

Vanessa Gomes da Costa

BSc in Biochemistry

**Biochemical and Biophysical Characterization of
Exoribonucleases from the Human Pathogen
*Mycobacterium tuberculosis***

Dissertation presented to obtain the master's degree in
Biochemistry for Health

Supervisors: Doctor Rute Gonçalves Matos & Doctor José Artur Brito

February 2021

Vanessa Gomes da Costa

BSc in Biochemistry

**Biochemical and Biophysical Characterization of
Exoribonucleases from the Human Pathogen
*Mycobacterium tuberculosis***

Dissertation presented to obtain the master's degree in
Biochemistry for Health

Jury:

President: Doctor Pedro Manuel Marques Henriques Matias

Arguer: Doctor Arsénio do Carmo Sales Mendes Fialho

Vogal: Doctor Ana Maria Varela Coelho

Supervisor: Doctor Rute Margarida Gonçalves Matos Luís

**Instituto de Tecnologia Química e Biológica António Xavier,
Universidade Nova de Lisboa**

February 2021

**Biochemical and Biophysical Characterization of Exoribonucleases from the
Human Pathogen *Mycobacterium tuberculosis***

Copyright © Vanessa Gomes da Costa, O Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa

O Instituto de Tecnologia Química e Biológica António Xavier e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

“Don’t just fly... Soar!”

- Dumbo

Acknowledgments

Sixteen months later, many working marathons later, a COVID-19 pandemic later, many smiles, laughs and unforgettable moments later, this incredible journey came to an end. Now, I would like to dedicate some special words, in my native language, to all the extraordinary people who have helped me along the way.

Primeiro, gostaria de agradecer às Professoras Cecília Arraiano e Margarida Archer por me terem recebido de braços abertos nos vossos laboratórios e por me terem facultado sempre, sem exceção, excelentes condições de trabalho.

Aos meus orientadores, José e Rutex, por toda a dedicação, paciência, otimismo e conhecimento que me transmitiram ao longo deste ano! Obrigada por me terem dado o privilégio de trabalhar ao vosso lado, por me ensinarem que existe sempre um plano B, C, D... e por me terem ajudado tanto a crescer no mundo da Ciência! Sem vocês não teria sido a mesma coisa.

Às equipas excecionais e ambos os laboratórios, que estiveram sempre dispostas a ajudar-me em qualquer altura e com as quais partilhei momentos incríveis, um obrigada gigante! Tenho um carinho muito especial por todos vocês! Cátia, Sandra, Maga, Susete, Patrícia, Susanex... Obrigada pela amizade! Por todos os almoços, gargalhadas, partilhas...

Aos meus amigos para a vida toda: Tati, Helena, SQUAD, Padrinho, Miúdo, Afis... Obrigada por fazerem parte da minha vida e por me fazerem sentir que estão sempre presentes, mesmo que distantes. Um obrigada especial a ti, Gonças, o meu companheiro de todas as horas, por me acompanhares neste caminho desde os tempos do penico na tola!

À minha família, os pilares da minha vida: OBRIGADA do tamanho do mundo!! Papá e Mamã, por me terem criado num nicho de amor incondicional, por todas as oportunidades, valores e educação que me deram ao longo da vida e por nunca me deixarem desistir dos meus sonhos. Puto, por toda a paciência que tens para aturar a resmunguice da mana e por todas as vezes que foste forte quando tive de ir para Lisboa. Avós e tios, por todo o amor, apoio, preocupação e fé que depositam em mim em todos os passos da minha vida.

Por fim, quero agradecer muiiiito às duas pessoas que mais aturam os meus dramas, panicanços e refilices do dia a dia. Pikachu, obrigada por fazeres parte de mim, por me apoiares, compreenderes e por seres muito mais do que uma melhor amiga desde sempre! Meu stitch, obrigada por teres entrado na minha vida, pelo amor, incentivo, paciência, tudo... Mil palavras não chegam para agradecer o bem que me fazes! Vocês são o meu porto de abrigo.

A todos vocês e a todos os outros que estão SEMPRE presentes...

Mais uma vez obrigada, do fundo do coração!

List of Communications

Oral Presentations:

Costa, V. G., Archer, M., Arraiano, C. M., Brito, J. A., Matos, R. G. “Targeting Mycobacterial Ribonucleases with an X-Ray Machine Gun”. IMPSG2020: International Meeting of the Portuguese Society of Genetics, 23th-24th January 2020, Universidade do Algarve, Faro, Portugal.

Costa, V. G., Archer, M., Arraiano, C. M., Brito, J. A., Matos, R. G. “Targeting Mycobacterial Ribonucleases with an X-Ray Machine Gun”. iMed Conference 12.0 - Astrazeneca Foundation Innovative Competition, 1st October 2020, Nova Medical School, Universidade Nova de Lisboa, Portugal (Selected for an oral presentation as one of the TOP 3 projects in Basic Research Competition of the Innovative Competition).

Poster Presentation:

Costa, V. G., Archer, M., Arraiano, C. M., Brito, J. A., Matos, R. G. “Targeting Mycobacterial Ribonucleases with an X-Ray Machine Gun”. Congress “II Simpósio de Investigação em Tuberculose e Micobactérias não Tuberculosas em Portugal”, 31th January 2020, Faculdade de Farmácia, Universidade de Lisboa, Portugal

Additional work - Publication as first co-author:

Over the past year, the scientific community from all fields of knowledge around the world have worked together tirelessly to confront the coronavirus outbreak. Cecília Arraiano’s Lab also started a new project in this area, in which I had the opportunity to participate with the knowledge acquired during the realization of the present master thesis. The project has given rise to the following publication, which is now submitted to an international Journal:

Saramago, M.* , Bárria, C.* , **Costa, V. G.***, Souza, C. S., Viegas, S. C., Domingues, S., Lousa, D., Soares, C. M., Arraiano, C. M., Matos, R. G. “New targets for drug design: Importance of nsp14/nsp10 complex formation for the 3’-5’ exoribonucleolytic activity on SARS-CoV-2”. bioRxiv (2021). doi:10.1101/2021.01.07.425745

* The first authors should be regarded as joint first authors.

Abstract

Mycobacterium tuberculosis, the causative agent of tuberculosis, remains the leading cause of mortality from a single infectious organism, infecting nearly one-third of the global population. Emergence of multidrug and extremely drug resistant strains of *M. tuberculosis* represents a serious health problem and emphasizes the need to improve prevention and control of the disease. Regulation of gene expression through RNA metabolism is a key mechanism for bacterial growth, division and rapid accommodation to environmental conditions. Ribonucleases ensure maturation, degradation and quality control of RNA in all domains of life. Particularly, ribonucleases belonging to the RNB-family of enzymes are often essential for viability of prokaryotes and are implicated in the establishment of virulence of several pathogens, making them interesting objects of study. In this work, several biochemical, biophysical and structural assays were assessed to characterize two putative exoribonucleases from the RNB-family of enzymes in *M. tuberculosis*, RNR1 and RNR2. Overexpression and purification of both wild-type and mutant proteins was performed and further *in vitro* activity, binding and helicase assays using synthetic RNA substrates were accomplished. In parallel, a biophysical characterization of the targets proceeded with several crystallization trials and protein stability tests. We have demonstrated that both RNases are 3' to 5' exoribonucleases with different degradation properties. The importance of highly conserved residues for catalysis was also unraveled. Moreover, we were able to identify improved buffer formulations that increase protein stability, possibly enhancing their propensity to crystallize. The information regarding RNA metabolism in *M. tuberculosis* is limited and RNB-family enzymes have not been previously characterized in this important human pathogen. Thus, a complete biochemical and structural knowledge of these ribonucleases is an approach to recognize their influence in *M. tuberculosis* metabolism, to better understand the post-transcriptional control in this pathogen and, ultimately, to study their effect in virulence.

Keywords: RNA degradation, Exoribonucleases, RNB-family, *Mycobacterium tuberculosis*, Activity assays, Thermal stability.

Resumo

Mycobacterium tuberculosis, o agente etiológico da tuberculose, mantém-se como a principal causa de mortalidade por um único patógeno, infectando quase um terço da população mundial. O surgimento de estirpes multirresistentes de *M. tuberculosis* constitui um problema sério de saúde e enfatiza a necessidade de prevenção e controlo da doença. A regulação da expressão génica via metabolismo RNA é um dos mecanismo-chave para a proliferação bacteriana, divisão celular e rápida adaptação às condições ambientais. As ribonucleases asseguram a maturação, degradação e controlo de qualidade do RNA em todos os domínios da vida. As que pertencem à família de enzimas RNB são particularmente essenciais para a viabilidade de procariotas e estão implicadas no estabelecimento de virulência em diversos patógenos, o que as torna interessantes objetos de estudo. Neste trabalho foram desenvolvidos diversos ensaios bioquímicos, biofísicos e estruturais por forma a caracterizar duas ribonucleases putativas da família de enzimas RNB em *M. tuberculosis*, a RNR1 e RNR2. Para tal, foi realizada a sobreexpressão e purificação das proteínas selvagens e mutantes e efetuaram-se ensaios *in vitro* de atividade, ligação e helicase, utilizando substratos sintéticos de RNA. Em paralelo, procedeu-se à caracterização biofísica, através de ensaios de cristalografia e testes de estabilidade proteica. Demonstrámos que ambas as RNases são exoribonucleases 3' para 5' com diferentes propriedades de degradação. Revelou-se também a importância de resíduos altamente conservados para catálise. Adicionalmente, conseguimos otimizar os tampões de modo a aumentar a estabilidade proteica, possivelmente melhorando a sua tendência para cristalizar. A informação relativa ao metabolismo do RNA em *M. tuberculosis* é limitada e a família de enzimas RNB ainda não foi caracterizada neste importante patógeno humano. Desta forma, um conhecimento completo destas ribonucleases a nível bioquímico e estrutural permitirá avaliar a sua influência no metabolismo de *M. tuberculosis*, compreender melhor os processos de controlo pós-transcricional e, consequentemente, o seu efeito na virulência deste organismo.

Keywords: Degradação do RNA, Exoribonucleases, Família RNB, *Mycobacterium tuberculosis*, Ensaios de atividade, Estabilidade térmica.

Table of Contents

Acknowledgments	V
List of Communications.....	VII
Oral Presentations.....	VII
Poster Presentation.....	VII
Additional work - Publication as first co-author	VII
Abstract	IX
Resumo	XI
Table of Contents	XIII
List of Figures	XV
List of Tables	XVII
Abbreviations.....	XIX
1. Introduction.....	1
1.1 Tuberculosis	1
1.1.1 <i>Mycobacterium tuberculosis</i> Features	2
1.1.3 Transmission and Pathogenesis of Tuberculosis.....	3
1.1.4 Available Treatments for Tuberculosis Disease	4
1.1.4.1 Multi-Drug Resistant Tuberculosis	4
1.2 RNA Metabolism in the control of gene expression	6
1.2.1 Prokaryotic mRNA stability	6
1.2.2 Ribonucleases in RNA metabolism.....	7
1.2.2.1 Mechanisms of RNA decay	8
1.2.2.2 Ribonucleases in <i>M. tuberculosis</i>	10
1.2.3 RNB-family of exoribonucleases.....	10
1.2.3.1 <i>E. coli</i> RNase II	12
1.2.3.2 <i>E. coli</i> RNase R	15
1.3 Thesis Motivation and Aims	20
2. Materials and Methods	21
2.1 Bacterial Strains.....	21
2.2 Methods	22
2.2.1 Bioinformatic Analysis.....	22
2.2.2 Plasmid Construction	22
2.2.3 Transformation of Competent Cells.....	23
2.2.4 Protein Overexpression.....	23
2.2.5 Cell Lysis	24

2.2.6	Protein Purification	24
2.2.6.1	Affinity chromatography	24
2.2.6.2	Size Exclusion chromatography.....	25
2.2.7	Protein Quantification	25
2.2.8	Western Blot Analysis.....	25
2.2.9	Radioactive Labeling of Oligonucleotides	26
2.2.10	Activity Assays	26
2.2.11	Electrophoretic mobility shift assay	27
2.2.12	Crystallization Trials	27
2.2.13	Protein Thermostability Assays	28
2.2.13.1	Thermal Shift Assay	28
2.2.13.2	Nano Differential Scanning Fluorimetry.....	28
3.	Results and Discussion.....	29
3.1	Target Validation and Small-scale Protein Purification.....	29
3.2	Biochemical Characterization of RNR1 and RNR2.....	34
3.2.1	Exoribonucleolytic Activity Assays	34
3.2.2	RNA binding ability	39
3.2.3	Helicase Activity	42
3.3	Crystallization and Thermostability Assays	43
3.3.1	Crystallization Trials	45
3.3.2	Thermostability Assay	47
3.3.2.1	RNR1 Thermostability	49
3.3.2.2	RNR2 Thermostability	53
4.	Conclusions and Future Perspectives	57
5.	Bibliography.....	59
6.	Supplementary Material.....	67
	Appendix A	67
	Appendix B	84

List of Figures

1. Introduction.....	1
Figure 1.1. Top 10 global causes of death.....	2
Figure 1.2. Schematic representation of mRNA degradation in <i>E. coli</i> and <i>B. subtilis</i>	9
Figure 1.3. The RNB-family of enzymes.....	11
Figure 1.4. <i>E. coli</i> RNase II protein structure	13
Figure 1.5. Model for RNA degradation by RNase II.....	15
Figure 1.6. <i>E. coli</i> RNase R protein structure.....	18
Figure 1.7. Working models for RNA binding, unwinding and degradation by RNase R.....	19
2. Material and Methods.....	21
Figure 2.1. Schematic diagram of pET15b_RNR1 and pET15b_RNR2, respectively.....	23
3. Results and Discussion.....	31
Figure 3.1 Phylogenetic tree between <i>E. coli</i> RNB-family members and the <i>M. tuberculosis</i> targets RNR1 and RNR2	30
Figure 3.2 RNR1 purification by affinity and size-exclusion chromatography	31
Figure 3.3. SDS-PAGE analysis of the purified proteins.....	32
Figure 3.4. SDS-PAGE and Western blot analysis of recombinant RNR1 and RNR2 after purification	33
Figure 3.5. Exoribonucleolytic activity of RNR1 and RNR2 on a ssRNA.....	35
Figure 3.6. Exoribonucleolytic activity of RNR1 and RNR2 wild-type and mutants on a ssRNA.....	36
Figure 3.7 Exoribonucleolytic activity of RNR1 and RNR2 wild-type and mutants on a double-stranded RNA.....	38
Figure 3.8. Electrophoretic mobility shift assay (EMSA) of wild type RNR1 and RNR2 and mutant proteins with a poly(A) substrate	40
Figure 3.9. Helicase activity of RNR2 mutant proteins in the presence of a double-stranded substrate.....	42
Figure 3.10. Large-scale protein purification of RNR1, RNR2 and RNR2_D267N by affinity chromatography followed by size-exclusion chromatography	44
Figure 3.11. Distinction between protein or salt crystals in a RNR2 crystallization hit.	46
Figure 3.12. Thermal Shift Assay Optimization Curves	48
Figure 3.13 Effect of buffer pH, ionic strength and additives on RNR1 thermal stability. ..	50
Figure 3.14 NanoDSF analysis of RNR1 samples.	52

Figure 3.15 Effect of buffer pH, ionic strength and additives on RNR2_D267N thermal stability	54
6. Supplementary Material.....	69
Figure S.1. Multiple sequence alignment colored by domains of proteins belonging to RNB-family.....	69
Figure S.2. Affinity chromatography of and <i>M. tuberculosis</i> and <i>E. coli</i> ribonucleases from the RNB-family.....	73
Figure S.3. Size-exclusion chromatography of <i>M. tuberculosis</i> and <i>E. coli</i> ribonucleases from the RNB-family.....	75
Figure S.4. Affinity data of the <i>M. tuberculosis</i> ribonucleases for binding to RNA	76

List of Tables

2. Material and Methods.....	21
Table 2.1. Nucleotide sequence of the RNA substrates used in the experimental work....	26
3. Results and Discussion	31
Table 3.1. Dissociation constant of RNR1 and RNR2 wild type and mutant proteins	40
6. Supplementary Material	69
Table S.1. TB drugs classification and their targets.....	67
Table S.2. Proteins used in the experimental procedures.	71
Table S.3. Composition of the RUBIC Buffer screen conditions.	76
Table S.4. Composition of the RUBIC Additive screen conditions.....	78
Table S.5. Midpoint temperatures of the protein-unfolding transition for RNR1 in the presence of each RUBIC Buffer and Additive Condition.	79
Table S.6. Composition of the RNR1 NanoDSF screen conditions.	79
Table S.7. Composition of the Durham pH screen conditions.....	80
Table S.8. Composition of the Durham Salt screen conditions.	81
Table S.9. Midpoint temperatures of the protein-unfolding transition for RNR2_D267N in the presence of each Durham pH and Salt Condition.	83

Abbreviations

A	Adenine
Amp	Ampicillin
APS	Ammonium persulfate
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
bp	Base pair
BSA	Bovine Serum Albumin
CAI	Codon Adaptation Index
CSD	Cold Shock Domain
C-terminal	Carboxi-terminal
ddH₂O	Double-distilled water
dsRNA	Double-stranded RNA
DTT	Dithiotreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediamine tetraacetic acid
e-value	Expect value
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LA	Luria-Bertani Agar
LB	Luria-Bertani Broth
MDR-TB	Multi-drug resistant Tuberculosis
Mg²⁺	Magnesium ion
mRNA	Messenger RNA
MTB	<i>Mycobacterium tuberculosis</i>
MTBC	<i>Mycobacterium tuberculosis</i> complex
N-terminal	Amino-terminal
nt	nucleotide
OB	Oligonucleotide-binding
OD₆₀₀	Optical Density at 600nm
PAA	Polyacrylamide
PAGE	Polyacrylamide Gel Electrophoresis
PMSF	Phenylmethylsulfonyl Fluoride
PNK	Polynucleotide kinase
RNA	Ribonucleic acid

RNase	Ribonuclease
rpm	Rotations per minute
SDS	Sodium Dodecyl Sulphate
ssRNA	Single-stranded RNA
TB	Tuberculosis
TBm	Terrific Broth Medium
TBE	Tris-borate-EDTA
TBS	Tris-buffered saline
TEMED	Tetramethylethylenediamine
T_m	Melting Temperature
TrB	Transfer Buffer
UV	Ultraviolet
V	Volts
WHO	World Health Organization
WT	Wild type
XDR-TB	Extensively-drug resistant Tuberculosis

1. Introduction

The human body comprises millions of harmless and helpful microorganisms essential to shape an effective host immune system¹. These organisms, known as the resident microbiota, are usually limited to certain areas of the body and they are able to co-exist in a symbiotic relationship with the human host, establishing unique metabolic traits and providing benefits mutually². However, if the immune system is weakened or if the resident microbiota enters a normally sterile part of the body, it may lead to disease.

Pathogens, contrarily to the resident microbiota, do not require the host to be immunocompromised to impair the normal function of the body. To successfully survive and multiply within a host, the pathogen must be able to enter, travel to the location where it can establish itself, evading or overcoming the host's immune response, and cause damage, leading to infection. In many cases, the cycle is completed when the pathogen exits in the host and is transmitted to a new host³.

The disease caused by the direct effect of a pathogen, either cellular or acellular, is known as an infectious disease. In contrast to a noncommunicable infectious disease, that cannot spread from one person to another, some infectious diseases are communicable, meaning that they are transmissible, directly or indirectly, from one person to another⁴. According to the last World Health Organization (WHO) report of 2016, more than half of all deaths in low-income countries were caused by infectious communicable diseases, namely tuberculosis, malaria and human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS)^{5,6}.

1.1 Tuberculosis

Tuberculosis (TB) is an airborne disease caused by members of the *Mycobacterium tuberculosis* complex (MTBC)⁷. The complex is comprised by several genetically related mycobacterial species, such as *Mycobacterium tuberculosis* (MTB), *Mycobacterium bovis*, *Mycobacterium africanum*, *Mycobacterium microti*, *Mycobacterium caprae*, *Mycobacterium pinnipedii*, *Mycobacterium canetti* and *Mycobacterium mungi*, that mainly differ among themselves in their epidemiology⁸. Although *M. tuberculosis* is the predominant causative agent of human TB worldwide, in certain geographical regions, *M. africanum* and *M. bovis* are also important agents of human disease^{6,9}.

Tuberculosis is still among the top 10 causes of death worldwide, presenting a major threat to human health. It is the world's leading cause of death from a single infectious organism,

ranking above human immunodeficiency viruses/acquired immunodeficiency syndrome (Figure 1.1)⁵. According to the most recent data available from WHO, in 2018 alone, TB was responsible for 1.5 million deaths worldwide and 10 million new cases of this disease were estimated, in all age groups and countries¹⁰.

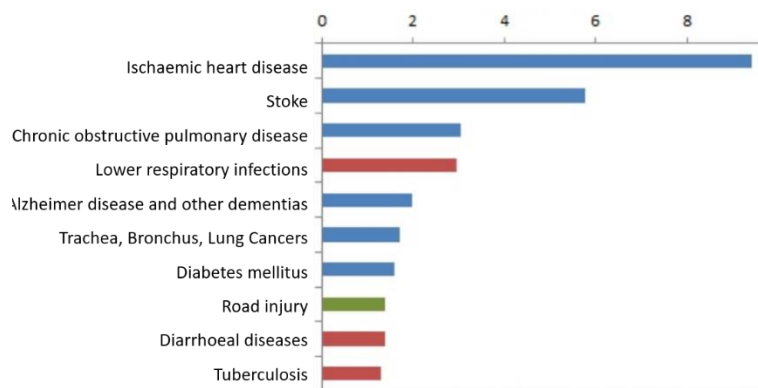


Figure 1.1. Top 10 global causes of death. Communicable, maternal, neonatal and nutrition conditions are depicted in red, noncommunicable diseases in blue and injuries in green. Adapted from World Health Organization, 2018⁵.

1.1.1 *Mycobacterium tuberculosis* Features

M. tuberculosis is the causative agent of most cases of tuberculosis. This pathogen is a nonmotile, rod-shaped bacterium, with a typical size of 0.5 to 3 μm and a generation time of approximately 15-20 hours. Although the growth behavior of *M. tuberculosis* strongly depends on the growth conditions¹¹, this pathogen is a facultative intracellular strictly aerobic parasite and, contrarily to many other bacteria, has an asymmetric division, producing two differently sized daughter cells^{12,13}.

Phylogenetically, *M. tuberculosis* is a Gram-positive bacterium that belongs to the rather diverse group of organisms with high GC content in their genomes¹⁴. Noteworthy, the mycobacterial cell wall has a unique and complex architecture composed of multiple covalently and noncovalently linked layers, which confers mechanical and chemical resistance to these organisms. The defining aspect of this structure is the outer membrane, the mycomembrane, that is rich in mycolic acid and represents 60 % of the mycobacterial cell wall forming the characteristic waxy coat of mycobacteria^{12,13}. This peculiar cell wall structure plays a fundamental role in virulence, providing this organism with an exceptionally strong impermeable barrier to several noxious compounds and drugs¹⁵. Interestingly, although classified as a gram-positive organism, *M. tuberculosis* possesses a cell envelope that shares notable features with Gram-negative cell walls.

Contrarily to other pathogens, *M. tuberculosis* does not produce any virulence factors such as exotoxins or endotoxins. Instead, to ensure its persistence in the host, this organism uses the host's macrophages as carriers to traverse mucosal barriers and reach sterile regions deep in the body. Damage is caused by uncontrolled and progressive chronic inflammation¹⁶.

1.1.3 Transmission and Pathogenesis of Tuberculosis

Humans are not uniformly susceptible to *M. tuberculosis* infection, which implies the existence of resistance mechanisms capable of operating differently according to each immune system. Contrarily to immunocompetent patients, that are often capable of clearing the invading organism completely before it causes any symptoms, in immunocompromised hosts, the immune system may fail to control the infection. This usually leads the disease to progress to an active symptomatic state, causing chronic cough with blood-tinged sputum, fever, night sweats, and weight loss^{6,17}.

The most common route of transmission of TB is through the air when an infected patient with active tuberculosis coughs, sneezes, or otherwise transmits saliva, releasing aerosol droplets containing pathogenic organisms inside^{6,16}. Infection occurs when a different individual inhales these droplet nuclei and they then reach the host's alveoli. Here, the resident macrophages quickly phagocytose and, most often, kill the mycobacterial invaders due to the innate immune responses. However, if the bacilli can survive this first line of defense, they start to actively replicate and the pathogen is able to use macrophages to diffuse through the lung epithelium to deeper tissues, infecting other cells¹⁸.

Thereafter, the adaptive immune system is activated resulting in granuloma formation^{9,19}. Granuloma is an organized aggregate of differentiated macrophages and other immune cells that can limit *M. tuberculosis* growth as well as it provides a survival niche from which the bacteria may disseminate to the newly arriving macrophages. As the inflammation process continues, the granuloma becomes fully organized, forming a barrier shell that keeps the bacteria under control. In this scenario, known as latent tuberculosis infection (Latent TB), non-apparent signs or symptoms of disease ensue, and dynamic equilibrium between the bacilli and host immune responses is established. Nonetheless, if any event that weakens cell-mediated immunity occurs, the pathogen can overcome the immune system and multiply, resulting in progression from latent to active disease (Active TB). In the last presented scenario, the infected macrophages can undergo necrosis forming a necrotic core that supports bacterial growth and transmission to the next host, resulting in dissemination^{9,18}.

The size of the infecting dose of *M. tuberculosis* and the immune status of the host are two important factors that determine the risk of progression from Latent TB to symptomatic disease²⁰.

1.1.4 Available Treatments for Tuberculosis Disease

The current treatment for TB depends on whether an individual has drug-sensitive TB, drug-resistant TB, clinically active TB or only latent TB. In all scenarios, the goals are not only to cure the patient but also to minimize the risk of death and disability, to prevent future relapses and to reduce transmission^{8,20}.

During tuberculosis infection, patients can harbour three different *M. tuberculosis* sub-populations: actively multiplying bacilli in open cavities (responsible for transmission), slowly intermittently growing bacilli, and dormant bacilli. The latter are present in lesions characterized by an acidic environment and under anaerobic conditions, such as in inflammatory tissues or within macrophages, and are unaffected by standard therapies¹⁵.

To date, more than twenty drugs have been developed for TB treatment, that should be used in combination for an adequate duration to achieve a durable cure and avoid the emergence of resistance¹⁰. Anti-TB drugs can be classified into five main groups, based on a hierarchy of safety and/or the effectiveness considerations²¹ (Table S.1). Each drug has a specific action on one or more bacillary populations and aims to target various aspects of *M. tuberculosis* biology, including inhibition of cell wall synthesis, protein synthesis, or nucleic acid synthesis^{15,22}. Regimens for the treatment of TB disease must contain multiple drugs to which the bacteria are susceptible and the whole process takes much longer than treating other types of bacterial infections. If treatment is not continued for a long enough duration, the surviving bacteria may cause the patient to become ill and infectious again, potentially with drug-resistant disease²³.

1.1.4.1 Multi-Drug Resistant Tuberculosis

Multi-drug resistant tuberculosis (MDR-TB) is caused by *M. tuberculosis* organisms that do not respond to drugs normally used to treat the disease, namely isoniazid and rifampicin, the two most powerful anti-TB drugs¹⁰. Generally, treatment of MDR-TB requires second-line drugs such as Kanamycin, which are less effective, more toxic and much more expensive than first-line drugs²¹. Globally in 2018, 3.4 % of new cases and 18 % of previously treated cases had MDR-TB¹⁰. Transmission of MDR-TB occurs in the same way as drug-susceptible TB and is no more infectious. However, delay in the recognition of drug resistance, or prolonged periods of infectiousness, may facilitate increased transmission. In some countries, it is becoming

increasingly difficult to treat MDR-TB since treatment options are limited, expensive and the recommended medicines are not always accessible. This usually leads to a misuse of available drugs that can ultimately develop extensively drug-resistant TB (XDR-TB)²³.

XDR-TB is a form of MDR-TB resistant to at least four of the core anti-TB drugs which increases the complexity of treatment^{10,24}. By the end of 2018, the average proportion of MDR-TB cases with XDR-TB was already ciphered at 6.2 %¹⁰. Regimens for XDR-TB usually require at least 20 months of four to six second-line anti-TB drugs and outcomes are frequently poor¹⁵. According to the latest available data from the WHO, the treatment success rates are 85 % for drug-susceptible TB, 56 % for MDR-TB and 39 % for XDR-TB, which raises concerns between the scientific community of a future TB epidemic¹⁰. The emergence and spread of MDR-TB and, more recently, XDR-TB remains a health security threat.

At present, the available antibiotics used to fight infections, as it occurs in anti-TB drugs, have as targets essential prokaryotic enzymes involved in various pathways such as cell wall synthesis, protein translation, RNA transcription or DNA replication. Despite the existing diversity, most of the antibiotics are derivatives of molecules discovered in the early 1900's and are nowadays losing their foothold as effective means of treating infections with the rise in antibiotic resistance^{25,26}. Thus, there is an urgent need to find other strategies and to amplify the drug development processes to other attractive cellular targets in order to successfully overcome the existing drug-resistance problems.

In the past years, the idea of targeting directly RNA with the ultimate goal of expanding the antibiotic arsenal is emerging as a new frontier in drug discovery²⁷. RNA is upstream of all biology and the existing differences between the molecular machinery that governs bacterial and eukaryotic messenger RNA (mRNA) degradation are fundamental to take advantage of these attractive drug targets²⁵. Ribonucleases (RNases) are important members of this machinery and thus carry an important therapeutic potential, reason why their extensive study and understanding are crucial.

1.2 RNA Metabolism in the control of gene expression

The central dogma of molecular biology describes the two-step process by which the genetic information flows within a biological system, transcription and translation. Briefly, the information encoded in DNA is transcribed to mRNA by RNA polymerases, and mRNA is then translated into protein by ribosomes²⁸. RNA is a central intermediate and a key player in the whole protein-making machinery and, nowadays, it is known that many of cellular mechanisms cannot be fully understood without a profound knowledge of RNA metabolism²⁹. This molecule plays an important role in regulating several cellular processes in all living organisms, from cell division and differentiation to growth, death, cell ageing and fast adaptation to changing environments³⁰.

The level of expression of a gene is determined primarily by the efficiency of transcription of DNA into mRNA, the stability of mRNA, and the frequency of its translation into proteins³¹. Therefore, every factor that either modifies mRNA or that alters its decay, plays a major role in controlling gene expression³². In prokaryotes, the continuous degradation and synthesis of mRNAs allows not only the metabolic changes required for cells to grow and divide but also a rapid adaptation to new environments and conditions^{29,33,34}.

1.2.1 Prokaryotic mRNA stability

The term “stability” is defined as the chemical lifetime of RNA species, period during which the original full-size transcript remains physically intact, that can be estimated by determining its half-life in the cell^{35,36}. Over the years, the development of transcriptome analysis has enabled the monitoring of the rate of degradation of intracellular RNA levels at the genomic scale³⁶.

The lifetime of each mRNA is unique and mRNA degradation is a vital process that enable bacteria to rapidly alter patterns of protein synthesis in an ever-changing environment without any alteration in transcription³⁷. In eukaryotes, mRNA half-lives ranging from 3 to more than 100 minutes have been observed, whereas in exponentially growing prokaryotic cells shorter half-lives were reported³⁶. For fast-growing bacteria, such as *Escherichia coli* and *Bacillus subtilis*, the average half-lives of bulk mRNA were in the range of 2–10 minutes. Similarly, in *Mycobacterium smegmatis*, a frequently used mycobacterial model, the bulk mRNA half-life was around 5.2 minutes³⁵. Curiously, in *M. tuberculosis* the mean half-life of the mRNA transcripts was determined to be 9.5 minutes, substantially longer than the global half-life of other prokaryotes studied to date^{35,38}. More than 80 % of mRNA transcripts had half-lives between 8 and 12 min,

and even the shortest half-lives in *M. tuberculosis* were longer than the average transcript in *E. coli*, indicating that the mRNA pool in *M. tuberculosis* is very stable³⁹.

Interestingly, some studies revealed that *M. tuberculosis* undergoes drastic remodeling of its transcriptome in response to several stress conditions inherent to its pathogenicity. During transmission between hosts, this organism is exposed to significant temperature variations that results in an increase of the mRNA half-life by more than 5 hours and, in response to hypoxic stress, the decay of *M. tuberculosis* mRNA is at least 15-fold slower than in aerobic conditions. These observations reflect an impressive rise in mRNA stability³⁹ and suggest that mRNA stabilization contributes to energy conservation in the energy-limited environments that *M. tuberculosis* encounters during infection⁴⁰.

The stability of mRNA involves numerous players and relies on several features. One of the main effectors are ribonucleases (RNases), whose diversity, structures, targets, and modes of action can vary significantly, providing multiple solutions for a similar issue³⁶.

1.2.2 Ribonucleases in RNA metabolism

RNases are present in all domains of life and play a central role in the control of gene expression^{31,33,41}. These proteins are involved in determining the levels of functional RNAs in a cell, in the quality control of transcripts and also allow the recycling of the ribonucleotides^{36,41,42}. They can act alone, or be part of RNA degradation complexes, namely, the degradosome (in prokaryotes) and the exosome (in eukaryotes). Some of the existing RNases in the cell are essential enzymes, while others have overlapping functions. However, all of them operate according to the requirements of growth, in adaptation to a specific environment, and carrying out surveillance destroying aberrant RNA molecules that in turn would produce detrimental proteins^{29,31,41}.

