IDENTIFICATION OF SMALLER NONCODING RNAS PRODUCED BY MYCOBACTERIUM TUBERCULOSIS IN INFECTED MACROPHAGES THAT REGULATE MTB GROWTH AND SURVIVAL

APPROVED BY SUPERVISORY COMMITTEE

Dustin Hancks, PhD

Nicolai van Oers, PhD

Sebastian Winter, PhD

Nicholas Conrad, PhD

Identification of smaller noncoding RNAs produced by *Mycobacterium tuberculosis* in infected macrophages that regulate Mtb growth and survival

ACKNOWLEDGEMENTS

I would like to begin by acknowledging my mentor Dr. Nicolai van Oers. He has been a great mentor throughout my grad life. He has been there for me whenever I needed help and that allowed me to build a skillset and confidence especially in my early years. Thanks to him I grew in terms of scientific and critical thinking, ethical conduct of experiments, articulating a thought and conveying them in a simple manner.

Second, I want to acknowledge my committee members Drs. Nicholas Conrad, Sebastian Winter and Dustin Hancks. You have always been very helpful and responsive. I am truly grateful that I had you in my committee. I would also like to acknowledge our collaborators in Europe, Drs. Przemyslaw Plocinski and Jarmila Hnilicova for sharing reagents and plasmids, which sped up my project. Also, thank you Dr. Didem Agac-Cobanoglu for carefully proofreading my thesis.

Next, I'd like to thank all the past and current members of the van Oers lab. They have been a great support not only in the technical terms, but also by providing a warm and friendly environment. I will miss my conversations with Dr. Igor Dozmorov and my tea breaks with Drs. Pratibha Bhalla and Angela Moses. Plus, I'd like to acknowledge all the precious members of the Core Facilities in UTSW including the Flow Core and the Microarray Core.

Finally, I'd like to thank my family for raising me this way, so that I turned out to be the person that I am now and thanks a lot to my friends in Dallas or elsewhere for their support whenever I needed.

Identification of smaller noncoding RNAs produced by *Mycobacterium tuberculosis* in infected macrophages that regulate Mtb growth and survival

by

FATMA SEVDE COSKUN

DISSERTATION

Presented to Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

December, 2021

Identification of smaller noncoding RNAs produced by *Mycobacterium tuberculosis* in infected macrophages that regulate Mtb growth and survival

Publication No:

Fatma Sevde Coskun, PhD.

The University of Texas Southwestern Medical Center at Dallas, 2021

Supervising Professor: Nicolai S.C. van Oers

It is estimated that one-third of the world's population is infected with *Mycobacterium tuberculosis* (Mtb). While much work has focused on the role of different proteins encoded by Mtb in pathogenesis, recent studies have revealed that Mtb also transcribes many noncoding RNAs whose functions remain poorly characterized. A subset includes small RNAs (sRNAs) between the sizes of 50-350 nts. The current study focused on the identification and characterization of miRNA-like sRNAs <50 nts produced by Mtb. A sRNA-centered RNA-sequencing approach was performed and a subset of Mtb-encoded smaller noncoding RNAs (sncRNAs) were identified. Thirty-five distinct Mtb-encoded sncRNAs were discovered, with most being induced in infected eukaryotic cells. Three sncRNAs, sncRNA-1, sncRNA-6, and sncRNA-8, predominated the read counts. They were contained in longer RNA transcripts with stable secondary RNA stem loops and structures like precursor microRNAs. My work established that sncRNA-1 positively regulates two mycobacterial transcripts involved in oleic acid biosynthesis. Loss- and gain- of-function approaches reveal that sncRNA-1 enhances Mtb growth and survival in nutrient-depleted cultures

as well as in infected macrophages. Given evidence that RNA processing enzymes were involved in the formation of the sncRNAs, different components of core RNA degradosome were characterized for their ability to process the precursor forms of the sncRNAs. My work revealed that PNPase degrades sncRNA-8 and preliminary evidence suggests that sncRNA-1 is also likely a target, which could be critical in the oleic acid deficient media. Overall, my study reveals that Mtb produces a set of sncRNAs in infected cells, with one modulating mycobacterial gene expression and mycobacterial pathogenicity coupled to oleic acid biogenesis. Future studies will address the functions of other sncRNAs and focus on the identification of sncRNA processing enzymes.

LIST OF PUBLICATIONS

1. sncRNA-1 Is a Small Noncoding RNA Produced by Mycobacterium tuberculosis in Infected Cells That Positively Regulates Genes Coupled to Oleic Acid Biosynthesis

   Fatma S. Coskun, Shashikant Srivastava, Prithvi Raj, Igor Dozmorov, Serkan Belkaya, Smriti Mehra, Nadia A. Golden, Allison N. Bucsan, Moti L. Chapagain, Edward K. Wakeland, Deepak Kaushal, Tawanda Gumbo and Nicolai S. C. van Oers

   Frontiers in Microbiology, Infectious Diseases, July 2020
   Front Microbiol. 2020; 11: 1631.

2. Small RNAs Asserting Big Roles in Mycobacteria

   Fatma S. Coskun, Przemyslaw Plocinski , Nicolai S.C. van Oers

   Noncoding RNA, October 2021
   Preprints 2021, 2021070044

3. FOXN1 compound heterozygous mutations cause selective thymic hypoplasia in humans

   Qiumei Du, Larry K Huynh, Fatma Coskun, Erika Molina, Matthew A King, Prithvi Raj, Shaheen Khan, Igor Dozmorov, Christine M Seroogy, Christian A Wysocki, Grace T Padron, Tyler R Yates, M Louise Markert, M Teresa de la Morena, Nicolai Sc van Oers

   The Journal of Clinical Investigation, September 2019
   J Clin Invest. 2019 Nov 1;129(11):4724-4738.

4. Manipulation of noncanonical IRE1-dependent MAPK signaling by a Vibrio agonist-antagonist effector pair

   Nicole J. De Nisco, Amanda K. Casey, Mohammed Kanchwala, Alexander E. Lafrance, Fatma S. Coskun, Lisa N. Kinch, Nick V. Grishin, Chao Xing, Kim Orth

   American Society for Microbiology Journals, February 2021
   mSystems. 2021 Feb 9;6(1):e00872-20.

5. The Antibiotic Efflux Protein TolC Is a Highly Evolvable Target under Colicin E1 or TLS Phage Selection

   Yusuf Talha Tamer, Ilona Gaszek, Marinelle Rodrigues, Fatma Sevde Coskun, Michael Farid, Andrew Y. Koh, William Russ, Erdal Toprak

   Molecular Biology and Evolution, June 2021
   Mol Biol Evol. 2021 Sep 27;38(10):4493-4504.

6. Seropositive IgG Responses to Endemic Human Coronaviruses in Children Provides Cross-Protection Against SARS-CoV-2 (in prep)

   Quan-Zhen Li, Fatma Coskun, Prithvi Raj, Patricia Pichilingue-Reto, Chengsong Zhu, Igor Dozmorov, Maurice Turk, Jeffrey S. Kahn, Rebecca S. Gruchalla, Christian A. Wysocki, and Nicolai S.C. van Oers

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

Figure 59. Identification of smaller noncoding RNAs produced by *Mycobacterium tuberculosis* in infected macrophages that regulate Mtb growth and survival is summarized.

