins from the column that had failed to elute in the previous steps are removed, would be expected to again show a very small amount *E. coli* proteins of a range in size as well as any RNase E that failed to elute during the previous step.

Purification by IMAC of the 146-824aa catalytically active RNase E variant showed a high level of expression and effective elution with 150mM imidazole, both indicated by the pronounced band in the E and [E] fractions (Fig. 5). The elution fraction shows a small amount of copurification, but overall effective separation of RNase E from *E. coli* proteins with minimal elution in wash and strip fractions. The predicted molecular weight of this variant is approximately 79kDa; however, the presumed RNase E runs to about 95-100kDa on a gel. A possible explanation for this is the presence of a
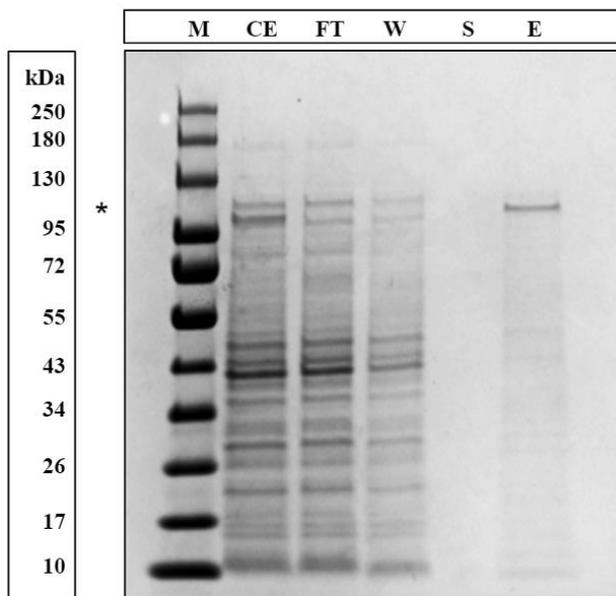
Figure 6 - SDS-PAGE 146 - 824aa Catalytically Dead Partial Truncated RNase E IMAC Fractions. Molecular weight standard (M), crude extract (CE), flowthrough (FT), wash (W), strip (S), and elute (E), were loaded and run on an SDS-PAGE denaturing gel. The position of the presumtive RNase E band is identified on the left by an asterisk (*).

Figure 7 - SDS PAGE Gel 331 - 824aa Catalytically Active Full N-Terminal Deletion RNase E IMAC Fractions. Molecular weight standard (M), crude extract (CE), flowthrough (FT), wash (W), strip (S), and elute (E), were loaded and run on an SDS-PAGE denaturing gel. The position of the presumptive RNase E band is identified on the left by an asterisk (*).

stretch of positively-charged arginine residues in the remaining N-terminal scaffold domain, slightly reducing the migration of the protein during electrophoresis.

Purification by IMAC of the 146-824aa catalytically dead variant showed less presence of RNase E and *E. coli* proteins, indicated by the generally fainter banding in all fractions (Fig. 6). This variant was expressed under suboptimal culture conditions, which might explain the low amount of all proteins. However, there was effective separation from unwanted *E. coli* proteins with low copurification, and minimal RNase E elution in the W fraction and no elution in the S fraction. The predicted molecular weight of the catalytically dead variant is 79kDa, and like the catalytically active version, this variant had an apparent migration of 95-100kDa.

Purification by IMAC of the 331-824aa full N-terminal deleted variant showed high expression, indicated by the pronounced band (Fig. 7). The E fraction showed effective separation of RNase E from unwanted *E. coli* proteins, with the exception of one band above the presumed RNase E band and another at the bottom of the gel. These bands might represent a non-specific binding of *E. coli* proteins to the RNase E variant. Moreover, there was minimal elution of RNase E in the W and S fractions. The predicted molecular weight of this variant is 58kDa and ran truer to size during electrophoresis than the 146-824aa

variants. Since the entire N-terminal scaffold domain has been deleted in this variant, there is no arginine string that might interfere with migration.

SEC Purification of RNase E Variants

After purification by IMAC, each RNase E variant was further refined by SEC. Since this method separates proteins on the basis of size, in analysis by SDS-PAGE there is expected to be a molecular weight gradient across the collected fractions, with early elution fractions containing large proteins and later fractions containing smaller proteins. Since all RNase E variants maintain the "zinc-link" sequence in the catalytic domain involved in dimer and tetramerization, there is a possibility that these variants will elute through the SEC column as either monomers, homodimers or homotetramers. Therefore, many fractions were collected to ensure recovery of RNase E. However, despite how the variants travel through the column, since SDS-PAGE is denaturing, only the monomeric form of the enzyme will be visible in analysis by gel electrophoresis.

Purification of the 146-824aa catalytically active variant by SEC showed very effective separation of proteins by size, with the majority of RNase E in fractions 45-51 (Fig. 8). There is evidence of a small amount of co-purification of *E. coli* proteins with RNase E. However, there is generally a large reduction in unwanted proteins between the IMAC elute and the SEC fractions containing RNase E. The presumed RNase E band again ran slightly higher than the predicted molecular weight (79kDa) and migrated to the 95-100 kDa position during separation by gel electrophoresis.
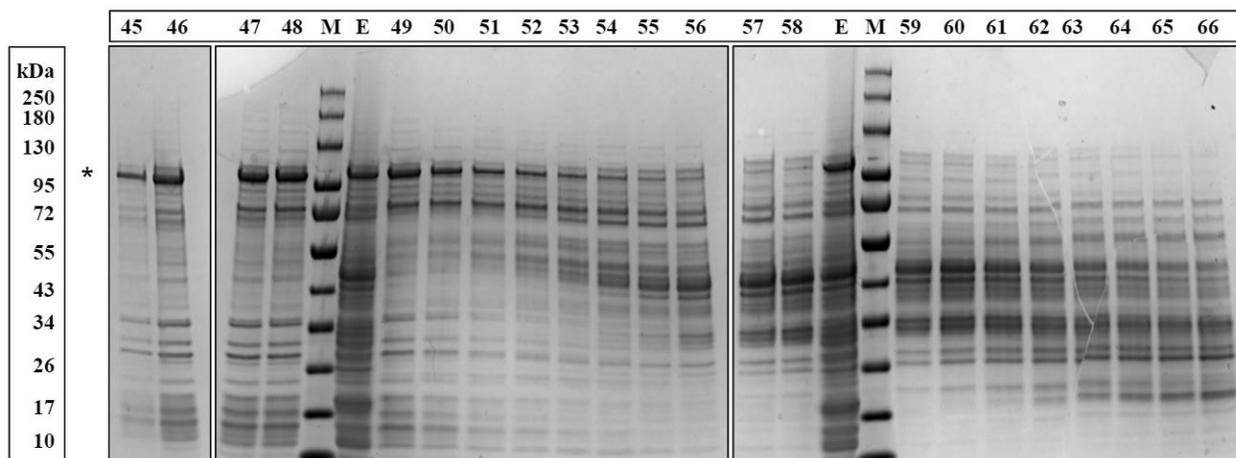
Figure 8 - SDS-PAGE Gel 146 - 824aa Catalytically Active Partial Truncated RNase E SEC Fractions. SEC fractions 45-66 were run on an SDS-PAGE denaturing gel with IMAC elute (E) fraction and a molecular weight standard (M). The position of the presumptive RNase E band is identified on the left by an asterisk (*).