RNases degrade ribonucleotides of all Archaea, Bacteria, and Eukarya members and have molecular sizes varying between 17 and 118 kDa⁴³. These enzymes can be divided into endoribonucleases, which cleave RNA molecules internally by digesting phosphodiester bonds in the RNA molecule, and exoribonucleases, which degrade RNA by the extremities, removing terminal nucleotides (nt) from the 3' end or the 5' end of the molecule that can be further reused in RNA synthesis^{29,41,44}.

1.2.2.1 Mechanisms of RNA decay

The action of each ribonuclease in RNA metabolism is influenced by the presence of specific sequences or structures in the RNA molecules that can act either as stabilizer or destabilizer elements to specific RNases. For example, the presence of ribosomes during active translation can hide some RNA *loci* that are vulnerable to degradation and even the addition of poly(A) tails can modulate the process due to the preference of several RNases for poly(A) stretches^{45,46}.

In bacteria, mRNAs can be degraded by several mechanisms that act independently but in parallel, targeting different sites with different efficiencies³⁵. Most of the bacterial organisms contain a low-specificity endonuclease that cleaves single-stranded RNA (RNase E/G, RNase Y, RNase J), a double-stranded RNA-targeting endonuclease (RNase III), at least one 3' exonuclease (PNPase, RNase II, RNase R) and an oligoribonuclease that degrades very short RNA fragments⁴⁷.

Noteworthy, the existing knowledge on the diverse mechanisms of mRNA decay and maturation in prokaryotes comes from studies with two model organisms, *E. coli* and *B. subtilis*, with their evolutionary distance providing a broad picture of potential pathways and enzymes involved in mRNA turnover³⁷. Interestingly, the major ribonucleases involved in the initiation of mRNA decay in the two organisms are very different, as shown in Figure 1.2³⁵.

In *E. coli*, the decay of the majority of transcripts starts with an endonucleolytic cleavage by RNase E that has a preference for 5'-monophosphorylated substrates or, although to a limited degree by RNase III, an enzyme that acts specifically on double-stranded structures (Figure 1.2 A). Following this initial endonucleolytic cleavage, breakdown products are ready for exonucleolytic digestion by any of the three main exonucleases in this bacterium, PNPase, RNase II and RNase R (Figure 1.2 A). An alternative pathway in the cell is the direct exoribonucleolytic degradation of full-length transcripts. In both pathways, exoribonucleolytic activity is promoted by the 3'-polyadenylation of substrates by PNPase and poly(A) polymerase I, the main polyadenylating enzyme in *E. coli* (Figure 1.2 A). The short fragments released by the exonucleolytic degradation are subsequently degraded to mononucleotides by oligoribonuclease^{36,48}.

In *B. subtilis*, a bacteria lacking RNase E, the RNA degradation mechanism is different. This organism possesses two paralogous endoribonucleases, RNase J1 and J2, that share functional homology with RNase E and exhibit equivalent endonucleolytic activity. However, RNase J1 catalyzes not only the endonucleolytic cleavage of an RNA substrate but also continues the degradation of the generated 5' end by switching to the 5' exonucleolytic mode. To date, RNase J1 is the only known ribonuclease in prokaryotes able to catalyze the exonucleolytic degradation

of RNA in the 5' to 3' direction^{36,49}. Regarding the mRNA degradation mechanism, transcripts can be degraded from the 5'-end through the 5'-3' exonuclease activity of RNase J1 after phosphate removal, or they can be first endonucleolytically cleaved either by RNase J1/RNase J2 or RNase Y (Figure 1.2 B). The breakdown products can be then further degraded by the 3'-5' exonucleases, PNPase and RNase R in unprotected 3' ends, or by the 5'-3' exonuclease activity of RNase J1, generating monophosphorylated 5'-ends (Figure 1.2 B). While RNase J1 is able to fully degrade its RNA substrates to mononucleotides, the final products released by both RNase R and PNPase need to be further degraded by an oligoribonuclease^{36,50}.

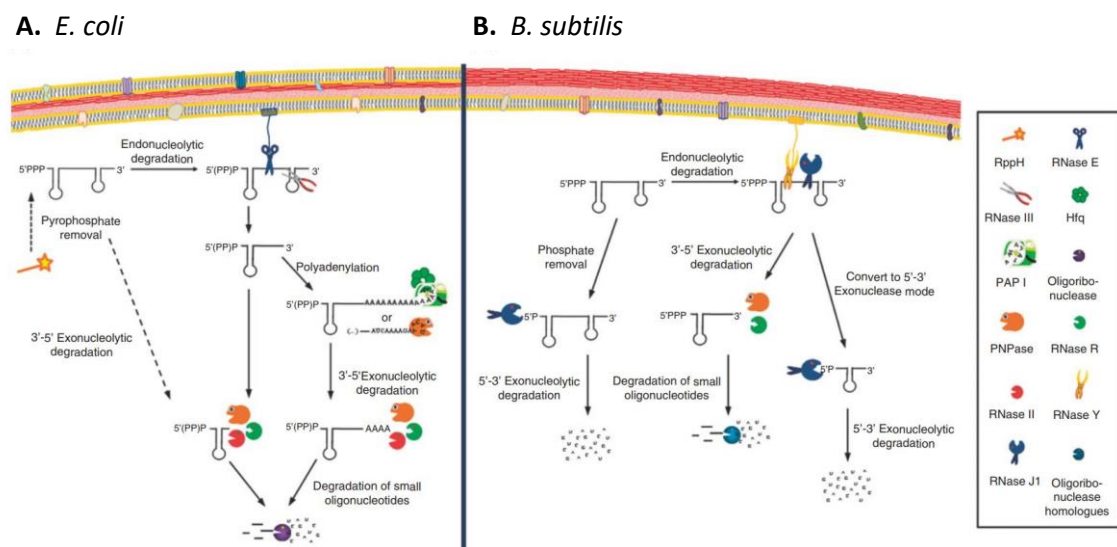


Figure 1.2. Schematic representation of mRNA degradation in *E. coli* (A) and *B. subtilis* (B)³⁶. These two model organisms possess different RNA-degradation enzymes but the strategy is similar: the process is initiated by an endonucleolytic cleavage followed by processive exoribonucleolysis of the generated fragments. The short fragments released are subsequently degraded to mononucleotides by an oligoribonuclease.

In *M. tuberculosis*, the underlying bacterial mRNA degradation mechanism remains to be understood although there is currently no doubt that it is an important aspect of *M. tuberculosis*'s intracellular life⁵¹. As previously mentioned, this organism is capable of both intra- and extra-pulmonary invasion and it has the remarkable ability to persist for years inside the host in a state of reversible dormancy. Such survival capacity requires a constant monitoring of a range of environmental *stimuli* and sometimes can require some drastic and rapid changes in gene expression^{51,52}.

In this context, there are certain processes that may contribute to the extended mRNA stability in this organism. *M. tuberculosis* ribonucleases could limit the rate of mRNA decay, either through slower enzyme kinetics or lower abundance compared with other bacteria. For

example, the initial destabilization step in *E. coli* is catalyzed by the endonuclease RNase E and the homolog to this gene in *M. tuberculosis* is kinetically equivalent to the enzyme in *E. coli*. However, the *M. tuberculosis* version is more specific about the cleavage sites, which can directly result in a slower global rate of mRNA decay. Alternatively, features of the MTB transcripts such as stabilizing secondary structure or post-transcriptional modifications can also contribute to diminished decay rates³⁹.

1.2.2.2 Ribonucleases in *M. tuberculosis*

The combined action of endo- and exoribonucleases is essential to control the balance between synthesis and degradation of transcripts in *M. tuberculosis*⁵². This pathogenic organism encodes a curious set of RNases with elements from both gram-negative and gram-positive species that are directly involved in various cellular processes, among which only six have been structurally characterized^{52–55}.

PNPase, RNase E and RNase J were identified as the major components of mycobacterial degradosome⁵⁶: PNPase is a phosphorolytic enzyme processing RNA in the 3'-to-5' direction, RNase E possesses an endoribonucleolytic activity and finally, RNase J comprises both endo and exoribonuclease activities^{56,57}. Mutations in PNPase, an exoribonuclease, were recently linked with *M. tuberculosis* resistance to the first-line antituberculosis drug pyrazinamide⁵⁸ which gives an important highlight to this type of proteins and justifies the importance of expanding the existing knowledge in this organism, especially regarding protein families for which no data or literature is available.

In this work, we will study RNases belonging to the RNB-family of enzymes, the principal hydrolytic 3'-5' exoribonucleases in bacterial cells, in order to gain some more insights into the mRNA decay process in *M. tuberculosis*⁵⁹.

1.2.3 RNB-family of exoribonucleases

Members of the RNB-family of enzymes are ubiquitous and play a wide spectrum of functions in the cell^{44,60,61}. In prokaryotes, these exoribonucleases have shown to be involved in RNA metabolism, viability, stress responses, competence and, more importantly, in some pathogenic organisms they play an important role in the establishment of virulence^{60,62,63}. In eukaryotes, proteins from this family can act independently or can be associated in a multiprotein complex called the exosome, a complex of exoribonucleases crucial for RNA metabolism, where the catalytic subunit can be a RNB-family member designated Rrp44/Dis3⁶³.

Mutations in the genes that encode for these proteins are linked with some abnormal chloroplast biogenesis, mitotic control and several diseases^{60,64}.

All the enzymes belonging to this family processively degrade RNA in the 3' to 5' direction with a hydrolytic activity releasing 5'-nucleotide monophosphates^{44,60,65} and are characterized by the presence of four domains: two cold-shock domains (CSD1 and CSD2) at the N-terminal region involved in RNA binding, a central RNB domain responsible for catalysis and a C-terminal S1 domain, also involved in RNA binding (Figure 1.3 A)^{61,62,66,67}. The RNB domain is highly conserved and contains four prominent sequence motifs (I to IV), with several invariant amino acids, exclusive of the RNase II-like proteins(Figure 1.3 B) ^{68,69}.

E. coli RNase II is the prototype of this family of enzymes, which also comprises RNase R and the eukaryotic Dis3-like proteins^{60,61,70}. All proteins possess a similar domain architecture, the most variable zone being the N-terminus, both in length and amino acid sequence (Figure 1.3). The bacterial RNase R contains a helix-turn-helix domain at the N-terminal region and a highly basic lysine-rich region after the S1 domain. Human DIS3 contains an additional endonucleolytic PIN domain preceded by a region with three conserved cysteines (CR3 region) at the N-terminal extension^{61,62}. The two other DIS3 homologues in the human genome, DIS3L and DIS3L2, differ in some structural characteristics. DIS3L possesses an inactive PIN domain and, similarly to Dis3, associates with the exosome complex. Contrarily to the other two enzymes, DIS3L2 lacks the PIN domain region, does not associate with the exosome and RNA uridylation was shown to stimulate its degradation activity⁷¹⁻⁷³.

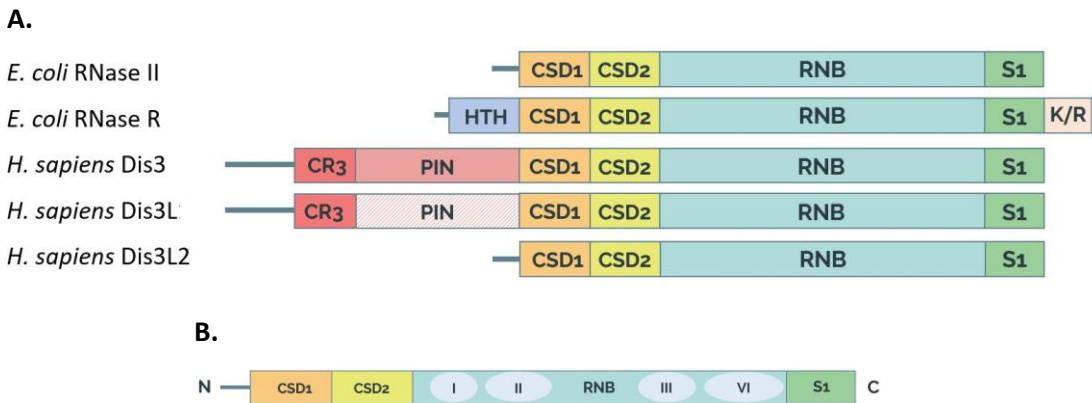


Figure 1.3. The RNB-family of enzymes. **A-** Schematic representation of RNB-family members which share a similar modular organization, with more pronounced differences in the N-terminal region. Adapted from Matos *et al.* (2011)^{41,61}; **B-** Linear representation of RNase II domains. The four sequence motifs of the RNB domain are depicted (I-IV)⁷⁴. Figures are not to scale.

The representative members of this family, *E. coli*'s RNase II and RNase R, possess different mechanisms of action and different capacity to degrade distinct RNA molecules. Their structure and catalytic properties will be further discussed below.

1.2.3.1 *E. coli* RNase II

E. coli RNase II is a 72 KDa protein encoded by the *rnb* gene and responsible for 90 % of the exoribonucleolytic activity in crude extracts, being the major hydrolytic enzyme in this organism^{41,44,60,75}. Its expression is regulated transcriptionally by two different promoters, P1 and P2; post-transcriptionally by ribonucleases such as PNPase, RNase E and RNase III^{41,60,76} that directly or indirectly interfere with *rnb* mRNA; and also post-translationally^{77,78} by the *gmr* gene (Gene Modulating RNase II), which is located downstream of the *rnb* gene and affects the stability of RNase II, allowing its regulation according to the growth and environmental conditions^{79,80}. In the past years, the thermosensitive *rnb500* allele has been useful to study this protein *in vivo*⁸¹.

As a member of the RNB-family, this enzyme processively degrades RNA in the 3' to 5' direction, meaning that it has the ability to degrade nucleotides from the 3' end without releasing the substrate during the cleavage process, generating as an end-product a 4 nt fragment. In the presence of RNA molecules with less than 10 nt, the activity of RNase II becomes distributive, dissociating from the RNA after every catalytic event⁶². RNase II activity is blocked by the presence of secondary structures in the RNA fragments and is sequence-independent, despite the marked preference for adenosine-rich regions in the 3' end^{60,62,75}. By rapidly degrading the poly(A) tails, RNase II can either accelerate the degradation of targeted RNA molecules or impede the binding of other exoribonucleases to the RNA molecule preventing degradation and acting as a protector of some messages^{41,46,80}.

The determination of the crystal structure of *E. coli*'s RNase II at 2.44 Å resolution (PDB entry: 2IX0⁶⁶), complemented by a thorough biochemical analysis, gave new insights into the mechanisms of catalysis, translocation and processivity of this important RNA-degrading enzyme. Interestingly, at the same time that the wild type RNase II structure was solved, the structure of the D209N point mutant was also determined with 2.74Å resolution (PDB entry: 2IX1⁶⁶). This natural mutant, originated by a spontaneous mutation that encodes an inactive form of the protein able to bind to RNA but unable to cleave it, was co-crystallized with a single-stranded 13 nt fragment, allowing the identification of the residues in contact with the RNA^{41,64,82} (Figure 1.4).

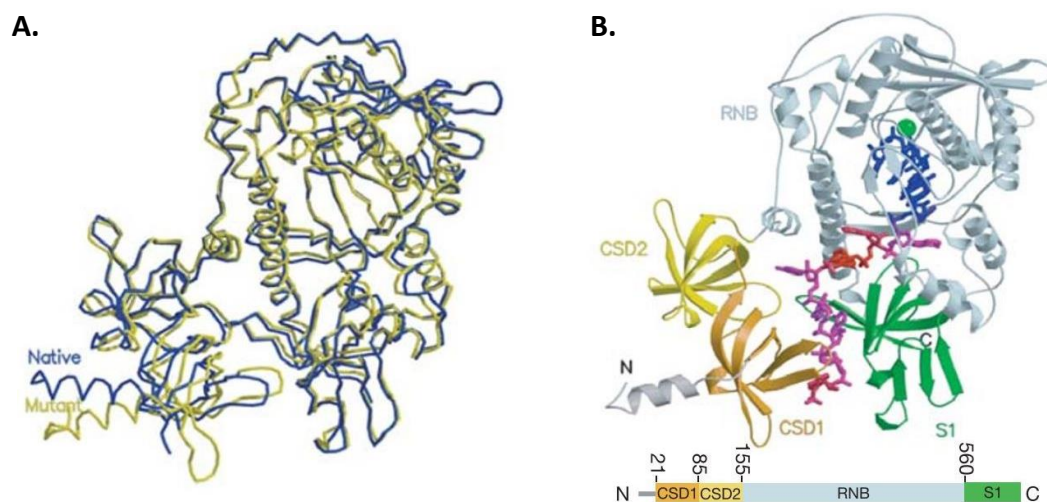


Figure 1.4. *E. coli* RNase II protein structure. **A-** Superposition of the polypeptide chains of native RNase II (blue) and its D209N derivative (yellow). **B-** Three-dimensional representation of RNase II–RNA complex, showing domains CSD1 (orange), CSD2 (yellow), RNB (light cyan) and S1 (green); the Mg ion is depicted as a green sphere and RNA is shown in sticks (blue to red); below the picture bottom, the respective linear representation of RNase II domains with numbers specifying the domain boundaries⁶⁶.

The RNA-binding domains of *E. coli*'s RNase II are grouped on one end side of the structure, while the active site lies on the other end of the molecule^{41,66}. Each domain is able to act independently and, regarding the RNA-binding capacity of the protein, S1 proved to be more critical than the N-terminal CDS1 and CDS2^{62,80}. Both CSDs and S1 domains display a similar fold consisting of five-stranded antiparallel β -barrel known as oligonucleotide-binding (OB) fold, commonly seen in nucleic acid-binding proteins⁸³. Interestingly, these OB motifs in CSD1 and CSD2 face the solvent, which may suggest their function in RNA recruitment⁶⁶.

RNA contacts the enzyme at two different and non-contiguous regions that act synergistically to provide a processive degradation: the anchoring region, formed by the three RNA-binding domains, and the catalytic region, buried inside the catalytic domain⁷⁹. In fact, the shortest RNA molecule able to contact with these two regions simultaneously is a 10 nt fragment, which is the minimum length necessary to maintain the processivity of the enzyme, as previously mentioned. Shorter RNA-fragments establish fewer interactions with the protein, which may explain why the enzyme becomes distributive⁶⁶. The last five nucleotides at the 3'-end of the RNA molecule are clamped between two aromatic residues Y253 and F358 of the catalytic region in the RNB domain. When the RNA molecule is shorter than 5 nt, the stacking no longer occurs, releasing a 4 nt fragment as the end product of this enzyme^{41,66}. In fact, Y253 is very important for the maintenance of the RNA “clamping” and a mutational analysis identified

this residue as responsible for setting the final end product in RNase II. Its substitution by an alanine changed the RNase II product from 4 to 10 nt^{41,60,74}. Access to the catalytic pocket is restricted to ssRNA owing to the steric hindrance at its entrance, which explains the inability of RNase II to degrade double-stranded RNA (dsRNA)⁶⁶.

Moreover, RNase II contains four highly conserved aspartic acids in the active site, namely D201, D207, D209, and D210, that were shown to be important to position the RNA molecule correctly and to promote the nucleophilic attack of the phosphodiester bond. These aspartates are not equivalent in their function, D209 being the most essential one^{41,60,70}. Although only one Mg²⁺ ion was found in both structures ligated by D201, D210 and water molecules, it was postulated the recruitment of a second Mg²⁺ for catalysis which could assume the function of the second Mg²⁺ in the two-metal-ion nucleotide cleavage mechanism^{66,84}. This second labile Mg²⁺ ion, coordinated by D209 and D207, promotes a nucleophilic attack by an activated water molecule. The D209N mutant impedes the coordination of the second Mg²⁺, explaining the mutant inactivity⁶⁰.

In the active site of the enzyme, another highly conserved residue is present, R500, which was proved to have a crucial role in RNA degradation. This residue was postulated to assist in catalysis by fixing the phosphodiester bond at the cleavage position and enhancing the susceptibility of the leaving phosphorous atom to a nucleophilic attack. Its substitution by an alanine practically lead to the inactivation of the enzyme, supporting the essential role of R500 in assisting catalysis^{41,60}. Glutamic acid 542 (E542), a residue in close proximity with the leaving nucleotide, also showed to be important for RNase II activity and it was proposed to help in the nucleotide elimination after the cleavage event, thus allowing the degradation process to continue. Surprisingly, when substituted by an alanine, the mutant protein became 110-fold more active and has a 20-fold higher RNA affinity when compared to the wild type protein, having being described as the RNase II “Super-Enzyme”^{64,74,85}.

Additionally, the three dimensional structure allowed to see specific amino acid residues interacting with ribose 2' oxygens⁶⁶, confirming that RNase II has a strict requirement for a ribose at the second and/or the fourth nucleotide from the 3'-end of the molecule to maintain its activity^{60,74}. RNase II is also capable of binding DNA molecules but not to hydrolyze them, meaning that DNA can act as an inhibitor of this enzyme^{41,44,64}.

The proposed model for RNase II-mediated RNA degradation is exemplified in Figure 1.5. Single-stranded RNA (ssRNA) binds at the interface between the OB and the RNB domains and is threaded into the catalytic cavity. In this cavity, tight packing of the five 3'-terminal nucleotides

occurs together with numerous RNA–protein hydrogen bonds. After cleavage of the phosphodiester bond, the protein promotes the nucleotide elimination by reinforcing an existing hydrogen bond and formation of a new hydrogen bond with E542. As the nucleotide leaves, the RNA translocates and a new phosphodiester bond gets positioned for the next cleavage event. When the RNA is shorter than 5 nt, the required packing of the bases does not occur, thus preventing the translocation of the RNA and generating a 4 nt fragment as the final product⁶⁶.

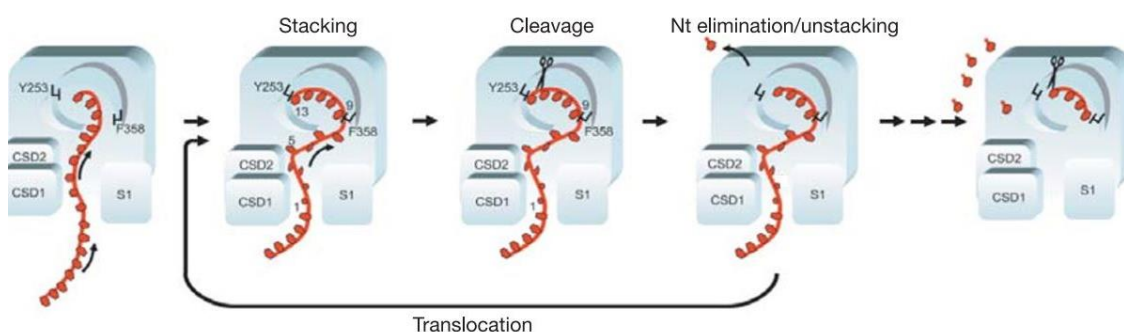


Figure 1.5. Model for RNA degradation by RNase II⁶⁶. ssRNA (red) is threaded into the catalytic cavity and clamped between Y253 and F358. The additional stabilization of RNA inside the cavity drives the RNA translocation after each cleavage, up to a final 4 nt fragment.

1.2.3.2 *E. coli* RNase R

E. coli RNase R, a 92 kDa protein encoded by the *rnr* gene, is a processive, 3′–5′ exoribonuclease that acts through a hydrolytic mechanism releasing nucleoside monophosphates^{68,86,87}. This protein shares 60 % sequence homology with *E. coli*'s RNase II. Degradation is also processive and sequence-independent but, unlike RNase II, RNase R releases a dinucleotide as the final product of digestion and has the remarkable feature of degrading both single- and double-stranded RNA without the help of a helicase^{68,86–88}. Like no other known 3′ to 5′ exoribonuclease, RNase R possesses an intrinsic unwinding activity, regardless of its inherent ribonuclease activity. Nevertheless, to properly degrade dsRNA, this enzyme requires a 3′ single-stranded overhang with least 7 nt in length to serve as a binding site from which degradation can be initiated^{87,89}.

RNase R is a cold-shock protein whose activity in the cell is modulated according to different environmental *stimuli* and growth conditions, such as cold shock or stationary phase of growth. This enzyme is regulated at the post-transcriptional level with the participation of RNase E that is involved in the processing of *rnr* transcripts and at the level of protein stability by the post-translational acetylation status of a single lysine residue, K544^{67,80,90,91}. In exponential phase, acetylation targets RNase R for degradation, while during several stress conditions, such as

stationary phase or adaptation to cold, the protein responsible for that acetylation is absent and thus, RNase R is stabilized being induced by several fold^{67,71}. Strikingly, this ribonuclease has also been implicated in virulence mechanisms of different microorganisms. The stress-induction of RNase R levels can be highly advantageous, enabling pathogens to adapt to environmental challenges imposed prior and during the infection process. Accordingly, most RNase R-deficient bacteria have been shown to be less virulent than the wild-type strains^{41,92}.

Regarding its structural features, RNase R is a multidomain protein that generally follows the modular arrangement found on RNB-family members, the catalytic RNB domain flanked by the three RNA-binding domains (CSD1, CSD2 and the S1 domains), with two additional domains: a K/R rich domain at the C-terminal region and a helix-turn-helix domain (HTH) at the N-terminus^{41,71} (Figure 1.3 A). These potential extra protein-RNA interactions might contribute to unwind secondary RNA structures and help to explain the impressive RNase R mode of action. In fact, the helicase activity associated with this protein is particularly important for the catalytic efficiency of RNase R⁹³.

Most of the residues interacting with the 3' end of RNA are conserved throughout this family, as previously mentioned for RNase II, and most of the critical residues for exoribonucleolytic activity are located in equivalent spatial positions. The highly conserved aspartate residues located in the active centre of RNase R, D272, D278, D280, D281, were shown to be responsible for catalysis and involved in the coordination of divalent metal ions, as in the other members of the RNB-family. In particular, RNase R mutant D280N proved to have no exoribonucleotidic activity but was still able to bind RNA efficiently, analogous to what was reported with the equivalent D209N mutant in RNase II, confirming the indispensable role of this aspartate residue in catalysis in this family of ribonucleases^{41,87,94,95}. Additionally, the clamping Y324 residue, the equivalent to Y253 residue from *E. coli* RNase II, was found to be responsible for setting the final end-product of RNase R. This residue is fully conserved in all domains of life and was proved to have the same function in several family members^{41,87}.

Despite the similarities, there is an evident divergence in substrate recognition between RNB-family enzymes, that can only be explained by giving a close look at the structural differences amongst these exoribonucleases⁴¹. For instance, comparing both RNase R and RNase II sequences, it was noteworthy that the presence of F358, an important clamping residue together with Y253, is exclusive to RNase II. Instead, RNase R presents a phenylalanine residue in the immediate downstream position, F429, perhaps with similar functionality in fixing the RNA, that could explain the existing differences in regard to RNA degradation⁹⁴. Moreover, the

amino acid residue R572, located in the nuclease domain channel of RNase R, strongly affects its catalytic properties. Mutation to a lysine, the equivalent residue found in RNase II, shown to greatly impair RNase R activity on structured RNAs^{41,96}. Finally, in contrast to what occurs in RNase II, it was showed that the RNB domain alone in RNase R is capable of degrading duplex RNA with or without a 3' overhang, although relatively ineffectively compared with the full-length protein. This suggests not only that this domain has all binding, unwinding and degrading properties, but also that other regions of the protein must also contribute to digestion of structured RNA.^{89,94,97,98}

Additionally, it was found that several RNases R of different bacteria possess two conserved ATP-binding Walker A and Walker B motifs, with the Walker A motif located in the C-terminal region of the protein and the Walker B motif present in the CSD2 domain in the N-terminal extension. In fact, the helicase activity present in RNase R is highly unusual in that the protein lacks almost all the motifs usually associated with RNA helicases, except for those abovementioned involved in ATP binding⁹⁸. It was proposed that the presence of duplex RNA at the entrance to the catalytic channel induces a conformational change that brings the Walker A and B sequences together to generate the ATP-binding site. Once bound, the presence of ATP stimulates binding of the strands of the RNA duplex to the CSD and S1 domains that results in partial strand separation⁹³.

Interestingly, the helicase activity of RNase R is only ATP-binding dependent and the hydrolysis of the molecule is not required. However, bacterial strains with a helicase-deficient RNase R due to mutations in the Walker motifs revealed growth deficiencies at low temperatures, which proves the importance of the ATP-binding site^{98,99}. All conformational change, ATP binding, and helicase activity are an integral part of the mechanism through which RNase R is able to digest structured RNA^{93,98}. Moreover, the S1 domain showed to have some amino acid residues with important roles in the degradation of structured RNA, particularly D716 and E717. Mutations in those residues render RNase R unable to bind ATP and to lose the unwinding ability⁹³.

Crystallization of full-length RNase R proved to be a challenging task. Currently, only a truncated structure at 1.85 Å resolution (PDB entry: 5XGU⁸⁷) lacking the N-terminal helix-turn-helix domain, the C-terminal K/R rich region, and with an A131V mutation (RNase R ΔHTH-K mutant) is available (Figure 1.6 A, 1.6 B). The truncation proved not to significantly affect the overall enzymatic activity and the spontaneous mutation enhanced crystallization, but the reason remains unclear since the mutated residue is located in a disordered loop^{71,87}. The

structure revealed a unique feature of the enzyme, showing that the exoribonuclease has not only one but two open channels, contrary to *E. coli* RNase II. A top channel between the S1 and CSD1 domains, with sufficient width for accommodating a duplex RNA, and a narrower side channel between the RNB and CSD1 domains. Additionally, a close look at the active site reveals a Mg^{2+} ion bound by residues D272 and D281 (Figure 1.6 C) that matches the same position of the Mg^{2+} ion in the crystal structures of RNase II, as expected⁸⁷.

Another important feature of RNase R structure is a tri-helix region located in the RNB domain (Figure 1.6 A,B) that was proved to play a crucial role in RNA unwinding activity without interfering in substrate binding and degradation, suggesting that it may serve as the 'wedge' to unwind duplex RNA⁸⁷. In fact, an analysis of several RNA-bound crystal structures of RNB-family members showed that the tri-helix region also interacts with RNA in Dis3L2 and Rrp44, two enzymes capable of degrading dsRNA, but not in RNase II, which argued for the importance of this region for the unwinding activity^{71,87}.

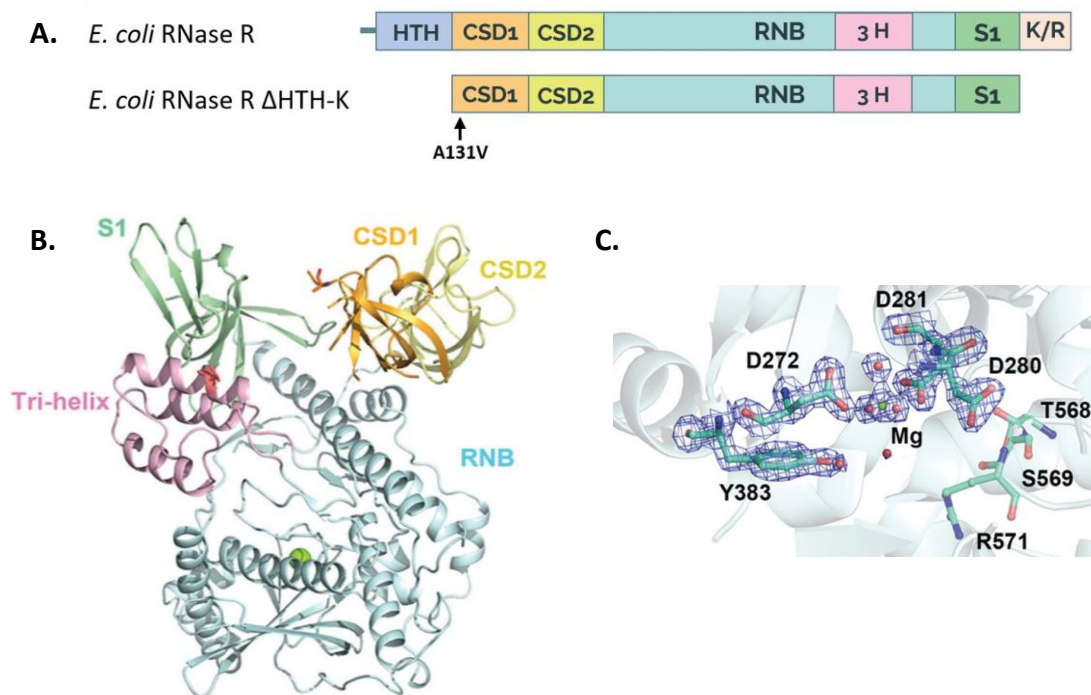


Figure 1.6. *E. coli* RNase R protein structure. **A** - Linear domain organization of full-length RNase R and the truncated RNase R HTH-K; **B**- Crystal structure of RNase R HTH-K. The RNB domain (cyan) is capped with S1 (green), CSD1 (orange) and CSD2 (yellow) domains. The Mg ion located in the active site is shown as a green sphere and the 3-helix region (3H) highlighted in pink. **C**- Active site in the RNB domain contoured at 1.9 σ (blue). The Mg ion is displayed as a red sphere⁸⁷.

Currently, two possible RNA unwinding and degradation models for RNase R have been proposed, based on the crystal structure determined, depending on which one of the two channels is responsible for threading the RNA to the active site or serve as an exit for the non-scissile 5' end that is unwound during degradation.