LIST OF TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 6

LIST OF ABBREVIATIONS

AM: Alveolar macrophages

AS: antisense

ATc: anhydrous tetracycline

BCG: Bacillus Calmette Guerin

BMDM: Bone marrow-derived macrophages

CFU: Colony forming units

CLR: C-type lectin receptors

CRISPRi: CRISPR interference

Ctl: Control

DrrS: DosR Regulated sRNA

dsRNA: double-stranded RNA

DEGs: Differentially expressed genes

E. coli: *Escherichia coli*

EDR: Extremely drug resistant

FAM: 6-Carboxyfluorescein

FcR: Fc receptors

GEO: Gene Expression Omnibus

Hfq: Host factor for phage Qb

IFNγ: Interferon gamma

IGR: Intergenic region

IL1β: Interleukin 1 beta

IPTG: Isopropyl β-D-1-thiogalactopyranoside

LNA: Locked nucleic acid

LNA-PI: Locked nucleic acid power inhibitor

LTBI: Latent tuberculosis infection

M. avium: *Mycobacterium avium*

M. bovis: *Mycobacterium bovis*

M. smegmatis: *Mycobacterium smegmatis*

MFE: Minimum Free Energy

miRNAs: microRNAs

MOI: Multiplicity of Infection

MrsI: Mycobacterial regulatory sRNA in Iron

Mtb: *Mycobacterium tuberculosis*

MTS: Mycobacterium tuberculosis sRNAs

ncRNAs: noncoding RNAs

NOS: Nitric oxide synthase

nt: nucleotide

nts: nucleotides

OADC: oleic albumin dextrose catalase

PAMP: Pathogen-associated molecular patterns

PTB: Pulmonary TB

RBS: Ribosomal binding site

RD1: Region of difference 1

RhlE: ATP-dependent RNA helicase

RNA-Seq: RNA-sequencing

RNAP: RNA polymerase

RNase III: Ribonuclease III, RNC

RNase E: Ribonuclease E, RNE

RNase J: Ribonuclease J, RNJ

RNI: Reactive nitrogen intermediates

ROI: Reactive oxygen intermediates

PNPase: Polynucleotide phosphorylase, PNP

PRR: Pattern-recognition receptors

sncRNAs: smaller noncoding RNAs

SR: Scavenger receptors

SRG: Stress responsive genes

sRNAs: small RNAs

TB: Tuberculosis

TLR: Toll-like receptors

TNFα: Tumor necrosis factor alpha

UTR: untranslated region

CHAPTER 1

Introduction

1.1 *Mycobacterium tuberculosis* as a pathogen

1.1.1 Tuberculosis

Tuberculosis (TB) is an infectious disease caused by an airborne pathogen called *Mycobacterium tuberculosis* (Mtb). Mtb was first isolated and cultured in 1882 (Tornheim & Dooley, 2019). Despite its long history, Mtb remains a major public health threat with 10.4 million new cases per year (Cambau & Drancourt, 2014). It is estimated that overall, 2.3 billion people are infected with about 2000 people dying daily; thus making Mtb the deadliest infectious agent until the emergence of SARS-Cov-2 in 2019 (Dheda, Barry, & Maartens, 2016; Simonsen & Viboud, 2021). More concerning for public health is that recent studies in South Africa report that a significant percentage of new cases is due to the transmission of extremely drug-resistant (EDR) or multiple drug-resistant (MDR) Mtb, for which a cocktail regimen with a longer treatment period is required for a favorable outcome (Conradie et al., 2020; Shah et al., 2017). Among those who survive TB, the majority are thought to develop a life-long latency. While the establishment of latency requires mycobacteria to adapt and develop a long-term relationship with the host within which it maintains its life, the molecular processes governing Mtb pathogenesis leading to latency remain incompletely understood (Behr, Edelstein, & Ramakrishnan, 2018; Supply et al., 2013). Therefore, more studies on Mtb are needed to eradicate TB.

1.1.2 Mtb pathogenicity

Since the beginning of research on TB in early 1900s, a rough understanding of TB pathogenesis has been developed. Mtb primarily infects the lungs. Once inhaled in air droplets, Mtb reaches to distal airways where it encounters innate immune cells. A successful infection in murine models

is completely dependent on the presence of alveolar macrophages (AM) which is followed by the recruitment of monocyte-derived macrophages and neutrophils (Cohen et al., 2018). Phagocytosis of the bacteria by AM initiates a cascade of cytokine and chemokine production including tumor necrosis factor alpha (TNFα) (Rojas, Olivier, Gros, Barrera, & Garcia, 1999). TNFα with other chemokines recruit CD4+ and CD8+ T cells which surround the infected cells that undergo necrosis (Roca, Whitworth, Redmond, Jones, & Ramakrishnan, 2019). Overall, they form a macrophage-rich structure termed granuloma that restricts bacterial dissemination (Guirado & Schlesinger, 2013). Given the primary role of alveolar macrophages in Mtb phagocytosis and restriction, it is crucial to study the biology of the events occurring inside the macrophages in the early phases of the infection to better understand the pathogenicity of Mtb.

1.1.3 Macrophages as a niche with environmental stress

Macrophages are the innate immune cells that encounter Mtb once they invade the bronchial lining. Pathogen-associated molecular patterns (PAMP) found in Mtb are recognized by pattern recognition receptors (PRR) expressed on as well as in macrophages. These PRRs include Toll-like receptors (TLR), C-type lectin receptors (CLR), Fc receptors (FcR) and scavenger receptors (SR) (Figure 1) (Queval, Brosch, & Simeone, 2017). Recognition of Mtb through such signaling induces phagocytosis of the bacteria followed by the maturation of the phagosome. In the case of infection with other bacteria, this is followed by the recruitment of ATPases to acidify the vacuole to pH:5, whereas in Mtb infection this is disrupted through multiple mechanisms which causes Mtb to reside in a pH:6.2 environment (Guirado, Schlesinger, & Kaplan, 2013). The less acidic environment enables Mtb to survive and divide inside macrophages. Several days post-infection, some mycobacteria escape the phagosome and enter the cytosol. This process requires genes residing in the region of difference 1 (RD1) (Houben et al., 2012). This 9.5 kb region contains six

genes (Rv3871–Rv3876), two of which are *esxA* and *esxB* (Mahairas, Sabo, Hickey, Singh, & Stover, 1996). These two genes are implicated in the pathogenesis of mycobacteria and are lost in the nonvirulent strains (Pym, Brodin, Brosch, Huerre, & Cole, 2002; Yang et al., 2014).

After Mtb uptake, macrophages get activated by proinflammatory cytokines such as TNFα, interleukin 1 beta (IL1β) and/or interferon gamma (IFNγ) (C. H. Wang & Kuo, 2001). This activation induces macrophage expression of nitric oxide synthase (NOS), which converts nitric oxide to reactive nitrogen intermediates (RNI) or reactive oxygen intermediates (ROI) causing oxidative stress (Nathan & Shiloh, 2000) (Figure 1). ROI and RNI can reduce the oxidation state of metals which in turn increases the DNA damage and spontaneous mutagenesis in bacteria (Touati, 2000). Mtb must circumvent such conditions for successful infection and that is achieved by the activation of stress-responsive genes (SRG) including superoxide dismutase, which break down reactive oxygen molecules (Kusunose, Ichihara, Noda, & Kusunose, 1976; Liao, Fan, & Bao, 2013). Another means by which macrophages attempt to restrict Mtb growth is by starving them of certain amino acids such as tryptophan (Russell, 2013). Mtb can upregulate tryptophan production in infected cells to counteract this mechanism (Y. J. Zhang et al., 2013). Supporting this, the loss of tryptophan production in Mtb renders Mtb susceptible to killing inside macrophages (Y. J. Zhang et al., 2013). In addition to Mtb protein coding-genes reported previously (Ehrt & Schnappinger, 2009; Voskuil, Bartek, Visconti, & Schoolnik, 2011), recent studies have implicated Mtb noncoding genes induced upon such stress, most of which remain uncharacterized (K. B. Arnvig et al., 2011; K. B. Arnvig & Young, 2009; Coskun et al., 2020; DiChiara et al., 2010; Pellin, Miotto, Ambrosi, Cirillo, & Di Serio, 2012).

1.2 Noncoding RNAs in eukaryotes and prokaryotes: microRNAs and small RNAs

1.2.1 Noncoding RNAs in humans: miRNAs

It is appreciated that only 1.5% of the human genome encodes for proteins, whereas it has an abundance of noncoding RNAs (ncRNAs) (Esteller, 2011). The ncRNAs include a class of RNAs called microRNAs (miRNAs). miRNAs are 21-25 nucleotide (nts) in length and negatively regulate target mRNAs post-transcriptionally. This activity is mediated by base pairing between the 7-8 nts seed sequence within the miRNA and the complementary sequence usually in the 3' untranslated region (UTR) of the target, or in certain cases in the coding region of the transcript (Agrawal et al., 2003; Bartel, 2009; Guo, Ingolia, Weissman, & Bartel, 2010). miRNAs have been implicated in many cellular processes such as proliferation, differentiation, development as well as stress response with their dysregulation causing various diseases (Ali Syeda, Langden, Munkhzul, Lee, & Song, 2020; Kalayinia, Arjmand, Maleki, Malakootian, & Singh, 2021; Lee & Dutta, 2009; Mendell & Olson, 2012; Petri, Malmevik, Fasching, Akerblom, & Jakobsson, 2014). Biogenesis of miRNAs begins with the processing of a primary miRNA transcript to a precursor miRNA, the latter characterized by the presence of a 60-120-nt hairpin structure (O'Brien, Hayder, Zayed, & Peng, 2018) (Figure 2). This is mediated by an enzyme called Drosha. The hairpin is further cleaved by an enzyme named Dicer yielding the 22-nt mature miRNA. Drosha and Dicer belong to the RNase III class of enzymes that have a double-stranded RNA binding domain and a catalytic domain (Court et al., 2013). After its maturation, the miRNA is bound by Argonaute proteins to form the RNA-induced silencing complex (RISC), which can specifically cleave the target mRNA (Agrawal et al., 2003). The specificity of the target is determined by the base pairing between the 8-nt seed sequence of the miRNA starting 2-7 nts of the 5' end of the miRNA and its complementary sequence in the 3' UTR of the target mRNA (Doench & Sharp, 2004; Kehl et al., 2017). A perfect base pairing between the miRNA and its target leads to the degradation of the

target, whereas an imperfect match suppresses the translation. This is a tightly controlled mechanism that requires multiple proteins acting together with RNAs to regulate gene expression.