Purification by SEC of the 146-824aa catalytically dead variant also showed effective separation by size, with the majority of RNase E eluting in fractions 48-54 (Fig. 9). Due to the

lower expression of proteins, it is difficult to determine the level of co-purification with *E. coli proteins* in the RNase E elute fractions. However, from what is visible on the gel analyzing fractions collected from this purification, there appears to be a good separation of RNase E from many lower molecular weight proteins, which are visible in the concentrated IMAC elute fraction and in higher fractions 59-64.
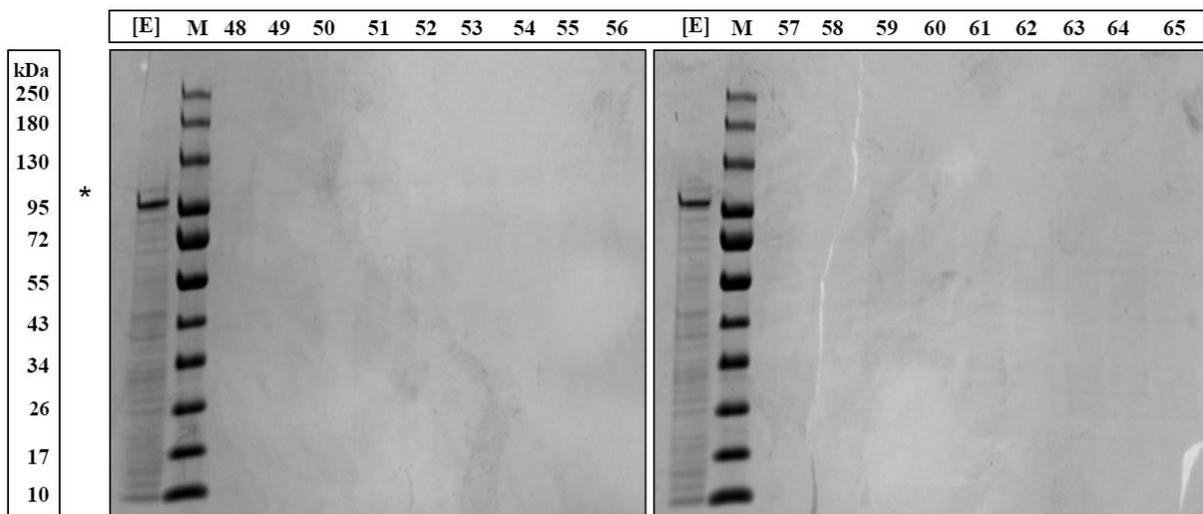
Figure 9 - **SDS-PAGE Gel 146 - 824aa Catalytically Dead Partial Truncated RNase E SEC Fractions.** SEC fractions 48-75 were run on an SDS-PAGE denaturing gel with IMAC elute (E) fraction and a molecular weight standard (M). The position of the presumptive RNase E band is identified on the left by an asterisk (*).
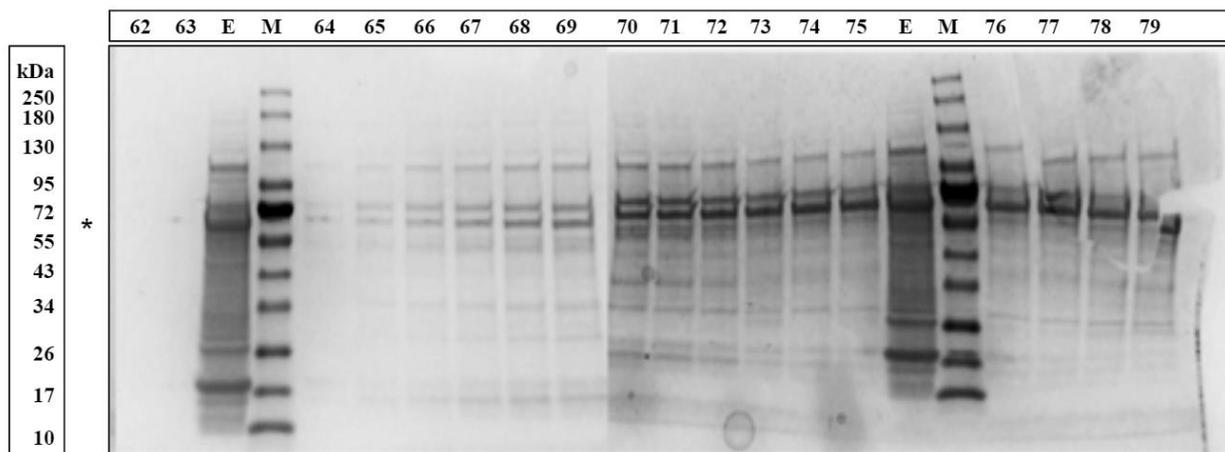
Figure 10 - **SDS PAGE Gel 331 - 824aa Catalytically Active Full N-Terminal Deleted RNase E SEC Fractions.** SEC fractions 62-79 were run on an SDS-PAGE denaturing gel with IMAC elute (E) fraction and a molecular weight standard (M). The position of the presumptive RNase E band is identified on the left by an asterisk (*).