The first model is based on the RNase II-RNA complex showed above, whereby the single-stranded overhang of structured RNA enters from the top channel and is threaded all the way into the active site in the RNB domain (Figure 1.7 A). The RNA duplex region is bound by the CSD1 and S1 domains in the top channel and, upon reaching the tri-helix wedge region of the RNB domain, the RNA duplex is unwound. RNA unwinding is driven by RNA hydrolysis, with RNA being continuously degraded in a processive manner. While the unwound 3' overhang is pulled into the active site for processive RNA degradation, the 5' non-scissile strand exits from the side channel⁸⁷.

The second model is based on the Rrp44-RNA crystal structure of Rrp44 bound with RNA, whereby the single-stranded overhang of duplex RNA is guided through the side channel into the active site and the RNA duplex region is bound between the CSD1 and RNB domains (Figure 1.7 B). Upon reaching the tri-helix wedge region of the RNB domain, the RNA is unwound and its 3' overhang is further threaded into the RNB active site for cleavage while the 5' end of the non-scissile strand exits via the top channel. It should be noted that the side channel in RNase R is not wide enough to accommodate a duplex RNA, however, the CSD1/2 domains may be flexible and change their conformation upon RNA binding⁸⁷.

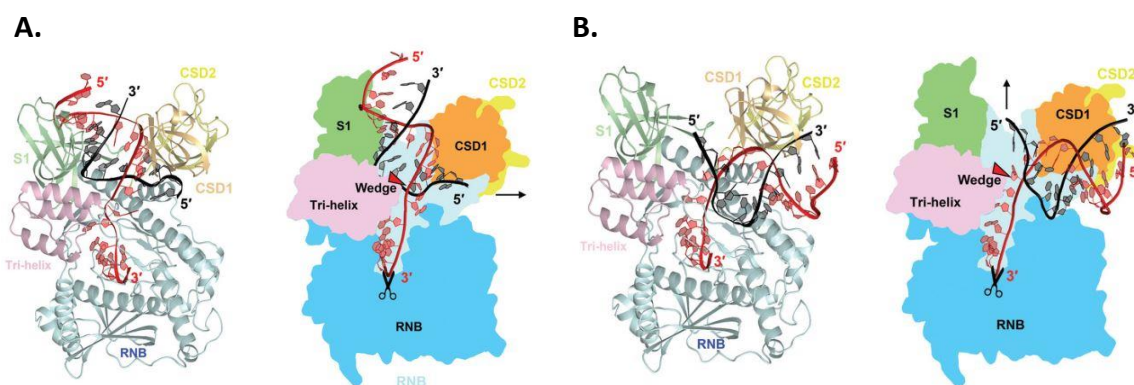


Figure 1.7. Working models for RNA binding, unwinding and degradation by RNase R. A- An RNase R-RNA model was built based on the crystal structure of RNase II bound with RNA (PDB entry: 2IX1). B- A second RNase R-RNA model was built based on the crystal structure of Rrp44 bound with RNA (PDB entry: 2VNU).

Taken together, all the regions combine in a way that makes RNase R unique amongst exoribonucleases and both models provide a plausible explanation for how RNA with extensive secondary structures can be degraded efficiently by this enzyme that concurrently unwinds and degrades RNA molecules. Nonetheless, a co-crystallization of RNase R with a structured RNA substrate would certainly allow for a refined model of its action, as happened with other members of the RNB-family of enzymes^{41,71}.

1.3 Thesis Motivation and Aims

Tuberculosis continues to be a major global health problem worldwide, despite all the years of scientific effort to identify effective treatments. The causative agent of this disease, *M. tuberculosis*, has the potential to adapt to a variety of different environmental challenges that require constant adjustments in RNA levels³⁹. Exoribonucleases from the RNB-family are involved in RNA metabolism and are known to be important for the establishment of virulence in an increasing number of pathogens¹⁰⁰.

Currently, there is no data or literature available regarding the RNB-family of enzymes in *M. tuberculosis*. The work here performed focused on the identification and characterization of two putative members of the RNB-family of enzymes in this human pathogen. Our main goal was to perform a thorough biochemical, biophysical and structural characterization of the proteins to better understand their action mechanism and binding mode. For that, the proteins were purified and several activity, binding and helicase assays were performed *in vitro* to properly characterize the degradation properties of the targets. In parallel, several crystallization trials and protein thermostability assays proceeded to attempt to determine the three-dimensional structures of these RNases. Point mutations were also studied to assess the influence of certain key residues in the properties under study. Ultimately, the objective of the work is to unravel novel potential antimicrobial targets that can possibly be useful to find new strategies to fight *Mycobacterium*.

2. Materials and Methods

All the media, buffers and solutions required for the experimental procedures were prepared as summarized in Table S.2.

2.1 Bacterial Strains

2.1.1 *Escherichia coli* DH5 α

Escherichia (*E.*) *coli* DH5 α competent cells are engineered to improve transformation efficiency and frequently used for general cloning applications. This particular strain carries a mutation on the *recA* gene that ensures a higher insert stability avoiding heterologous recombination and a mutation in the *endA* gene that increases the yield and quality of DNA preventing plasmid digestion.

Genotype: F⁻ Φ 80/*lacZ* Δ M15, Δ (*lacZ*YA-*argF*), U169, *recA1*, *endA1*, *hsdR17* (*r*_K⁻, *m*_K⁺), *phoA*, *supE44* λ ⁻ *thi-1* *gyrA96* *relA1*

Source: *Invitrogen*TM

2.1.2 *Escherichia coli* BL21 (DE3) Δ *rn*b Δ *rn*r⁸⁶

E. coli BL21(DE3) competent cells are engineered for protein expression purposes using the T7 expression system. This strain carries the DE3 lysogen and lacks the genes encoding for *E. coli* RNase II and *E. coli* RNase R. It is deficient in some proteases and OmpT which increases stability of the expression protein.

Genotype: F⁻, *ompT*, *hsdSB* (*r*_B⁻, *m*_B⁻), *gal*, *dcm*, λ (DE3), Δ *rn*b Δ *rn*r

Source: Kindly provided by Murray P. Deutscher, Department of Biochemistry and Molecular Biology, University of Miami School of Medicine, Miami, Florida.

2.2 Methods

2.2.1 Bioinformatic Analysis

A protein BLAST (Basic Local Alignment Search Tool) ¹⁰¹ using *E. coli* RNase II (Uniprot ID P30850) and *E. coli* RNase R (Uniprot ID P21499) as query sequences was performed restricting taxonomically the search for *M. tuberculosis* Lehmann and Neumann 1896 (taxid:1773) strain. The parameter considered to validate a BLAST search result was the Expect (E) value. The two selected hits possessed accession numbers SGC48612.1 and SGC96063.1. The first, named hereinafter RNR1, presented an E-value of 0.0, corresponding to $\approx 66\%$ of identity with the query sequence. The second hit, named henceforward RNR2, also yielded an E-value of 0.0, for $\approx 74\%$ of identity with the query sequence. Additionally, a multiple sequence alignment using *Clustal Omega*¹⁰² was performed (Figure S.1) to confirm the existence of sequence similarity between both targets.

2.2.2 Plasmid Construction

The vectors used in the present experimental work, containing the genes that encode for *M. tuberculosis* RNR1, RNR2, RNR1_D211N and RNR2_D285N have been designed and optimized by *GenScript's OptimumGene*¹⁰³ in order to achieve the highest possible levels of productivity in our bacterial expression system, *E. coli* competent cells. The native genes possess several rare codons from MTB that could reduce the efficiency of translation or even disengage the translational machinery in *E. coli* mainly due to the metabolic differences existing between both organisms. Thus, the codon usage bias in *E. coli* was increased by the manufacturers and the Codon Adaptation Index (CAI) was upgraded. An additional mutant was obtained during the experimental work while attempting to produce the other mutations, RNR2_D267N (data not shown). The optimized gene sequences were cloned into pET15b vector between *NcoI* and *BamHI* restriction sites, under the control of a T7 promoter, allowing the expression of His6-tagged RNR1 and RNR2 fusion proteins (Table S.3). The vectors also contain the ampicillin (*amp*) resistant gene and the *lacI* gene to allow expression upon IPTG induction (Figure 2.1).

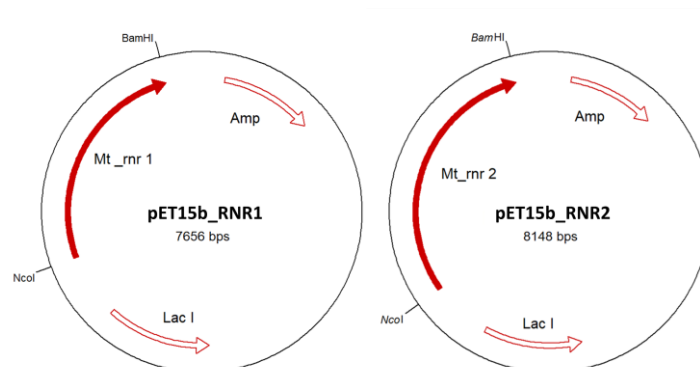


Figure 2.1. Schematic diagrams of pET15b_RNR1 and pET15b_RNR2, respectively. RNR1, RNR2, *amp* and *lacI* regions are represented by arrows pointing the direction of transcription. *NcoI* and *BamHI* restriction sites are indicated in the figure. The image was produced with Clone Manager software¹⁰⁴.

2.2.3 Transformation of Competent Cells

E. coli DH5 α and BL21(DE3) $\Delta rnb \Delta rnr$ competent cells were transformed with the previously mentioned synthetic plasmids using the heat shock method: 100 ng of the plasmid DNA were mixed with 50 μ L of competent cells, incubated on ice for 30 minutes, placed in a water bath at 42°C for 1 minute and then placed back on ice for an additional minute. Afterwards, to recover from the shock, 400 μ L of LB medium was added to the mixture followed by incubation at 37°C for 1 hour. The cells were then plated on LA selective media (supplemented with 100 μ g/mL Amp) and the growth of isolated colonies was verified after an overnight incubation at 37°C. A single colony from each plate was picked to grow overnight in LB and a small volume of this culture was adjusted to a final concentration of 10 % (v/v) glycerol to be kept at -80°C as stock for further use.

2.2.4 Protein Overexpression

Expression tests with RNR1 and RNR2 were performed in different growth conditions (16°C, 30°C and 37°C in LB or TBm) to determine the best protein expression conditions (data not shown). For protein overexpression, cells were grown in 600 mL of TBm for RNR1 variants and in 1000 mL of LB for RNR2 variants, supplemented with 100 μ g/mL ampicillin at 30°C until a OD_{600nm} of 0.5. At this point, protein expression was induced by the addition of 0.5 mM IPTG and bacteria were grown for further 16 hours at 30°C (RNR1) or 16°C (RNR2). Cell cultures were centrifuged at 5000 rpm for 10 minutes at 15°C and the cell pellets were stored at -20°C.

2.2.5 Cell Lysis

Following protein overexpression, the stored pellets were thawed, resuspended on the lowest possible volume of buffer A and supplemented with 0.5 mM phenylmethylsulfonyl fluoride (PMSF). Samples were divided into different tubes containing Glass Beads (*Sigma-Aldrich*) and cell lysis was achieved by mechanical disruption in a FastPrep-24™ grinder instrument (*MP Biomedicals*) set to 6.5 m/s for 2 cycles of 60 seconds each. The crude extracts were centrifuged at 15000 rpm for 30 minutes at 4°C and 0.5 µl of Benzonase 250 U/µL (*Sigma-Aldrich*) was added to the lysate and incubated on ice for at least 30 minutes. The clarified extracts containing the protein of interest were kept on ice for further protein purification.

2.2.6 Protein Purification¹⁰⁵

2.2.6.1 Affinity chromatography

Purification of the recombinant proteins was performed by histidine affinity chromatography using 5 mL HisTrap HP columns (*GE Healthcare*) on an ÄKTA Start system (*GE Healthcare*), following the manufacturer's instructions. For each protein, a different column previously washed with water and equilibrated with buffer A, was used. The clarified extracts were loaded into the columns using a flow rate of 1.0 mL/min to allow the proper binding of the recombinant proteins to the column. Columns were washed with buffer A until constant UV absorbance was observed and protein elution was achieved by increasing the percentage of buffer B passing in the column (from 20 mM to 500 mM of imidazole). SDS-PAGE analysis was performed on a small volume of each collected fraction by mixing it with 2 µL of protein loading buffer 6X and boiled in water for 10 minutes. These samples were loaded into a pre-cast Novex™ 4-12 % Tris-Glycine SDS-PAGE (*Invitrogen*) and run at a constant voltage of 180 V. The gels were then stained with Blue Safe protein stain (*NZYTech*) at RT and scanned using the Image Lab Software bundled with the Gel Doc EZ Gel Imaging System (*BioRad*). Henceforth, whenever this technique is mentioned, the executed procedure was similar to the previously described. After SDS-PAGE analysis, similar fractions containing the purified protein were pooled together and concentrated up to 500 µL by centrifugation at 4000 rpm and 4°C using a Vivaspin™ 6 MWCO 50 sample concentrators (*GE Healthcare*).

2.2.6.2 Size Exclusion chromatography

Protein purification was complemented with a size exclusion chromatography polishing step. The concentrated samples were loaded into a Superdex 200 Increase 10/300 GL column (GE Healthcare), previously washed with 2 volumes of water and equilibrated with 2 volumes of buffer C on an ÄKTApurifier UPC 10 System (GE Healthcare). Protein elution was followed analyzing the UV graph in the UNICORN software (GE Healthcare) and protein collection was performed in microcentrifuge tubes. The collected fractions were visualized in a pre-cast Novex™ 4-12 % Tris-Glycine SDS-PAGE stained with the Bluesafe protein stain (NZYTech) at RT. Selected fractions containing the protein of interest were combined and concentrated by centrifugation at 4000 rpm, 4°C using Vivaspin™ 6 MWCO 50 sample concentrators (GE Healthcare). After concentration, 50 % (v/v) glycerol was added to the samples intended to be used in the activity assays before storage at -20°C. The remaining purified proteins were shock frozen in liquid N₂ and kept at -80°C for further use in stability and crystallization screenings.

Alongside, *E. coli* RNase II and *E. coli* RNase R were also expressed and purified (Figure S.2 and S.3) following the optimized protocol described by Matos *et. al*¹⁰⁵ to be used as a control in the *in vitro* studies.

2.2.7 Protein Quantification

Protein quantification was performed using the Bradford method. Briefly, 1 mL of Bradford reagent was mixed with 10 µL of each protein diluted to 1:10 and 1:20 ratios and incubated for 5 minutes at room temperature. The absorbance was then measured in a Ultrospec 3100 pro UV/Visible Spectrophotometer (GE Healthcare) at a wavelength of 595 nm. Protein concentration was calculated by interpolation of the results into a calibration curve prepared with known concentrations of Bovine Serum Albumin (BSA) diluted in water.

2.2.8 Western Blot Analysis

A western blot was performed to detect and identify the proteins of interest in the different fractions. Known amounts of the purified proteins were loaded into a pre-cast Novex™ 4-12 % Tris-Glycine SDS-PAGE and run as previously described. The gel was then washed in 1X Transfer Buffer (TrB) during approximately 20 minutes, before being transferred to a nitrocellulose membrane equilibrated in 1X TrB using a Mini Blot Module (Thermo Fisher Scientific). Protein transfer proceeded at 10V for 60 minutes. When finished, the membrane was incubated overnight in a blocking solution (5 % (w/v) non-fat milk diluted in 1x TBS) at 4°C under constant

shaking and washed 3 times with 1X TBS. The primary antibody solution (Anti-His diluted 1: 10000 in 1X TBS) was applied over the membrane for 2 hours at 4°C with gentle agitation and the membrane was again washed three times for 10 minutes with 1X TBS. Afterwards, the secondary antibody solution was added for another 2 hours (Anti-Mouse diluted 1: 10000 in 1X TBS) and the washing steps were repeated. Chemiluminescent detection was performed using Western Lightning Plus-ECL (*PerkinElmer*) reagent according to the manufacturer's recommendations. Signals were visualized and photographed with ChemiDoc™ XRS+ System (*Bio-Rad*).

2.2.9 Radioactive Labeling of Oligonucleotides¹⁰⁵

RNA Poly(A) and 30-mer substrates (Table 2.1) were radioactively labeled at the 5' end using the enzyme T4 polynucleotide kinase (PNK). Reactions were carried out in a final volume of 20 µL, by mixing 10 µM of each nucleic acid with 2 µL of $\gamma^{32}\text{P}$ -ATP, 1 µL of PNK, 2 µL of 10X PNK Buffer and nuclease-free water. After one-hour incubation at 37°C, PNK was inactivated at 80°C for 5 minutes and the substrates were purified to remove the unincorporated nucleotides using MicroSpin G-25 columns (*GE Healthcare*) as specified by the manufacturer.

Table 2.1. Nucleotide sequence of the RNA substrates used in the experimental work.

Substrates	Sequence
Poly(A)	5' AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3'
30-mer	5' CCCGACACCAACCACUAAAAAAAAAAAAA 3'
16-mer	5' AGUGGUUGGUGUCGGG 3'

2.2.10 Activity Assays¹⁰⁵

Synthetic radioactive labeled molecules were denatured at 100°C for 5 minutes and incubated at 37°C for at least 30 minutes to ensure that the proper native folding was acquired. Double-stranded RNA (dsRNA) substrate was obtained by incubating the labeled 30-mer oligoribonucleotide with the complementary nonradioactive 16-mer in a 1:5 (mol:mol) ratio at 100°C for 5 minutes and at 37°C for at least 1 hour. Exoribonucleolytic reactions were carried out in a final volume of 10 µL containing 25 nM of RNA substrate, activity buffer (20 mM Tris-HCl pH 8, 100 mM KCl, 1 mM MgCl_2 , 2 mM DTT) and different concentrations of each enzyme. Reactions were started by the addition of the enzyme and incubation at 37°C. Samples of 2 µL were withdrawn at different time points (1, 5, 10 and 20 minutes) and mixed with 1 µL of

formamide-containing dye supplemented with 10 mM EDTA to stop the reaction. In parallel, a control reaction was assembled containing the RNA substrate and the activity buffer but lacking the enzyme. Reaction products were denatured for 5 minutes at 100°C and resolved in a 20 % PAA (polyacrylamide) 19:1/7 M urea gel at 1750 V until bromophenol blue has migrated approximately 22 cm from the well. The gel was exposed on a phosphor screen for 24 hours and the radioactive substrates were analyzed by phosphorimaging using a FLA-5100 scanner (*FujiFilm*) and the respective software.

2.2.11 Electrophoretic mobility shift assay (EMSA)¹⁰⁵

The labeled poly(A) substrate was denatured at 100°C for 5 minutes and incubated at 37°C for 30 minutes. Reactions were carried out in a final volume of 10 µL containing 25 nM RNA substrate, binding buffer (10 mM Tris-HCl pH 8.0, 2 mM DTT, 100 mM KCl and 50 mM EDTA) and increasing concentrations of protein. Reactions were incubated for 10 minutes at 37°C followed by the addition of 2 µL of loading buffer 6X. The samples were then subjected to UV crosslinking (600 mJ, 2 min, 254 nm) using an UVC 500 CrossLinker (*Amersham Biosciences*), and the RNA-protein complexes were analyzed in a 6 % non-denaturing PAA gel. The run proceeded at 200 V in a cold room for 4 hours. The gel was exposed on a phosphor screen for 24 hours and the radioactive complexes were analyzed by phosphorimaging using a FLA-5100 scanner (*FujiFilm*) and the respective software. Afterwards, the radioactive signal in each bound and unbound band was quantitated and the percentage of bound RNA in each reaction was plotted versus the protein concentration. The data was fit with a binding equation using a nonlinear regression in Prism software to obtain the K_D value (Figure S.4).

2.2.12 Crystallization Trials

Purified proteins were initially screened in various crystallization conditions using two commercially available screens: JCSG+™ and BCS (*Molecular Dimensions*). The first crystallization trials were done with RNR1 in two different concentrations, 10 mg/mL and 15 mg/mL, and RNR2 at a concentration of 10 mg/mL. Experiments were performed using the vapour-diffusion sitting-drop technique with a Mosquito (*TTP LabTech*) robot dispensing system and 96-well flat bottom iQ plates (*TTP LabTech*). The drop ratio used was 1:1 of protein to precipitant solutions in a final drop volume of 200 nL, equilibrated against a reservoir of 40 µL crystallization screen solution. All plates were kept at 20°C and visualized under the microscope regularly.

2.2.13 Protein Thermostability Assays

2.2.13.1 Thermal Shift Assay^{106,107}

Protein melting temperature (T_m) determination was performed by monitoring protein unfolding using the fluoroprobe SYPRO Orange (*Sigma-Aldrich*). An initial step of optimization was performed varying the amount of dye and protein added to each condition in order to determine the ideal concentration of both components that should be used for an optimal assay performance. A range of final concentrations from 1X to 5X of SYPRO Orange, and 1 to 5 μg of protein was tested. The best fluorescence signal-to-noise ratio was obtained using 5 μg of protein and 5X dye as final assay concentrations for RNR1, and 3 μg of protein combined with 1X Sypro Orange for RNR2. These conditions were used for the remaining work.

Commercially available screenings, RUBIC and Durham Screens (*Molecular Dimensions*), were prepared according to manufacturer's instructions either in 96 or 384-well plates. The plates were sealed with optically clear adhesive sheets (*Applied Biosystems™*) and centrifuged in a swinging-bucket rotor for 1 minute at 1000 rpm immediately before the assays were performed to remove possible air bubbles. Thermal denaturation was carried out in a QuantStudio 7 Flex Real-Time PCR system (*Applied Biosystems™*), heating the samples using a temperature gradient of 1°C/min from 20°C to 95°C, followed by the fluorescence reading. At the end of the assay, the melting curves were analysed and the T_m 's calculated using the Protein Thermal Shift™ Software v1.4 (*Applied Biosystems™*).

2.2.13.2 Nano Differential Scanning Fluorimetry (NanoDSF)

Based on the best results obtained from the thermal shift assay, a few conditions were selected to perform the NanoDSF assay. Samples were set in a total volume of 20 μl containing 2 μL of protein and 18 μL of a solution previously prepared with the desired buffer, ionic strength and additive. Approximately 10 μL from each mixture was loaded to a capillary by touching one end of the capillary into the solution. Capillaries were placed in a Prometheus NT.48 instrument (*NanoTemper Technologies*), and thermal denaturation was monitored from 20°C to 95°C with stepwise increments of 1°C per minute, followed by the fluorescence reading. The results were analysed using the PR.ThermControl software (*NanoTemper Technologies*).

3. Results and Discussion

This section is divided into three main parts. The first part includes a target validation step to confirm the presence of both ribonucleases under study in the *M. tuberculosis* genome, as well as the purification of all protein variants. In the second part, the exoribonucleotidic activity and binding capacity of all these ribonucleases was tested using *in vitro* assays with different synthetic RNA substrates. Finally, several crystallization trials were performed followed by additional differential scanning fluorimetry experiments to assess the most suitable sample buffer conditions for the target ribonucleases.

3.1 Target Validation and Small-scale Protein Purification

Members of the RNB-family of enzymes are present in all domains of life and play important roles in RNA metabolism. *E. coli* RNase II is the prototype of this family of enzymes, which also comprises RNase R⁴¹. Members of this family have a similar modular organization presenting a central very conserved RNB domain responsible for the catalytic activity, two N-terminal cold-shock domains (CSD1 and CSD2) and an S1 domain at the C-terminal region involved in nonspecific RNA binding¹⁰⁸. RNase R-like enzyme are more heterogeneous in their size and composition and can comprise an additional helix-turn-helix domain and a lysine-rich region¹⁰⁹.

Homologues of this family were searched in *M. tuberculosis* genome using the BLAST¹⁰¹ protein tool and two candidates have been selected, hereafter referred to as RNR1 and RNR2. We were able to prove the existing homology amongst the targets taking into account the commonly used relation that considers two amino acid sequences to be homologous if they share more than 30 % identity over their entire lengths¹¹⁰. RNR1 shares ≈ 66 % identity with *E. coli* RNase II, while RNR2 proved to be ≈ 74 % identical to *E. coli* RNase R; both these values are well above the consensus cut-off of 30 % sequence identity. Furthermore, with a multiple sequence alignment (Figure S.1) it was possible to understand that *M. tuberculosis* proteins comprised the typical domain organization previously described and also the conserved residues characteristic of proteins belonging to the RNB-family, namely the four highly conserved aspartate residues important for catalysis and a conserved SPIRKY motif in the RNB domain, highlighted in grey in Figure S.1. The alignment also suggested that while RNR1 is similar to RNase II, RNR2 resembles a RNase R-like protein. This was confirmed through a phylogenetic analysis (Figure 1.1), where the evolutionary relationships between homologous sequences was determined, giving insights into the functional specificity of members of the RNB-family¹¹¹.

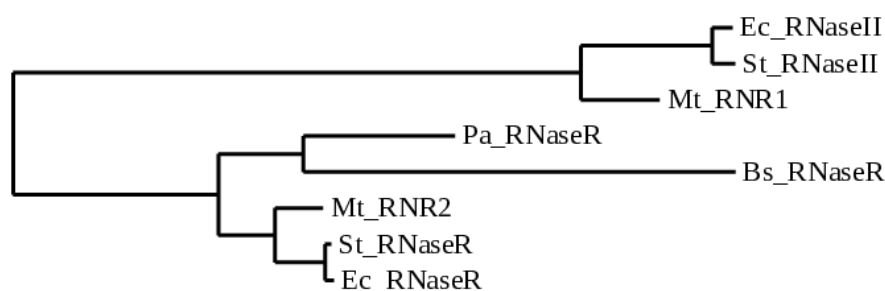


Figure 3.1 Phylogenetic tree between RNB-family members. The tree was constructed using the Phylogeny.fr software¹¹² based on protein sequences available with the following accession numbers: *E. coli* RNase II (P30850); *E. coli* RNase R (P21499); *M. tuberculosis* RNR1 (SGC48612.1); *M. tuberculosis* RNR2 (SGC96063.1); *Salmonella typhimurium* RNase II (SGSC1412); *Salmonella typhimurium* RNase R (Q8XF68); *Pseudomonas aeruginosa* RNase R (AOA0H2ZI34); *B. subtilis* RNase R (O32231).

Moreover, previous *in vitro* biochemical assays have highlighted the importance of the D209 residue for the activity of RNase II⁸². It is known that its single substitution to an asparagine residue leads to a complete inactivation of the enzyme without compromising the binding ability, which enabled the determination of the structure of the *E. coli* RNase II bound to RNA^{66,82,85}. This knowledge was already applied to effectively design the corresponding inactive mutants in homologs from other organisms and, more importantly, those mutants were proved to be valuable in the determination of their three-dimensional structures¹⁰⁸. We have managed to identify the corresponding conserved aspartate in RNR1 and RNR2 and constructed the corresponding single-residue mutants, RNR1_D211N and RNR2_D285N. While attempting to produce these mutations, a mutation in D267 residue from RNR2 (RNR2_D267N) was obtained. Considering that this residue is also conserved (Figure S.1), we have decided to also proceed with the analysis of this mutant. Our goal is to understand the influence of those residues in the catalytic activity of the enzyme and their importance in the RNA degradation mechanism. The inactive mutants can also be useful to obtain the RNA-bound three-dimensional structures of the proteins.

The genes were cloned into pET-15b and protein overproduction was performed in an RNase II and RNase R-less strain derivative of *E. coli* BL21(DE3)⁸⁶. This strain was used instead of a normal *E. coli* BL21(DE3) strain to prevent contaminations with *E. coli* RNB-family enzymes. Protein expression conditions were optimized for each ribonuclease and protein purification was accomplished as described in the Materials and Methods section. RNR1 purification results are present in Figure 3.2 as an example, while the remaining results are shown in the Supplementary Materials section (Figure S.2, Figure S.3).

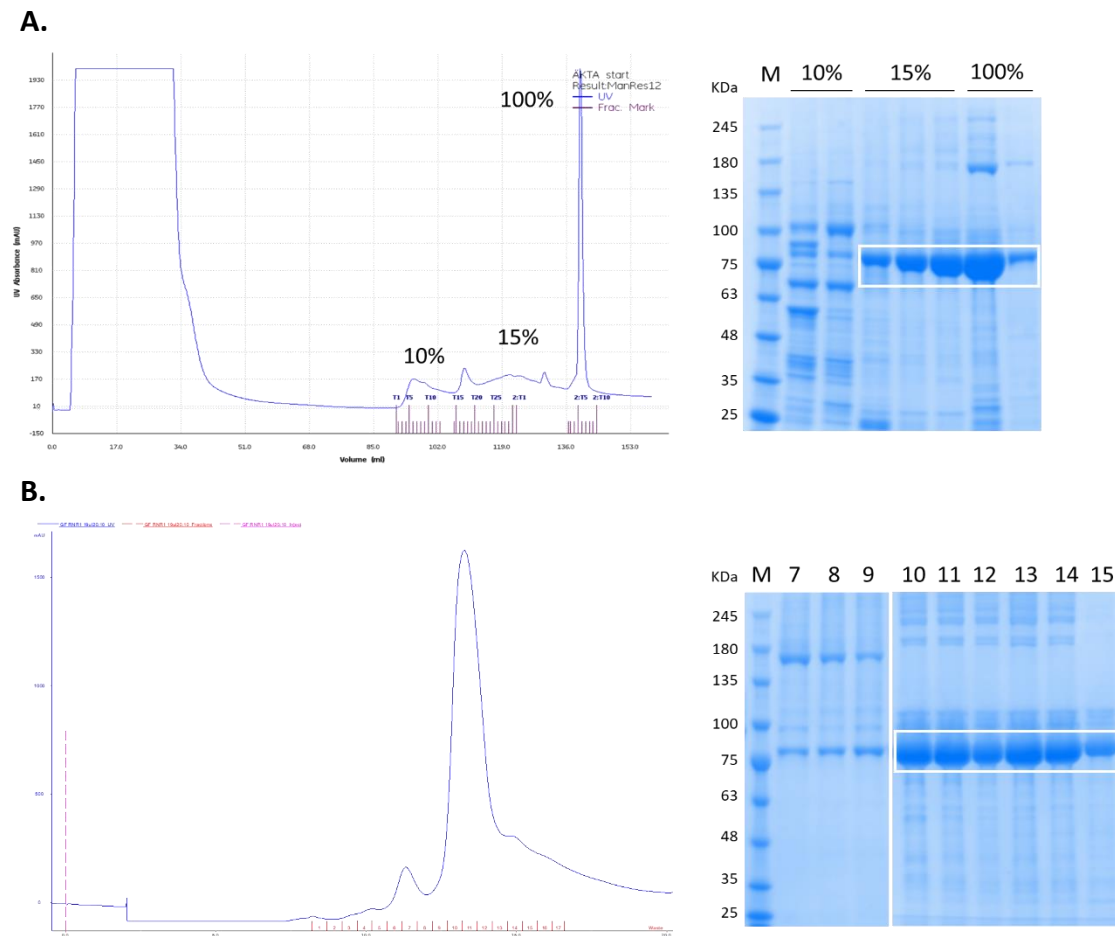


Figure 3.2 RNR1 purification by affinity and size-exclusion chromatography. **A, left panel-** Elution profile using Ni Sepharose High Performance column on an ÄKTA Start system. A gradient up to 500 mM imidazole was used and each peak corresponds to a different percentage of Buffer B passing in the column. **B, left panel-** Size-exclusion chromatogram of RNR1 using a Superdex 200 Increase 10/300 GL on an ÄKTA Purifier UPC 10 GE system. Sample injection is depicted in pink, UV absorbance in blue and the collected fractions in red. **A and B, Right panel-** 4-12 % SDS-PAGE analysis of the protein fractions collected during the corresponding chromatography. Each well was loaded with the same sample volume (10 μ L) and the marker used (lane “M”) was the NZYColour Protein Marker II (*Nzytech*). The bands corresponding to the target protein are highlighted by a white box.

The first protein purification step applied was an affinity chromatography. In this technique, the presence of surface-exposed histidine residues or other complex-forming amino acids can lead to unwanted binding of untagged host cell proteins to the HisTrap column. However, the binding affinity of these contaminants is often lower than that of the recombinant His-tagged proteins, and thus it is possible to reduce the unspecific bindings by optimizing the percentage of imidazole used to elute the protein. First, this competitive agent was added in low concentration in the binding buffer (Buffer A_RNR) to reduce the initial binding of contaminant proteins. Then, to determine the optimal elution conditions for each protein under study, an increasing imidazole concentration was applied rising the percentage of the sample buffer

(Buffer B_RNB) passing in the column, from 5 % to 100 %. Samples were collected and analysed by SDS-PAGE in each stage to define the best protocol to implement in further purifications. The results (Figure 3.2 and S.2) demonstrated that the optimal imidazole concentration is protein-dependent, and it must be determined case by case in order to increase the final purity and purification yield of the target proteins, reducing the amount of contaminants in the fractions of interest. Furthermore, it was notorious the formation of a white precipitate between purification steps, especially in RNR2 wild-type and mutant experiments. Thus, all the samples were centrifuged to eliminate the precipitate before the size-exclusion chromatography.