1.2.2 Noncoding RNAs in bacteria: small RNAs

Like eukaryotes, prokaryotes including bacteria also produce noncoding RNAs. Among these are diverse numbers of small RNAs (sRNAs) between 50-300 nts. Our knowledge on sRNAs is derived mostly from studies on Gram-negative species such as *Escherichia coli* (E. coli) and *Salmonella enterica* and the Gram-positive strain *Staphylococcus aureus* (Barnhill et al., 2019; Mika & Hengge, 2013; Pichon & Felden, 2005). The structure and function of the bacterial sRNAs remain ill-defined as they do not show characteristics that are conserved across multiple species. Broadly, sRNAs are trans-acting regulatory elements which interact with target mRNA or DNA to regulate gene expression (Figure 3) (Carrier, Lalaouna, & Masse, 2018). The sRNAs appear to be involved in rapid adaptation to stress as they can respond to environmental changes by skipping translation steps into proteins or micropeptides.

The bacterial sRNAs are heterogenous in size, ranging between 50-300 nts. As most have been identified with conventional RNA-Sequencing (RNA-seq) procedures, those sRNAs less than 50 nts in length are often not characterized. The genomic location of small RNAs also varies, which again complicates the identification and characterization of such RNAs. The various sRNAs can be transcribed from operons oriented in the antisense direction to a coding gene or from intergenic regions (Gottesman & Storz, 2011). The RNA transcribed from the antisense strand tend to act on the sense RNA by base pairing with a perfect complementarity (Carrier et al., 2018). The RNAs transcribed from intergenic regions have broader specificity than the antisense RNAs as they tend to interact with different targets through the 7-12 nts seed sequence with imperfect complementarity (Carrier et al., 2018).

sRNA can positively or negatively affect transcription and/or translation using different mechanisms such as altering the RNA conformation of the target transcript, modulating protein binding to the RNA and/or interacting with the DNA (Carrier et al., 2018; Chen, Morita, & Gottesman, 2019; Gimpel & Brantl, 2017; Nejman-Falenczyk et al., 2015; Zapf et al., 2019) (Figure 3). For sRNA-mRNA interactions, downregulation is mediated by blocking ribosomal binding and/or processivity. In some bacteria, this involves the chaperone host factor for phage Qb (Hfq) (Kajitani & Ishihama, 1991; Kajitani, Kato, Wada, Inokuchi, & Ishihama, 1994; Tsui, Leung, & Winkler, 1994; Vogel & Luisi, 2011). In E. coli, for example, Hfq recruits RNase E to the site of the sRNA/mRNA complex, cleaving the target mRNA (Afonyushkin, Vecerek, Moll, Blasi, & Kaberdin, 2005; Morita, Maki, & Aiba, 2005). The upregulation of a transcript by a sRNA occurs by stabilization of the target mRNA, preventing transcription degradation. The mechanistic requirements for sRNA-DNAs interactions remain poorly understood.

1.3 Mycobacterial small RNAs

Like most bacterial species, noncoding RNAs are also produced by diverse mycobacterial species, including pathogenic strains such as Mtb. About 11% of the Mtb transcriptome may comprise non-coding RNAs that are designated as sRNAs (K. Arnvig & Young, 2012; K. B. Arnvig et al., 2011; Coskun et al., 2020). Interestingly, the mycobacterial strains lack an obvious Hfq homologue, which suggests the existence of either distinct or divergent protein-sRNA processing systems (Jousselin, Metzinger, & Felden, 2009).

Initial RNA screening approaches to find low molecular weight RNAs in Mtb identified 4 antisense (ASdes, ASpks, AS1726, and AS1890) and 5 intergenic (B11, B55, C8, F6, and G2) sRNAs between 38-238 nts (Table 1) (K. B. Arnvig & Young, 2009). Among these, ASdes, ASpks, AS1726, and AS1890 were also identified in Bacillus Calmette Guerin (BCG) mycobacterial

cultures (DiChiara et al., 2010). Moreover, ASdes was detected in the plasma taken from 15 of 27 TB patients and 6 of 24 healthy individuals vaccinated with BCG (Fu et al., 2018). Following these early studies, RNA-seq based approaches have been employed to detect Mtb-encoded sRNAs, which revealed 20 more intergenic sRNAs (Table 1) (K. B. Arnvig et al., 2011). Three of these named as MTS2823, MTS0997 and MRS1338 have been further characterized, MTS2823 being the most abundant sRNA in log phase Mtb cultures (K. B. Arnvig et al., 2011; Ignatov et al., 2015). In addition 37 sRNAs have been identified in *Mycobacterium bovis* (M. bovis) (Table 1) (DiChiara et al., 2010). Among these, the expression of 19 have been verified with northern blotting including Mcr6, Mcr11, and Mcr14 which were the same as the C8, MTS0997, and F6 sRNAs, respectively. Later studies confirmed the presence of Mcr3, Mcr4, Mcr8, Mcr9, Mcr11, and Mcr14 in M. bovis and B11, B55, C8, F6, G2, and ASdes in Mtb (Pellin et al., 2012). The most recent study on Mtb-encoded sRNAs characterized those induced in more extreme stress conditions such as oxidative stress, pH stress, membrane stress, or nutrient or iron starvation conditions (Table 1) (Gerrick et al., 2018). To summarize, multiple studies have identified >189 sRNAs produced by Mtb in various conditions, which are all named according to the new nomenclature used for Mtb-encoded sRNAs (Lamichhane, Arnvig, & McDonough, 2013).

1.4 Functions of mycobacterial small RNAs

Some of the sRNAs identified in Mtb have been functionally characterized. Among these are B11/6C, MTS1338/DrrS, MTS2823/Ms1, MTS0097/Mcr11, ncRv11846/MrsI, the predicted structures of which are shown in Figure 4. One approach to study the function of a sRNA has been the use of mycobacterial overexpression vectors. Overexpression of B11 (93 nts) regulates genes coupled to DNA replication and protein secretion such as *panD* and *dnaB* (Figure 5A). It forms 2 stem loops through 6 cytidines, hence nicknamed as 6C, which is also conserved in other bacteria

(Mai et al., 2019; Weinberg et al., 2007). Overexpression of MTS1338 (117 nts) regulates genes related to Mtb growth (Figure 5B) (Ignatov et al., 2015). It was named as DosR Regulated sRNA (DrrS) as it was induced by a master regulator DosR in Mtb. MTS0997 (131 nts) was shown to upregulate genes involved in fatty acid metabolism including Rv3282, *fadA3*, and *lipB* (Figure 5C) (Girardin & McDonough, 2020). ncRv11846 (106 nts), which is an ortholog of *E. coli* sRNA RhyB, is also involved in Mtb metabolism (Gottesman, 2004). Nicknamed as Mycobacterial regulatory sRNA in Iron (MrsI), it is induced upon iron starvation to suppress non-essential iron containing genes such as *hypF* and *bfrA*, (Figure 5D) (Gerrick et al., 2018). MTS2823 (300 nts), which has a homology to 6S sRNA in other bacteria, has been further characterized in *Mycobacterium smegmatis* (*M. smegmatis*) (Hnilicova et al., 2014). In *E. coli*, 6S sRNA mimics an open promoter and reduces the activity of RNA polymerase (Cavanagh & Wassarman, 2014). Studies have shown that MTS2823 sequesters RNAP in mycobacteria (Figure 5E) (Sikova et al., 2019). Overall, these studies have implicated critical roles for sRNAs in Mtb.

1.5 miRNA-like small RNAs in mycobacteria

Most of the RNA screens with Mtb focused on RNAs >50 nts. However, RNAs <50 nts are known in other bacteria, including *Salmonella* (Dang et al., 2019; Gu et al., 2017; Kang, Choi, Lee, Hong, & Lee, 2013; Shmaryahu, Carrasco, & Valenzuela, 2014). Sal-1 is an sRNA generated from the 5' end of the ribosomal RNA of *Salmonella* identified in infected epithelial cells (Gu et al., 2017). Produced using the eukaryotic miRNA processing enzymes, it regulates inducible nitric oxide synthase (iNOS). Therefore, in cells infected with *Salmonella* lacking Sal-1, there was reduced bacterial load (Gu et al., 2017). Sal-1 is 22 nts and resembles a eukaryotic miRNA in terms of its structure and being produced from a precursor that contains a hairpin. One such miRNA-like RNA has been identified in mycobacteria to date. This small 23-nt RNA was discovered in

Mycobacterium marinum (Furuse et al., 2014). Interacting with the Argonaute protein in eukaryotes, this RNA transcript also resembles a miRNA. However, its targets have not been identified (Furuse et al., 2014). A more recent study has identified 6 miRNA-like RNAs produced by Mtb in infected patients (Chakrabarty et al., 2019). They were all 22 nts, which was consistent with the size of eukaryotic miRNAs, but they had extremely high GC content (86%-100%) which was unusual for eukaryotic miRNAs that tend to have a GC content closer to 50% (B. Wang, 2013).