Purification by SEC of the 331-824aa catalytically active full N-terminal deletion variant showed elution of RNase E over a relatively larger range of fractions, from fractions 63-79 (Fig. 10). This difference may be attributed to an equipment error during this purification, where the

fraction collector measured fractions of 5 drops (~150μL) instead of the programmed 10 drops (~300μL). Despite this error, there was visible reduction of unwanted *E. coli* proteins in the SEC RNase E elute fractions relative to the IMAC elute.

RNase E Activity Assay

To determine if cleavage sites previously mapped *in vivo* for the MSMEG_4626 mRNA transcript (Martini et. al, 2019) can be correlated to the activity of RNase E and to address putative autoregulation of the transcript, an RNase E activity assay was performed using two of the puried variants. There are four cleavage sites mapped along the 5' UTR + 45nts coding sequence (281nts), which after cleavage would produce predicted fragment sizes of 41nts, 102nts, 92nts, 27nts, and 45nts (Fig. 11).
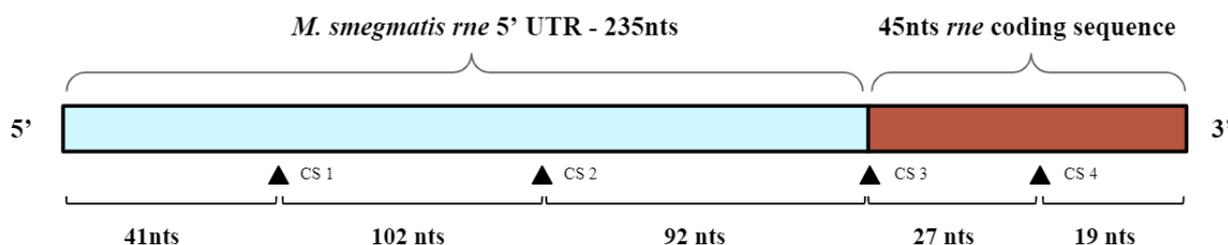
Figure 11 - *M.smegmatis rne* **5' UTR + 45nts Coding Sequence with Cleavage Sites.** RNase E activity assay was preformed using an *in vitro* transcribed *M. smegmatis rne* 5' UTR + 45nts *rne* coding sequence. Four either high or medium confidence cleavage sites within this 281 nucleotide sequence have been previously identified *in vivo* (Martini et. al, 2019).

This sequence, generated by *in vitro* transcription, served as the substrate for the activity assay. It would be expected that an RNA sample incubated with the catalytically active variant would show several specific cleavage products (Fig. 11). Conversely, the RNA samples incubated with the catalytically dead variant or with no enzyme are expected to show no evidence of specific cleavage products.

Digested RNA samples were separated by gel electrophoresis for visualization of relative fragment sizes and abundance, seen in Figure 12. Since there was visible cleavage in the RNA sample incubated with the 146-824aa catalytically dead variant and a small level of degradation in the no RNase E control, only bands that are unique in the catalytically active RNase E treatment were considered related to the activity of the enzyme. Four such bands were identified. There appeared to be a slight shift upwards in migration of the RNA substrate on the gel, with the undigested fragment of 281nts migrating slightly above the 300nt position on the ladder. The estimated fragment sizes of the novel bands in the catalytically active treatment are 120nts, 90nts, 70nts, 45nts. Three of these fragments sizes are within reasonable range of the fragment sizes predicted from cleavage sites mapped *in vivo*. Cleavage at site 1 (Fig. 11) would result in fragment
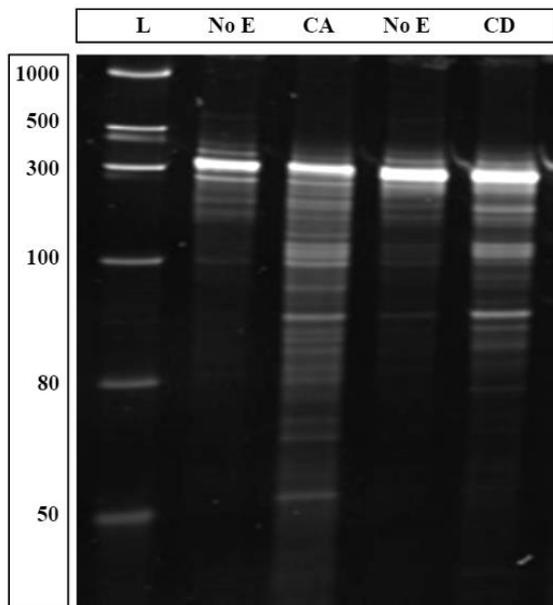
Figure 12 - RNase E Activity Assay With Native *M.smegmatis rne* 5' UTR + 45nts Coding Sequence. Preliminary RNase E activity assay with *rne* transcript incubated with 146-824aa catalytically active and 146-824aa catalytically dead RNase E variants. From left to right on the gel, samples run include ladder (L), no enzyme (No E), catalytically active (CA), no enzyme (NoE), and catalytically dead (CD).

sizes of 41nts and 241nts, which may correlate to the approximately 45nt fragment seen *in vitro*. Cleavage at sites 2 and 3 would result in fragment sizes of 141nts and 46nts, which again may correlate to the 45nts fragment visible in the assay. Cleavage at sites 2 and 4 would result in fragment sizes of 141, 119, and 19nts, which may correlate to the 120nts fragment seen on the gel. Information regarding these cleavage sites *in vivo* as well as a comprehensive list of cut site combinations and resulting fragment sizes can be found in Supplemental Material S8 and S9, respectively. In general, the results of this preliminary RNase E *in vitro* activity assay are indicative that cleavage sites mapped *in vivo* along the *M. smegmatis rne* transcript +45nts coding sequence may be related to the activity of RNase E.

DISCUSSION

One of the major achievements of this research was the development of an effective method to purify *M. smegmatis* RNase E by IMAC and SEC for *in vitro* studies, which to our knowledge has not been done previously. The 146-824aa catalytically active variant expressed well and was effectively purified both by IMAC and SEC. There was a significant reduction in unwanted proteins with a minimal amount of co-purification with *E. coli* proteins. Relatively speaking, purification by SEC of the 146-824aa catalytically active variant was the most successful, with effective and high resolution separation by size. The only RNase E variant that showed low expression was the 146-824aa catalytically dead variant. Since this variant was grown under suboptimal culture conditions, it is possible cell growth and/or protein expression was impeded. For optimal growth and expression, it is important to ensure proper aeration of liquid cultures. It is unlikely, but there is a possibility the point mutations in the catalytic domain might have led this variant to be toxic to *E. coli*. Both of these situations might explain the slow growth in culture and low protein expression that was seen. With regards to the 331-824aa variant, there was a high level

of expression and good purification by IMAC, but a less effective purification by SEC. This variant eluted from the SEC column later and over a wider range of fractions relative to the two 146-824aa variants, which may be due in part to an equipment error that resulted in a collection of fractions that were half the intended volume. However, a later elution is still to be expected with a smaller molecular weight. There is also a very pronounced doublet that co-purifies with the presumed RNase E band that is not present in either of the other purifications. It is possible that a full deletion of the N-terminal scaffolding domain exposes residues in the catalytic domain that may tightly associate with a particular *E. coli* protein, which would get pulled down during purification.