Subsequently, the concentrated proteins were further purified by size-exclusion chromatography using a Superdex 200 Increase 10/300 GL column to remove the remaining contaminants. This system is suitable for small-scale and high-resolution analysis applications and possesses a fractionation range between 10 and 600 kDa, an interval that comprises the molecular weight of our proteins. Figure 3.2 and Figure S.3 include the chromatograms obtained during this step and the analysis by SDS-PAGE of some relevant fractions. For RNR1, RNR2_D267N and RNR2_D285N the highest peak of the chromatogram, between 10 and 15 mL of the column volume, corresponds to the protein of interest as confirmed by the appearance of bands with the expected molecular weight in the SDS-PAGE. The smaller peaks are contaminants that are being separated from the protein of interest, as desired. On the contrary, although it is not possible to distinguish a single high-peak for RNR1_D211N and RNR2, the SDS-PAGE analysis revealed that the smaller peaks visible at 15 mL correspond to the target proteins. Thereby, all the proteins were successfully purified (Figure 3.3) and further used in the activity assays.

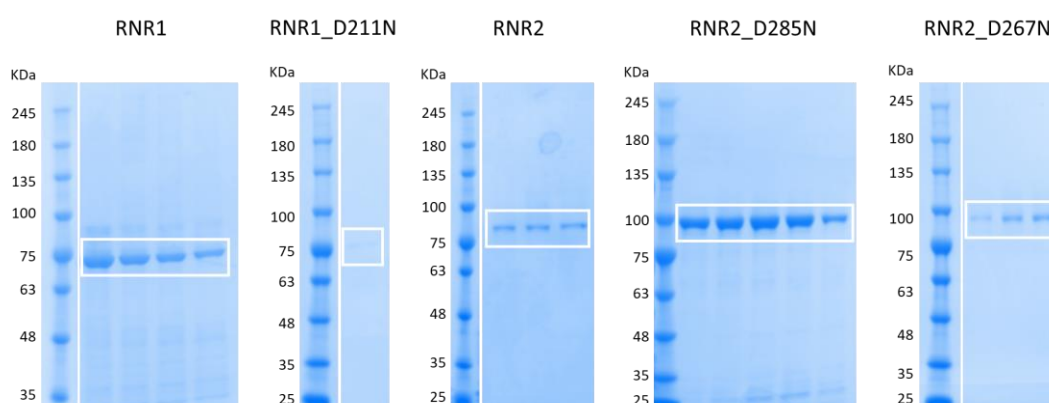


Figure 3.3. SDS-PAGE analysis of the purified proteins. 4-12 % SDS-PAGE analysis of samples containing the protein obtained after both purification steps. Each well was loaded with the same sample volume (10 μ L) and the marker used was the NZYColour Protein Marker II (Nzytech). The bands corresponding to the target protein are highlighted by white boxes.

After the wild-type purification, a gel was run with 1 μ g and 4 μ g of protein and the results show that both enzymes were migrating according to the expected size (Figure 3.4). However, in the wells containing 4 μ g of protein it was possible to observe bands that apparently do not correspond to the protein of interest. Even in the samples containing only 1 μ g, several bands with higher molecular weight than the target protein were visible. Thus, a western blot analysis was performed to confirm the identity of the samples and it was demonstrated that most of those bands correspond to the target RNases, particularly regarding RNR2 since the transfer of RNR1 to the membrane was not very efficient. Lower molecular weight bands can be protein degradation products, while the upper bands can correspond to the protein of interest bound to different nucleic acids, that migrate in the gel with a higher molecular mass than expected. Moreover, this migration pattern could also be a result of a non-homogeneous sample containing the protein in different oligomeric conformations that may not have denatured completely.

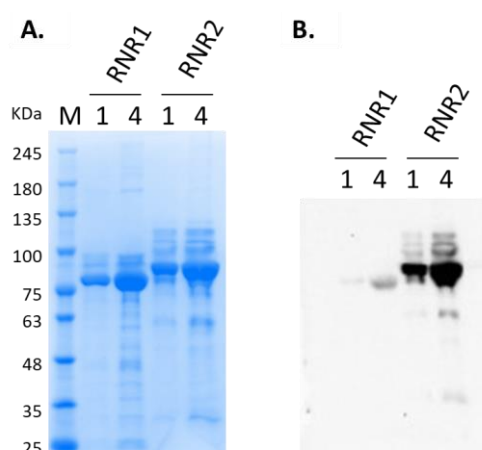


Figure 3.4. SDS-PAGE and Western blot analysis of recombinant RNR1 and RNR2 after purification. **A-** Samples containing 1 and 4 μ g of protein were separated by 4-12 % SDS-PAGE. Each well was loaded with the same sample volume (10mL) and the marker used (lane “M”) was the NZYColour Protein Marker II (Nzytech). **B-** Similar samples were analysed by western blot. The recombinant proteins were detected using a anti-His tag monoclonal antibody as primary antibody.

3.2 Biochemical Characterization of RNR1 and RNR2

After protein purification, several biochemical properties of the ribonucleases of interest were studied, namely the ribonucleolytic activity, RNA-binding capability and helicase potential. The enzymatic activity of a ribonuclease, similar to all kinds of enzymes, is influenced by the substrate. Thus, using distinct synthetic substrates in the assays, it is possible to mimic the actual *in vivo* enzyme substrates and reveal important features of their behavior in the living organisms⁴⁴.

In this work, two main substrates were used: a single-stranded poly(A) polymer and a double-stranded oligomer with a single-stranded extension in the 3' end. The first was chosen based on the well-known preference of *E. coli* RNase II for poly(A) stretches, while the second was selected to evaluate more specific requirements of the enzyme, such as single-stranded constraints.

3.2.1 Exoribonucleolytic Activity Assays

For each protein under study, *in vitro* activity assays were performed using both RNA substrates above mentioned. Initially, an experiment using the poly(A) ssRNA substrate and different concentrations of the wild-type proteins was performed to evaluate the activity of the enzymes (Figure 3.5). The conditions previously optimized for *E. coli* RNase II¹⁰⁵ were applied, using magnesium as a cofactor.

In some conditions tested, it was possible to visualize the substrate disappearing and the emergence of smaller fragments, showing that both RNR1 and RNR2 were able to cleave the ssRNA oligonucleotide. RNR1 proved to be active when present at a concentration as low as 1 nM and, as expected, it is possible to see an increase in the formation of products as the protein concentration increases. This ribonuclease has demonstrated a marked formation of a major 4 nt degradation product, especially in the assay with higher enzyme concentration (50 nM). The assays with lower enzyme concentrations (1 nM and 10 nM) revealed the existence of several intermediate reaction products in solution, with 6 nt and 8 nt-long, formed prior to the complete degradation down to 4 nt. RNR2 exhibited a minimal ribonucleolytic activity at 10 nM and only in the highest concentration tested (50 nM) the enzyme was able to completely degrade the substrate. This enzyme released a mixture of 2 nt and 4 nt fragments, the 2 nt being the predominant product. Together, these results show that both proteins are indeed 3' to 5' exoribonucleases.

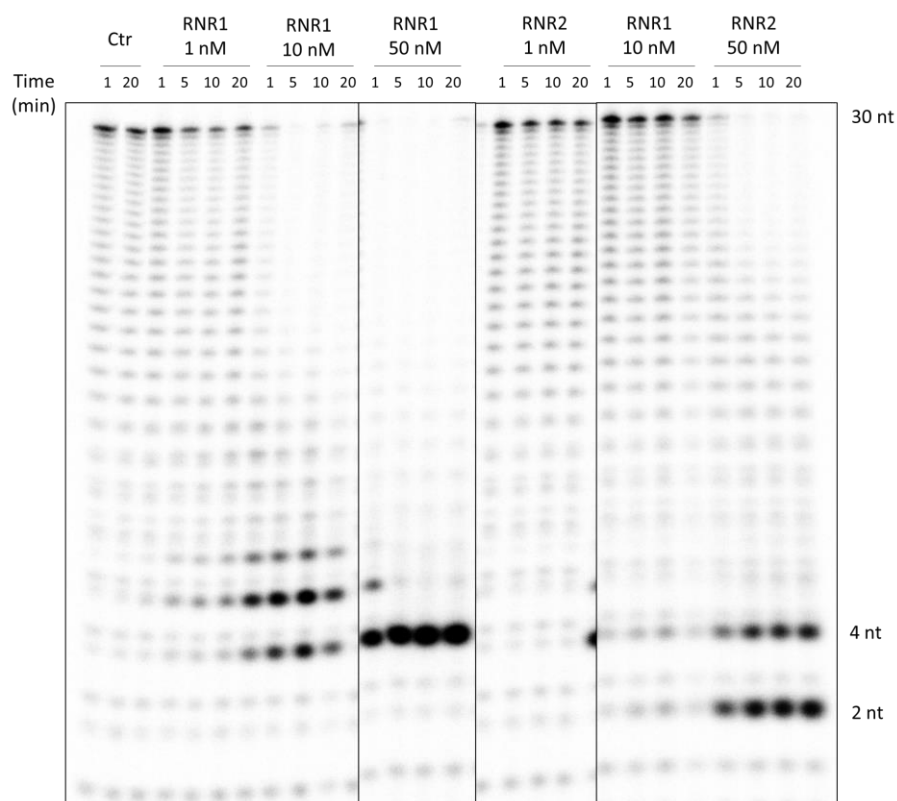


Figure 3.5. Exoribonucleolytic activity of RNR1 and RNR2 on a ssRNA. Different protein concentrations (1, 10 and 50 nM) were incubated at 37°C with 25 nM poly(A) substrate. Samples were taken at indicated time points between 1 and 20 minutes and reaction products were analysed in a denaturing 20 % PAA gel with 7M urea. Control reaction mixtures without enzyme (Ctr) were incubated at the maximum reaction time for each protein. Numbers to the right indicate sizes of the major reaction products.

Subsequently, the assay with the Poly(A) substrate was repeated testing all the recombinant proteins and the homologous *E. coli* enzymes in parallel to properly compare the degradation patterns (Figure 3.6). RNR1 demonstrated a degradation profile similar to the *E. coli* RNase II enzyme, corroborating the hypothesis of it being an RNase II-like enzyme, while RNR2 proved to be an RNase R-like enzyme, since the products released by the mycobacterial enzyme are identical to the ones formed by *E. coli* RNase R.

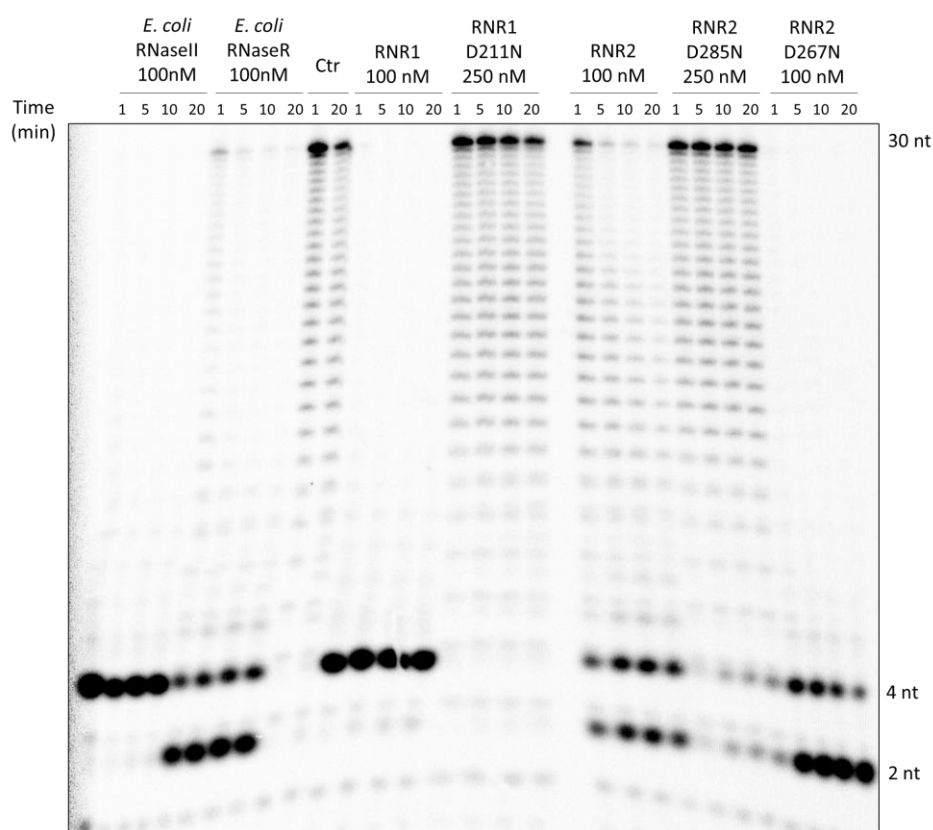


Figure 3.6. Exoribonucleolytic activity of RNR1 and RNR2 wild-type and mutants on a ssRNA. Enzymes were incubated at 37°C with 25 nM poly(A) substrate. The mutants and the protein concentrations used in each reaction are specified on top. Samples were taken at indicated time points between 1 and 20 minutes and the reaction products were analysed in a denaturing 20 % PAA gel with 7M urea. *E. coli* RNase II and *E. coli* RNase R were used for comparison. Control reaction mixtures without enzyme (Ctr) were incubated at the maximum reaction time for each protein. Numbers to the right indicate sizes of the major reaction products.

Interestingly, and similarly to what was observed for *E. coli* RNase II, the final degradation product of RNR1 is a 4 nt molecule, while for *E. coli* RNase R and RNR2 is always a mixture of fragments with 2 nt and 4 nt, even when higher protein concentrations are used. A possible explanation for the differences regarding the end-product between RNR1 and RNR2 may rely on the aromatic residues responsible for clamping the substrate in the catalytic pocket. In *E. coli* RNase II, the two clamping residues proved to be Y253 and F358⁶⁶; analysing the multiple sequence alignment produced (Figure S.1), it is possible to understand that those residues are conserved in RNR1, corresponding to Y255 and F360. As such, we can infer that this enzyme comprises a degradation mechanism similar to the one of *E. coli* RNase II, hence releasing the 4 nt fragment as end-product. Regarding the RNR2 protein sequence, it is clear that the corresponding tyrosine residue is also present, contrary to the phenylalanine that is not

conserved in all RNB-family members. Instead, there is a F359 in the position immediately downstream of the equivalent residue in RNR1 that can possibly perform the same function, in parallel to what was observed for *E. coli* RNase R¹¹³. For these enzymes, it was postulated that some 4 nt fragments are released when they are still partially “clamped” in the active site while others bind the catalytic cavity more efficiently and are degraded up to a 2 nt fragment. However, only the complete three-dimensional structure of an RNaseR-like protein with a bounded RNA fragment can clarify this question.

Moreover, the substitution of the strictly conserved aspartic acid by an asparagine residue (mutants RNR1_D211N and RNR2_D285N) completely inhibited the activity of RNR1 and rendered RNR2 almost inactive on this ssRNA substrate. The dramatic effect in catalysis caused by the exchange of only one residue strongly suggests the crucial role of this amino acid in these enzymes, similarly to what was observed for other family members¹⁰⁸. Based on what is known about *E. coli* RNase II, we postulate that one of the carboxyl groups of the aspartic acid is involved in the coordination of one of the magnesium ions required for the two-metal-ion nucleotide cleavage mechanism and, in its absence, a catalytically inactive protein is rendered⁶⁶. Thus, the residual degradation activity visible in the RNR2_D285N mutant was not expected and possibly occurred due to some contamination during the purification process or in the preparation of the activity assays. Additional experiments are needed to confirm this result.

As mentioned previously, the RNR2_D267N mutant was also purified and tested in the activity assays. This mutation does not alter the degradation profile of the enzyme, contrary to what was observed with the other variants. In fact, RNR2_D267N appears to be more active than the wild-type protein, degrading the substrate faster and more efficiently when the same concentration is used (100 nM). After analysing the crystal structures of *E. coli* RNase II and *E. coli* RNase R (PDB: 2IX0 and 5XGU), it was possible to see that the corresponding aspartic acid in this organism is placed in a disordered loop, far from the Mg²⁺ ions and from the catalytic pocket, which explains why this residue does not inhibit the exoribonucleolytic activity of the enzyme. The determination of the three-dimensional structure, together with the biochemical data, may clarify the impact of this residue in the increased catalytic activity of RNR2.

We also tested the ability of RNR1 and RNR2 for cleaving a structured RNA molecule, the 16-30ds RNA substrate (Figure 3.6). Noteworthy, the RNR1 activity was impaired by the presence of secondary structures, similar to what was reported for *E. coli* RNase II⁶². Both proteins are able to degrade only the 3' single stranded tail of the substrate, releasing a product 4 to 7 nt-long before the duplex region⁴¹. The RNR1_D211N mutant, again demonstrated to be completely inactive.

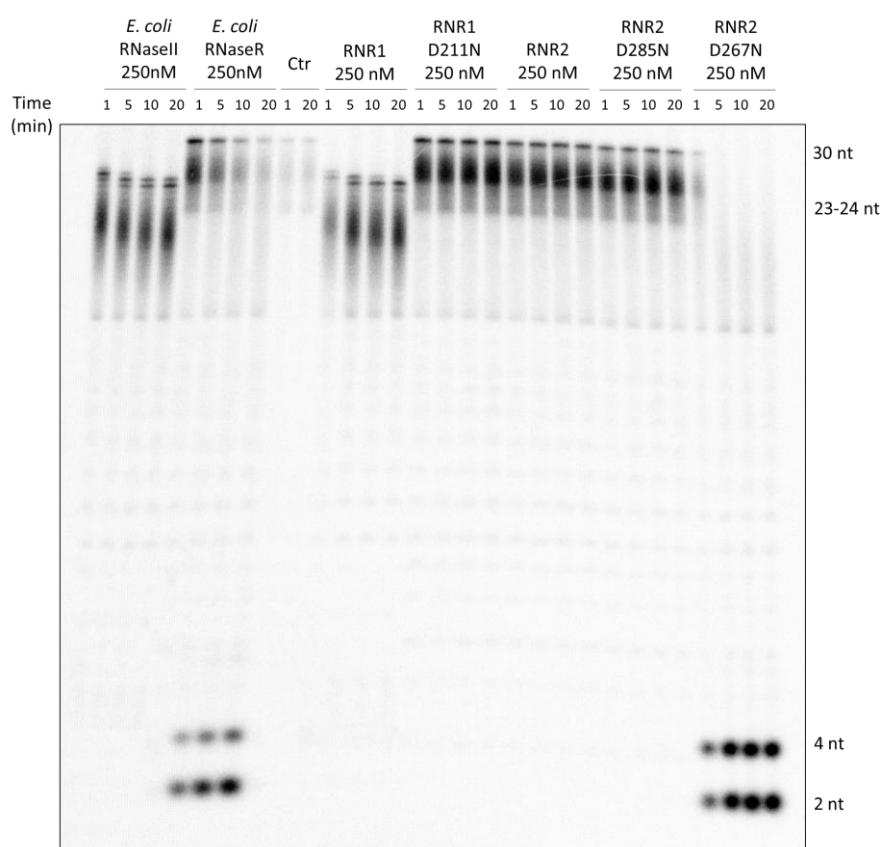


Figure 3.7 Exoribonucleolytic activity of RNR1 and RNR2 wild-type and mutants on a double-stranded RNA. A 30-mer oligoribonucleotide was hybridized to the complementary 16-mer oligodeoxyribonucleotide and the corresponding double-stranded molecule (16–30ds) with a 3' overhang was obtained. Enzymes were incubated at 37°C with 25 nM 16–30ds substrate. The mutants and the protein concentrations used in each reaction are specified on top. Samples were taken at indicated time points between 1 and 20 minutes and reaction products were analysed in a denaturing 20 % PAA gel with 7 M urea. *E. coli* RNase II and *E. coli* RNase R were used for comparison. Control reaction mixtures without enzyme (Ctr) were incubated at the maximum reaction time for each protein. Numbers to the right indicate sizes of the major reaction products.

It is known that RNase R-like enzymes have the remarkable feature of degrading double-stranded RNA without requiring to be associated with a helicase to overcome such structures⁹³. *E. coli* RNase R was used as a control in the assay showed in Figure 3.7 and, as expected, it can degrade this substrate releasing a 2 nt fragment as end-product. This degradation pattern is not observed in RNR2, contrary to what we were expecting. However, the mutant RNR2_D267N is able to completely degrade the double-stranded RNA only 5 minutes after the beginning of the reaction, which can actually indicate that, although not detectable, possibly the wild-type protein also possesses this ability. Protein instability, and the fact that it was one of the first proteins to be purified, may account for an alteration in the protein's behaviour. For further experiments, a higher protein concentration and a protein batch with less storage time needs

to be tested, especially considering that after a certain period of time the exoribonucleolytic activity can decrease. Furthermore, the data presented above demonstrates that the D211N and D285N amino acid substitutions in RNR1 and RNR2, respectively, lead to the loss of the exoribonucleolytic activity of these proteins, also against the double stranded RNA. This shows that both residues are crucial for the enzymatic activity itself, regardless of the substrate used in the experiment.

Collectively, the previous results show that we were able to successfully purify and identify two ribonucleases of *M. tuberculosis*, RNR1 and RNR2, which belong to the RNB-family of enzymes and present different substrate specificities. Both enzymes are able to efficiently degrade ssRNA in the 3' to 5' direction in a processive and sequence-independent way. However, while RNR2 is capable of degrading highly structured RNA releasing a 2 nt fragment as end product, RNR1 activity is impaired by the presence of secondary structures and when degrading single-stranded substrates, it releases a 4 nt fragment as a final product. These degradation profiles demonstrate that RNR1 is an RNase II-like protein whereas RNR2 is a RNase R- like enzyme.

3.2.2 RNA binding ability

Our next aim was not only to determine the RNA binding ability of both *M. tuberculosis* ribonucleases, but also to evaluate the contribution of each mutated residue in such ability. For this purpose, electrophoretic mobility shift assays (EMSA) were performed using the poly(A) substrate and increasing protein concentrations as described in Materials and Methods section (Figure 3.8). The binding affinity of the RNA-protein interactions were quantitatively measured by determining the equilibrium dissociation constant (K_D) of each protein (Table 3.1, Figure S.4), that corresponds to the concentration at which 50 % of the RNA substrate is in a complex with the ribonuclease.

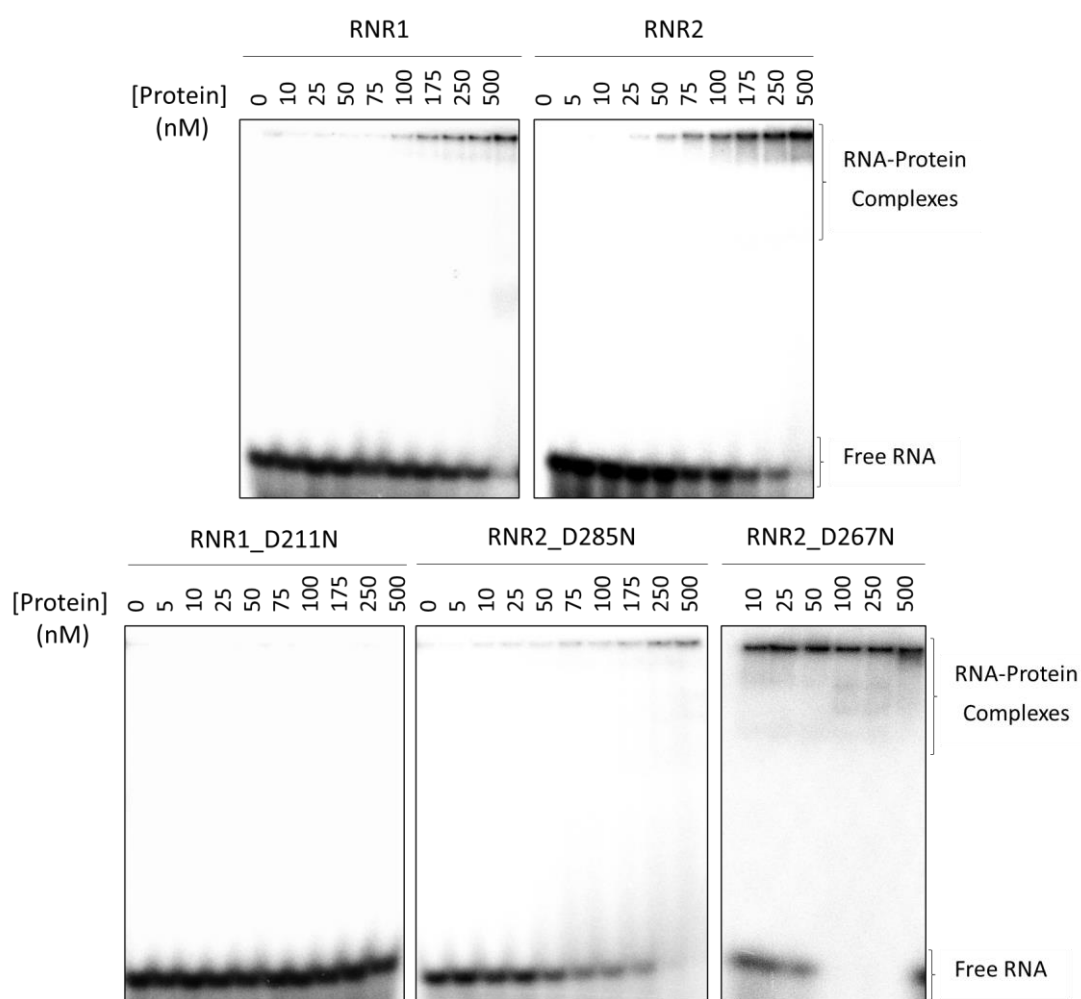


Figure 3.8. Electrophoretic mobility shift assay (EMSA) of wild type RNR1 and RNR2 and mutant proteins with a poly(A) substrate. The 30-mer poly(A) RNA was incubated for 10 minutes at 37°C with increasing amounts of each protein at the concentrations indicated. A control reaction without enzyme added was performed in all experiments. Binding reactions were applied on a 10 % non-denaturing PAA gel. A description of the reaction products is indicated on the right side of the gels.

Table 3.1. Dissociation constant of RNR1 and RNR2 wild type and mutant proteins. The constants were determined according to a specific binding with Hill slope equation. Measurements of the wild-type proteins were performed in duplicate. RNR1_D211N dissociation constant was not determined (ND). We were unable to measure replicates for the RNR2 variants, thus, an approximate value for the constant is shown.

Protein name	Dissociation Constant (K_D)
RNR1	212.0 ± 3.2 nM
RNR2	122.4 ± 15.9 nM
RNR1_D211N	ND
RNR2_D285N	~ 362.3 nM
RNR2_D267N	< 50 nM

The gels of Figure 3.8 demonstrate that all the proteins were able to bind the poly(A) oligoribonucleotide, except the RNR1_D211N mutant. There are two possible explanations for this outcome: firstly, the D211N amino acid substitution may have caused a conformational change in the protein that precluded the binding of the RNA; secondly, knowing that we had problems during the purification steps, the sample might be compromised. This second explanation seems more reliable taking into account the results obtained for RNR2 and the analogous mutant, and what was already extensively studied for the *E. coli* homologous proteins⁸⁵. Nevertheless, to draw any valid conclusion, the protein must be purified again, and the assays have to be repeated. Regarding the results obtained for the remaining variants, it was possible to see that all proteins were able to form RNA-protein complexes. In the lower concentrations tested, only unbound RNA is visible in the bottom of the gels; upon increasing the protein concentration, a shift in migration is detected resulting from the formation of those complexes, which start to appear at different protein concentrations for each ribonuclease.

RNR1 and RNR2-RNA complexes are detectable when the proteins are present, respectively, in concentrations above 100 nM and 25 nM. The K_D values determined were 212.0 ± 3.2 nM for RNR1 and 122.4 ± 15.9 nM for RNR2. These results confirm that both wild-type ribonucleases are able to bind RNA efficiently and that RNR2 has a slightly higher affinity for RNA. A similar result was observed for the RNR2_D285N which was previously shown to be an inactive form of the enzyme. This indicates that, despite the mutation, the protein is still able to bind RNA although with lower affinity for the substrate when compared to the wild-type enzyme. The presence of smears along the gel lanes suggests the formation of less stable RNA-protein complexes that have dissociated during the run. The observation is also supported by the K_D value determined, 362.2 nM, that reflects a 3-fold reduction on the binding ability. However, this result needs to be interpreted cautiously not only because it is based in just one experiment but also for the reason that, if the dissociation of the complexes happened during gel running, the results of K_D determined are being underestimated. Finally, for the RNR2_D267N variant, RNA-protein complexes were detected immediately at 10 nM of protein. The K_D value was not possible to be determined in the conditions tested and lower protein concentrations needed to be considered to perform a proper measurement. Thus, the K_D value is presented as an upper limit of the affinity of the interaction, $K_D < 50$ nM. At that concentration no unbound RNA is visible and only RNA-protein complexes were detected, meaning that the equilibrium constant had already been exceeded. The results indicate that this single mutant has a very high affinity for the poly(A) substrate, suggesting that the aspartic acid residue is somehow preventing the establishment of a tight complex between the protein and the

oligoribonucleotide, diminishing the degradation efficiency. This hypothesis also explains the increased exoribonucleotidic activity detected in the *in vitro* assays previously shown.

3.2.3 Helicase Activity

Helicases are motor proteins that unwind double-stranded nucleic acids by utilizing free energy from ATP hydrolysis¹¹⁴. RNase R, although it shares the family’s typical RNB catalytic domain, comprises distinctive features from RNase II. The most remarkable is its ability to degrade structured RNA substrates by making use of an intrinsic helicase activity⁷¹. For that, this enzyme only requires a 3’ single-stranded overhang at least 5 nt long to serve as a binding site from which degradation can be initiated^{41,96}. In this part of the work, considering that RNR2 behaves like an RNase R protein, the main objective was to test the intrinsic unwinding activity of this enzyme, carrying out *in vitro* helicase assays using both D267N and D285N mutants. With this in mind, the enzymes were incubated with the 16-30ds RNA in an ATP-containing condition and the reactions were followed over time. The chelating agent EDTA was added to the experiment with the D267N mutant to ensure that the exoribonuclease activity of this protein was blocked.

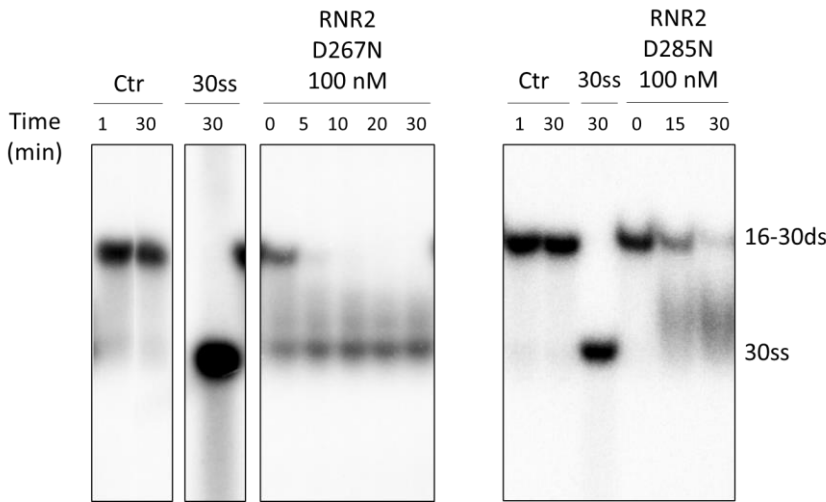


Figure 3.9. Helicase activity of RNR2 mutant proteins in the presence of a double-stranded substrate. A 30-mer oligoribonucleotide was hybridized to the complementary 16-mer oligodeoxyribonucleotide and the corresponding double-stranded molecule (16–30ds) with a 3’ overhang was obtained. RNR2_D267N and RNR2_D285N helicase activity was assayed in the presence of ATP and with 25 nM 16-30ds substrate at 37°C. EDTA was added to the reaction mixture to block the exoribonuclease activity. Protein concentrations used in each reaction are specified on top. Samples were taken at indicated time points between 1 and 30 minutes and reaction products were analyzed in a 20 % PAA gel. Control reaction mixtures without enzyme (Ctr) and a control with heat-denatured substrate showing ssRNA (30-mer) were incubated at the maximum reaction time for each protein. A description of the reaction products is indicated on the right side of the gels.

The results presented in Figure 3.9 clearly show the intrinsic RNA helicase activity of both RNR2 mutants. It is possible to observe that, at the beginning of the reaction (0 minutes) the substrate was correctly hybridized, with a migration pattern similar to the 16-30ds control. As the reaction proceeds, the dsRNA was unwound by both enzymes as visible by the emergence of bands migrating the same as the 30ss control. However, the D285N mutant unwinds the dsRNA substrate to a significantly lesser extent than the D267N mutant protein. The RNR2_D267N surprisingly showed a very efficient and rapid unwinding activity with this substrate, possible related with the increased activity of this mutant against the dsRNA, that needs to be explored.

Moreover, since the exoribonucleolytic activity was precluded by the addition of a chelating agent, we can conclude that RNR2 itself exhibits an intrinsic helicase activity *in vitro*, which is independent of its RNase activity. Previous studies on *E. coli* RNase R have shown that D716 and E717, located in the S1 domain, are essential for the helicase activity of the protein and are specifically involved against the structured RNA substrate⁹³. Analysing the multiple sequence alignment, it was possible to confirm that the corresponding amino acid residues also exist in the RNR2 protein sequence, D634 and E635, contrary to what is visible in RNR1 and *E. coli* RNase II. In sum, these results shed light on a unique property of this RNase R-like enzyme, making it distinct from other 3' to 5' processing exoribonucleases in *M. tuberculosis*, such as PNPase, RNase H or RNase PH. Although the precise cellular function of this ribonuclease in the pathogen has yet to be understood, RNR2 can have an important role in the adaptation to environmental changes and during transmission between hosts.