1.6 Summary of aims

The purpose of this study is to expand our knowledge on miRNA-like sRNAs produced by Mtb to reveal their potential roles in pathogenicity (Figure 6). sRNAs produced by Mtb in infected cells were comprehensively analyzed. Among these sRNAs, miRNA-like RNAs later named as smaller noncoding RNAs (sncRNAs) were identified with a screen described in Chapter 2. This analysis revealed 35 Mtb-encoded sncRNAs which are induced in infected cells. Functional studies utilizing mycobacterial overexpression vectors followed by transcriptome analyses have focused on three of these; named sncRNA-1, sncRNA-6, and sncRNA-8 described in Chapter 3. In-depth functional analysis of sncRNA-1 is discussed in Chapter 4 where its role in oleic acid biogenesis was defined using gain- and loss-of-function approaches. Overall, sncRNA-1 positively regulates genes involved in oleic acid production. In Chapter 5, putative enzymes that could be coupled to the processing of precursor sncRNAs into the mature sncRNA species were screened for those that could regulate sncRNA-8 and sncRNA-1. In this manner, one enzyme called PNPase, was identified as a sncRNA degradation enzyme. This would suggest that the role of PNPase could be critical in certain environments such as in oleic acid deficient media where the expression of sncRNA-1 is critical. Previous studies have linked PNPase activity to bacterial metabolism in

bacteria as its activity is suppressed by metabolic intermediates in *E. coli* (Cameron, Matz, & De Lay, 2018; Del Favero et al., 2008; Nurmohamed et al., 2011), but the interplay between sncRNA-1 and PNPase in oleic acid-deficient media remains unexplored in Mtb. Given the role of sncRNA-1 in oleic acid metabolism, elucidating the role of PNPase in Mtb could be critical to determine its potential as a drug target. Plus, the enzymes involved in sncRNA processing remain unidentified, which could also be targeted for treatment of TB. Future studies will focus on the identification of an enzyme or enzymes that mediate the processing of the precursor of sncRNAs to their mature forms.

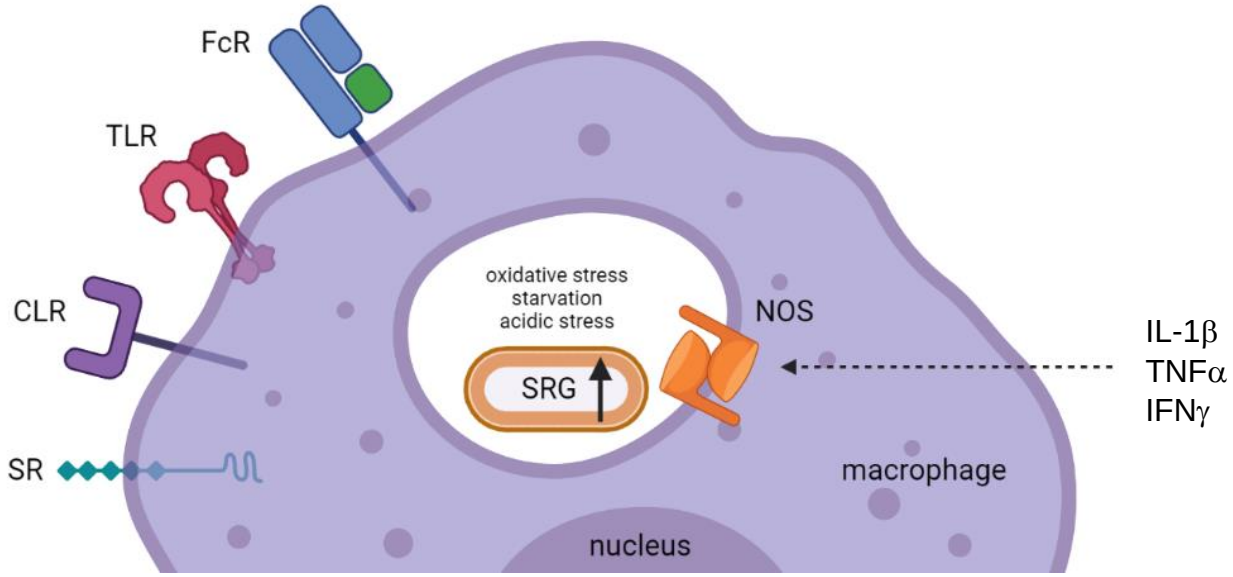


Figure 1. Uptake of Mtb by macrophages induces a cascade of stress-related responses both in macrophages and in Mtb. Pattern-recognition receptors including Toll-like receptors (TLR), C-type lectin receptors (CLR), Fc receptors (FcR) and scavenger receptors (SR) are sites of entry for Mtb. TLR2-mediated uptake induces the pro-inflammatory cytokines through the nuclear factor kappa beta pathway. MARCO, a type of SR and FcR gamma cooperates with TLR2. On the other hand, uptake with the mannose receptor which is a type of CLR induces anti-inflammatory cytokines. This complex signaling cascade results in the maturation of phagosomes and activation of macrophages by proinflammatory cytokines such as TNFα, interleukin 1 beta (IL1β) and interferon gamma (IFNγ). An enzyme in macrophages called nitric oxide synthase (NOS) which converts nitric oxide to reactive nitrogen intermediates (RNI) or reactive oxygen intermediates (ROI) causing an oxidative stress gets induced. In response, Mtb upregulates stress responsive genes (SRG) to survive inside the phagosomes.

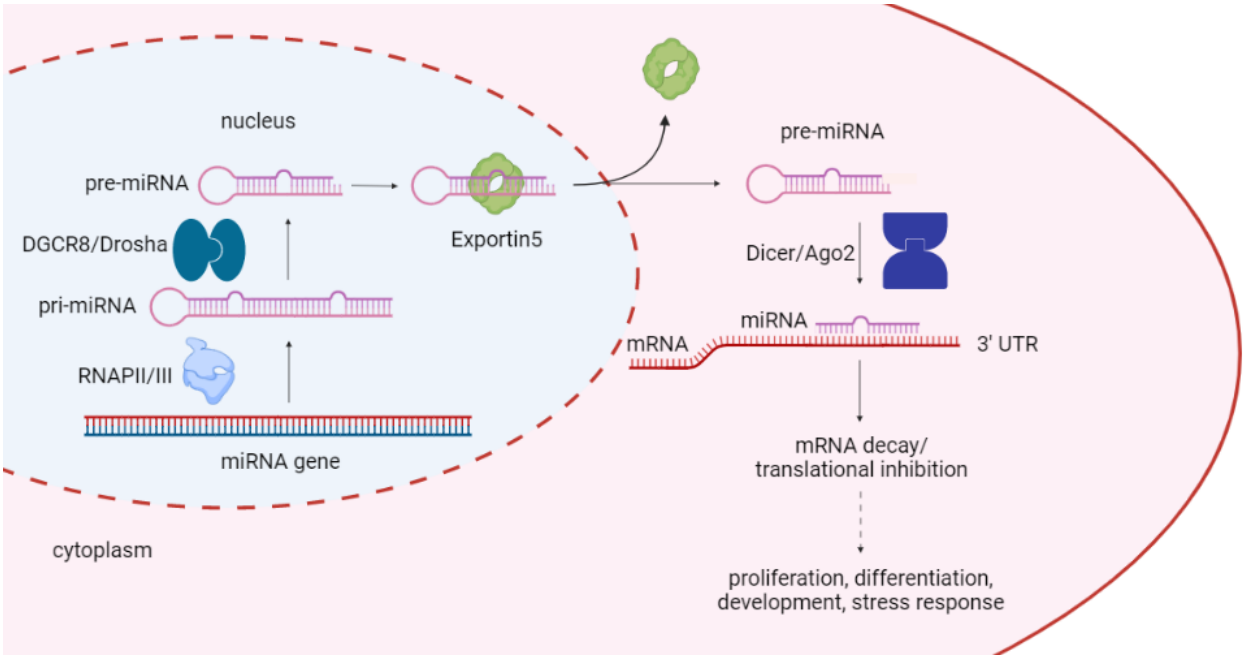


Figure 2. miRNA biogenesis occurs in multiple steps in distinct cellular compartments. miRNAs are produced by RNAP II/III as the primary miRNA (pri-miRNA). pri-miRNA is processed by DGCR8/Drosha complex to produce precursor miRNA (pre-miRNA) in the nucleus. Pre-miRNA is transporter out of the nucleus by Exportin5. It is then handed over to Dicer/Ago2 complex which mediates the interaction between miRNA and its target. This interaction leads to the decay or translational inhibition of the target mRNA, which affects many processes including proliferation, differentiation, development, and stress response.