Other possible explanations for the presence of some of the smaller molecular weight proteins is incomplete translation or partial degradation of RNase E. In the construct design, the HIS tag sequence was placed upstream of the *rne* coding sequence, and when expressed would be present at the N-terminus. Thus, there is a possibility that some of the smaller molecular weight proteins might be truncated versions of RNase E that were not fully translated during expression but would still bind to the IMAC column via HIS tag during purification. Equally likely, RNase E that has been partially degraded but maintains the HIS tag would bind to the IMAC column. What comes more to question, however, are bands of higher molecular weight that co-purify with RNase E, which cannot be explained by either of these possibilities.

An interesting observation during analysis of RNase E elute fractions from IMAC and SEC was migration of the 146-824aa variants during SDS-PAGE 10-20kDa higher than their predicted molecular weights. A likely explanation for this is the intact arginine rich RNA binding region present in the remaining N-terminal scaffolding domain of these variants which might have interfered with migration. This hypothesis is supported by evidence of the 331-824aa variant, which lacks this N-terminal scaffolding domain, running true to size when separated on a gel. This possibility has also been reported by other groups working with purified RNase E from other species.

Given the major differences in the cleavage patterns identified *in vivo* for the *M. smegmatis* and *E. coli* transcriptomes*,* this research aimed to investigate whether *M. smegmatis* RNase E exhibits cleavage activity *in vitro* that correlates with the *in vivo* mapped cleavage sites. The preliminary RNase E activity assay, where the native *M. smegmatis rne* 5' UTR + 45nts coding sequence was degraded *in vitro* by the146-824aa catalytically active variant gives evidence to suggest that cleavage sites mapped *in vivo* can in fact be correlated to the activity of RNase E *in vitro*.

There are several points of consideration for the results of this assay, the first being evidence of cleavage activity in the catalytically dead treatment in addition to the non-specific degradation seen in the no enzyme control. One plausible explanation for this is that the two D $\rightarrow$ R mutations in the catalytic domain do not knock out activity in mycobacterial RNase E as they

do in *E. coli*. Even though the catalytic residues in *E. coli* are conserved in mycobacteria, there may be a difference in the mechanism of RNA cleavage or in residues that facilitate cleavage in mycobacteria. A second and perhaps more likely cause for this evidence of cleavage is simply co-purification of *E. coli* RNases with *M. smegmatis* RNase E during the purification process, either associating with the remaining scaffolding domain or as heterodimers formed through the "Zinc-link" motif in the catalytic domain. There appear to be proteins of smaller molecular weight present in the analysis of SEC fractions containing RNase E that disappear along with the enzyme, suggesting the presence of these smaller *E. coli* proteins might be related to the presence of RNase E. However, it is still unknown the true relationship between these proteins and RNase E. There is also evidence to support this hypothesis in the results of the RNase E activity assay. If activity was not abolished in the 146-824aa catalytically dead variant, then there would theoretically be no difference in cleavage products as compared to the 146-824aa catalytically active variant. Since there are visible fragments unique to catalytically active, this is evidence to suggest that there is a difference particularly in *M. smegmatis* RNase E activity between the two treatments. An interesting and useful next step could be determination of the capacity of mycobacterial RNase E to form heterodimers or heterotetramers with *E. coli* RNase E. In principle, since the same "Zinc-link" motif identified as required for homotetramerization of RNase E in *E. coli* (Callaghan, 2005) is conserved in mycobacteria, and is intact in all three of the variants, it would make sense that heterodimerization or heterotetramerization is possible.

Despite visible activity in the 146-824aa catalytically dead treatment, there was evidence of cleavage unique to the 146-824aa catalytically active treatment, and several of these fragment sizes are consistent with predicted fragment sizes from the cleavage pattern mapped *in vivo*. Moreover, these results, which suggest cleavage at sites 2 and 3 (producing RNA fragment sizes of 141nts, 92nts and 19nts *in vivo*), are consistent with the findings of another research project conducted in the Shell Lab involving the *M. smegmatis rne* transcript. Karina Franca, through a primer extension experiment using an *rne* 5' UTR + 45nts CDS + mCherry construct, also isolated dominant fragments which correspond to cleavage *in vivo* at sites 2 and 3. Together, these results strongly suggest that the *rne* transcript is autoregulated *in vivo* at least in part through cleavage by RNase E dominant at sites 2 and 3 along the 5'UTR + coding sequence. Another step that might be taken to confirm these results is mapping the 5' ends of RNA fragments from the RNase E activity assay to determine if the same C-specific cleavage is found *in vitro* to match what was observed *in vivo*. If this specificity is found, together this evidence would be strongly indicative that mycobacterial RNase E exhibits cleavage activity that is different from what has been characterized for *E. coli* RNase E.

In general, these preliminary results support the hypothesis that RNase E is responsible for the C-specific cleavage pattern observed *in vivo* for *M. smegmatis* and is involved in autoregulation of the *rne* transcript. However, more activity assays should be done in the future to optimize the

conditions of the experiment and to ensure reproducibility of results. It is also important to consider that there are many additional intracellular factors that influence RNA degradation *in vivo* that are impossible with research tools available today to fully replicate *in vitro*. So, while this study yields a promising finding regarding mycobacterial transcript regulation, much more research must follow to fully classify the process and molecular players involved in RNA abundance and decay in mycobacteria.