3.3 Crystallization and Thermostability Assays

In the last part of the thesis, several stability assays and crystallization trials were performed with the ultimate goal of determining the three-dimensional structure of both enzymes, which is absolutely relevant to identify unique features of the *M. tuberculosis* ribonucleases under study. The most critical issue was the availability of high-quality samples. For crystallography applications, proteins must be purified in a quantity and quality suitable to perform several stability assays and to test numerous crystallization conditions. Ideally, the procedures should be reproducible, and the protein samples have to be soluble, in stable conditions, folded properly and free of critical contaminants¹¹⁵. In this work, the main problem we had to face was the precipitation of the proteins along the purification protocol, most of the times leading to insufficient final purification yields to carry out any further structural studies. After several attempts, we have managed to obtain good quality samples of both RNR1, RNR2 and

RNR2_D267N (due to time constraints we were not able to upscale the purification for RNR1_D211N and RNR2_D285N) to proceed with crystallization trials (Figure 3.10). Following the large-scale purification, a SDS-Page gel was run with 1 and 4 μ g of the purified proteins. The lanes containing 1 μ g of protein were useful to observe if the enzymes were migrating according to correct size without creating any smear or if there was no splitting into closely positioned bands in the gel. On the other hand, with the 4 μ g experiment, the aim was to detect any existent contaminant in the sample.

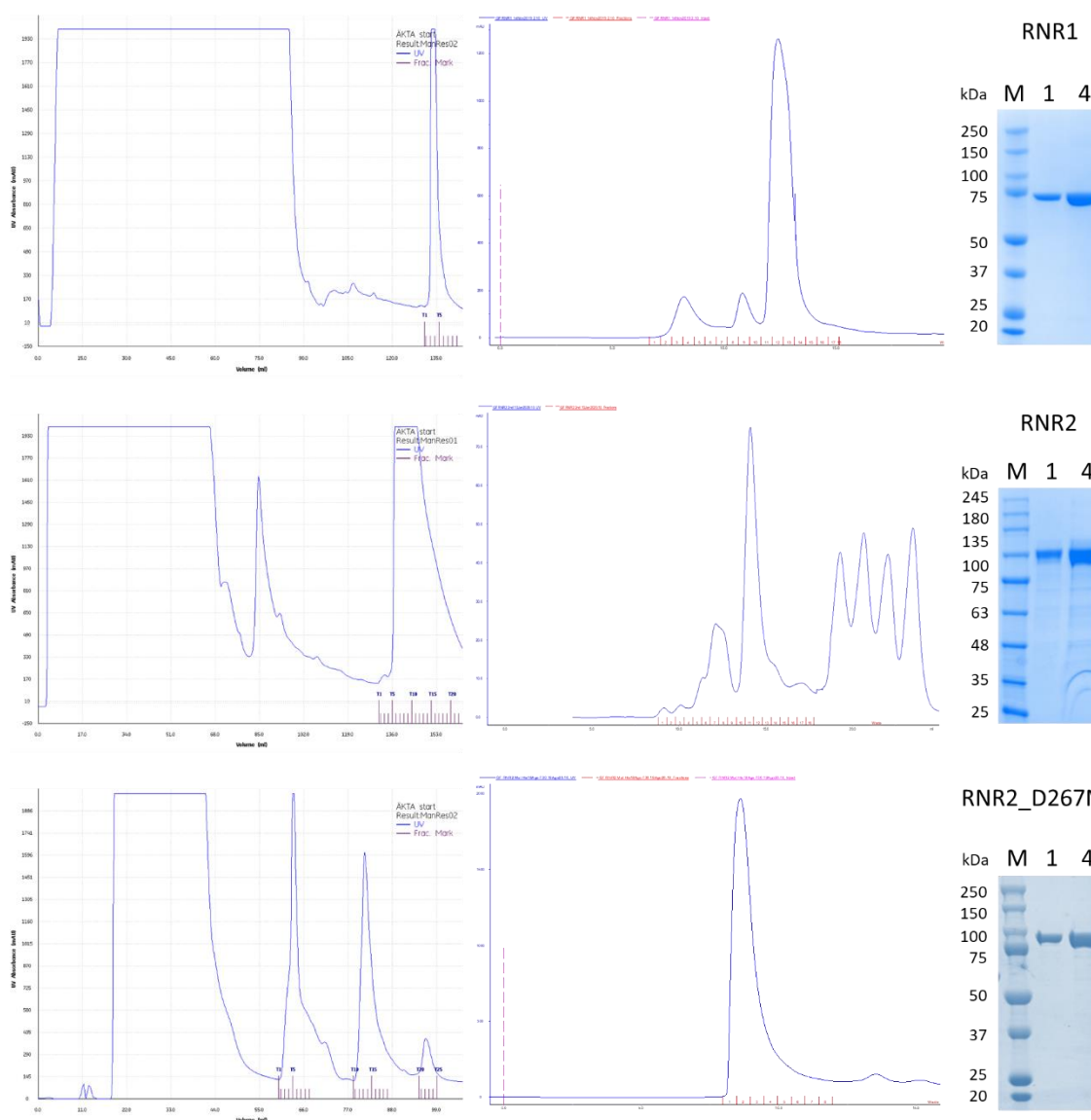


Figure 3.10. Large-scale protein purification of RNR1, RNR2 and RNR2_D267N by affinity chromatography followed by size-exclusion chromatography. **Left panel-** Elution profile using Ni Sepharose High Performance columns on an ÄKTA Start system. A gradient up to 500 mM imidazole was used. **Middle panel-** Chromatogram of each protein using a Superdex 200 Increase 10/300 GL on an ÄKTA Purifier UPC 10 GE system. Sample injection is depicted in pink, UV absorbance in blue and the collected fractions in red. **Right panel-** 4-12 % SDS-PAGE analysis of samples containing 1.0 and 4.0 μ g of protein obtained after both purification steps. Each well was loaded with the same sample volume (10 μ L) and the marker used (lane “M”) was the NZYColour Protein Marker II (Nzytech).

3.3.1 Crystallization Trials

When starting a crystallization experiment, screening conditions should cover a large part of the chemical space to enhance crystal formation. Thus, for the crystallization project of both ribonucleases, two commercially available screenings were initially carried out with the purified proteins RNR1 and RNR2: JCSG+, a sparse matrix protein crystallization screen designed to maximize the coverage of the crystallization parameter space (buffers, salts, precipitants and pH); and BCS, a grid screen formulated to include various PEG mixtures (called “smears”).

The plates were all visualized frequently. Generally, most of the conditions tested for RNR1 resulted in the formation of a heavy precipitate in the drops, which indicates that the biochemical conditions, with protein concentrations of 10 and 15 mg.mL⁻¹, have driven the protein sample into a state of supersaturation many times beyond the level required for growth of single crystals¹¹⁶. Crystalline needles were observed in a condition of the BCS Screen containing 28 % (v/v) PEG Smear Broad, 0.1 M Sodium Phosphate pH 6.2 and 0.2 M Sodium Chloride. Considering the high salt concentration present in the drop, those needles were likely to be salt crystals rather than the desired protein crystals. However, since differentiating between both is virtually impossible using solely standard light microscopy methods, a SpectroLight 600 (*Xtal Concepts*) instrument equipped with a UV-light source was used. It is known that proteins can show epifluorescence at 340 nm due to tryptophan residues. Thus, when in the presence of protein crystals, the image consists of a clearly blue tryptophan fluorescence in a the black background. On the contrary, when illuminated with UV-light, salt crystals become invisible and completely merged with the background, making it possible to distinguish between them¹¹⁷. This test was applied to the RNR1 drop previously mentioned (data not shown) and the visible crystalline material were, in fact, sodium chloride.

For RNR2 trials there was less protein available, hence the two same crystallization screens were tested but only with one protein concentration (10 mg.mL⁻¹). The preference was for testing more conditions rather than different concentrations of protein as they provide more relevant information regarding the promising crystallization conditions. In this case, two days after the assay was set up, drops were mostly clear or showing light precipitates. Clear drops in isolation provide limited information since they can either indicate that the sample is completely in solution or in a metastable supersaturated state. The first thermodynamic state indicates that the protein is far from crystallizing, but the latter must be considered a central point for optimization¹¹⁶. However, since this was the predominant outcome in both screenings, it was probably an undersaturated condition. Interestingly, one condition of the JCSG+ screen stood out, only composed by 20 % (v/v) ethanol and 0.1 M Tris-HCl pH 8.5. The drop appearance

consisted of a small amorphous precipitate in the center with dispersed transparent clusters in the remaining area that could possibly be protein crystals. To confirm, the intrinsic fluorescence imaging was tested similarly to the previously described for RNR1 and the results (Figure 3.11) showed an overall blue illumination of the drop, meaning that none of the visible structures correspond to protein crystals.

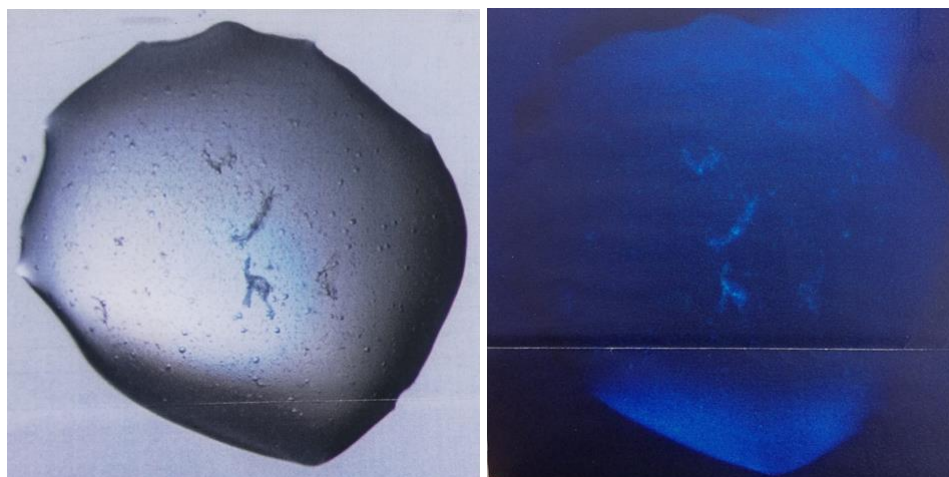


Figure 3.11. Distinction between protein or salt crystals in a RNR2 crystallization hit. Crystallization condition C8 of the JCBS+ screen (20 % (v/v) ethanol and 0.1 M Tris pH 8.5). Drop volume of 200 nL with a drop ratio of protein and precipitant solution of 1:1. The protein concentration used was 10 mg/ml in the buffer C_RNR. **Left-** Crystallization drop illuminated with a white light. **Right-** Crystallization drop illuminated with a UV light (blue) in a SpectroLight 600 (*Xtal concepts*).

Together, the results obtained for the two proteins were without success since no crystals were observed in any condition. Ribonucleases from the RNB-family are very dynamic proteins and their domains, especially the terminal ones are very mobile to allow the entrance of the substrate, which can difficult the crystallization process. Although structural information is known for *E. coli* RNase II, regarding RNase R, only a truncated structure is available. The lack of a complete structural model for a RNase R-like protein is limiting the understanding of the structural arrangement of the different domains and their role in the RNA degradation mechanism. Ideally, the best targets to try to crystallize are the two inactive mutants RNR1_D211N and RNR2_D285N since there is some possibility of forming protein crystals with the RNA substrate bound, which will provide more detailed information regarding the degradation mechanism.

In both cases, the amount of protein available was the limiting factor, precluding the execution of more crystallization trials. Therefore, thermostability studies were performed, in order to find a different buffer formulation where the protein would be more stable in solution, preventing the formation of precipitate during protein purification and, simultaneously increasing the likelihood of obtaining protein crystals.

3.3.2 Thermostability Assay

The thermal stability of a protein can be measured by heating the sample while monitoring its structural integrity upon fluorescence emission. Plotting the increase in fluorescence signal with increasing temperature generates a melting curve, where the inflection point is known as the melting temperature (T_m)¹⁰⁶. This measure reflects the stability of the protein in that particular experimental condition and can be used to identify previously unknown intramolecular and intermolecular interactions involving the protein of interest. As such, the expression and purification of a protein can be significantly improved by adding stabilizing buffers or ligands that reduce the propensity of proteins to unfold during purification and storage. A thermostability assay analysis also provides useful indicators for X-ray crystallography, since a higher T_m value generally indicates higher propensity for proteins to crystallize.^{106,118}

In order to determine the ideal purification and storage conditions for both RNR1 and RNR2, two different complementary techniques were used. First, a thermal shift assay (TSA) was performed. This assay offers a quick, easy and temperature-based analysis to assess the changes in the protein melting temperature under different conditions. The process involved incubation of the natively folded protein with SYPRO Orange, an environmentally sensitive fluorescent dye, and the monitoring of the fluorescence increase upon thermal protein unfolding, caused by interaction of the dye with the exposed hydrophobic core of the denatured protein^{119,120}. Second, a high throughput label-free NanoDSF assay was set up to confirm the results obtained on the TSA. In this case, the intrinsic protein fluorescence generated from the aromatic sidechains of tyrosine and tryptophan residues was examined. Upon unfolding, the environment of those residues alters, and become more exposed to the solvent, causing changes in the fluorescence intensity. In contrast to TSA, nanoDSF measurements are performed with unmodified proteins close to their native state, allowing high-precision and high-resolution results¹²⁰.

Protein sample homogeneity is potentially the most important factor for successful crystallization. By screening sparse matrix conditions, encompassing different buffer systems that cover a wide range of pH, additives, and salt concentrations, optimal buffer components can be identified for each individual protein. Ideally, since protein unfolding is a cooperative process, the unfolding of a small protein region could induce the immediate unfolding of the remaining protein core and the thermostability data should reflect higher transition slopes, in parallel with a higher T_m ¹²¹. Thus, the data derived from both thermostability methods was analyzed taking into consideration that the ideal curve profile presents a low baseline

fluorescence, a clean and unique sigmoidal transition during denaturation and a considerable shift in the melting temperature relatively to the initial purification buffer.

The methods described were performed for both RNR1 and RNR2_D267N. The mutant protein was used instead of the wild type because, along the experimental work, none of the purified samples obtained for the wild type RNR2 possessed the purity and concentration requirements needed for these techniques. This mutant proved to be more stable in solution, which favored its large-scale production (Figure 3.10).

A preliminary TSA screen was implemented in order to choose the best signal-to-noise ratio for both enzymes, where different dye concentrations were combined with different amounts of protein. The different RNR1 and RNR2_D267N denaturation curves were analysed, the protein melting temperature for each assay was calculated and a condition was chosen to proceed with the following assays (Figure 3.12).

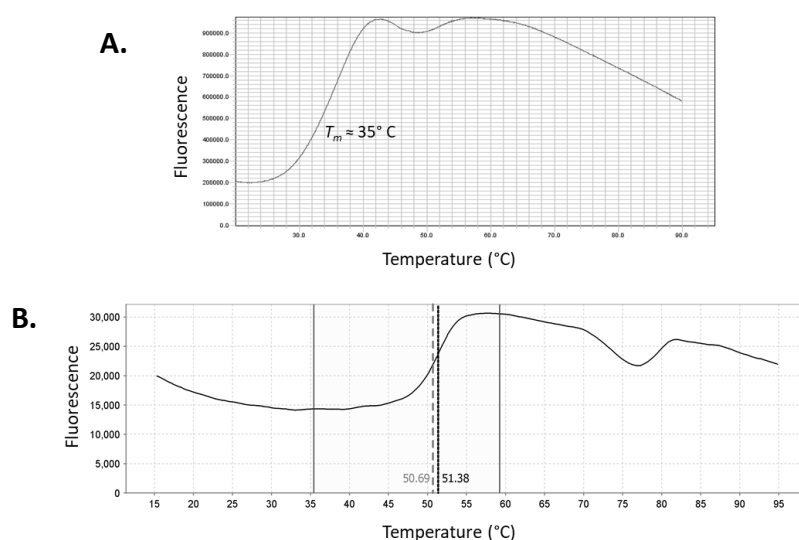


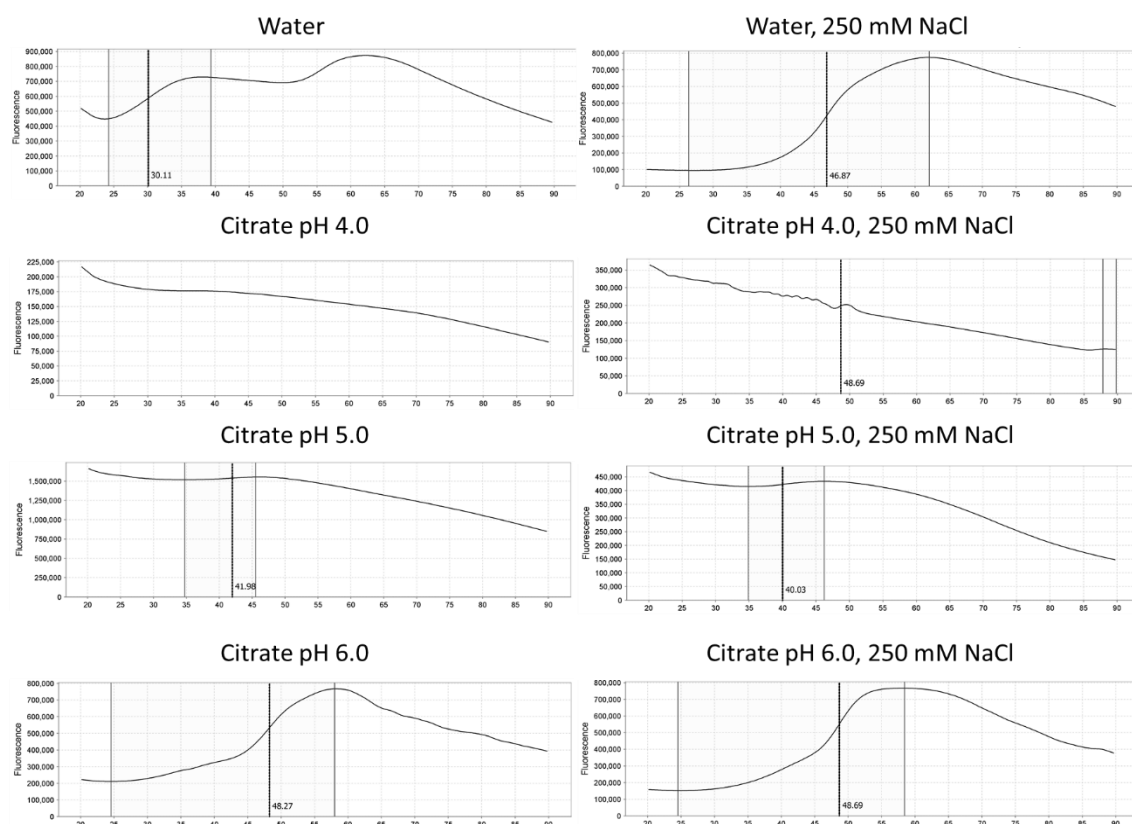
Figure 3.12. Thermal Shift Assay Optimization Curves. **A-** The unfolding of RNR1 was monitored using 5x SYPRO Orange (*Sigma Aldrich*) and 5 μg of protein. Data was collected at 1°C intervals from 20°C through 95°C . **B-** The unfolding of RNR2_D267N was monitored using 1x SYPRO Orange (*Sigma Aldrich*) 3 μg of protein. Data was collected at 1°C intervals from 15°C through 95°C . Melting profiles and respective melting temperatures are presented.

3.3.2.1 RNR1 Thermostability

The optimization screen for RNR1 showed that this protein, when diluted into 10 mM Tris-HCl pH 8.0 and 50 mM KCl (Buffer C_RNB), displays a T_m of only 35°C and a melting curve with two clearly distinct transitions (Figure 3.12). This could be an indicator that some protein domains are unfolding first than the others, possibly the N-terminal and C-terminal ones. The aim was to identify buffers presenting only a unique melting transition, meaning that the protein is structurally homogenous in that buffer, enhancing the chances of crystallization.

For this protein, a two-step approach was applied using a pair of RUBIC screens. First, RUBIC Buffer screen was used to search for global parameters, and it has been designed in such a way that it is possible to discern global stability trends according to pH, salt concentration, buffer type and concentration (Table S.4, Table S.6). Second, protein-specific additives were explored using the RUBIC Additive screen containing small molecules that can affect the folding, aggregation state and stability of the protein (Table S.5, Table S.6).

RUBIC Buffer Screen



RUBIC Additive Screen

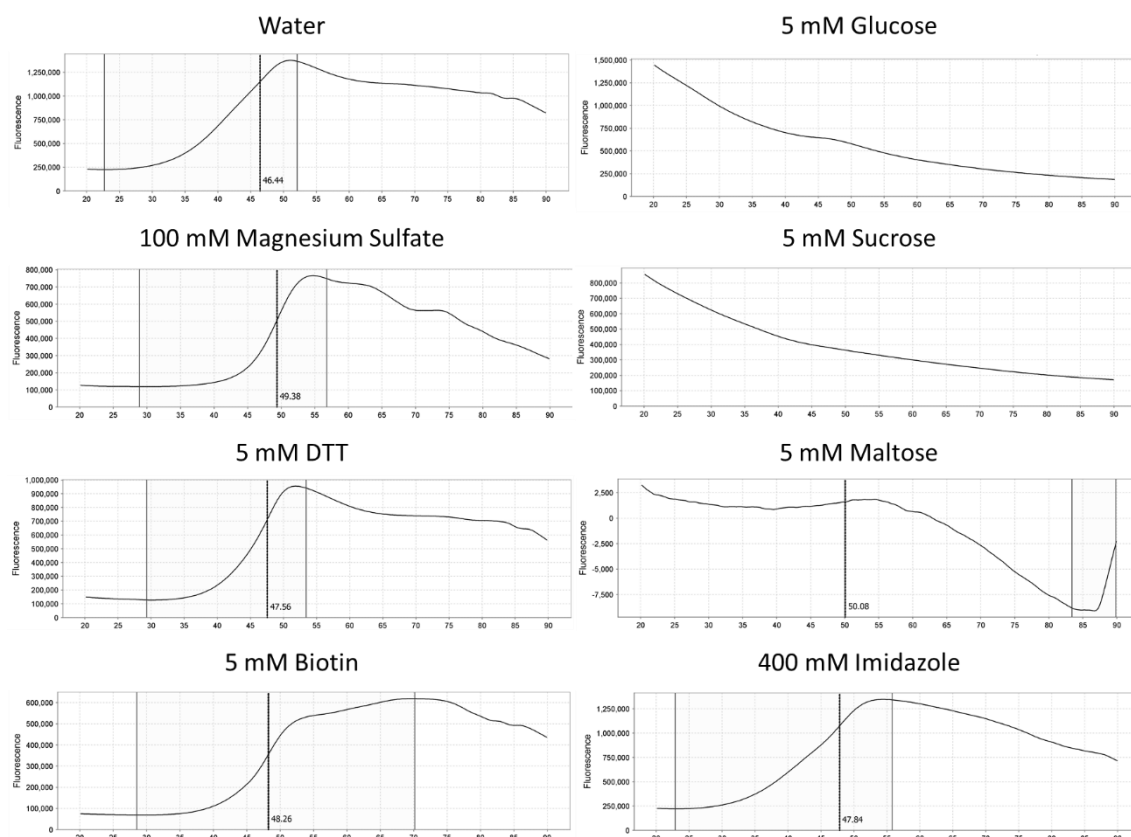


Figure 3.13 Effect of buffer pH, ionic strength and additives on RNR1 thermal stability. RUBIC Buffer and Additive screens (*Molecular Dimensions*) were used to determine RNR1 ideal purification and storage conditions performing a Protein Thermal Shift assay. The unfolding of RNR1 was monitored using SYPRO Orange (*Sigma Aldrich*). The protein amount in all the assays was fixed to 5 μ g. Data was collected at 1°C intervals from 20°C through 95°C. Melting profiles and respective melting temperatures for some selected conditions (indicated above each graph) are presented.

Results present in Figure 3.13 demonstrate that there is a general trend of increasing stability with increasing pH values. Conditions with pH's below 6.0 showed a strong destabilizing effect, with melting curves demonstrating a high initial fluorescence and almost zero transition, indicating that, at the beginning of the assay the protein is probably already unfolded. With pH 6.0, the T_m value increases by around 10°C. It was also possible to study the effect of ionic strength in RNR1 thermal stability. In the presence of solely 250 mM NaCl (i.e., no buffer, just water), it is visible only not an increase in the melting temperature of the protein, but also an improvement in the shape of the curve, presenting only one sharp transition, which indicates that this salt also possesses a determinant effect in protein stabilization. Ultimately, at a certain point, the pH and the ionic strength guarantee a good transition and the other factors do not have much influence. However, as a rule of thumb, T_m should be above 50°C and preferably

above 60°C for crystallization trials, and all the values obtained in this assay were still below what is considered to be acceptable, although some conditions significantly increased the T_m by over 10°C, and even changed the shape of the curve (from double to single transitions).

A detailed analysis of the additive screen (Figure 3.13) revealed that magnesium sulfate, DTT and also biotin may have some influence in stabilizing the protein. Although none caused a significant thermal shift, the transitions were clean and presenting only one melting transition. Thus, testing these additives combined with the conditions identified in the previous assay, a cumulative effect can be discovered. In addition, during the purification of the proteins under study, numerous problems with protein denaturation, precipitation or aggregation were faced. This procedure involved an affinity chromatography step where a high concentration of imidazole was used to elute the protein from the column, which could be interfering with both RNR1 and RNR2 folded state. In this additive assay, we have proved that this compound does not affect RNR1 stability and thus, it is not necessary to modify this purification step. Interestingly, all the four carbohydrates tested resulted in a complete protein denaturation.

Afterwards, several combinations with the best results obtained from thermal shift assay were tested in a NanoDSF assay (Figure 3.14, Table S.7). The most outstanding result was obtained when RNR1 was exposed in a Bis-Tris pH 6.5 buffer conditions without any NaCl concentration. In these conditions, the T_m value reduced, and denaturation occurred over a long temperature range generating gradual melting curves, normally visualized in less stable proteins. It was also possible to understand that, regardless of the buffer, biotin induced a decrease by almost 5°C of the T_m in all the conditions tested. The remaining conditions showed similar results among them, with a maximum observed shift of nearly 5°C in the melting temperature. Lastly, neither DTT nor MgSO₄ have generated considerable improvements in the transitions observed but also did not cause any destabilization. Those two additives were also tested separately (data not shown) and the results were similar to those observed when used simultaneously (Figure 3.14), meaning that no cumulative effect occurred. In future experiments, DTT and biotin will be used as additives in the crystallization assays, whereas MgSO₄ will be kept in the purification buffer of this protein, BisTris pH 6.5 or Citrate 6.0.

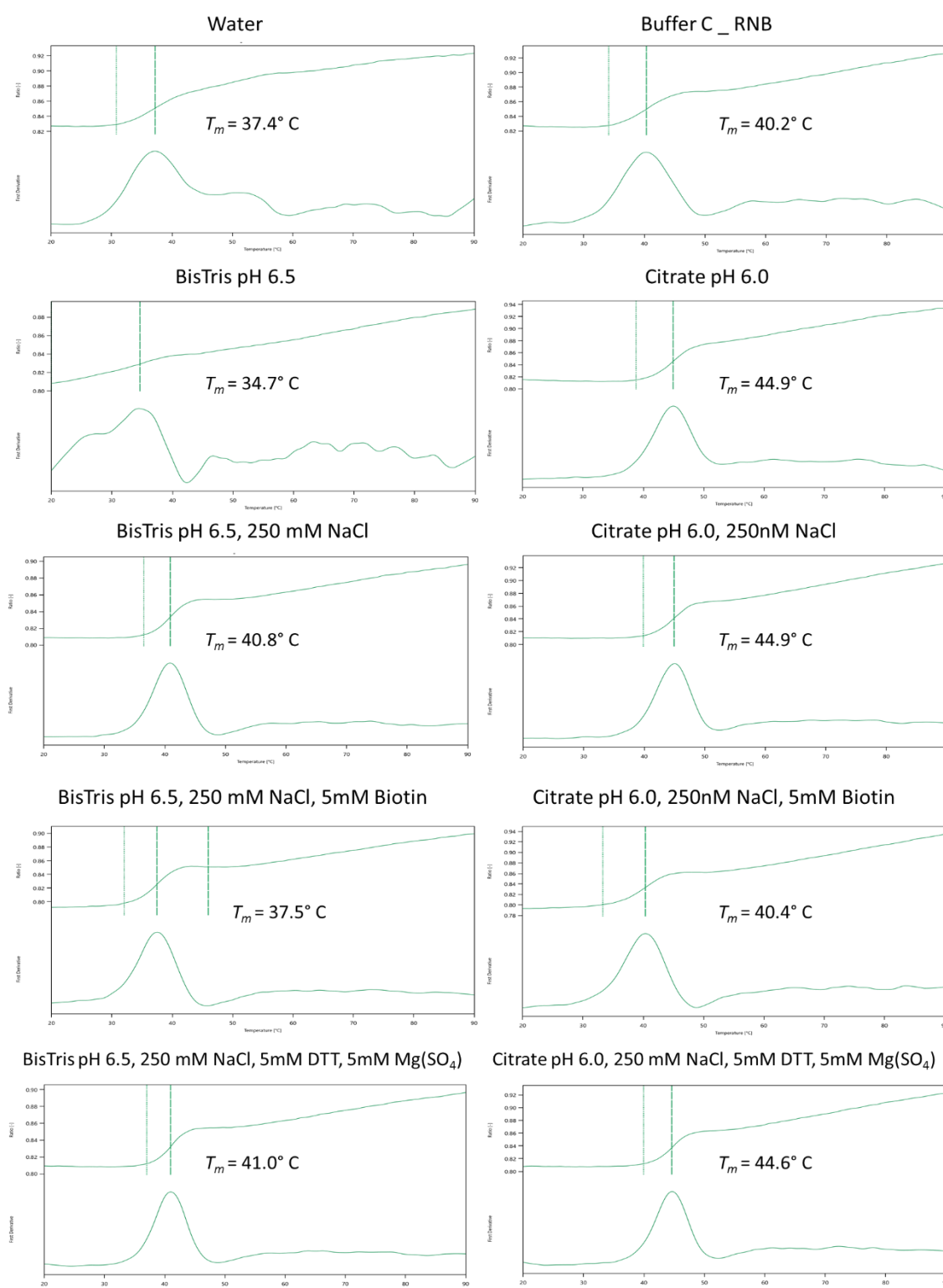


Figure 3.14 NanoDSF analysis of RNR1 samples. Thermal unfolding of RNR1 was assessed by monitoring changes of tryptophan fluorescence at 330 nm and 350 nm wavelengths. The protein concentration in all the assays was fixed to 5 μ g. Data was collected at 1°C intervals from 20°C through 95°C. Each graph shows the Boltzmann (top) and derivative (bottom) melting profiles, as well as the respective melting temperatures for some selected conditions (indicated above each graph).

3.3.2.2 RNR2 Thermostability

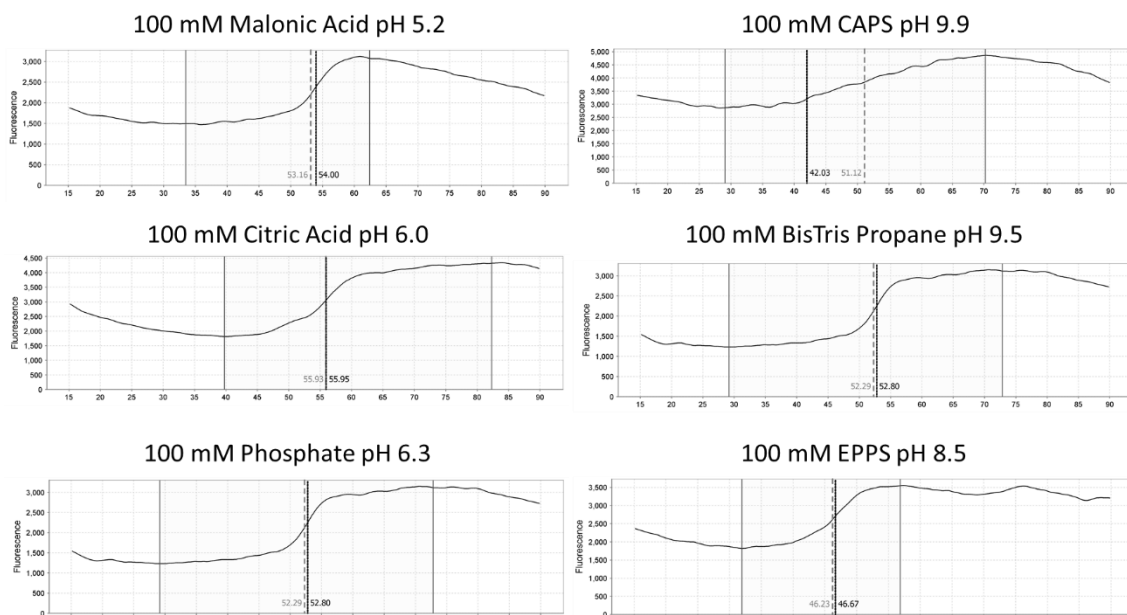
Purification of the mutant protein D267N was achieved taking into account the TSA and NanoDSF results obtained for RNR1, which immediately was reflected in the quality of the curve obtained in the preliminary test (Figure 3.12). A single melting transition was visible with an increase in the fluorescence signal upon denaturation, reaching a maximal value when the protein is completely unfolded. Then, the signal decreases corresponding to the dye-protein complexes dissociation.