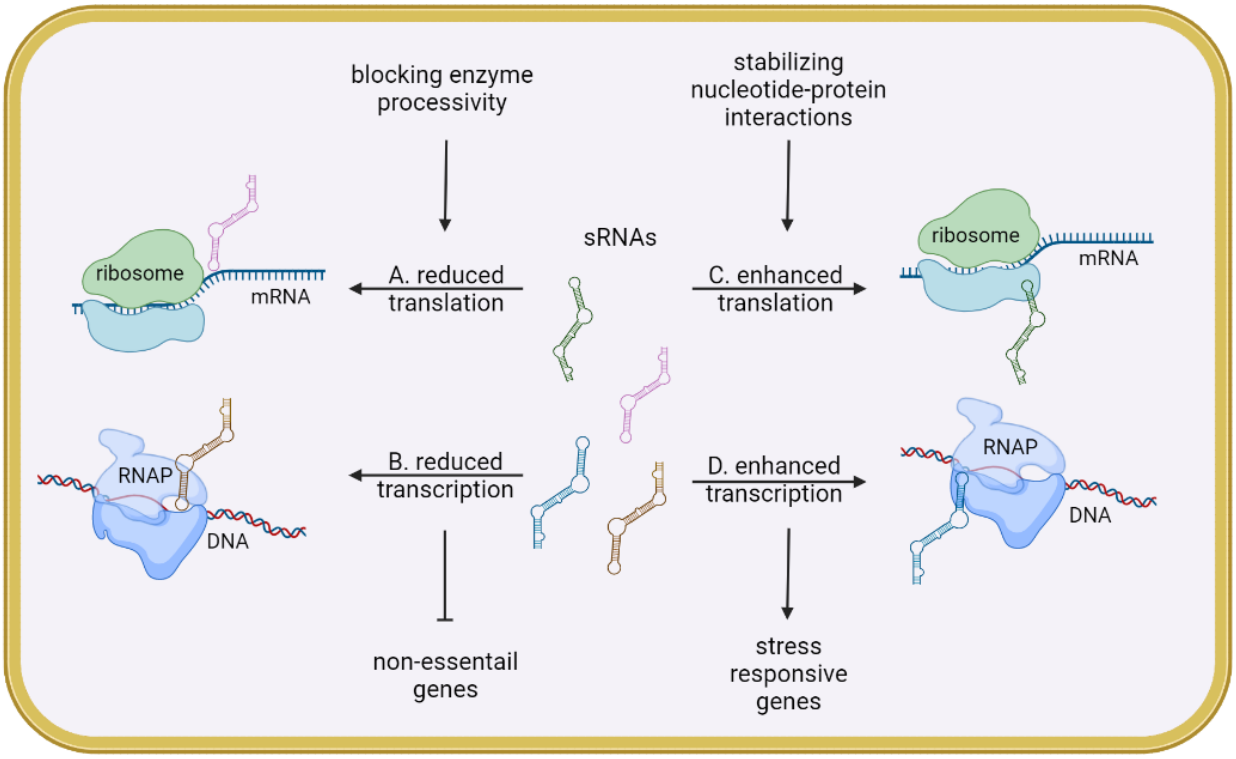


Figure 3. sRNAs can interact with RNA/DNA targets through multiple mechanisms. sRNA mediated gene regulation can be at the translational or transcriptional level to regulate the gene expression negatively or positively. mRNA is shown as an ssRNA bound to the ribosome and DNA is shown during RNA polymerization. An sRNA can reduce the (A) translation or (B) transcription of unessential genes to downregulate them by blocking enzyme processivity. Alternatively, an sRNA can enhance the (C) translation or (D) transcription of stress responsive genes to upregulate them by stabilizing nucleotide-protein complexes.

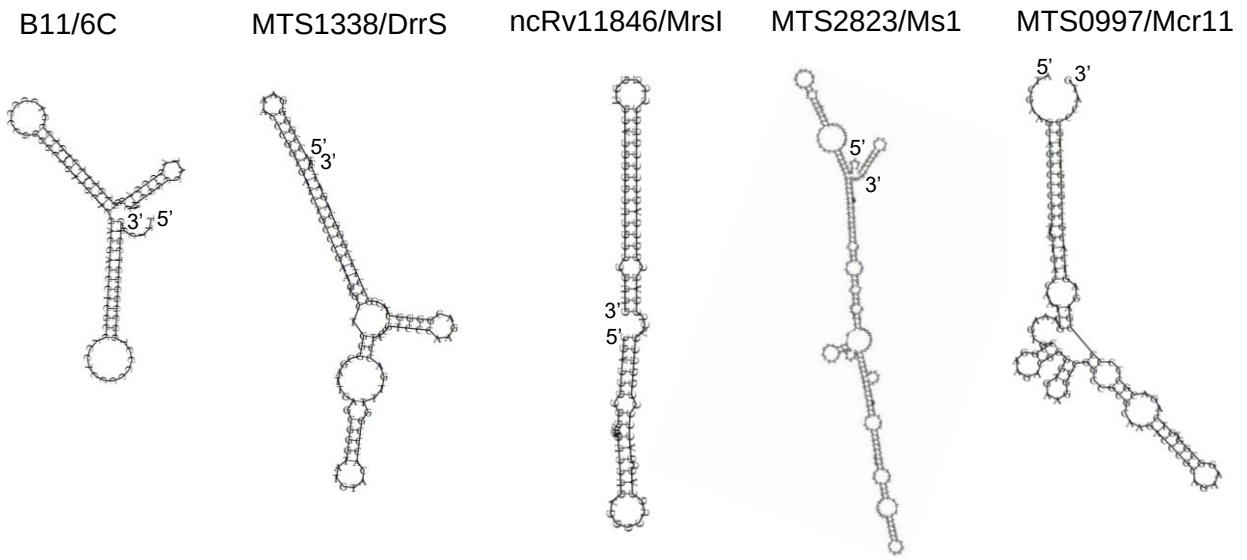


Figure 4. Characterized mycobacterial small RNAs including B11/6C, MTS1338/DrrS, ncRv11846/MrsI, MTS2823/Ms1 and MTS0997/Mcr11. The secondary structures were predicted using RNAfold online tool and sRNA sequences on Mycobrowser available at (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) and (https://mycobrowser.epfl.ch/), respectively.

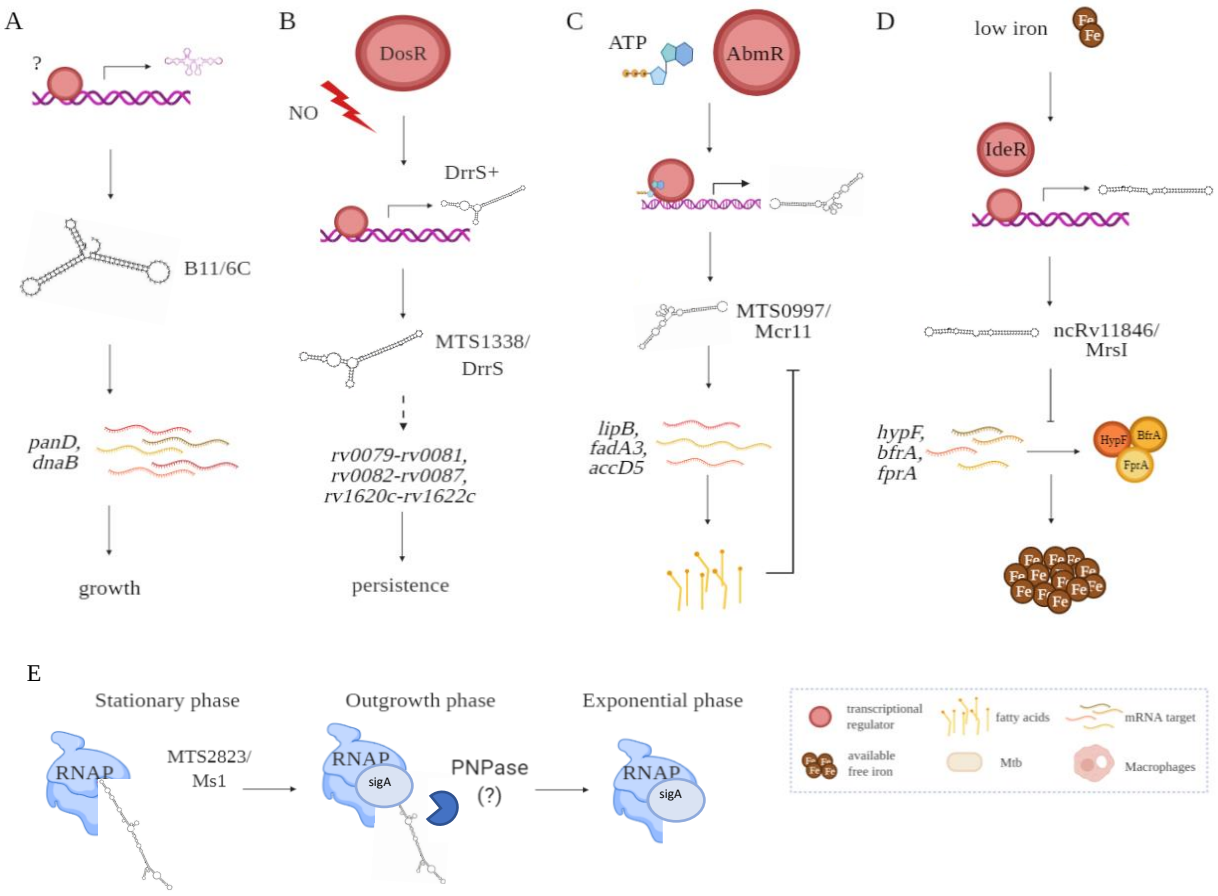


Figure 5. Characterized Mtb small RNAs including B11/6C, MTS1338/DrrS, MTS0997/Mcr11, ncRv11846/MrsI, and MTS2823/Ms1 have various functions in mycobacteria. (A) B11/6C induces genes couples to replication. (B) MTS1338/DrrS is produced by DosR transcription factor, as a larger transcript called DrrS+, which gets processed to yield MTS1338/DrrS. It then regulates genes involved in growth. (C) MTS0997/Mcr11 is induced by AmbR and regulates the fatty acid metabolism in Mtb. (D) ncRv11846/MrsI is induced upon iron starvation and suppress the translation of iron storing proteins. (E) MTS2823/Ms1 sequesters RNA Polymerase in the stationary phase and gets degraded upon entry to the outgrowth phase, which releases RNA Polymerase.