CONCLUSION

In summary, an effective method was designed and validated to purify *M. smegmatis* RNase E for *in vitro* studies, and evidence was collected to suggest that cleavage sites mapped *in vivo* for the native *M. smegmatis rne* 5' UTR + 45nts coding sequence may be attributed to the activity of RNase E. Moreover, the results of these experiments complement several other projects carried out in the Shell Lab, and excitingly come together to support an interesting and relevant finding in the topic of mycobacterial transcript regulation. Due to the circumstances surrounding COVID-19, only one RNase E activity assay was performed using two of the engineered variants. In the future, more activity assays should be run using all three variants to confirm reproducibility of results and to investigate whether further truncation of the N-terminus has an impact on RNase E activity. Mapping the 5' ends of the RNA fragments from *in vitro* activity assays could also be done to determine if the identities of the nucleotides at the +1 positions are the same that were found for predicted fragments from cleavage *in vivo*. Moreover, using a buffer for activity assays that is supplemented with Mn^{2+} instead of Mg^{2+} could be done to investigate if the same metal ion coordination preference that was found in *E. coli* is also found in *M .smegmatis*.

REFERENCES

Bandyra, K., Wandzik, J., Luisi, B. (2018). Substrate Recognition and Autoinhibition in the Central Ribonuclease RNase E. *Molecular Cell, 72*(2), 275 - 285. DOI: 10.1016/j.molcel.2018.08.039

Berstein, J., Lin, P., Cohen, S., & Lin-Chao, S. (2004). Global Analysis of *Escherichia coli* RNA Degradosome Function Using DNA Microarrays. *Proceedings of the National Academy of Sciences of the United States of America, 101*(9), 2758 - 2763. DOI:10.1073/pnas.0308747101

Callaghan, A., Marcaida, M., Stead, J., McDowall, K., Scott, W., Luisi, B. (2005). Structure of *Escherichia coli* RNase E Catalytic Domain and Implications for RNA Turnover. *Nature*, *437,* 1187 - 1191. DOI: 10.1038/nature04084

Carpousis, A. (2002). The *Escherichia coli* RNA Degradosome: Structure, Function and Relationship to Other Ribonucleolytic Multienzyme Complexes. *Biochemical Society Transactions, 30(2)*, 150 - 155. DOI: 10.1042/bst0300150

Carpousis, A. (2007). The RNA Degradosome of *Escherichia coli:* an mRNA-Degrading Machine Assembled on RNase E. *Annual Review of Microbiology, 61,* 71 - 87. DOI: 10.1146/annurev.micro.61.080706.093440

Caruthers, J., Feng, Y., McKay, D., & Cohen, S. (2006). Retention of Core Catalytic Functions by a Conserved Minimal Ribonuclease E Peptide That Lacks the Domain Required for Tetramer Formation. *Journal of Biological Chemistry, 281*(37), 27046 - 27051. DOI: 10.1074/jbc.M602467200

Chao, Y., Li, L., Girodat, D., Förstner, K., Said, N., Corcoran, C., Śmiga, M., Papenfort, K., Reinhardt, R., Weiden, H., Luisi, B., & Vogel, J. (2017). In Vivo Cleavage Map Illuminates the Central Role of RNase E in Coding and Non-coding RNA Pathways. *Molecular Cell, 65*(1), 39 - 51. DOI: 10.1016/j.molcel.2016.11.002

Csanadi, A., Faludi, I., Miczak, A. (2009). MSMEG_4626 Ribonuclease from *Mycobacterium smegmatis. Molecular Biology Reports, 36*(8), 2341 - 2344. DOI: 10.1007/s11033-009-9454-1

Deana, A., Celesnik, H., & Belasco, J. (2008). The Bacterial Enzyme RppH Triggers Messenger RNA Degradation by 5' Pyrophosphate Removal. *Nature, 451,* 355 - 358. DOI: 10.1038/nature06475

Ehlers, S., & Schaible, U. E. (2013). The Granuloma in Tuberculosis: Dynamics of a Host-pathogen Collusion. *Frontiers in Immunology*, *3*, 411. DOI:10.3389/fimmu.2012.00411

Esquerré, T., Laguerre, S., Turlan, C., Carpousis, A., Girbal, L., Cocaign-Bousquet, M. (2014). Dual role of Transcription and Transcript Stability in the Regulation of Gene Expression in *Escherichia coli* Cells Cultured on Glucose at Different Growth Rates. *Nucleic Acids Research, 42*(1). 2460 - 2472. DOI: 10.1093/nar/gkt1150

Górna, M., Carpousis, A., Luisi, B. (2012) From Conformational Chaos to Robust Regulation: The Structure and Function of the Multi-Enzyme RNA Degradosome. *Quarterly Review of Biophysics, 45*(2), 105 - 145. DOI: 10.1017/S003358351100014X

Kiran, D., Podell, B. K., Chambers, M., & Basaraba, R. J. (2016). Host-directed Therapy Targeting the Mycobacterium tuberculosis Granuloma: A Review. *Seminars in Immunopathology*, *38*(2), 167–183. DOI: 10.1007/s00281-015-0537-x

Kovacs, L., Csanadi, A., Megyeri, K., Kaberdin, V., Miczac, A. (2005) Mycobacterial RNase E-Associated Proteins. *Microbiology and Immunology, 49*(11), 1003 - 1005. DOI: abs/10.1111/j.1348-0421.2005.tb03697.x

Mackie, G. (1998). RNase E is a 5'-end-dependent Endonuclease. *Nature, 395,* 720 - 723. DOI: 10.1038/27246

Mackie, G. (2013). RNase E: At the Interface of Bacterial RNA Processing and Decay. *Nature Reviews Microbiology, 11,* 45 - 57. DOI: 10.1038/nrmicro2930

McDowall, K., Lin-Chao, S., Cohen, S. (1994). A + U Content Rather Than a Particular Nucleotide Order Determines the Specificity of RNA Cleavage. *Journal of Biological Chemistry, 269*(14), 10790 - 10796. DOI:

Park, M., Satta, G., & Kon, O. M. (2019). An Update on Multidrug-Resistant Tuberculosis . *Clinical Medicine (London, England)*, *19*(2), 135–139. DOI:10.7861/clinmedicine.19-2-135

Płociński, P., Macios, M., Houghton, J., Niemiec, E., Płocińska, R., Brzostek, A., Slomka, M., Dziadek, J., Young, D., Dziembowski, A. (2019). Proteomic and Transcriptomic Experiments Reveal an Essential Role of RNA Degradosome Complexes in Shaping the Transcriptome of *Mycobacterium tuberculosis. Nucleic Acids Research, 47*(11), 5892 - 5905. DOI: 10.1093/nar/gkz251