The best condition obtained in this first screen was used to prepare two additional screenings. Moreover, the results obtained previously for RNR1 have demonstrated that the most influencing factors in the stability of these ribonucleases were clearly the buffer pH value and salt concentration. Thus, instead of the RUBIC buffer and additive screens, for RNR2_D267N two more specific screenings were tested: the Durham pH Screen where the pH values (4 to 11) of different buffers (total of 28) overlap, making it easier to deconvolute the effects of pH and type of buffer on protein stability (Table S.8, Table S.10); and the Durham Salt Screen which contains commonly laboratory salts, chelating and reducing agents, giving a general indication on the preference of a protein sample for the environments with high/low ionic strengths (Table S.6, Table S.9).

Contrary to the preliminary experiments (Figure 3.12), that revealed a melting curve with a sharp transition between 20000 and 50000 fluorescence intensity, in the Durham screens all the transitions visualized were at least one order of magnitude smaller which consequently have increased the noise visible in each curve making the results difficult to interpret. Thus, instead of choosing only some of the best conditions, a matrix was created with the most promising results to be directly validated in the nanoDSF assay.

Malonic acid pH 5.2, citric acid pH 6.0 and sodium phosphate 6.8 were the most stabilizing buffers, whereas CAPPS pH 9.9, Bis-Tris propane pH 9.5 and EPPS pH 8.5 significantly destabilized the protein, probably due to the basic pH and not the intrinsic property of the buffer (Figure 3.15). Moreover, despite the noisy transitions, a T_m shift was visible in some conditions of the salt screen. Na₂ malonate and Na₃ fluoride shifted the melting temperature by more than 15°C and 10°C, respectively, comparing with the control. The protein also proved to be stable when in presence of NaCl, NaF, Mg(SO₄) and Na(SO₄). Additionally, RNR2_D267N was notably more stable the higher the salt concentration in solution. However, for successful crystallization trials it is important to keep the ionic strength as low as possible to prevent the protein from becoming too soluble which can easily inhibit crystallization¹²².

Durham pH Screen



Durham Salt Screen

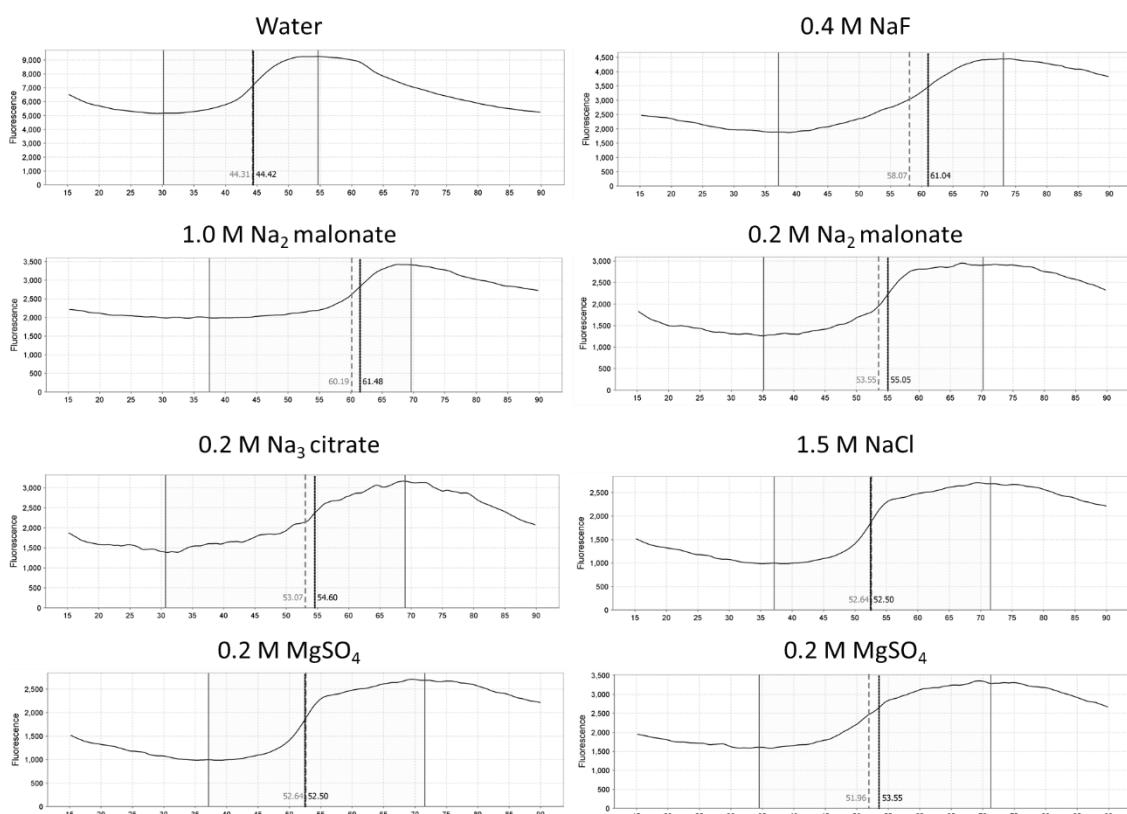


Figure 3.15 Effect of buffer pH, ionic strength and additives on RNR2_D267N thermal stability. Durham pH and salt screens (*Molecular Dimensions*) were used to determine RNR2 ideal purification and storage conditions performing a Protein Thermal Shift assay. The unfolding of RNR2_D267N was monitored using SYPRO Orange (*Sigma Aldrich*). The protein amount in all the assays was fixed to 3µg. Data was collected at 1°C intervals from 20°C through 95°C. Melting profiles and the respective melting temperatures for some selected conditions (indicated above each graph) are presented.

NanoDSF offers a great number of advantages over traditional fluorimetry approaches. Most importantly, for this particular protein it enables the adjustment of the excitation power, from 1 to 100 %, until the fluorescence intensity is within the limits of the detector. Ideally, the percentage should be the lowest possible value to obtain a detection range between 10,000 and 15,000 counts. A first trial with RNR2_D267N was performed containing 3 µg of protein separately diluted in water and in the purification buffer (Buffer C_RNR) and the raw fluorescence values obtained with 100 % of excitation power were below 1000 counts in both conditions. A few hours later the same test was completed using high sensitivity capillaries and the maximum value of counts obtained was 650 in a sample where the protein amount was duplicated. A reasonable explanation for these observations is the gradual denaturation of the protein between procedures. Even so, the test was run and upon denaturation a blueshift of the fluorescence was observed, meaning that the aromatic residues were already exposed on the surface of the protein, which can indicate protein denaturation. In the future, this assay has to be repeated with a freshly purified protein and eventually with a lower starting temperature. The results will be carefully analyzed to find the best storing conditions for this ribonuclease and to set up further crystallization trials.

4. Conclusions and Future Perspectives

RNA availability and stability has long been emerged as a major player controlling the expression of virulence factors that allow pathogenic bacteria to infect their hosts. A full comprehension of the molecular details of RNA metabolism in *M. tuberculosis* is important to have a more complete understanding about this organism pathobiology, physiology and regulation of the response mechanisms, which will be essential to develop new strategies to control this wide-spread pathogen.

The work carried out in this thesis is focused on the functional and structural characterization of RNB-family members in *M. tuberculosis*, RNR1 and RNR2. In an initial analysis, two orthologues of this family were found in the genomic sequences of the pathogen. Wild type and point mutants of the proteins were overexpressed, purified and different biochemical and structural assays were performed to properly characterize both targets.

Altogether, our results show that the mycobacterial proteins are, in fact, 3' to 5' hydrolytic exoribonucleases belonging to the RNB-family, in agreement with the predictions made by sequence comparison. RNR1 proved to be an RNase II-like protein, while RNR2 showed to behave like RNase R. Both are able to efficiently bind RNA, forming stable RNA-protein complexes. It was also possible to identify key amino acid residues essential for the exoribonucleolytic activity of the enzymes (D211 and D285) but which apparently do not interfere with their ability to bind to RNA. Interestingly, RNR2 revealed a very efficient and fast unwinding property, comprising both helicase and exoribonuclease activities. Moreover, after several crystallization trials, buffer optimization was performed by TSA and nanoDSF. We have managed to increase the melting temperature of the proteins and found conditions where it presents a sharp transition curve, a sign of improved protein thermostability. Additionally, the mutant RNR2_D267N proved to be more active, to have some interesting unwinding properties and showed to be more stable in solution. The second property can be relevant to enhance the crystallization process taking us one step further to determine the first complete structure of an RNase R-like protein.

In a near future, to better characterize the degradation properties of the two ribonucleases, it would be interesting to test different substrates and to understand if catalysis can occur in the presence of other divalent ions. The influence of D267 residue in the helicase activity of RNR2 can also be explored. Based on the results obtained in the thermostability tests, additional crystallization screenings can be carried out to attempt to obtain good quality protein crystals

and determine the desired three-dimensional structures. Ultimately, it would be important to perform some *in vivo* tests in order to address the role of each protein in *M. tuberculosis* biology.

In conclusion, the study of other bacteria rather than the model organisms *E. coli* and *B. subtilis* is increasingly more relevant to understand the presumably important role of RNA metabolism in pathogenicity. The detailed knowledge of the crystal structure and enzymatic properties of RNR1 and RNR2 is fundamental to aid in rational drug design and provide a strong practical advantage to develop novel therapeutic strategies. These enzymes represent potential targets for developing new antibiotics, which are particularly needed given the increasing prevalence of MDR-TB and XDR-TB worldwide.

The knowledge obtained during the realization of this master thesis and the acquired skills on biochemical and biophysical characterization of ribonucleases, allowed me to participate in a project aiming the study of these proteins as potential targets to develop effective drugs to treat SARS-CoV-2 infection (Appendix B).

5. Bibliography

1. Chiu, L. *et al.* Protective microbiota: From localized to long-reaching co-immunity. *Frontiers in Immunology* 8, (2017).
2. Institute of Medicine. What You Need to Know About Infectious Disease. *Washington DC: The National Academic Press* (2011).
3. Alberts B, Johnson A & Lewis J. *Molecular Biology of the Cell*. (2002).
4. Engelkirk, P. G. & Duben-Engelkirk, J. *Laboratory Diagnosis of Infectious Diseases: Essentials of Diagnostic Microbiology*.
5. World Health Organization. WHO - The top 10 causes of death. 24 Maggio 1–7 (2018). Available at: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>. (Accessed: 1st May 2020)
6. Holmes K.K., Bertozzi S, Bloom B. R. Major Infectious Diseases. 3rd edition. *Washington DC: The International Bank for Reconstruction and Development / The World Bank* (2017).
7. Frothingham, R., Hills, H. G. & Wilson, K. H. Extensive DNA sequence conservation throughout the *Mycobacterium tuberculosis* complex. *J. Clin. Microbiol.* 32, 1639–1643 (1994).
8. Sinha, P. *et al.* Differentiation of *Mycobacterium tuberculosis* complex from non-tubercular mycobacteria by nested multiplex PCR targeting IS6110, MTP40 and 32kD alpha antigen encoding gene fragments. *BMC Infect. Dis.* 16, 123 (2016).
9. Cambier, C. J., Falkow, S. & Ramakrishnan, L. Host evasion and exploitation schemes of *Mycobacterium tuberculosis*. *Cell* 159, 1497–1509 (2014).
10. World Health Organization. WHO | Global tuberculosis report 2019. *WHO* (2019).
11. Knechel, N. A. Tuberculosis: Pathophysiology, clinical features, and diagnosis. *Crit. Care Nurse* 29, 34–43 (2009).
12. Raffetseder, J. Interplay of human macrophages and *Mycobacterium tuberculosis* phenotypes. *Linköping University Electronic Press* (2016).
13. Dulberger, C. L., Rubin, E. J. & Boutte, C. C. The mycobacterial cell envelope — a moving target. *Nature Reviews Microbiology* 18, 47–59 (2020).
14. Guinn, K. M. & Rubin, E. J. Tuberculosis: Just the FAQs. *MBio* 8, (2017).
15. Machado, D., Girardini, M., Viveiros, M. & Pieroni, M. Challenging the drug-likeness dogma for new drug discovery in Tuberculosis. *Frontiers in Microbiology* 9, (2018).
16. Smith, I. *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. *Clinical Microbiology Reviews* 16, 463–496 (2003).
17. Dheda, K., Barry, C. E. & Maartens, G. Tuberculosis. *The Lancet* 387, 1211–1226 (2016).
18. Delogu, G., Sali, M. & Fadda, G. The biology of *Mycobacterium tuberculosis* infection. *Mediterranean Journal of Hematology and Infectious Diseases* 5, (2013).

19. Ehlers, S. & Schaible, U. E. The granuloma in tuberculosis: Dynamics of a host-pathogen collusion. *Frontiers in Immunology* 3, (2012).
20. Harries, A. D. & Dye, C. Tuberculosis. *Annals of Tropical Medicine and Parasitology* 100, 415–431 (2006).
21. World Health Organization. Target Regimen Profiles for TB Treatment. *WHO* (2016).
22. Laghari, M. *et al.* Nanoformulations and clinical trial candidates as probably effective and safe therapy for tuberculosis. *Tropical Journal of Pharmaceutical Research* 15, 201–211 (2016).
23. Centers of Disease Control and Prevention. *Core Curriculum on Tuberculosis: What the Clinician Should Know*. (2013).
24. Tiberi, S. *et al.* Classifying new anti-tuberculosis drugs: rationale and future perspectives. *Int. J. Infect. Dis.* 56, 181–184 (2017).
25. Eidem, T. M., Roux, C. M. & Dunman, P. M. RNA decay: A novel therapeutic target in bacteria. *Wiley Interdiscip. Rev. RNA* 3, 443–454 (2012).
26. Colameco, S. & Elliot, M. A. Non-coding RNAs as antibiotic targets. *Biochem. Pharmacol.* (2016). doi:10.1016/j.bcp.2016.12.015
27. Hong, W., Zeng, J. & Xie, J. Antibiotic drugs targeting bacterial RNAs. *Acta Pharm. Sin. B* 4, 258–265 (2014).
28. Li, G. W. & Xie, X. S. Central dogma at the single-molecule level in living cells. *Nature* 475, 308–315 (2011).
29. Arraiano, C. M. *et al.* The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiology Reviews* 34, 883–923 (2010).
30. RNA Society. What is RNA. Available at: <https://www.rnasociety.org/what-is-rna>. (Accessed: 2nd May 2020)
31. Regnier, P. & Arraiano, C. M. *Degradation of mRNA in bacteria: emergence of ubiquitous features*. *BioEssays* 22, (John Wiley & Sons, Inc, 2000).
32. Arraiano, C. M. Post-transcriptional control of gene expression: bacterial mRNA degradation. *World Journal of Microbiology and Biotechnology* 9, 421–432 (1993).
33. Arraiano, C. M. & Maquat, L. E. Post-transcriptional control of gene expression: Effectors of mRNA decay. in *Molecular Microbiology* 49, 267–276 (2003).
34. Jester, B. C., Romby, P. & Lioliou, E. When ribonucleases come into play in pathogens: A survey of gram-positive bacteria. *International Journal of Microbiology* (2012). doi:10.1155/2012/592196
35. Laalami, S., Zig, L. & Putzer, H. Initiation of mRNA decay in bacteria. *Cellular and Molecular Life Sciences* 71, 1799–1828 (2014).
36. Silva, I. J. *et al.* Importance and key events of prokaryotic RNA decay: The ultimate fate of an RNA molecule. *Wiley Interdisciplinary Reviews: RNA* 2, 818–836 (2011).

37. Laalami, S. & Putzer, H. mRNA degradation and maturation in prokaryotes: The global players. *Biomolecular Concepts* 2, 491–506 (2011).
38. Anderson, K. L. & Dunman, P. M. Messenger RNA turnover processes in *Escherichia coli*, *Bacillus subtilis*, and emerging studies in *Staphylococcus aureus*. *Int. J. Microbiol.* (2009). doi:10.1155/2009/525491
39. Rustad, T. R. *et al.* Global analysis of mRNA stability in *Mycobacterium tuberculosis*. *Nucleic Acids Res.* 41, 509–517 (2013).
40. Vargas-Blanco, D. A., Zhou, Y., Gregory Zamalloa, L., Antonelli, T. & Shell, S. S. mRNA degradation rates are coupled to metabolic status in *Mycobacterium smegmatis*. *MBio* 10, (2019).
41. Matos, R. G. *et al.* Structure and Degradation Mechanisms of 3' to 5' Exoribonucleases. in 193–222 (2011). doi:10.1007/978-3-642-21078-5_8
42. Matos, R. G. *et al.* The role of ribonucleases and sRNAs in the virulence of foodborne pathogens. *Frontiers in Microbiology* 8, (2017).
43. Kanwar, S. Ribonucleases and their Applications. *J. Adv. Biotechnol. Bioeng.* 4, 17–26 (2016).
44. Arraiano, C. M., Barbas, A. & Amblar, M. Characterizing Ribonucleases *In Vitro*. Examples of Synergies between Biochemical and Structural Analysis. *Methods in Enzymology* 447, 131–160 (2008).
45. Arraiano, C. M. *et al.* The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiology Reviews* 34, 883–923 (2010).
46. Dreyfus, M. & Régnier, P. The poly(A) tail of mRNAs: Bodyguard in eukaryotes, scavenger in bacteria. *Cell* 111, 611–613 (2002).
47. Hör, J., Gorski, S. A. & Vogel, J. Bacterial RNA Biology on a Genome Scale. *Molecular Cell* 70, 785–799 (2018).
48. Deutscher, M. P. Degradation of RNA in bacteria: Comparison of mRNA and stable RNA. *Nucleic Acids Res.* 34, 659–666 (2006).
49. Mäder, U., Zig, L., Kretschmer, J., Homuth, G. & Putzer, H. mRNA processing by RNases J1 and J2 affects *Bacillus subtilis* gene expression on a global scale. *Mol. Microbiol.* 70, 183–196 (2008).
50. Condon, C. RNA Processing and Degradation in *Bacillus subtilis*. *Microbiol. Mol. Biol. Rev.* 67, 157–174 (2003).
51. Schwenk, S. & Arnvig, K. B. Regulatory RNA in *Mycobacterium tuberculosis*, back to basics. *Pathog. Dis.* 76, 35 (2018).
52. Romano, M., Squeglia, F. & Berisio, R. Structure and Function of RNase AS: A Novel Virulence Factor From *Mycobacterium tuberculosis*. *Curr. Med. Chem.* 22, 1745–1756 (2015).

53. Sharrock, A., Ruthe, A., Andrews, E. S. V., Arcus, V. A. & Hicks, J. L. VapC proteins from *Mycobacterium tuberculosis* share ribonuclease sequence specificity but differ in regulation and toxicity. *PLoS One* 13, (2018).
54. Min, A. B. *et al.* The crystal structure of the Rv0301-Rv0300 VapBC-3 toxin - Antitoxin complex from *M. tuberculosis* reveals a Mg²⁺ ion in the active site and a putative RNA-binding site. *Protein Sci.* 21, 1754–1767 (2012).
55. Watkins, H. A. & Baker, E. N. Structural and functional characterization of an RNase HI domain from the bifunctional protein Rv2228c from *Mycobacterium tuberculosis*. *J. Bacteriol.* 192, 2878–2886 (2010).
56. Płociński, P. *et al.* Proteomic and transcriptomic experiments reveal an essential role of RNA degradosome complexes in shaping the transcriptome of *Mycobacterium tuberculosis*. *Nucleic Acids Res.* 47, 5892–5905 (2019).
57. Zeller, M. E. *et al.* Quaternary structure and biochemical properties of mycobacterial RNase E/G. *Biochem. J.* 403, 207–215 (2007).
58. Płociński, P. *et al.* Proteomic and transcriptomic experiments reveal an essential role of RNA degradosome complexes in shaping the transcriptome of *Mycobacterium tuberculosis*. *Nucleic Acids Research* 47, (2019).
59. Hui, M. P., Foley, P. L. & Belasco, J. G. Messenger RNA Degradation in Bacterial Cells. *Annu. Rev. Genet.* 48, 537–559 (2014).
60. Arraiano, C. M., Matos, R. G. & Barbas, A. RNase II: The finer details of the *modus operandi* of a molecular killer. *RNA Biology* 7, (2010).
61. Reis, F. P., Pobre, V., Silva, I. J., Malecki, M. & Arraiano, C. M. The RNase II/RNB family of exoribonucleases: Putting the ‘Dis’ in disease. *Wiley Interdiscip. Rev. RNA* 4, 607–615 (2013).
62. Amblar, M., Barbas, A., Fialho, A. M. & Arraiano, C. M. Characterization of the Functional Domains of Escherichia coli RNase II. *J. Mol. Biol.* 360, 921–933 (2006).
63. Matos, R. G., Bárria, C., Pobre, V., Andrade, J. M. & Arraiano, C. M. Exoribonucleases as modulators of virulence in pathogenic bacteria. *Frontiers in cellular and infection microbiology* 2, 65 (2012).
64. Arraiano, C. M., Mauxion, F., Viegas, S. C., Matos, R. G. & Séraphin, B. Intracellular ribonucleases involved in transcript processing and decay: Precision tools for RNA. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1829, 491–513 (2013).
65. Reis, F. P., Pobre, V., Silva, I. J., Malecki, M. & Arraiano, C. M. The RNase II/RNB family of exoribonucleases: Putting the ‘Dis’ in disease. *Wiley Interdiscip. Rev. RNA* 4, 607–615 (2013).
66. Frazão, C. *et al.* Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature* 443, 110–114 (2006).
67. Matos, R. G. *et al.* The importance of proteins of the RNase II/RNB-family in pathogenic bacteria. *Frontiers in Cellular and Infection Microbiology* 4, (2014).

68. Mian, I. S. Comparative sequence analysis of ribonucleases HII, III, II PH and D. *Nucleic Acids Res.* 25, 3187–95 (1997).
69. Zuo, Y. & Deutscher, M. P. *Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Research* 29, (2001).
70. Barbas, A. *et al.* New insights into the mechanism of RNA degradation by ribonuclease II: Identification of the residue responsible for setting the RNase II end product. *J. Biol. Chem.* 283, 13070–13076 (2008).
71. dos Santos, R. F. *et al.* Major 3'–5' Exoribonucleases in the Metabolism of Coding and Non-coding RNA. in *Progress in Molecular Biology and Translational Science* 159, 101–155 (Elsevier B.V., 2018).
72. Robinson, S. R. *et al.* Dis3 isoforms vary in their endoribonuclease activity and are differentially expressed within haematological cancers. *Biochem. J.* 475, 2091–2105 (2018).
73. Robinson, S. R., Oliver, A. W., Chevassut, T. J. & Newbury, S. F. The 3' to 5' exoribonuclease DIS3: From structure and mechanisms to biological functions and role in human disease. *Biomolecules* 5, 1515–1539 (2015).
74. Barbas, A. *et al.* Determination of key residues for catalysis and RNA cleavage specificity. One mutation turns RNase II Into a 'super-enzyme'. *J. Biol. Chem.* 284, 20486–20498 (2009).
75. Deutscher, M. P. & Reuven, N. B. Enzymatic basis for hydrolytic versus phosphorolytic mRNA degradation in *Escherichia coli* and *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* 88, 3277–3280 (1991).
76. Zilhão, R., Plumbridge, J., Hajnsdorf, E., Régnier, P. & Arraiano, C. M. *Escherichia coli* RNase II: Characterization of the promoters involved in the transcription of rnb. *Microbiology* 142, 367–375 (1996).
77. Zilhão, R., Cairrão, F., Régnier, P. & Arraiano, C. M. PNPase modulates RNase II expression in *Escherichia coli*: Implications for mRNA decay and cell metabolism. *Mol. Microbiol.* 20, 1033–1042 (1996).
78. Zilhão, R., Régnier, P. & Maria Arraiano, C. The role of endonucleases in the expression of ribonuclease II in *Escherichia coli*. *FEMS Microbiol. Lett.* 130, 237–244 (1995).
79. Cairrão, F., Chora, Â., Zilhão, R., Carpousis, A. J. & Arraiano, C. M. RNase II levels change according to the growth conditions: Characterization of *gmr*, a new *Escherichia coli* gene involved in the modulation of RNase II. *Mol. Microbiol.* 39, 1550–1561 (2001).
80. Andrade, J. M., Pobre, V., Silva, I. J., Domingues, S. & Arraiano, C. M. The Role of 3'-5' Exoribonucleases in RNA Degradation. *Progress in Molecular Biology and Translational Science* 85, 187–229 (2009).
81. Reis, F. P., Bárria, C., Gomez-Puertas, P., Gomes, C. M. & Arraiano, C. M. Identification of temperature-sensitive mutations and characterization of thermolabile RNaseII variants. *FEBS Lett.* 593, 352–360 (2019).

82. Amblar, M. & Arraiano, C. M. A single mutation in *Escherichia coli* ribonuclease II inactivates the enzyme without affecting RNA binding. *FEBS J.* 272, 363–374 (2005).
83. Theobald, D. L., Mitton-Fry, R. M. & Wuttke, D. S. Nucleic acid recognition by OB-fold proteins. *Annual Review of Biophysics and Biomolecular Structure* 32, 115–133 (2003).
84. Steitz, T. A. & Steitz, J. A. A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. U. S. A.* 90, 6498–6502 (1993).
85. Matos, R. G., Barbas, A. & Arraiano, C. M. Comparison of EMSA and SPR for the characterization of RNA-RNase II complexes. *Protein J.* 29, 394–397 (2010).
86. Cheng, Z. F. & Deutscher, M. P. Purification and characterization of the *Escherichia coli* exoribonuclease RNase R. Comparison with RNase II. *J. Biol. Chem.* 277, 21624–21629 (2002).
87. Chu, L. Y. *et al.* Structural insights into RNA unwinding and degradation by RNase R. *Nucleic Acids Res.* 45, 12015–12024 (2017).
88. Matos, R. G. *et al.* The importance of proteins of the RNase II/RNB-family in pathogenic bacteria. *Front. Cell. Infect. Microbiol.* 4, 68 (2014).
89. Vincent, H. A. & Deutscher, M. P. The roles of individual domains of RNase R in substrate binding and exoribonuclease activity the nuclease domain is sufficient for digestion of structured RNA. *J. Biol. Chem.* 284, 486–494 (2009).
90. Cairrão, F. & Arraiano, C. M. The role of endoribonucleases in the regulation of RNase R. *Biochem. Biophys. Res. Commun.* 343, 731–737 (2006).
91. Cairrão, F., Cruz, A., Mori, H. & Arraiano, C. M. Cold shock induction of RNase R and its role in the maturation of the quality control mediator SsrA/tmRNA. *Mol. Microbiol.* 50, 1349–1360 (2003).
92. Tobe, T., Sasakawa, C., Okada, N., Honma, Y. & Yoshikawa, M. vacB, a novel chromosomal gene required for expression of virulence genes on the large plasmid of *Shigella flexneri*. *J. Bacteriol.* 174, 6359–6367 (1992).
93. Hossain, S. T., Malhotra, A. & Deutscher, M. P. How RNase R degrades structured RNA: Role of the helicase activity and the S1 domain. *J. Biol. Chem.* 291, 7877–7887 (2016).
94. Matos, R. G., Barbas, A. & Arraiano, C. M. RNase R mutants elucidate the catalysis of structured RNA: RNA-binding domains select the RNAs targeted for degradation. *Biochem. J.* 423, 291–301 (2009).
95. Awano, N. *et al.* *Escherichia coli* RNase R has dual activities, helicase and RNase. *J. Bacteriol.* 192, 1344–1352 (2010).
96. Vincent, H. A. & Deutscher, M. P. Insights into How RNase R Degrades Structured RNA: Analysis of the Nuclease Domain. *J. Mol. Biol.* 387, 570–583 (2009).
97. Matos, R. G., Barbas, A., Gómez-Puertas, P. & Arraiano, C. M. Swapping the domains of exoribonucleases RNase II and RNase R: Conferring upon RNase II the ability to degrade ds RNA. *Proteins Struct. Funct. Bioinforma.* 79, 1853–1867 (2011).

98. Hossain, S. T., Malhotra, A. & Deutscher, M. P. The helicase activity of ribonuclease R is essential for efficient nuclease activity. *J. Biol. Chem.* 290, 15697–15706 (2015).
99. Tofajjen Hossain, S. & Deutscher, M. P. Helicase activity plays a crucial role for RNase R function *in vivo* and for RNA metabolism. *J. Biol. Chem.* 291, 9438–9443 (2016).
100. Matos, R. G. *et al.* The importance of proteins of the RNase II/RNB-family in pathogenic bacteria. *Frontiers in Cellular and Infection Microbiology* 4, 68 (2014).
101. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410 (1990).
102. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 47, W636–W641 (2019).
103. GenScript. OptimumGene™ - Codon Optimization. (2015). Available at: www.genscript.com. (Accessed: 7th September 2020)
104. Scientific & Educational Software. Clone Manager 9 Professional Edition. (2015).
105. Matos, R. G., Viegas, S. C. & Arraiano, C. M. In Vitro Characterization of the Prokaryotic Counterparts of the Exosome Complex. *Methods Mol. Biol.* 2062, 47–61 (2020).
106. Boivin, S., Kozak, S. & Meijers, R. Optimization of protein purification and characterization using Thermofluor screens. *Protein Expression and Purification* 91, 192–206 (2013).
107. Newman, J. Novel buffer systems for macromolecular crystallization. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 60, 610–612 (2004).
108. Viegas, S. C., Matos, R. G. & Arraiano, C. M. The Bacterial Counterparts of the Eukaryotic Exosome: An Evolutionary Perspective. *Methods Mol. Biol.* 2062, 37–46 (2020).
109. Haddad, N. *et al.* The RNase R from *campylobacter jejuni* has unique features and is involved in the first steps of infection. *J. Biol. Chem.* 289, 27814–27824 (2014).
110. Pearson, W. R. An introduction to sequence similarity ('homology') searching. *Curr. Protoc. Bioinforma.* 0 3, (2013).
111. Palidwor, G., Reynaud, E. G. & Andrade-Navarro, M. A. Taxonomic colouring of phylogenetic trees of protein sequences. *BMC Bioinformatics* 7, 79 (2006).
112. Dereeper, A. *et al.* Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 36, (2008).
113. Matos, R. G., Barbas, A. & Arraiano, C. M. RNase R mutants elucidate the catalysis of structured RNA: RNA-binding domains select the RNAs targeted for degradation. *Biochem. J.* 423, 291–301 (2009).
114. Tuteja, N. & Tuteja, R. Helicases as molecular motors: An insight. *Phys. A Stat. Mech. its Appl.* 372, 70–83 (2006).
115. Kim, Y. *et al.* Chapter 3 High-Throughput Protein Purification for X-Ray Crystallography and NMR. *Advances in Protein Chemistry and Structural Biology* 75, 85–105 (2008).

116. Luft, J. R., Wolfley, J. R. & Snell, E. H. What's in a drop? Correlating observations and outcomes to guide macromolecular crystallization experiments. *Cryst. Growth Des.* 11, 651–663 (2011).
117. Meyer, A., Betzel, C. & Pusey, M. Latest methods of fluorescence-based protein crystal identification. *Acta Crystallogr. Sect. F:Structural Biol. Commun.* 71, 121–131 (2015).
118. Ericsson, U. B., Hallberg, B. M., DeTitta, G. T., Dekker, N. & Nordlund, P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal. Biochem.* 357, 289–298 (2006).
119. Lo, M. C. *et al.* Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal. Biochem.* 332, 153–159 (2004).
120. Bruce, D., Cardew, E., Freitag-Pohl, S. & Pohl, E. How to Stabilize Protein: Stability Screens for Thermal Shift Assays and Nano Differential Scanning Fluorimetry in the Virus-X Project. *J. Vis. Exp* 58666 (2019). doi:10.3791/58666
121. Eftink, M. R. The use of fluorescence methods to monitor unfolding transitions in proteins. *Biophys. J.* 66, 482–501 (1994).
122. McPherson, A. & Gavira, J. A. Introduction to protein crystallization. *Acta Crystallographica Section F:Structural Biology Communications* 70, 2–20 (2014).
123. Rendon, A. *et al.* Classification of drugs to treat multidrug-resistant tuberculosis (MDR-TB): Evidence and perspectives. *Journal of Thoracic Disease* 8, 2666–2671 (2016).

6. Supplementary Material

Appendix A:

Table S.1. TB drugs classification and their targets (WHO 2011). Group 1 for the first-line oral anti-TB drugs; Group 2 for injectable anti-TB drugs; Group 3 for fluoroquinolones; Group 4 for oral bacteriostatic second-line anti-TB drugs; Group 5 - Newer drugs with increasing evidence ^{22,24,123}.