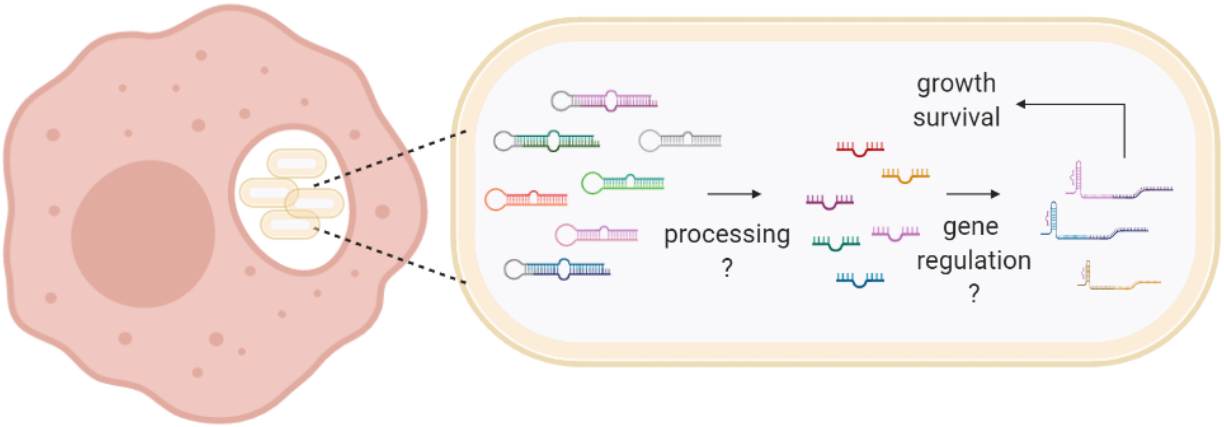


Figure 6. Summary of the aims. Mtb encodes miRNA-like small RNAs, or smaller noncoding RNAs (sncRNAs) in infected macrophages. These are produced as precursor sncRNAs (pre-sncRNAs) which are larger transcripts with predicted hairpin loop structure. Pre-sncRNAs undergo processing to yield sncRNAs. sncRNAs are involved in gene regulation in Mtb to support mycobacterial growth and survival. This thesis identifies sncRNAs produced by Mtb in infected cells and their role in pathogenesis. sncRNA identification is detailed in Chapter 2. The functional assessment of sncRNA-1, sncRNA-6 and sncRNA-8 is presented in Chapter 3. In-depth functional characterization of sncRNA-1 is presented in Chapter 4 and data on sncRNA processing is shown in Chapter 5.

Table 1. List of Mtb sRNA identified by Arnvig *et al*., DiChiara *et al*., and Gerrick *et al*.

Name	Northern or PCR size	Location	Surrounding genes	Expression
B11/6C (candidate_1603)	93	4099386-4099478 (-)	Rv3660c-Rv3661	H_2O_2 and pH=5
B55 (candidate_84)	61	704187-704247(+)	Rv0609A-Rv0610c	H_2O_2 and Mitomycin C
C8 (Mcr6, candidate_1621)	58,70,128	4168154-4168281(-)	Rv3722c-Rv3723	Stationary phase cultures
F6 (Mcr14, candidate_29)	38,58,102	293604-293705(+)	fadA2-fadE5	H_2O_2 and pH=5
G2 (candidate_1269)	67, 214, 229	1914962-1915190(-)	tyrS-IprJ	Stationary phase cultures
ASdes (candidate_121)	48, 63, 68, 83, 94, 109, 149, 169, 195	918264-918458(+)	within desA1	Stationary phase cultures
ASpks	78, 89, 91, 102, 129, 142, 162	2299745-2299886(+)	within pks12	H202
AS1726	61, 77, 85, 110, 213	1952291-1952503(-)	within Rv1726	Stationary phase cultures
AS1890	63, 109, 191, 238	2139419-2139656(+)	within Rv1890	Stationary phase cultures
MTS2823/Ms1	250,300	4100669-4100968(+)	Rv3661-Rv3662c	*in vivo*

MTS1338/DrrS	100	1960667-1960783(+)	Rv1733c-Rv1735c	NO, stationary phase, *in vivo*
MTS0997/Mcr11 (candidate_1693)	115	1413094-1413224(-)	Rv1264-Rv1265	*in vivo*, stationary phase, low pH or hypoxia
Mcr1	>300	2029043-2029087(TBD)	PPE26-PPE27	Stationary phase cultures
Mcr2	120	1108857-1108824(TBD)	Rv0967-Rv0968	Stationary phase cultures
Mcr3 (candidate_190)	118	1471619-1471742(+)	murA-rrs	Stationary phase cultures
Mcr4 (candidate_1314)	200-250	2137148-2137103(TBD)	fbpB-Rv1887	Stationary phase cultures
Mcr5	80	2437823-2437866(-)	within Rv2175c	Stationary phase cultures
Mcr7	350-400	2692172-2692521(+)	Rv2395-PE_PGRS41	Stationary phase cultures
Mcr8 (candidate_1935)	200	4073966-4073908(TBD)	Rv3661-3662c	Stationary phase cultures
Mcr9 (candidate_1502)	66-82	3317634-3317517(TBD)	ilvB1-cfp6	Stationary phase cultures
Mcr10	120	1283693-1283815(+)	within Rv1157c	Stationary phase cultures

Mcr12	118	1228436-1228381(TBD)	Rv1072-Rv1073	Stationary phase cultures
Mcr13	311	4315154-4315215(TBD)	Rv3866-Rv3867	Stationary phase cultures
Mcr15	>300	1535417-1535716(-)	Rv1363c-Rv1364c	Stationary phase cultures
Mcr16	100	2517032-2517134(-)	within fabD	Stationary phase cultures
Mcr17	82-90	2905457-2905402(TBD)	within Rv2613c	Stationary phase cultures
Mcr18	82	3466287-3466332(TBD)	within nuoC	Stationary phase cultures
Mcr19	66-82	575033-575069(+)	within Rv0485	Stationary phase cultures
ncRv11846/MrsI	100	2096766-2096867(+)	blal-Rv1847	iron starvation, oxidative stress, and membrane stress

CHAPTER 2

Discovery of Smaller Noncoding RNA in *Mycobacterium tuberculosis*

2.1 Introduction

The causative agent of Tuberculosis, *Mycobacterium tuberculosis* (Mtb) has a complex genome comprising 4.4 million base pairs long with a 65% GC content, which is considered above average, and 4000 open reading frames for which functional annotation remains limited (Cole et al., 1998). Moreover, recent screen-based studies have revealed that Mtb transcriptome has an abundance of noncoding RNAs which are known as small RNAs (sRNAs) (K. B. Arnvig et al., 2011; K. B. Arnvig & Young, 2009; Coskun et al., 2020; DiChiara et al., 2010; Pellin et al., 2012). sRNAs are post transcriptional regulators of the gene expression in bacteria, which can act in a positive or negative manner. Certain sRNAs are differentially expressed upon entry into the stationary phase or in response to environmental stress (K. B. Arnvig et al., 2011; Coskun et al., 2020; Gerrick et al., 2018). These early observations have been intriguing as they implied that sRNAs may be functionally needed for Mtb survival under such harsh environmental conditions.