Richards, J. & Belasco, J. (2015) Distinct Requirements for 5' Monophosphate-assisted RNA Cleavage by *Escherichia coli* RNase E and RNase G. *The Journal of Biological Chemistry 291*(10), 5038 - 5048. DOI: 10.1074/jbc.M115.702555

Rustad, T. R., Minch, K. J., Brabant, W., Winkler, J. K., Reiss, D. J., Baliga, N. S., & Sherman, D. R. (2013). Global Analysis of mRNA Stability in Mycobacterium tuberculosis. *Nucleic Acids Research*, *41*(1), 509–517. DOI:10.1093/nar/gks1019

Thompson, K., Zong. J., Mackie, G. (2015). Altering the Divalent Metal Ion Preference of RNase E. *Journal of Bacteriology*, *197*(3), 477 - 482. DOI: 10.1128/JB.02372-14

World Health Organization. (2017). Tuberculosis (TB). Retrieved from https://www.who.int/news-room/fact-sheets/detail/tuberculosis

World Health Organization. (2019). *Global Tuberculosis Report 2019*. Geneva. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/329368/9789241565714-eng.pdf?ua=1

Zeller, M., Csanadi, A., Miczak, A., Rose, T., Bizebard, T., & Kaberdin, V. (2007). Quaternary Structure and Biochemical Properties of Mycobacterial RNase E/G. *Biochemical Journal, 403,* 207 - 215, DOI: 10.1042/BJ20061530

SUPPLEMENTAL MATERIAL

Solutions and Buffers

S1 - 2X IMAC Buffer Concentrations and Amounts. The table below summarizes the components, their stock concentrations, final concentrations needed to create the 2X IMAC buffer, and volume to add.

Component	Stock Concentration	Final Concentration in 2X IMAC Buffer	Volume to Add
Tris-HCl pH 7.9	1M	40 mM	10 mL
NaCl	3M	300 mM	25 mL
Glycerol	50%	10%	50 mL
IGEPAL	10%	0.02%	500μL
DI H2O	-	-	165mL
		Final Volume	250mL

S2 - 1X IMAC Buffer with 10mM Imidazole. The table below summarizes the components, their stock concentrations, final concentrations in the 1X IMAC buffer, and volume to add.

Component	Stock Concentration	Final Concentration in 1X IMAC Buffer	Volume to Add
2X IMAC Buffer	2X	1X	10 mL
DI H2O	-	-	10mL
Imidazole	2M	10mM	100μL
Protease Inhibitor	100X	1X	200μL
		Final Volume	20mL

S3 - IMAC Purification Buffer Recipe. The table below summarizes the buffer components and volumes of each to add to create the three buffers (wash, elute, and strip) used during purification by IMAC.

Buffer	1X IMAC Buffer (mL)	Imidazole Concentration in Buffer	Volume 2M Imidazole to add (μL)	Protease Inhibitor Concentration in Buffer	Volume 100X Protease Inhibitor to add (μL)	Volume DI H2O to add (mL)	Final Volume (mL)
Wash	2	20mM	40	1X	40	2	4
Elution	4	150mM	600	1X	80	3.7	8
Strip	2	500mM	1000	1X	40	1	4

S4 - SEC Buffer Concentrations. The table below summarizes the components, stock concentrations, final concentrations, and the volume of each to add to create the SEC buffer.

Component	Stock Concentration	Final Concentration in SEC Buffer	Volume to Add
Tris-HCl pH 7.9	1M	20mM	20mL
NaCl	3M	150mM	50mL
Glycerol	50%	5%	200mL
IGEPAL	10%	0.01%	1mL
EDTA	500mM	1mM	2mL
DTT	1M	1mM	1mL
DI H2O	-	-	726mL
		Final Volume	1L

S5 - Protocol. Purification of *M. smegmatis* RNase E by IMAC and SEC.
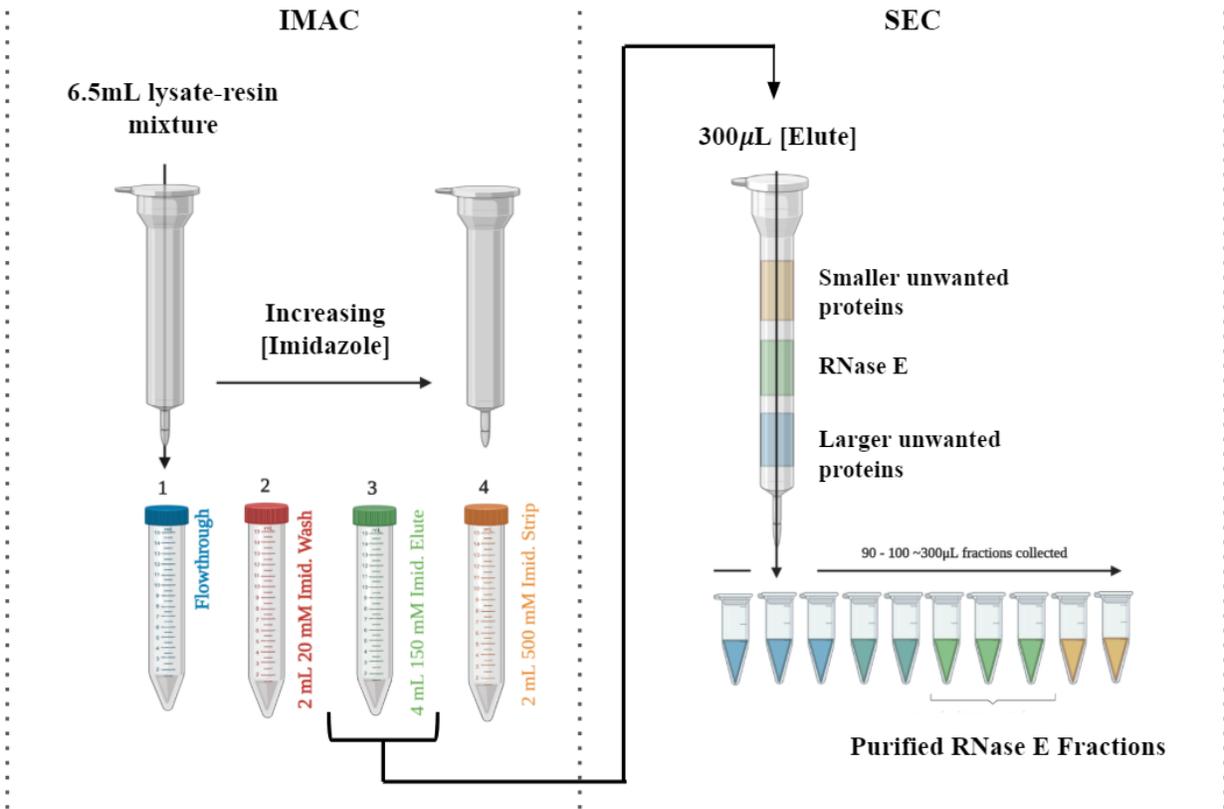
Clarification of Cell Lysate