Group	Drug	Targeted Site	Inhibited target
1	Isoniazid	Enoyl-acyl carrier protein reductase	Mycolic acids biosynthesis
	Rifampicin	Subunit of DNA-dependent RNA polymerase	RNA synthesis
	Pyrazinamide	S1 component of 30S ribosomal subunit	Translation
	Ethambutol	Arabinosyl transferases	Arabinogalactan biosynthesis
2	Streptomycin	S12 and 16S rRNA components of 30S ribosomal subunit	Protein synthesis
	Kanamycin	30S ribosomal subunit	Protein synthesis
	Amikacin	30S ribosomal subunit	Protein synthesis
	Capreomycin	Inter bridge B2a between 30S and 50S ribosomal subunit	Protein synthesis
3	Ofloxacin	DNA gyrase/topoisomerase IV	DNA supercoiling
	Levofloxacin	DNA gyrase/topoisomerase IV	DNA supercoiling
	Moxifloxacin	DNA gyrase/topoisomerase IV	DNA supercoiling
	Gatifloxacin	DNA gyrase/topoisomerase IV	DNA supercoiling
4	p-Aminosalicylic acid	Dihydropteroate synthase	Folate biosynthesis
	Cycloserine	D-alanine racemase	Peptidoglycan synthesis
	Ethionamide/ Prothionamide	Enoyl-acyl carrier protein reductase (InhA)	Mycolic acids biosynthesis
5	Linezolid	23S portion of the 50S subunit	Protein synthesis
	Bedaquiline	Subunit c of ATP synthase	Proton pump
	Pretomanid	Poorly understood	Mycolic acids biosynthesis
	Clofazimine	Guanine bases of DNA	Bacterial proliferation
	Amoxicillin	Penicillin-binding proteins	Cell wall synthesis
	Imipenem	Penicillin-binding proteins	Cell wall synthesis
	Meropenem	Penicillin-binding proteins	Cell wall synthesis
	Thioacetazone	Not Available	Not Available
	Clarithromycin	23S portion of the 50S subunit	Protein synthesis

Mt_rnr2	MSVEKQSNDPYREREAEKYDSPIASREFILETVRRHTKPVSREDIAQELGITGEEAQEAL	60
Ec_RNaseR	-----MSQDPFQEREAEKYANPIPSREFILEHLLTKREKSPASRDELAVELHIEGEEQLEGL	55
Mt_rnr1	-----	0
Ec_RNaseII	-----	0
Mt_rnr2	RRRLRAMERDGGQLVFTRRQCYALPERLDLIKGTVIGHRDGFGFLRVEGQKDDLYIPQDQM	120
Ec_RNaseR	RRRLRAMERDGGQLVFTRRQCYALPERLDLVKGTVIGHRDGYGFLRVEGRKDDLYLSSEQM	115
Mt_rnr1	-----MFQDNPLLAQLKQ--QLHDKTLRVEGIVKSTDKGYGFLEVDGQKS-YFIPPPQM	51
Ec_RNaseII	-----MFQDNPLLAQLKQ--QLHSQTPRAEGVVKATEKGFGLFLEVDQKS-YFIPPPQM * :*. * : * * . : * * . :*. * :*. * : *	51
Mt_rnr2	KLCIHGDVILAQVMGADRKGRREARVVRVTEPKTGQIVGRYFTEE--GMGFVVPDDSRRLC	178
Ec_RNaseR	KTCIHGDQVLAQPLGADRKGRREARIVRVLPKTSQIVGRYFTEA--GVGFVVPDDSRRLS	173
Mt_rnr1	KKVMHGDRTAAIHTVNNKDVAE--PEALITPFLDRFVGRVQRKENDDRLWIIPDHPLLK	109
Ec_RNaseII	KKVMHGDRIIAVIHSEKERESAE--PEELVEPFLTRFVGKVGQK--NDRLAIVPDHPLLK * :*** : * . : * : * :*** : : . :*** *	107
Mt_rnr2	FDILIPKE--SINGARMGNIVVAELTRPGKRVQALGV-ITEVLGEKMGTSIAVDIALRK	235
Ec_RNaseR	FDILIPPD--QIMGARMGFVVVVELTQRPTRRTKAVGK-IVEVLGDNMGTMGMAVDIALRT	230
Mt_rnr1	DSIPCRPAKGVDFEFRQGDWAVAEMRRHPLKGD RSFHAETG FITDADDHYAPWWVTLTR	169
Ec_RNaseII	DAIPCAARGLNHEFKEGDWAVAEMRRHPLKGD RSFYAELTQYITFGDDHFVPWWVTLAR * : * : * : * : * : * : * : * : *	167
Mt_rnr2	HEIPFTWPAQVERQVADIREEVPESAKKGVRDLRDLPLVTIDGEDARDFDDAVHCEAKRG	295
Ec_RNaseR	HEIPYIWPQAVEQVQVAGLKEEVEEAKAGVRDLRDLPLVTIDGEDARDFDDAVYCEKKRG	290
Mt_rnr1	HQHALKEPEPEVETM-----ADDNLPRTDLTHLDFVTIDASTEDMDDALYITTGDN	221
Ec_RNaseII	HNLEKEAPDGVATEM-----LDEGLVREDLTALDFVTIDASTEDMDDALFAKALPD * : * * : . * * * : * * . : * : * : *	219
Mt_rnr2	GGWRLWVAIADVSYVVRPQTALDDEARSRGTSVYFPSQVVPMLPEVLSNGVCSLNPQVDR	355
Ec_RNaseR	GGWRLWVAIADVSYVVRPSTPLDREARNRGTSVYFPSQVVPMLPEVLSNGLCSLNPQVDR	350
Mt_rnr1	NQLQLTIAIADPTSYIPAGSELDAIAHKRAFTNYLPGFNIPMLPRNLSDDLCSLHPHVRR	281
Ec_RNaseII	DKLQLIVAIADPTAWIAEGSKLDKAAKIRAFNTYLPGFNIPMLPRELSDDLCSLRANEVR . : * : * * : : : * * * : * . : * * . : * * . : * * : *	279
Mt_rnr2	LCMVCEMTISAKGKLSG-YRFYEALMSSSHARLTYTKVWKILEGDEELREHYAPLVKHLEE	414
Ec_RNaseR	LCMVCEMTVSSKGRLTG-YKFYEAVMSSSHARLTYTKVWHILQGDQDLREQYAPLVKHLEE	409
Mt_rnr1	PALVCQSVNEDGSLSDDAVFFAAWVESKAKLAYDDVSDLEGVADRVTDETIRRRQISL	341
Ec_RNaseII	PVLACRMTLSADGTIEDNIEFFAATIESKAKLVYDQVSDWLENTGDWQPESEIAEQVRL :. * : * . * : * * : * : * * . * . * : * : * : *	339
Mt_rnr2	LHKLYKALAEARDERGAI AFES EAKFI FNAEKRIERIEPVL RND AHKLIEECMIMANVA	474
Ec_RNaseR	LHNLYKVLDKAREERGGISFESE EAKFI FNAERRIERIEQTQRND AHKLIEECMILANIS	469
Mt_rnr1	LQDMYQRRHHWRETHALVFKDRPDYR FILDENG VVDIVAEQRRSANRIVEEAMITANIC	401
Ec_RNaseII	LAQICQRRGEWRHNHALVFKDRPDYR FILGEKGEVLDIVAEPRRIANRIVEEAMIAANIC * . : : . * . : : : : * * . : : * * . : * : * * * *	399
Mt_rnr2	AARFVEKNNEPALYRVHDPKEDSVINLRTVFNELGLTLPGGLTPEPKDYARVMLEVAER	534
Ec_RNaseR	AARFVEKAKEPALFRIHDKPSTEAITSF RSVLAELGLELPGGNKPEPRDYAELLESVADR	529
Mt_rnr1	AAQVLDKKLGFGVYNVHTGLDPLTISQAVALLAENGINFSAEELLTLDGFCRLRRELDNQ	461
Ec_RNaseII	AARVLRDKLGFGIYNVHMGFDPANADALAALLKTHGLHVDAAEVLTLDGFCCLRRELDQA ** : . . : * : . : * : * : * : * : *	459
Mt_rnr2	PDHELLQTMLLRSMKQAIYDPDNRGHFGGLALPAYAHFTSPIRRYPDL SLHRAIKYLLAKE	594
Ec_RNaseR	PDAEMLQTMLLRSMKQAIYDPENRGHFGGLALQSYAHFTSPIRRYPDL TLHRAIKYLLAKE	589
Mt_rnr1	PTQ-FLDSRIRRFQTFAEVRMEKGP HFGGLGLEYATWTSPIRKYS DIVNHRLLKAIAGQ	520
Ec_RNaseII	PTG-FLDSRIRRFQSFAEISTEPGPHFGGLGLEAYATWTSPIRKYGDMINHRLLKAVIKGE * : * : : * . * : * * * . * * : * * * * * : * * : * : *	518

Mt_rnr2	RGEG--KSQTGAWHYDMQSMQLQGEHTSMTERRADEATRVDWLKCDFMQDQVGN--VF	650
Ec_RNaseR	QGHQGN TTETGGYHYSMEEMQLQGHC SMAERRADEATRVDWLKCDFMLDQVGN--VF	647
Mt_rnr1	SAEKPAE-----ALAEHLAERRRANRMAERDVG DWLYARYLQPYAGTDTPF	566
Ec_RNaseII	TATRPQD-----EITVQMAERRRLNRMAERDVG DWLYARFLKDKAGTDTRF	564
	. : : . * * ** . ** . : : . * . *	
Mt_rnr2	TGLITSVTGFGFFVRLNDL----FIDGLVHVSTLDNDYYRYDNVQRLIGESSSNVYRLG	706
Ec_RNaseR	KGVISSVTGFGFFVRLDDL----FIDGLVHVSSLDNDYYRFDQVGQRLMGESSGQTYRLG	703
Mt_rnr1	PAEIIDITRGGVRVRLTDNGAVAFIPGTFIHP-VKDELQCSQETGSVL--IKGEVRYRLN	623
Ec_RNaseII	AAEIVDISRGGMRVRLVDNGAIAFIPAPFLHA-VRDELVCSQENGTVQ--IKGETVYKVT	621
	. * . : : * . ** * ** . . : : : * . . * : :	
Mt_rnr2	DEVEIRVEAVHMDERTIDFALISTTRRARNEGKTARTRAKAGEKGDKRRDGKKGDPRDK	766
Ec_RNaseR	DRVEVRVEAVNMDERKIDFSLISSERAPRNVGKTAREKAKKGDAGKK--GGKRRQVGKKV	761
Mt_rnr1	DIIPVKIEEVKMETRSILARPLCSGITD-----	651
Ec_RNaseII	DVIDVTIAEVRMETRSIIARPVA-----	644
	* : : : * . * : * . * : .	
Mt_rnr2	NFEPDSAFRKKKKDDGPKAPKAPK--GKKSARTKKIEANIKAKRAKKKAEK-	815
Ec_RNaseR	NFEPDSAFRGEKKTKPKAAKKDARKAKKPSAKTQKIAAATKAKRAAKKKVAE	813
Mt_rnr1	-----	651
Ec_RNaseII	-----	644

Figure S.1. Multiple sequence alignment colored by domains of proteins belonging to RNB-family. *E. coli* RNase II, *E. coli* RNase R, *M. tuberculosis* RNR1 and *M. tuberculosis* RNR2 amino acid sequences aligned using the Clustal Omega software. HTH domain is represented in purple, CSD1 in dark-orange, CSD2 in yellow, RNB in blue, S1 in green and K/R rich region in light-orange. Some conserved amino acid residues are highlighted in grey.

Table S.2. Culture media, solutions and gels used in the experimental media.

Luria Agar (LA)	Luria Broth (LB)
10 g Tryptone	10 g Tryptone
5 g Yeast Extract	5 g Yeast Extract
10 g NaCl	10 g NaCl
10 g Agar	ddH ₂ O up to 1000 mL and autoclave
ddH ₂ O up to 1000 mL and autoclave	
Terrific Broth (TBm)	Ampicillin (Amp)
50.8 g Terrific Broth	100 mg Amp per mL ddH ₂ O
4 mL Glycerol	Sterile filtered using 0.22 µm syringe
ddH ₂ O up to 1000 mL and autoclave	Aliquot and store at -20°C
	100 µg/mL added in all culture media

TBE 10x

121.1 g Tris Base
61.8 g Boric Acid
7.4 g EDTA
ddH₂O to 1000 mL

TBS 10X

24.2 g Tris Base
80 g NaCl
Milli-Q H₂O up to 1000 mL

SDS-PAGE Running Buffer 10X

30 g of Tris Base
144 g of Glycine
10 g of SDS
ddH₂O up to 1000 mL

Buffer A_RNR

20 mM Bis-Tris pH 6.0
250 mM NaCl
5 mM MgSO₄
20 mM Imidazole
Filter 0.45 µm and Degas

Buffer B_RNR

20 mM Bis-Tris pH 6.0
250 mM NaCl
5 mM MgSO₄
250 mM Imidazole
Filter 0.45 µm and Degas

Transfer Buffer 10X

1.6 g SDS
32 g Tris base
16 g Glycine
Milli-Q H₂O up to 400 mL

Transfer Buffer 1X

25 mL Transfer Buffer 10X
50 mL Methanol
Milli-Q H₂O up to 250 mL

Protein Loading Buffer 6X

10 (w/v) SDS
10 mM β-mercaptoethanol
20 % (v/v) Glycerol
0.2 M Tris-HCl pH 6.8
0.05 % (w/v) Bromophenol Blue

Buffer A_RNB

20 mM Tris-HCl pH 8.0
500 mM NaCl
20 mM Imidazole
Filter 0.45 µm and Degas

Buffer B_RNB

20 mM Tris-HCl pH 8.0
500 mM NaCl
500 mM Imidazole
Filter 0.45 µm and Degas

Buffer C_RNR

20 mM Bis-Tris pH 6.0
 250 mM NaCl
 5 mM MgSO₄
 Filter 0.45 µm and Degas

Buffer C_RNB

10 mM Tris-HCl pH 8.0
 50 mM KCl
 Filter 0.45 µm and Degas

Activity Buffer 5x

100 mM Tris-HCl pH 8.0
 500 mM KCl
 5 mM MgCl₂
 5 mM DTT

Binding Buffer 5x

10 mM Tris-HCl pH 8.0
 2 mM DTT
 100 mM KCl
 25 mM EDTA

STOP Solution

95 % (w/v) Formamide
 0.025 % (w/v) Xylene Cyanol
 0.025 % (w/v) Bromophenol Blue
 5 mM EDTA
 0.025 % (w/v) SDS

Loading Buffer 6X

0,25 % (w/v) Bromophenol Blue
 0,25 % (w/v) Xylene Cyanol
 30 % (v/v) Glycerol

20 % Denaturing PAA gel

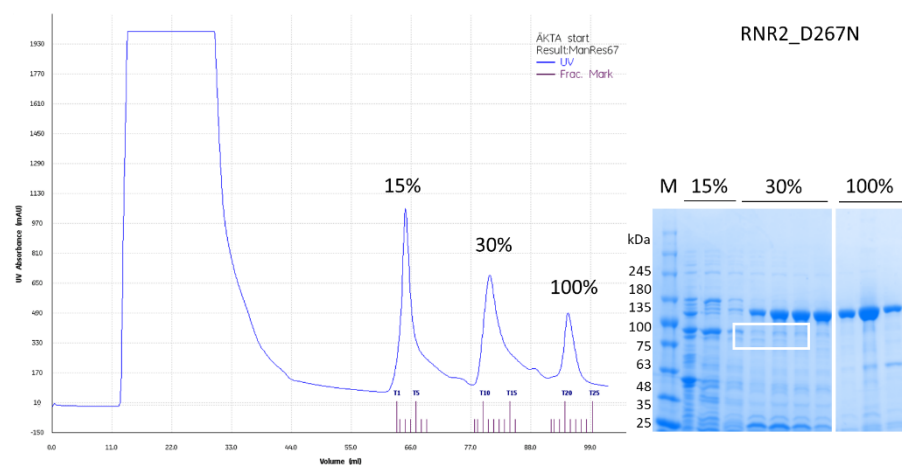
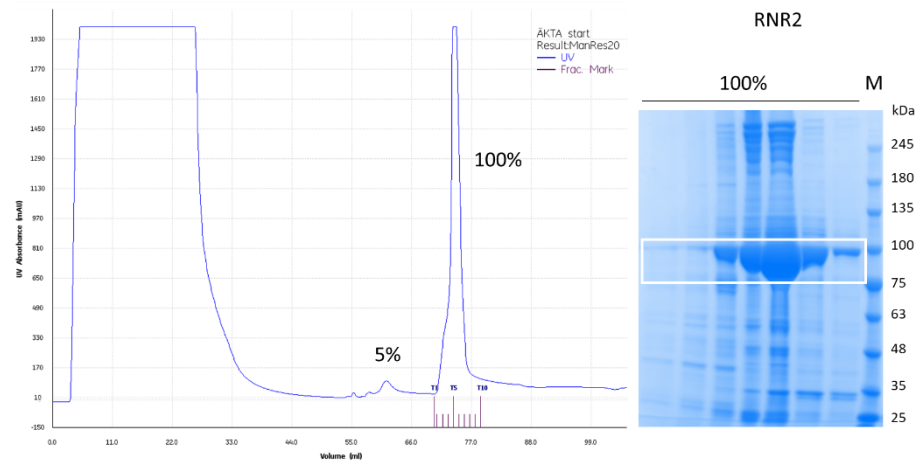
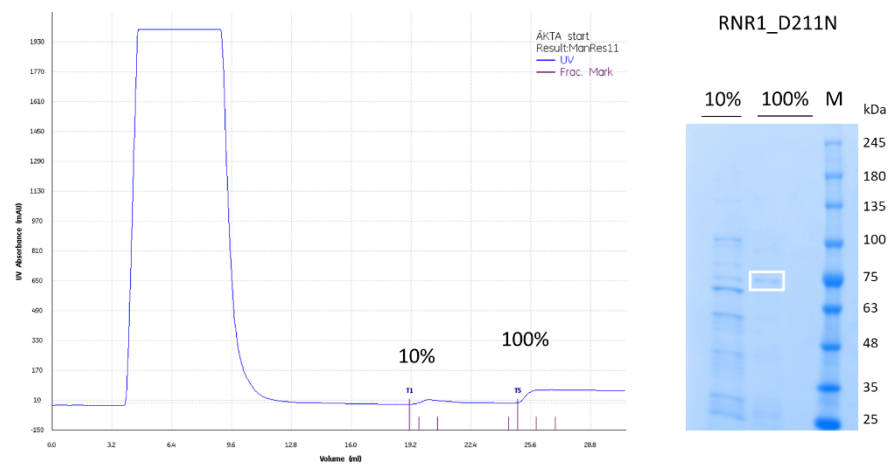
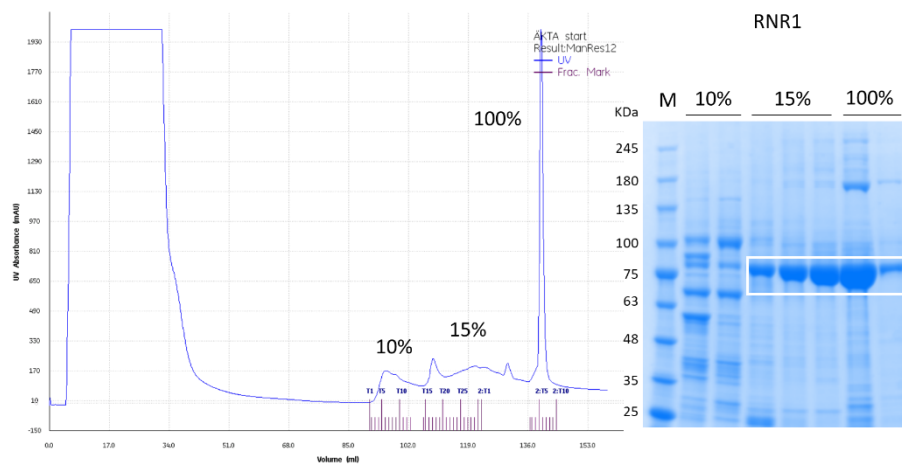
35 mL 40 % (w/v) Acrylamide/Bis 19:1
 7 mL TBE 10X
 29.4 g Urea
 Dissolve at 60°C, ddH₂O up to 70 mL
 300 µL APS 10 % (w/v)
 50 µL TEMED

6 % Non-denaturing PAA gel

7,5 mL 40 % (w/v) Acrylamide/Bis 19:1
 5 mL TBE 10X
 ddH₂O up to 50 mL
 300 µL APS 10 % (w/v)
 50 µL TEMED

Table S.3. Proteins used in the experimental procedures.

Protein name	Original Strain	Total of amino acid residues	Molecular weight (kDa)	Isoelectric Point
RNase II	<i>E. coli</i>	644	72.4	5.30
RNase R	<i>E. coli</i>	813	92.1	8.84
RNR1	<i>M. tuberculosis</i>	669	75.6	5.64
RNR2	<i>M. tuberculosis</i>	833	94.7	8.42



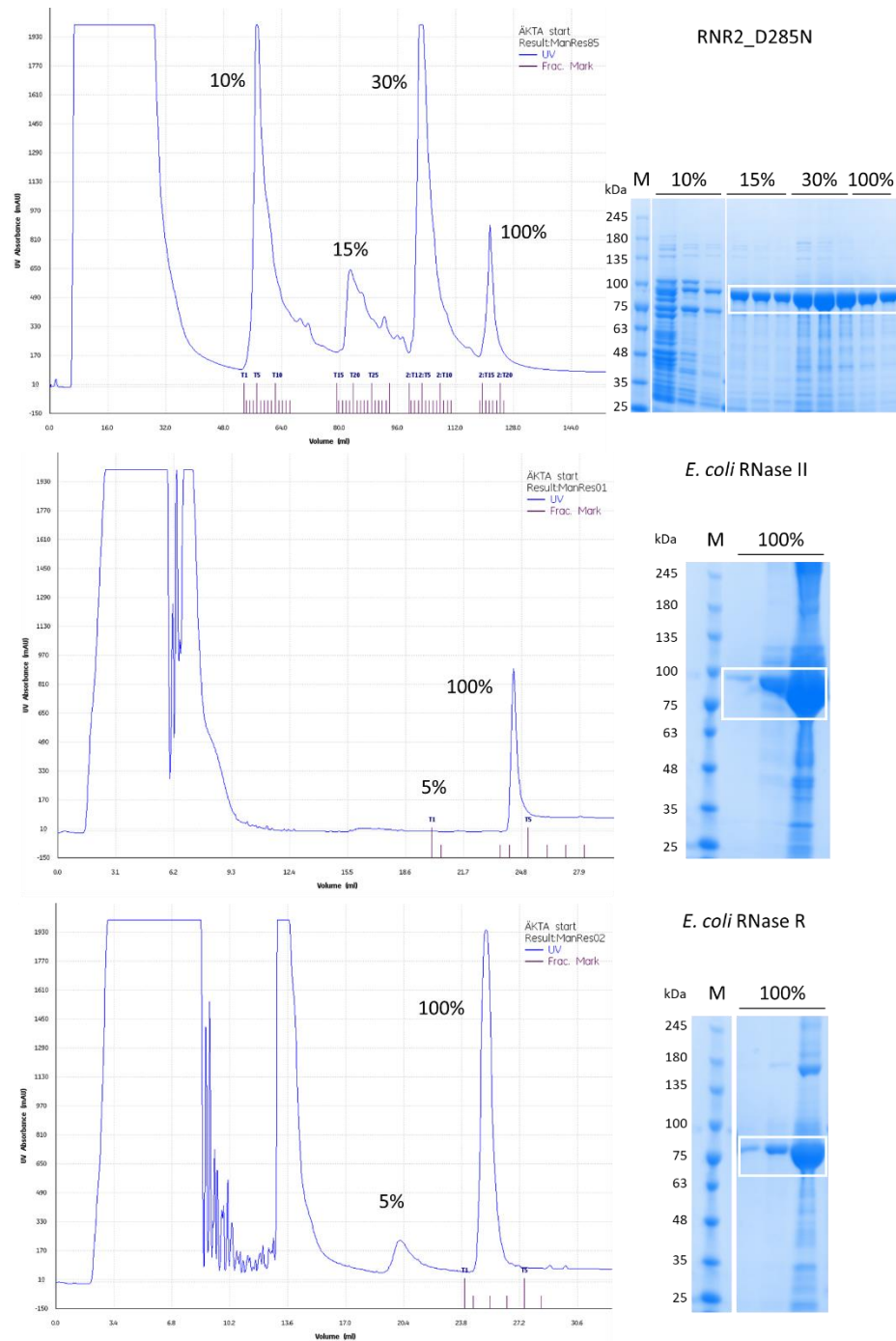
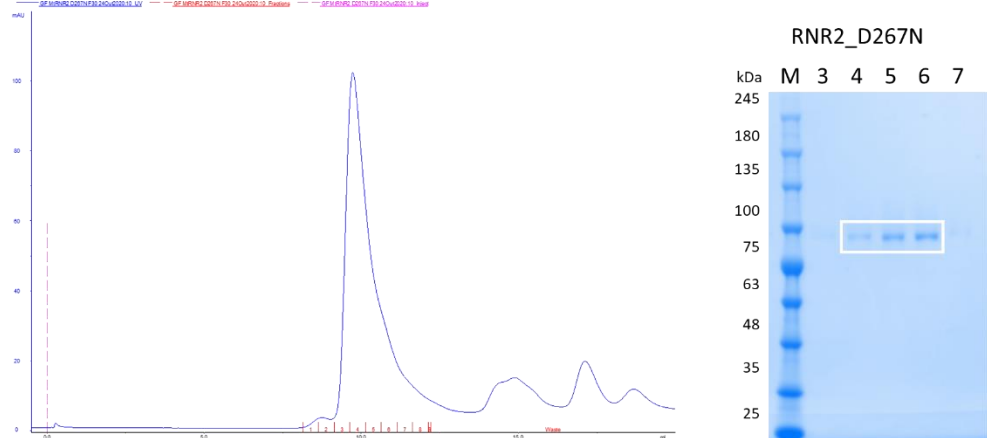
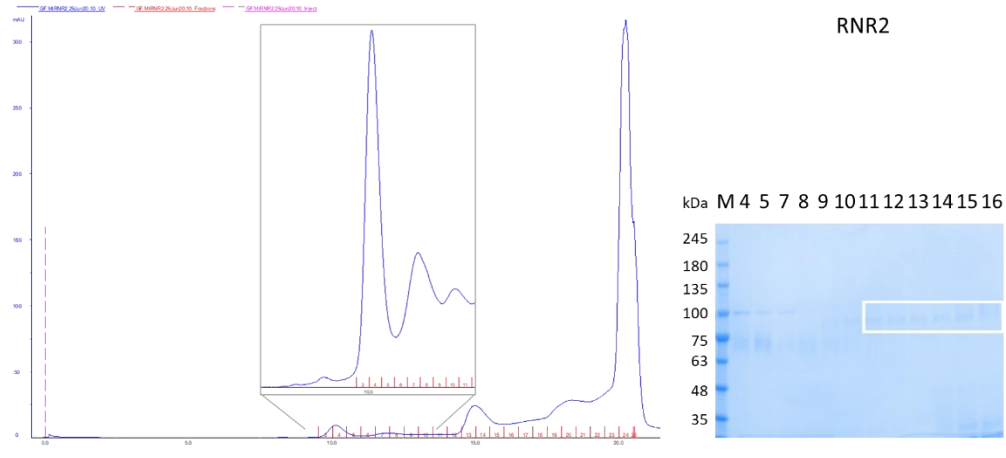
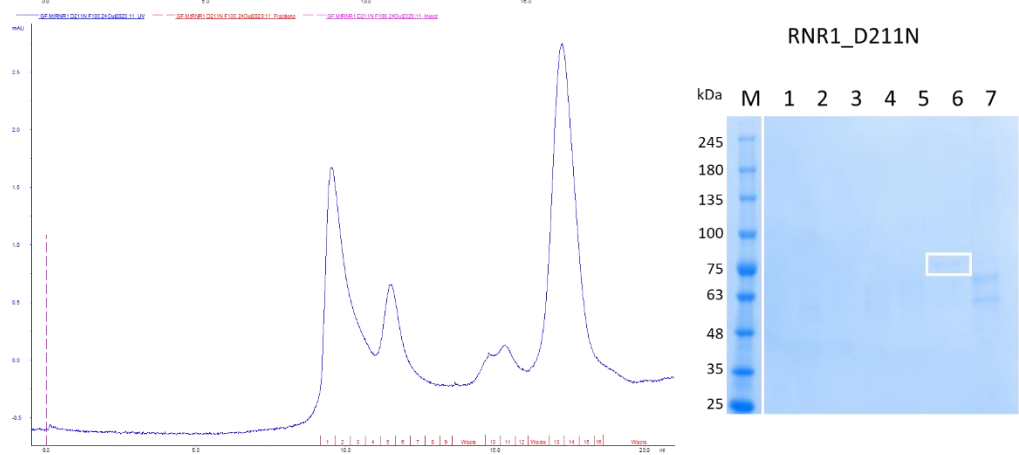
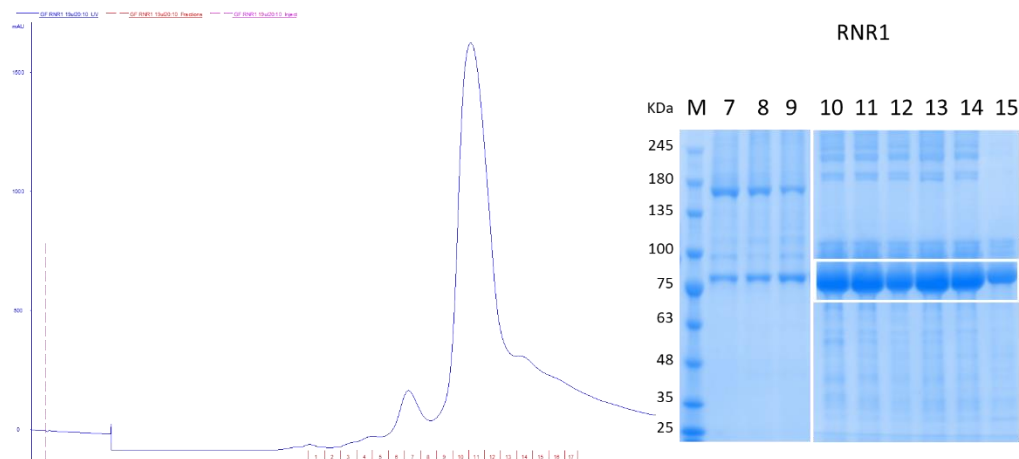


Figure S.2. Affinity chromatography of and *M. tuberculosis* and *E. coli* ribonucleases from the RNB-family. Left panel- Elution profile using Ni Sepharose High Performance columns on an ÄKTA Start system. A gradient up to 500 mM imidazole was used and each peak corresponds to a different percentage of Buffer B passing in the column. Right panel- 4-12 % SDS-PAGE analysis of the protein fractions collected during the affinity chromatography. Each well was loaded with the same sample volume (10µL) and the marker used (lane “M”) was the NZYColour Protein Marker II (Nzytech). The bands corresponding to the target protein are highlighted in white.



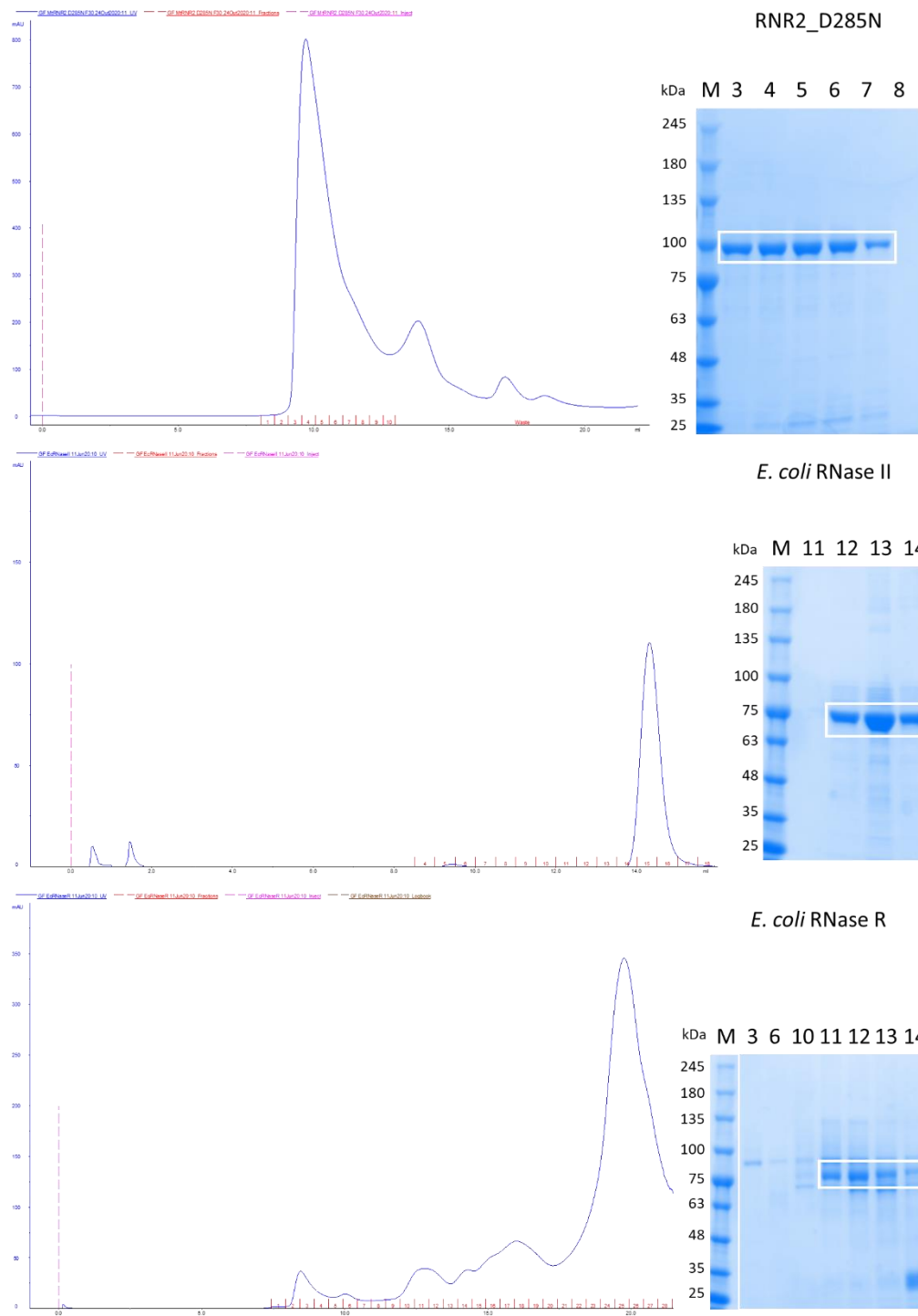


Figure S.3. Size-exclusion chromatography of *M. tuberculosis* and *E. coli* ribonucleases from the RNB-family. Left panel- Chromatogram of each protein using a Superdex 200 Increase 10/300 GL on an ÄKTA Purifier UPC 10 GE system. Sample injection is depicted in pink, UV absorbance in blue and the collected fractions in red. **Right panel-** 4-12 % SDS-PAGE analysis of the protein fractions collected during the size-exclusion chromatography. Each well was loaded with the same sample volume (10mL) and the marker used (lane “M”) was the NZYColour Protein Marker II (Nzytech). The bands corresponding to the target protein are highlighted in white.