Despite these promising findings, Mtb-encoded sRNAs lack consensus features. Structurally, most remain uncharacterized. They have a wide range of sizes between 50-350 nts. However, smaller RNAs <50 nts do exist in other bacteria (Gu et al., 2017; Zhao et al., 2017). For instance, *Salmonell*a expresses such a smaller RNA called Sal-1, which is processed within the host and makes *Salmonella* more susceptible to killing by the host when it is knocked out (Gu et al., 2017). Processing of Sal-1 exploits host's enzymes involved in the processing of microRNAs (miRNAs). miRNAs are between 20-22 nts, but unlike bacterial sRNAs they primarily regulate their target in a negative manner (O'Brien et al., 2018). They are processed from a larger precursor transcript with the help of an enzymatic machinery which enables their interaction with the target. The

specificity of their target is determined between the seed sequence of a miRNA and the 3' UTR of its target mRNA. Some studies have attempted to identify miRNA like RNAs in Mtb (Chakrabarty et al., 2019; Furuse et al., 2014). Like Sal-1, one study identified an Mtb-encoded miRNA like RNA which exploits the host's miRNA processing machinery for its maturation (Furuse et al., 2014). The longer transcript comprising this RNA has characteristic features of a eukaryotic precursor miRNA species, including a stem-loop. However, the function and targets of the Mtb-encoded miRNA remain unknown. A more recent study reported on six Mtb-encoded miRNA-like sRNA isolated from the sera of patients infected with Mtb. Their functions also remain unknown (Chakrabarty et al., 2019).

In this chapter, RNA-sequencing (RNA-seq) was used to identify RNAs that are miRNA like based on size and structure criteria produced by Mtb in infected macrophages. To distinguish these from eukaryotic miRNA, they were named as smaller noncoding RNAs (sncRNAs). A data analysis was performed with a focus on RNAs <30 nts with surrounding sequences that potentially formed hairpin loop structures. Thirty-five such sncRNAs were identified. Their expression increased over the course of 6-day infection period. Three of them with the highest expression level and suggested structural stability - sncRNA-1, sncRNA-6 and sncRNA-8 - were analyzed in more detail including expression validation using a miRNA-specific quantitative real time polymerase chain reaction (q-RT-PCR). Moreover, sncRNA-8 which had the highest expression level in infected cells was detected in the lungs of *Rhesus macaques* infected with Mtb.

2.2 Materials and Methods

2.2.1 Bacterial strains, cell cultures and their growth conditions

Mtb H37Rv (ATCC #27294) was grown in 7H9 media supplemented with 10% oleic albumin dextrose catalase (OADC) (Remel, Thermo-Fisher) and 5% glycerol (G9012, Sigma- Aldrich).

The auxotroph mutant of Mtb H37Rv 6206 (Δ*panCD*, Δ*leuC*) was kindly provided by Dr. William Jacobs (Albert Einstein College of Medicine) (Sampson et al., 2004). Mtb H37Rv 6206 was grown in the 7H9 supplemented with 0.05% tyloxapol (T0307, Sigma-Aldrich), 0.2% casamino acids (223050, BD Bacto™) and 24 µg/ml pantothenate (D-Pantothenic acid hemicalcium salt, Sigma-Aldrich) and 80 ug/ml leucine (Sigma-Aldrich). For the screening of smaller noncoding RNAs, the human-derived THP-1 monocyte/macrophage cell line, obtained from American Type Culture Collection (ATCC TIB-202^{TM}), was used. The THP-1 cells were cultured in RPMI 1640 medium supplemented with 20% FBS (Hyclone or Atlanta Biologicals). Prior to infections, the cells were stimulated with PMA (final concentration 10^{-9} M) for 24 hrs. to establish adherence to the plates. Bone marrow-derived macrophages (BMDM) were obtained from the femur, tibia, and spinal column of the C57BL/6 mice (Hu et al., 2015). Macrophages were differentiated using macrophage media containing 30% of L929 supernatant (ATCC CCL-1^{TM}) mixed with 70% of RPMI containing 10% FCS (Weischenfeldt & Porse, 2008). Briefly, red blood cells were lysed, and the remaining cells were cultured in macrophage media overnight. Non-adherent cells were collected and cultured for four days in macrophage media. On day 4, cells were detached from the culture plates and stained with an antibody detecting CD11b (Catalog # ab24874, Abcam).

2.2.2 Primate studies and ethical considerations

Rhesus macaques (*Macaca mulatta*) of Indian lineage were employed in this study. The animals were 4-8 years old and weighed 5-10 kg. Following infection, the animals were housed in a BSLIII setting within the Regional Biosafety Laboratory at the Tulane National Primate Research Center and euthanized around 12-weeks post infection. All macaques used in this study were individually housed in the ABSLIII facility at the TNPRC and maintained in accordance with the "Guide for the Care and Use of Laboratory Animals" of the Institute of Laboratory Resources, National

Research Council. Animal care facilities were accredited by the American Association for Accreditation of Laboratory Animal Care and licensed by the U. S. Department of Agriculture. All clinical procedures were carried out under the direction of a veterinarian. Necropsy was performed in a dedicated BSL3 necropsy laboratory. All surgery, fluid collections, and clinical exams was performed while primates were under sedation with Ketamine (10 mg/kg). The standard method of euthanasia for NHPs which is anesthesia with Ketamine hydrochloride (10mg/kg) followed by an overdose with sodium pentobarbital was used.

2.2.3 RNA analysis of Primate lung tissues

Macaque primates (*Macaca mulatta*) were infected with 10 colony forming units (CFU) of the Mtb strain CDC1551 (primate GP50 and JD72) or with 500 CFU of the Erdman strain (primate KN59 and KC64) via a head-only aerosol method. Latency was defined as tuberculous skin test conversion, with no accompanying signs of disease or any disease on serial chest X ray, and normal laboratory tests such as blood C-reactive protein, up to 21 weeks after infection. Lung biopsies were isolated, and small RNAs were extracted with the Qiagen miRNeasy kit following an RNA enrichment protocol provided by the manufacturers (Qiagen, Valencia, CA). sncRNA-8 specific primers were used for miRNA-specific q-RT-PCR assay (Exiqon Inc., now part of Qiagen Inc.).

2.2.4 RNA isolation, library preparation and RNA-sequencing

PMA activated THP-1 cells, either uninfected or 3- and 6- days post-infection with Mtb H37Rv (MOI=10:1) were lysed in a Qiazol lysis buffer, followed by bead beating with glass beads. Total RNA was extracted with the Qiagen miRNeasy kit following the manufacturers' instructions (Qiagen, Valencia, CA). Samples were first treated with DNase (Ambion-Thermo-Fisher). Where required, the RNA was concentrated with an RNA Pure and Concentrate kit (Zymo Research, CA). RNA quality and quantity were determined using a NanoDrop 2000/2000c spectrophotometer

and/or Bioanalyzer (Thermo Fisher Scientific Inc.). Five µg of total RNA was processed using a small RNA discovery platform using an Illumina high throughput sequencer (LC Sciences, Houston, TX). A total of 9 samples were sequenced; 3 uninfected cultures (ctl_D3), 3 from macrophages infected for 3 days (RV-THP_D3), and 3 samples from a 6-day culture (RV-THP_D6). A small RNA library was generated from samples using the Illumina Truseq™ Small RNA Preparation kit according to Illumina's TruSeq™ Small RNA Sample Preparation Guide (Catalog # RS-930-1012, Part # 15004197 Rev. B, January 2011). The purified cDNA library was used for cluster generation on Illumina's Cluster Station and then sequenced on Illumina GAIIx following vendor's instruction for running the instrument. Raw sequencing reads (40 nts) were obtained using Illumina's Sequencing Control Studio software version 2.8 (SCS v2.8) following real-time sequencing image analysis and base-calling by Illumina's Real-Time Analysis version 1.8.70 (RTA v1.8.70). The extracted sequencing reads were used in the standard data analysis. A proprietary pipeline script, ACGT101-miR v4.2 (LC Sciences), was used for sequencing data analysis (Li et al., 2010; Meyer et al., 2011; Wei, Liu, Feng, & Chang, 2011). The raw reads were mapped to the human genome first. Next, the sequences were mapped to miRNAs and pre-miRNAs with miRbase (Figure 7).

Species screened included hsa, ptr, ppy, ppa, ggo, ssy, mml, mne, pbi, age, lla, sla, lca, mmu, rno, cgr, bta, oar, eca, oan, cfa, ssc, mdo, meu, and sha. The reads that did not map to the human genome (nohit in the first step) were subsequently mapped to the Mtb H37Rv, as well as to all the miRNAs species in that miRbase database. Criteria for identifying and annotating miRNAs was undertaken by LC Sciences, using more rigorous algorithms developed for plant miRNAs; i.e. (i) the miRNA/miRNA* sequences should be on opposite stem-arms that can form a duplex with 2-nt 3' overhangs; (ii) ≤4 mismatched miRNA bases between the miRNA and the other arm of the hairpin,

which includes the miRNA*, and (iii) ≤1 asymmetric bulge and only ≤2 bases, especially in the miRNA/miRNA* region. The raw data has been uploaded to Gene Expression Omnibus (GEO) with the accession number GSE146228.