1. For timing purposes, it is best to make the 2X IMAC buffer, 1X IMAC buffer, wash, elution, and strip buffers for IMAC purification before beginning the lysis step. (Buffer recipes found in Supplemental Material S1 - S4).
2. Thaw pellets on ice and resuspend each in approximately 5mL of 1X IMAC buffer. Shake and swirl vigorously to disrupt the pellet, then split the total lysate volume between two 50mL conical tubes.
3. Use a Tissue Tearor to mechanically homogenize the cell lysate. This is most easily done if there is one person lysing with the Tissue Tearor and another keeping track of time and cycle number. With the tissue Tearor set to 15, perform 10 cycles of 15 seconds on, and 60 seconds off while incubating the lysate on ice.
4. After both 50mL tubes of lysate have been homogenized, save $500\mu L$ for analysis. This is the crude extract fraction.
5. Distribute the remaining lysate into 1.5mL Eppendorf tubes and centrifuge at maximum speed (14,000 RCF) for 15 minutes at 4°C to pellet cell debris. Pipette off the supernatant from each tube, being careful to not disturb the pellets, into a new 50mL conical tube. This is the clarified lysate fraction.

Immobilized Metal Affinity Chromatography (IMAC)

6. Prepare the Ni-NTA resin. If there is a high volume of lysate, purification by IMAC can be performed in duplicate to save time. Resuspend the resin and pipette just under 2mL from in the storage buffer, spin at the lowest possible speed (800 RCF) for 1 minute at 4°C.
7. Pipette off the supernatant and resuspend in 2mL of 10 mM imidazole 1X IMAC buffer to wash the resin.. Spin again at lowest speed for 1 minute and then pipette off and discard the supernatant.
8. Slowly add the clarified lysate. Resuspend the washed Ni-NTA a little at a time using the clarified lysate and transferring it back to the 50 mL conical tube.
9. Incubate the resin-lysate mixture rocking at room temperature for 30 minutes.
10. Prepare the IMAC column. Fill a polypropylene column with DI water and using a glass pipette or transfer pipette, push the white filter disc down into the column. Once the filter is secured at the bottom, make sure to return the plastic stopper to the bottom. Use a ring stand and clamps to position the column and a tube rack to position conical tubes for collection underneath.
11. Pour in the resin-lysate mixture into the plastic column and open the stopper, collecting the flow through in the plastic conical tube. The resin should pack to a height of about 1cm. If the flow through is eluting slowly or stops, a glass pipette can be used to kick up and repack the resin to encourage the lysate to flow through. This is the flow through fraction.
12. Apply the wash buffer. Add 2mL of the wash buffer (20mM imidazole) to the column and collect the flow through in a 15mL conical tube. This is the wash fraction.
13. Apply the elution buffer. Add 4mL of elution buffer (150mM imidazole) to the column and collect the flow through in a 15mL conical tube. This is the elution fraction.
14. Apply the strip buffer. Add 2mL of strip buffer (500mM imidazole) to the column and collect the flow through in a 15mL conical tube. This is the strip fraction.
15. Save all collected fractions at 4°C until analysis.

Concentration of Elute Fraction

16. Save $500\mu L$ of the elute fraction prior to concentration for analysis.
17. Add approximately $500\mu L$ of elute each to four Microcon PL-30 (30,000 NMWL) protein concentrator columns. Be sure not to overfill the columns.
18. Centrifuge at maximum speed (14,000 RCF) for 15 minutes at 4°C. This step will concentrate the $500\mu L$ of elute to around $250\mu L$.
19. Pipette off and discard the liquid in the collection tube. Add an additional $250\mu L$ (again being cautious not to overfill the column) to each concentrator and centrifuge at maximum speed for an additional 15 minutes. Repeat this process of removing flow through, adding more eluent, and centrifuging until all of the eluent has been added to the concentrator columns. This can take up to several hours depending on the purity of the elute fraction.
20. After another 1-2 centrifugation cycles, combine all of the remaining eluent from the four concentrators into one new concentrator. Be careful to transfer the liquid inside the concentrator and not the flow through in the collector.
21. Concentrate this final volume down to between 300-400μL. This should take around 10 minutes at maximum speed. Save the concentrated elute at 4°C until SEC can be performed.

Size Exclusion Column Setup

Preparation of the size exclusion column must be done at least one day in advance of purification, to allow time for equilibration of the column with the SEC buffer.

22. Assemble the size exclusion column. Screw the stopcock onto the bottom of the column and fasten the column to a ring stand in a 4°C deli case fridge. The column should be completely vertical, as any tilt to the column will cause the resin to pack unevenly.
23. Add DI water to the column until it reaches 2 cm from the bottom of the column. Remove all visible bubbles from the water by gently opening the stopcock. Once all bubbles have been removed close the stopcock.
24. Resuspend the Sephacryl S-200 to form a slurry by gently rocking back and forth. Then, pour 80 mL of the slurry into a graduated cylinder. Using an automatic pipette, pipette the slurry into the column. This must be done carefully to avoid introducing any bubbles, which may interrupt elution or cause the resin to crack.
25. Place a 1L beaker under the column to collect the buffer. Then, add a small volume of SEC buffer (Supplemental Material S5) to the column using a transfer pipette.
26. Open the stopcock and allow the buffer to flow through the resin for approximately 4 hours. Once the resin has settled and there is a visible line at the top, close the stopcock and add SEC buffer carefully so that it enters the wider opening at the top of the column.
27. Add SEC buffer to another 1 L beaker. Place the opening of the tubing connected to the Masterflex C/L pump to the SEC buffer, careful to avoid introducing bubbles into the line.
28. Attach the pump's column adaptor by twisting the ring on the adaptor to tightly fit the column. The adaptor should be attached just above the resin bedding and attached in one slow downward motion. Proper attachment of the adaptor can be checked with a light tug.
29. Turn on the pump at the slowest setting, open the stopcock and allow SEC buffer to run through the column overnight. Before leaving, watch the tubing connected to the pump for bubbles.