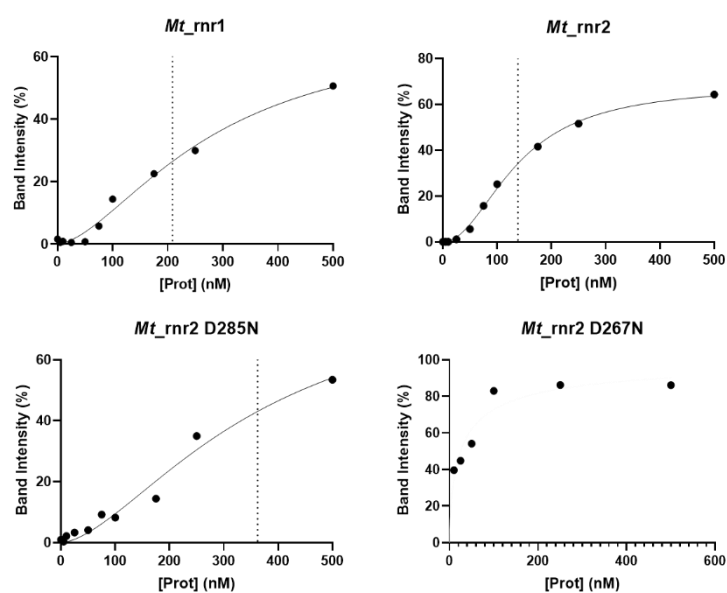


Figure S.4. Affinity data of the *M. tuberculosis* ribonucleases for binding to RNA. The data are plotted as the fraction of bound RNA versus molar concentration of each protein. Curves were calculated by fitting a non-linear regression model to the experimental data points using the GraphPad Prism software.

Table S.4. Composition of the RUBIC Buffer screen conditions.

Concentration during assay		Concentration during assay	
A1	100 % Ultrapure water	E1	100 mM SPG 4.0
A2	100 mM Citrate 4.0	E2	100 mM SPG 4.5
A3	100 mM Sodium acetate 4.5	E3	100 mM SPG 5.0
A4	100 mM Citrate 5.0	E4	100 mM SPG 5.5
A5	100 mM MES 6.0	E5	100 mM SPG 6.0
A6	100 mM Potassium phosphate 6.0	E6	100 mM SPG 6.5
A7	100 mM Citrate 6.0	E7	100 mM SPG 7.0
A8	100 mM Bis-Tris 6.5	E8	100 mM SPG 7.5
A9	100 mM MES 6.5	E9	100 mM SPG 8.0
A10	100 mM Sodium phosphate 7.0	E10	100 mM SPG 8.5
A11	100 mM Potassium phosphate 7.0	E11	100 mM SPG 9.0
A12	100 mM HEPES 7.0	E12	100 mM SPG 10.0
B1	100 mM MOPS 7.0	F1	100 mM HEPES 7.5
B2	100 mM Ammonium acetate 7.3	F2	20 mM HEPES 7.5
B3	100 mM Tris-HCl 7.5	F3	50 mM HEPES 7.5
B4	100 mM Sodium phosphate 7.5	F4	125 mM HEPES 7.5
B5	100 mM Imidazole 7.5	F5	250 mM Sodium phosphate 7.5
B6	100 mM HEPES 8.0	F6	20 mM Sodium phosphate 7.5
B7	100 mM Tris-HCl 8.0	F7	50 mM Sodium phosphate 7.5
B8	100 mM Tricine 8.0	F8	125 mM Sodium phosphate 7.5
B9	100 mM BICINE 8.0	F9	250 mM Tris-HCl 8.0
B10	100 mM BICINE 8.5	F10	20 mM Tris-HCl 8.0
B11	100 mM Tris-HCl 8.5	F11	50 mM Tris-HCl 8.0
B12	100 mM CHES 9.0	F12	125 mM Tris-HCl 8.0
C1	250 mM Sodium Chloride	G1	250 mM HEPES 7.5 / 50 mM Sodium Chloride

C2	100 mM Citrate 4.0 / 250 mM Sodium Chloride	G2	50 mM HEPES 7.5 / 125 mM Sodium Chloride
C3	100 mM Sodium acetate 4.5 / 250 mM Sodium Chloride	G3	50 mM HEPES 7.5 / 250 mM Sodium Chloride
C4	100 mM Citrate 5.0 / 250 mM Sodium Chloride	G4	50 mM HEPES 7.5 / 500 mM Sodium Chloride
C5	100 mM MES 6.0 / 250 mM Sodium Chloride	G5	50 mM HEPES 7.5 / 750 mM Sodium Chloride
C6	100 mM Potassium phosphate 6.0 / 250 mM Sodium Chloride	G6	50 mM HEPES 7.5 / 1000 mM Sodium Chloride
C7	100 mM Citrate 6.0 / 250 mM Sodium Chloride	G7	50 mM Tris-HCl 8.0 / 50 mM Sodium Chloride
C8	100 mM Bis-Tris 6.5 / 250 mM Sodium Chloride	G8	50 mM Tris-HCl 8.0 / 125 mM Sodium Chloride
C9	100 mM MES 6.5 / 250 mM Sodium Chloride	G9	50 mM Tris-HCl 8.0 / 250 mM Sodium Chloride
C10	100 mM Sodium phosphate 7.0 / 250 mM Sodium Chloride	G10	50 mM Tris-HCl 8.0 / 500 mM Sodium Chloride
C11	100 mM Potassium phosphate 7.0 / 250 mM Sodium Chloride	G11	50 mM Tris-HCl 8.0 / 750 mM Sodium Chloride
C12	100 mM HEPES 7.0 / 250 mM Sodium Chloride	G12	50 mM Tris-HCl 8.0 / 1000 mM Sodium Chloride
D1	100 mM MOPS 7.0 / 250 mM Sodium Chloride	H1	50 mM MES/Bis-Tris 6.0
D2	100 mM Ammonium acetate 7.3 / 250 mM Sodium Chloride	H2	50 mM MES/Imidazole 6.5
D3	100 mM Tris-HCl 7.5 / 250 mM Sodium Chloride	H3	50 mM Bis-Tris/PIPES 6.5
D4	100 mM Sodium phosphate 7.5 / 250 mM Sodium Chloride	H4	50 mM MOPS/Bis-Tris propane 7.0
D5	100 mM Imidazole 7.5 / 250 mM Sodium Chloride	H5	50 mM Phosphate/Citrate 7.5
D6	100 mM HEPES 8.0 / 250 mM Sodium Chloride	H6	50 mM MOPS/Sodium HEPES 7.5
D7	100 mM Tris-HCl 8.0 / 250 mM Sodium Chloride	H7	0,1 mM BICINE/Tris 8.5
D8	100 mM Tricine 8.0 / 250 mM Sodium Chloride	H8	50 mM Imidazole 7.5 / 250 mM Sodium Chloride
D9	100 mM BICINE 8.0 / 250 mM Sodium Chloride	H9	125 mM Imidazole 7.5 / 250 mM Sodium Chloride
D10	100 mM BICINE 8.5 / 250 mM Sodium Chloride	H10	250 mM Imidazole 7.5 / 250 mM Sodium Chloride
D11	100 mM Tris-HCl 8.5 / 250 mM Sodium Chloride	H11	350 mM Imidazole 7.5 / 250 mM Sodium Chloride
D12	100 mM CHES 9.0 / 250 mM Sodium Chloride	H12	500 mM Imidazole 7.5 / 250 mM Sodium Chloride

Table S.5. Composition of the RUBIC Additive screen conditions.

Concentration during assay		Concentration during assay	
A1	100 % Ultrapure water	E1	1 mM CHAPS
A2	100 mM Sodium acetate trihydrate	E2	1 mM CHAPSO
A3	100 mM Calcium acetate hydrate	E3	1 mM OG
A4	100 mM Potassium acetate	E4	1 mM DM
A5	100 mM Ammonium acetate	E5	1 mM DDM
A6	100 mM Sodium sulfate	E6	25 mM Monosaccharides mix
A7	100 mM Magnesium sulfate	E7	25 mM D-Glucose
A8	100 mM Potassium sulfate	E8	25 mM Sucrose
A9	100 mM Ammonium sulfate	E9	25 mM Maltose
A10	100 mM Sodium phosphate monobasic	E10	50 mM Carboxylic acids mix
A11	100 mM Sodium phosphate dibasic	E11	50 mM L-Proline
A12	100 mM Potassium phosphate	E12	50 mM Glycine
B1	100 mM Potassium phosphate dibasic	F1	50 mM L-Glutamic acid monosodium salt
B2	100 mM Sodium tartrate dibasic	F2	500 mM L-Glutamic acid monosodium salt
B3	100 mM Sodium citrate tribasic	F3	50 mM L-Arginine
B4	100 mM Sodium malonate dibasic	F4	500 mM L-Arginine
B5	100 mM Sodium nitrate	F5	50 mM L-Glutamic acid /50 mM L-Arginine
B6	100 mM Sodium formate	F6	500 mM L-Glutamic acid /500 mM L-Arginine
B7	100 mM Potassium formate	F7	50 mM Gly-Gly-Gly
B8	100 mM Sodium fluoride	F8	5 mM Oxaloacetic acid
B9	100 mM Potassium fluoride	F9	5 % v/v Dimethyl sulfoxide
B10	100 mM Ammonium fluoride	F10	5 % v/v Ethylene glycol
B11	100 mM Lithium chloride	F11	5 % v/v Glycerol
B12	100 mM Sodium chloride	F12	20 % v/v Glycerol
C1	100 mM Potassium chloride	G1	5 % v/v PEG 400
C2	100 mM Ammonium chloride	G2	5 % w/v PEG 1000
C3	100 mM Sodium iodide	G3	5 % w/v PEG 3350
C4	100 mM Potassium iodide	G4	5 mM DTT
C5	100 mM Sodium bromide	G5	5 mM TCEP
C6	1 mM Magnesium chloride	G6	5 mM Biotin
C7	1 mM Calcium chloride	G7	5 mM Betaine hydrochloride
C8	1 mM Manganese(II) chloride	G8	5 mM Coenzyme A
C9	1 mM Nickel(II) chloride	G9	5 mM Nicotinic acid
C10	1 mM Iron(III) chloride	G10	1 mM Spermidine
C11	1 mM Zinc chloride	G11	1 mM Spermine tetrahydrochloride
C12	1 mM Cobalt(II) chloride hexahydrate	G12	1 mM Sarcosine
D1	5 mM EDTA	H1	20 µM Deoxyribonucleic acid
D2	5 mM EGTA	H2	1 mM ATP/ 1mM Magnesium chloride
D3	0.1 M Urea	H3	1 mM ATPyS/ 1mM Magnesium chloride
D4	0.5 M Urea	H4	1 mM cAMP/ 1mM Magnesium chloride
D5	1 M Urea	H5	1 mM GTP/ 1mM Magnesium chloride
D6	2 M Urea	H6	1 mM GTPyS/ 1mM Magnesium chloride
D7	4 M Urea	H7	1 mM cGMP/ 1mM Magnesium chloride
D8	150 mM Guanidine hydrochloride	H8	1 mM NADH/ 1mM Magnesium chloride
D9	500 mM Guanidine hydrochloride	H9	1 mM NADPH/ 1mM Magnesium chloride
D10	1 mM NDSB 195	H10	5 mM Polyethyleneimine 800
D11	1 mM NDSB 201	H11	200 mM Imidazole
D12	1 mM Fos-Choline-12	H12	400 mM Imidazole

Table S.6. Midpoint temperatures of the protein-unfolding transition for RNR1 in the presence of each RUBIC Buffer and Additive condition. Melting temperature values (T_m) are presented in degrees Celsius (°C). In a few cases, where samples did not produce interpretable curves with a defined denaturation transition, the values were interpreted as denatured and are not included in T_m comparisons.

RUBIC Buffer Screen												
	1	2	3	4	5	6	7	8	9	10	11	12
A	30.11	-	38.70	41.98	41.42	45.47	48.69	33.46	43.10	48.27	48.83	47.01
B	45.06	37.51	46.59	46.87	32.34	49.25	47.71	51.48	52.18	39.47	47.15	36.67
C	46.87	-	-	-	39.89	46.73	48.27	48.13	47.99	46.31	48.13	49.25
D	48.69	47.01	47.43	46.73	45.89	47.43	46.73	47.85	45.89	45.47	46.45	43.94
E	-	-	42.68	41.98	46.87	48.83	48.69	47.57	47.43	46.17	45.47	40.73
F	31.09	-	49.53	-	39.47	48.55	46.31	46.45	36.12	39.47	46.31	41.84
G	48.69	34.58	47.71	47.15	43.10	47.99	41.84	38.77	46.73	44.22	47.43	46.73
H	41.70	29.13	42.26	31.76	42.96	35.56	34.72	39.61	38.63	36.81	41.14	39.61
RUBIC Additive Screen												
	1	2	3	4	5	6	7	8	9	10	11	12
A	46.44	41.97	-	38.61	39.87	42.67	49.38	40.15	41.69	41.69	43.93	39.59
B	48.96	45.88	50.22	42.39	37.77	36.24	45.88	47.84	52.73	47.56	47.28	47.84
C	46.44	48.40	47.98	41.83	47.42	46.16	50.36	50.50	49.10	47.98	51.61	51.48
D	50.78	50.36	50.92	51.34	45.32	48.12	47.28	44.62	46.30	-	48.12	49.10
E	48.40	48.54	38.19	44.62	51.20	-	-	-	49.66	47.72	46.17	48.40
F	40.99	48.26	36.93	46.72	34.00	41.41	40.99	39.73	38.33	37.91	47.70	47.42
G	47.42	47.28	42.25	47.56	46.72	48.26	40.57	41.27	39.73	47.70	50.64	52.17
H	50.9	51.20	50.22	48.54	46.86	49.80	48.12	42.67	45.18	45.32	46.86	47.84

Table S.7. Composition of the RNR1 NanoDSF screen conditions.

NanoDSF Experimental Conditions	
20 mM citrate pH 6.0	20 mM Bis-Tris pH 6.5
20 mM citrate pH 6.0 5 mM MgSO ₄	20 mM Bis-Tris pH 6.5 5 mM MgSO ₄
20 mM citrate pH 6.0 5 mM DTT	20 mM Bis-Tris pH 6.5 5 mM DTT
20 mM citrate pH 6.0 5 mM Biotin	20 mM Bis-Tris pH 6.5 5 mM Biotin
20 mM citrate pH 6.0 5 mM MgSO ₄ 5 mM DTT	20 mM Bis-Tris pH 6.5 5 mM MgSO ₄ 5 mM DTT
20 mM citrate pH 6.0 5 mM MgSO ₄ 5 mM Biotin	20 mM Bis-Tris pH 6.5 5 mM MgSO ₄ 5 mM Biotin
20 mM citrate pH 6.0 5 mM DTT 5 mM Biotin	20 mM Bis-Tris pH 6.5 5 mM DTT 5 mM Biotin

20 mM citrate pH 6.0 5 mM DTT 5 mM Biotin 5 mM MgSO ₄	20 mM Bis-Tris pH 6.5 5 mM DTT 5 mM Biotin 5 mM MgSO ₄
20 mM citrate pH 6.0 250 mM NaCl	20 mM Bis-Tris pH 6,5 250 mM NaCl
20 mM citrate pH 6.0 5 mM MgSO ₄ 250 mM NaCl	20 mM Bis-Tris pH 6,5 5 mM MgSO ₄ 250 mM NaCl
20 mM citrate pH 6.0 5 mM DTT 250 mM NaCl	20 mM Bis-Tris pH 6.5 5 mM DTT 250 mM NaCl
20 mM citrate pH 6.0 5 mM Biotin 250 mM NaCl	20 mM Bis-Tris pH 6.5 5 mM Biotin 250 mM NaCl
20 mM citrate pH 6.0 5 mM MgSO ₄ 5 mM DTT 250 mM NaCl	20 mM Bis-Tris pH 6.5 5 mM MgSO ₄ 5 mM DTT 250 mM NaCl
20 mM citrate pH 6.0 5 mM MgSO ₄ 5 mM Biotin 250 mM NaCl	20 mM Bis-Tris pH 6.5 5 mM MgSO ₄ 5 mM Biotin 250 mM NaCl
20 mM citrate pH 6.0 5 mM DTT 5 mM Biotin 250 mM NaCl	20 mM Bis-Tris pH 6.5 5 mM DTT 5 mM Biotin 250 mM NaCl
20 mM citrate pH 6.0 5 mM DTT 5 mM Biotin 5 mM MgSO ₄ 250 mM NaCl	20 mM Bis-Tris pH 6.5 5 mM DTT 5 mM Biotin 5 mM MgSO ₄ 250 mM NaCl

Table S.8. Composition of the Durham pH screen conditions.

Concentration during assay		Concentration during assay	
A1	100 % Ultrapure water	E1	100 mM Sodium phosphate pH 6.3
A2	100 % Ultrapure water	E2	100 mM Sodium phosphate pH 6.8
A3	4 M Urea	E3	100 mM Sodium phosphate pH 7.3
A4	100 mM Citrate pH 4.1	E4	100 mM PIPES pH 6.3
A5	100 mM Citrate pH 4.6	E5	100 mM PIPES pH 6.8
A6	100 mM Citrate pH 5.1	E6	100 mM PIPES pH 7.3
A7	100 mM Acetic acid pH 4.2	E7	100 mM Imidazole pH 6.6
A8	100 mM Acetic acid pH 4.7	E8	100 mM Imidazole pH 7.1
A9	100 mM Acetic acid pH 5.2	E9	100 mM Imidazole pH 7.6
A10	100 mM Succinic acid pH 4.4	E10	100 mM MOPS pH 6.6
A11	100 mM Succinic acid pH 4.9	E11	100 mM MOPS pH 7.1
A12	100 mM Succinic acid pH 5.4	E12	100 mM MOPS pH 7.6
B1	100 mM DL-Malic acid pH 4.3	F1	100 mM Bis-Tris propane pH 6.6

B2	100 mM DL-Malic acid pH 4.8	F2	100 mM Bis-Tris propane pH 7.1
B3	100 mM DL-Malic acid pH 5.3	F3	100 mM Bis-Tris propane pH 7.6
B4	100 mM L-Tartaric acid pH 4.3	F4	100 mM HEPES pH 7.0
B5	100 mM L-Tartaric acid pH 4.8	F5	100 mM HEPES pH 7.5
B6	100 mM L-Tartaric acid pH 5.3	F6	100 mM HEPES pH 8.0
B7	100 mM Propionic acid pH 4.3	F7	100 mM Tricine pH 7.5
B8	100 mM Propionic acid pH 4.8	F8	100 mM Tricine pH 8.0
B9	100 mM Propionic acid pH 5.3	F9	100 mM Tricine pH 8.5
B10	100 mM Malonic acid pH 5.2	F10	100 mM EPPS pH 7.5
B11	100 mM Malonic acid pH 5.7	F11	100 mM EPPS pH 8.0
B12	100 mM Malonic acid pH 6.2	F12	100 mM EPPS pH 8.5
C1	100 mM Citrate pH 5.5	G1	100 mM Tris pH 7.7
C2	100 mM Citrate pH 6.0	G2	100 mM Tris pH 8.2
C3	100 mM Citrate pH 6.5	G3	100 mM Tris pH 8.7
C4	100 mM Succinic acid pH 5.6	G4	100 mM BICINE pH 7.7
C5	100 mM Succinic acid pH 6.1	G5	100 mM BICINE pH 8.2
C6	100 mM Succinic acid pH 6.6	G6	100 mM BICINE pH 8.7
C7	100 mM MES pH 5.6	G7	100 mM TAPS pH 7.9
C8	100 mM MES pH 6.1	G8	100 mM TAPS pH 8.4
C9	100 mM MES pH 6.6	G9	100 mM TAPS pH 8.9
C10	100 mM Maleic acid pH 5.7	G10	100 mM Bis-Tris propane pH 8.5
C11	100 mM Maleic acid pH 6.2	G11	100 mM Bis-Tris propane pH 9.0
C12	100 mM Maleic acid pH 6.7	G12	100 mM Bis-Tris propane pH 9.5
D1	100 mM Sodium cacodylate pH 5.7	H1	100 mM Boric acid pH 8.6
D2	100 mM Sodium cacodylate pH 6.2	H2	100 mM Boric acid pH 9.1
D3	100 mM Sodium cacodylate pH 6.7	H3	100 mM Boric acid pH 9.6
D4	100 mM ADA pH 6.1	H4	100 mM CHES pH 8.8
D5	100 mM ADA pH 6.6	H5	100 mM CHES pH 9.3
D6	100 mM ADA pH 7.1	H6	100 mM CHES pH 9.8
D7	100 mM Bis-Tris pH 6.1	H7	100 mM Glycine pH 9.2
D8	100 mM Bis-Tris pH 6.6	H8	100 mM Glycine pH 9.7
D9	100 mM Bis-Tris pH 7.1	H9	100 mM Glycine pH 10.2
D10	100 mM ACES pH 6.3	H10	100 mM CAPS pH 9.9
D11	100 mM ACES pH 6.8	H11	100 mM CAPS pH 10.4
D12	100 mM ACES pH 7.3	H12	100 mM CAPS pH 10.9

Table S.9. Composition of the Durham Salt screen conditions.

Concentration during assay		Concentration during assay	
A1	Ultrapure water	E1	0.5 M Lithium chloride
A2	Ultrapure water	E2	0.2 M Lithium chloride
A3	4 M Urea	E3	0.5 M Rubidium chloride
A4	3 M Guanidine hydrochloride	E4	0.2 M Rubidium chloride
A5	1 M Guanidine hydrochloride	E5	0.5 M Cesium chloride
A6	0.8 M Guanidine hydrochloride	E6	0.2 M Cesium chloride
A7	0.6 M Guanidine hydrochloride	E7	0.4 M Sodium fluoride
A8	0.4 M Guanidine hydrochloride	E8	0.1 M Sodium fluoride
A9	0.2 M Guanidine hydrochloride	E9	1.5 M Sodium bromide

A10	0.005 M Guanidine hydrochloride	E10	0.4 M Sodium bromide
A11	0.5 M Sodium citrate tribasic	E11	0.1 M Sodium bromide
A12	0.2 M Sodium citrate tribasic	E12	0.4 M Sodium iodide
B1	1.5 M Sodium malonate dibasic	F1	0.1 M Sodium iodide
B2	1 M Sodium malonate dibasic	F2	0.4 M Magnesium chloride
B3	0.8 M Sodium malonate dibasic	F3	5 mM Magnesium chloride
B4	0.6 M Sodium malonate dibasic	F4	5 mM Calcium chloride dihydrate
B5	0.4 M Sodium malonate dibasic	F5	5 mM Strontium chloride hexahydrate
B6	0.2 M Sodium malonate dibasic	F6	1 mM Zinc chloride
B7	1.5 M Ammonium sulfate	F7	0.1 mM Zinc chloride
B8	1 M Ammonium sulfate	F8	1 mM Nickel(II) chloride hexahydrate
B9	0.8 M Ammonium sulfate	F9	0.1 mM Nickel(II) chloride
B10	0.6 M Ammonium sulfate	F10	5 mM Manganese(II) chloride
B11	0.4 M Ammonium sulfate	F11	0.5 mM Manganese(II) chloride
B12	0.2 M Ammonium sulfate	F12	1 mM Cobalt(II) chloride
C1	1.5 M Sodium chloride	G1	0.1 mM Cobalt(II) chloride
C2	1 M Sodium chloride	G2	1 mM Copper(II) sulfate
C3	0.8 M Sodium chloride	G3	0.1 mM Copper(II) sulfate
C4	0.6 M Sodium chloride	G4	1 mM Cadmium sulfate 8/3-hydrate
C5	0.4 M Sodium chloride	G5	5 mM EDTA pH 7.5
C6	0.2 M Sodium chloride	G6	5 mM EGTA pH 7.5
C7	1.5 M Ammonium chloride	G7	2 mM I3C
C8	1 M Ammonium chloride	G8	2 mM Lanthanum(III) nitrate
C9	0.8 M Ammonium chloride	G9	2 mM Praseodymium(III) chloride
C10	0.6 M Ammonium chloride	G10	2 mM Neodymium(III) chloride
C11	0.4 M Ammonium chloride	G11	2 mM Samarium(III) chloride
C12	0.2 M Ammonium chloride	G12	2 mM Europium(III) chloride
D1	1 M Magnesium sulfate	H1	2 mM Gadolinium(III) chloride
D2	0.8 M Magnesium sulfate	H2	2 mM Dysprosium(III) chloride
D3	0.6 M Magnesium sulfate	H3	2 mM Holmium(III) chloride
D4	0.4 M Magnesium sulfate	H4	2 mM Ytterbium(III) chloride
D5	0.2 M Magnesium sulfate	H5	2 mM Lutetium(III) chloride
D6	1 M Sodium sulfate	H6	5 mM Sodium phosphate dibasic
D7	0.8 M Sodium sulfate	H7	5 mM Sodium orthovanadate
D8	0.6 M Sodium sulfate	H8	5 mM Sodium tungstate
D9	0.4 M Sodium sulfate	H9	5 mM Sodium molybdate
D10	0.2 M Sodium sulfate	H10	5 mM DTT
D11	0.5 M Potassium chloride	H11	5 mM TCEP pH 7.0
D12	0.2 M Potassium chloride	H12	5 mM β -Mercaptoethanol (β -ME)

Table S.10. Midpoint temperatures of the protein-unfolding transition for RNR2_D267N in the presence of each Durham pH and Salt condition. Melting temperature values (T_m) are presented in degrees Celsius (°C). In a few cases, where samples did not produce interpretable curves with a defined denaturation transition, the values were interpreted as denatured and are not included in T_m comparisons.

Durham pH Screen												
	1	2	3	4	5	6	7	8	9	10	11	12
A	46.8 2	44.4 2	- -	45.4 7	53.1 0	59.8 4	45.7 7	46.0 7	48.1 6	46.5 2	48.7 6	50.1 1
B	46.2 2	54.3 0	52.9 5	46.2 2	53.8 5	51.7 6	45.0 2	47.8 6	47.4 1	54.0 0	53.2 5	53.5 5
C	- -	55.9 5	- -	51.4 6	51.4 6	52.6 5	51.7 6	48.3 1	48.7 6	49.2 1	48.1 6	49.5 1
D	49.6 6	- -	51.1 6	51.4 6	52.2 0	52.6 5	49.6 6	50.1 1	51.3 1	46.6 7	46.6 7	48.3 1
E	52.9 5	52.8 0	52.5 0	50.5 6	51.6 1	50.7 1	48.3 1	47.1 1	49.5 1	48.0 1	48.7 6	49.5 1
F	49.0 6	51.4 6	50.4 1	48.4 6	50.2 6	49.6 6	51.4 6	50.4 1	48.1 6	49.8 1	48.6 1	47.2 6
G	47.5 6	48.7 6	47.7 1	50.5 6	49.8 1	48.7 6	49.8 1	48.9 1	46.8 2	51.6 1	47.7 4	44.7 2
H	47.8 6	47.3 7	46.9 7	48.6 1	46.3 7	43.5 7	48.0 1	48.6 1	48.6 1	- -	- -	- -
Durham Salt Screen												
	1	2	3	4	5	6	7	8	9	10	11	12
h	44.5 7	47.5 1	46.8 2	- -	- -	41.1 3	46.9 7	44.1 2	48.0 1	46.9 7	57.5 9	54.6 0
B	63.4 3	61.4 8	60.8 9	59.5 4	57.2 9	55.0 5	54.4 5	53.2 5	53.5 5	53.4 0	49.8 1	49.5 1
C	52.5 0	51.7 6	51.0 1	51.3 1	49.9 6	49.2 1	45.4 7	46.9 7	47.7 1	47.7 1	48.1 6	48.6 1
D	55.3 3	54.0 8	53.3 7	51.1 1	51.3 1	57.7 4	56.6 9	56.1 0	55.0 5	53.5 5	48.6 1	48.0 1
E	51.0 1	49.9 6	50.1 1	49.2 1	48.9 1	49.2 1	61.0 4	60.5 6	45.1 7	46.9 7	48.9 1	55.9 5
F	34.3 9	48.9 1	44.1 2	46.9 7	45.4 7	57.5 9	42.7 7	48.6 1	45.1 7	46.0 7	45.3 1	- -
G	46.0 7	40.6 8	41.8 8	56.1 0	53.1 0	47.8 6	50.2 6	55.0 5	43.5 2	57.2 9	48.0 1	46.5 2
H	53.4 0	54.9 0	42.4 2	48.7 6	49.6 6	48.6 1	47.2 6	47.2 6	45.9 2	45.6 2	47.2 6	45.6 2

Appendix B:

New targets for drug design: Importance of nsp14/nsp10 complex formation for the 3'-5' exoribonucleolytic activity on SARS-CoV-2

Margarida Saramago*^Ψ, Cátia Bárria*, Vanessa Costa*, Caio S. Souza, Sandra C. Viegas, Susana Domingues, Diana Lousa, Cláudio M Soares, Cecília M Arraiano^Ψ, Rute G. Matos^Ψ

Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Apartado 127, 2781-901 Oeiras, Portugal

*The first three authors should be regarded as joint first authors

^Ψ These authors are joint corresponding authors

Abstract:

SARS-CoV-2 virus has triggered a global pandemic with devastating consequences for health-care and social-economic systems. Thus, the understanding of the more fundamental aspects of SARS-CoV-2 is of extreme importance. In this work, we have focused our attention on the viral ribonuclease (RNase) nsp14, since this protein was considered one of the most interferon antagonists from SARS-CoV-2, and affects viral replication. This RNase is a multifunctional protein that harbors two distinct activities, an N-terminal 3'-to-5' exoribonuclease (ExoN) and a C-terminal N7-methyltransferase (N7-MTase), both with critical roles in coronaviruses life cycle. Namely, nsp14 ExoN knockout mutants are non-viable, indicating that nsp14 ExoN has a very important role in SARS-CoV-2 RNA synthesis, showing up as a prominent target for the development of antiviral drugs. Nsp14 ExoN activity is stimulated through the interaction with the nsp10 protein, which has a pleiotropic function during viral replication. In this study, we have performed the first biochemical characterization of the complex nsp14-nsp10 from SARS-CoV-2. Here we confirm the 3'-5' exoribonuclease and MTase activities of nsp14 in this new Coronavirus, and the critical role of nsp10 in upregulating the nsp14 ExoN activity in vitro. Furthermore we demonstrate the nsp14 N7-MTase activity of SARS-CoV-2 and show that in this virus nsp14 N7-MTase activity is functionally independent of the ExoN activity. The nsp14 MTase activity also seems to be independent of the presence of nsp10 cofactor, contrarily to nsp14 ExoN. Until now, there is no available structure for the SARS-CoV-2 nsp14-nsp10 complex. As such, we have modelled the SARS-CoV-2 nsp14-nsp10 complex based on the 3D structure of the

complex from SARS-CoV (PDB ID 5C8S). Mutations of conserved residues on the nsp10 surface are shown to have a strong impact on ExoN activity. We have managed to map key nsp10 residues involved in its interaction with nsp14, all of which are also shown to be essential for stimulation of the ExoN activity by nsp10. This reinforces the idea that a stable interaction between nsp10 and nsp14 is strictly required for the nsp14-mediated ExoN activity of SARS-CoV-2, as observed for SARS-CoV. We have also studied the role of conserved DEDD catalytic residues of SARS-CoV-2 nsp14 ExoN. Our results show that motif I of ExoN domain is essential for the nsp14 function contrasting the functionality of these conserved catalytic residues in SARS-CoV, and MERS. The striking differences here revealed can have important implications regarding the specific pathogenesis of SARS-CoV-2. The nsp10-nsp14 interface is a recognized attractive target for antivirals against SARS-CoV-2 and other coronaviruses. This work has unravelled a basis for discovering inhibitors targeting the specific amino acids here reported, in order to disrupt the assembly of this complex and interfere with coronaviruses replication.