Size Exclusion Chromatography

30. Turn off the pump and close the stopcock. Remove the column adaptor by loosening it and slowly pulling up to avoid disrupting the resin bed. Place the column adapter in the 1L beaker of the SEC buffer to avoid introducing bubbles to the column adapter tubing.
31. Using a micropipette, remove excess buffer from the top of the column to around 1 cm above the top of the resin bed. Open the stopcock to allow the remaining buffer to flow into the resin. Close the stopcock as soon as the buffer has fully entered the resin.
32. Pipette $300 \mu L$ of the concentrated IMAC eluent carefully on top of the packed resin by placing the pipette tip against the side of the column and letting it run down the sides.
33. Open the stopcock to flow the eluent into the resin. Close the stopcock immediately once the eluent has completely entered the resin.
34. Chase the concentrated eluent with buffer by adding 1.5mL of SEC buffer to the top of the resin. Then, open the stopcock, count out exactly 12 drops, and immediately close the stopcock.
35. Add buffer to the wide part of the column slowly and re-attach the column adapter with the same technique described above. Attach the waste tube to the fraction distiller.
36. Turn on the fraction distiller, set the fraction size to 10 drops, and insert the waste tube into the first fraction. Turn the pump on to the slowest speed and open the stopcock.

37. ~300μL fractions will be collected automatically. Once all fractions have been collected, transfer the contents of each vial to a 1.5mL Eppendorf tube and save at 4°C for analysis.

S6 - T7 MSMEG_4626 5' UTR + 45nts Coding Sequence G-Block. The following represents the T7 *rne* 5' UTR +45nts coding sequence gblock used for *in vitro* transcription.

```
TAATACGACTCACTATAGGGTGAGCGGCCTCAACCGGGGCTGAGCACCCGAATGAAGACTTC
ATTATGCTGAGTGATATCCCACTCGCCGGAGTTGGCCCCGACTCGTGGGCTTACTTCTGAAG
```

T7 promoter   RNase E 5'UTR   CS 1

70   80   90   100   110   120

```
GACGTGCGCGACCACCGCAAACCGGTGCGGCGCCGACCGTCCAGCCATCATTCAGCCAGGCA
CTGCACGCGCTGGTGGCGTTTGGCCACGCCGCGGCTGGCAGGTCGGTAGTAAGTCGGTCCGT
```

RNase E 5'UTR

130   140   150   160   170   180

```
CCGGGTGGGTTCGACAGCGGACCGCTCGGAGCGAGACCACAACAAGTCCGGGGAACGTAGCC
GGCCCACCCAAGCTGTCGCCTGGCGAGCCTCGCTCTGGTGTTGTTCAGGCCCCTTGCATCGG
```

RNase E 5'UTR   CS 2

190   200   210   220   230   240

```
CGGGCCAATAGCGGAAGCCCTCGCGTGGCCGCGTCGTACGACGCGCCCGGGGGCTTGAGGAG
GCCCGGTTATCGCCTTCGGGAGCGCACCGGCGCAGCATGCTGCGCGGGCCCCCGAACTCCTC
```

RNase E 5'UTR

250   260   270   280   290   300   310

```
AATTCGTGGCCGAAGATGCCCATACCGAAGACCTATCAACCCAGACTCCC
TTAAGCACCGGCTTCTACGGGTATGGCTTCTGGATAGTTGGGTCTGAGGG
```

CS 4

CS 3   RNase E first 45 nt

320   330   340   350   360

S7 - Protocol. RNase E Activity Assay.

1. Based on the concentration of substrate from *in vitro* transcription, calculate the appropriate volume of RNA substrate to achieve 360ng total. Pipette this volume of substrate into a 200μL PCR tube and bring the final volume to 8μL by adding the appropriate volume of 1.25X reaction buffer. Repeat this for the total number of reactions to be run.
2. Using a heat block or thermal cycler, heat the samples to 65°C for 3 minutes to denature any secondary structure. Then, reduce the temperature to 37°C and incubate for an additional minute.
3. Quickly pipette 2μL of the appropriate purified enzyme (80ng) to each reaction tube and incubate at 37°C for 60 minutes.
4. After 60 minutes, pipette 1μL of 0.1M EDTA solution to each reaction tube to stop the reaction. Then, pipette 11μL of 2X Loading Buffer II into each reaction tube as well.
5. Incubate at 65°C for an additional 3 minutes immediately before analysis to ensure no secondary structure affects the run during separation by gel electrophoresis.

S8 - MSMEG_4626 5' UTR + 45nts Coding Sequence Cleavage Sites. This table summarizes the cleavage sites mapped *in vivo* for the MSMEG_4626 5' UTR + 45nts coding sequence and their number designation for *in vitro* RNase E activity assay reference.

Cleavage Site *in vitro*	Cleavage Site Position Relative to TSS	Position *in vivo*	Sense Strand
1	+41	4713195	-
2	+143	4713093	-
3	+235	4713001	-
4	+262	4712974	-

S9 - Comprehensive List of Cleavage Site Combinations and Fragment Sizes. This table summarizes the possible combinations of all cleavage sites mapped along the MSMEG_4626 5' UTR + 45nts coding sequence under the circumstances of 1, 2, 3, and 4 cuts by RNase E.

Number of Cuts	Cleavage Site Combination	Fragment 1	Fragment 2	Fragment 3	Fragment 4	Fragment 5
1	1	41	240			
1	2	143	138			
1	3	235	46			
1	4	262	19			
2	1 + 2	41	102	138		
2	1 + 3	41	194	46		
2	1 + 4	41	221	19		
2	2 + 3	143	92	46		
2	2 + 4	143	119	19		
2	3 + 4	235	27	19		
3	1 + 2 + 3	41	102	92	46	
3	1 + 2 + 4	41	102	119	19	
3	1 + 3 + 4	41	194	27	19	
3	2 + 3 + 4	143	92	27	19	
4	1 + 2 + 3 + 4	41	102	92	27	19