2.2.5 Copy number normalization and multiple comparisons

A modified global normalization is used to correct copy numbers among different samples assuming that there is a subset of sequences that do not change significantly across all samples and in each sample the reading variations occur in the same proportion. First, a common set of sequences was found among all samples. Then, a reference data set was constructed. Each data in the reference set was the copy number median value of a corresponding common sequence of all samples. 2-based logarithm transformation on copy numbers of all samples and reference data set was done and the differences between individual sample and the reference data set was calculated. A subset of sequences selected with a cut off less than 4-fold change from the reference set was formed. Linear regression between individual samples and reference set on the subset sequence was performed to derive linear equations such that $y=a\ x$ +b where a and b are the slop and interception, respectively, of the derived line, x is log2 (copy #) of the reference set, and y is the expected log2(copy #) of the sample on a corresponding sequence. Mid values of the reference set were calculated. The corresponding expected log2(copy #) of a sample was calculated. Logarithmic correction factor of a sample was determined. Then, the arithmetic correction factor was derived. Copy numbers of individual samples were corrected by multiplying corresponding arithmetic correction factor to original copy numbers. For comparisons of multiple samples derived from the small RNA-sequencing screen, raw reads of multiple samples were combined. The number of read copies from each sample was tracked during mapping and normalized for comparison. Reads may be mapped to multiple entries of the reference database and the number

of the read copies was divided by the number of mapped entries. Normalization of sequence counts in each sample (or data set) was achieved by dividing the counts by a library size parameter of the corresponding sample. The library size parameter was a median value of the ratio between the counts a specific sample and a pseudo-reference sample. A count number in the pseudo-reference sample was the count geometric mean across all samples.

2.2.6 sncRNA structure prediction

The RNAFold web server was used for MFE (Minimum Free Energy) RNA secondary structure predictions. The Institute for Theoretical Chemistry at the University of Vienna maintains this website (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi).

2.2.7 Infection Assays and q-RT-PCR

THP-1 cells and BMDMs were prepared as described in section 2.2.1. After confirming positive staining for the CD11b marker with flow cytometry, the BMDMs were plated for infection experiments in RPMI media containing 10% FCS. Macrophages were infected with wild-type Mtb H37Rv or the auxotroph Mtb H37Rv at MOI of 3:1. Three hours. post-infection, the THP-1 cells or BMDMs were washed and cultured for indicated time courses. RNA was isolated 3 hours, 3 days and 6 days post infection as described in section 2.2.2. Custom designed locked nucleic acid miRNA primers were used to detect the sncRNAs with a miRNA-specific q-RT-PCR assay, miRCURY^{TM} locked nucleic acid (LNA)^{TM} miRNA PCR system (Exiqon Inc., now part of Qiagen Inc.). For CFU enumeration, the infected cells were lysed in PBS containing SDS (0.05% v/v). Ten-fold serial dilutions were made in the same buffer, and samples were spread on 7H10 agar plates containing 10% OADC and other supplements required for the auxotroph mutant to determine the total bacterial burden.

2.3 Results

2.3.1 Smaller noncoding RNA identification in Mtb H37Rv

To identify Mtb-encoded sncRNAs, RNA was isolated from THP-1 infected with Mtb H37Rv at MOI of 10:1, at 3 days or 6 days post infection or uninfected THP-1 cells 3 days after the initiation of the experiment. Small RNAs (<350 nts) were sequenced and analyzed using the pipeline in Figure 7. In short, the normalized RNA-seq reads were initially mapped to the human genome (HumanRefSeqBuilt37; GRCh37p13). This analysis revealed 13.5 and 7.1 million human encoded sRNA reads pre- and post-infection, respectively. Then, among these, those <30 nts that had surrounding sequences with potential stem-loop structures were filtered for. This analysis classified 15% and 30% of sRNAs pre- and post-infection, respectively as eukaryotic miRNAs. Among these, 48 were significantly up- or down-regulated as shown in Figure 8. Gene ontology analysis on predicted targets of these suggested that they regulate genes involved in the immune system, metabolism, and development as shown in Table 2. Because of our interest in Mtb-encoded RNAs, later analyses have focused on sRNAs that did not map to the human genome. Screening them against the Mtb H37Rv genome (NC_000962) revealed 2.8 million Mtb-encoded sRNA reads. Of these, only 0.15% (4200 total reads for sRNAs) were classified as miRNA-like RNAs based on the size and structure criteria defined for plant and animal microRNAs. The overall read count for sRNAs represented 35 distinct sRNA sequences. Each of these 35 distinct sRNAs were significantly upregulated over the infection period. They were named as sncRNA-1 to sncRNA-35 based on their descending p-values as shown in Table 3.

2.3.2 Commonalities among sncRNAs

Overall, the sncRNAs identified were between 18-30 nts with an average GC content of 50%. This is much lower than the 65% GC content of the Mtb genome. They were allocated throughout Mtb genome, but mostly found in the intergenic regions except for sncRNA-3 and sncRNA-33. Of note,

sRNAs transcribed from the intergenic regions have multiple gene targets in bacteria. Following are some common features of sncRNAs based on their location. sncRNA-4, sncRNA-5, sncRNA-6, and sncRNA-26 were mapped to multiple locations. sncRNA-1, sncRNA-22, sncRNA-28 and sncRNA-30 were located near the known virulence factors including *esxA*, *esxB*, and *espI*. sncRNA-3, sncRNA-8, sncRNA-16, and sncRNA-33 located between rrs (16S) and *murA*, a cell wall component. In support of our study, earlier studies on sRNAs uncovered larger transcripts in which sncRNA-6 (ncRv10685), sncRNA-8 (Mcr3), sncRNA-17 (MTS1082), sncRNA-19 (Mcr11), and sncRNA-21 (Mcr14) were contained (K. B. Arnvig et al., 2011; K. B. Arnvig & Young, 2009; DiChiara et al., 2010). These features of sncRNAs have been summarized in Table 4.

The sequence reads were predominated by the sequence of sncRNA-1, sncRNA-3, sncRNA-4, sncRNA-5, sncRNA-6, sncRNA-8, sncRNA-26. The changes in the expression of these seven most abundant sncRNAs are shown in Figure 9. As this screen was initially designed to identify *hsa*-encoded small regulatory RNAs, the expression of Mtb-encoded sncRNAs independent of infection had not been assessed with the RNA-seq approach. In any case, the expression of the seven sncRNAs increased from 3-days to 6 days- post infection in our RNA-seq analysis as shown in Figure 9. Among these seven sncRNAs sncRNA-1, sncRNA-6, and sncRNA-8 were analyzed in more detail because they predominated the sequence reads and had smaller minimum free energy (MFE) which indicate a more stable secondary structure (Figure 10). sncRNA-1 is located between *esxA* and *espI*. This is a major pathogenicity locus in Mtb called region of difference 1 (RD1) and is lacked in nonpathogenic mycobacteria. sncRNA-6 is located at two sites, one being between *tuf*, an iron-regulated elongation factor and Rv0686, involved in cell wall biogenesis. sncRNA-6 is defined with a sequence of 21 nts. In addition, there were shorter sequence reads that

contained a portion of sncRNA-6. Each of these sequences lacked 2 nts in the 3' end and have 19-nt length. The 19-nt shorter versions of sncRNA-6 aligned to 4 different regions in Mtb genome. sncRNA-8 is found between *rrs* (16S RNA) and *murA*, a gene involved in cell wall formation.

2.3.3 microRNA-specific q-RT-PCR validation of Mtb-encoded sncRNAs

To verify the RNA-seq data, a q-RT-PCR based approach was employed which assessed the levels of three sncRNAs in infected cells. The miRCURY^{TM} locked nucleic acid (LNA)^{TM} miRNA PCR system was used, which is a highly sensitive and specific assay for the quantification of miRNAs (Blondal et al., 2013; Chousalkar, Cheetham, & Roberts, 2009; Feng et al., 2018). The total RNA was isolated from THP-1 cells or BMDMs, either uninfected or infected with either the wild-type Mtb H37Rv or the auxotroph mutant Mtb H37Rv 6206 at MOI of 3:1 at multiple time points matching the RNA-seq assay. The auxotroph mutant Mtb H37Rv 6206 was derived from the wild-type Mtb H37Rv and lacks the *panCD* and *leuC* regions in its genome (Sampson et al., 2004). However, this mutant has similar growth and infectivity characteristics as the wild-type Mtb H37Rv when grown in the media supplemented with pantothenate and leucine (leu^+ pan^+ macrophage media). Therefore, all the infection assays were done in supplemented media. Moreover, this mutant only requires BSL2 plus biosafety conditions which was established in our laboratory. In these assays, sncRNA-1, sncRNA-6, and sncRNA-8 were consistently detected when either of the macrophage cells types or mycobacteria strains used (Figure 11).

The expression of these three sncRNAs increased between 3- to 30-fold during the 6-day infection period relative to 5S RNA with primers specific for Mtb 5S differing from *hsa*-5S were used (Figure 11A& B). These three sncRNAs were not detected in uninfected macrophages. The initial screen did not allow us to determine the levels of sncRNAs independent of infection due to how it was designed. However, this is an important parameter to investigate as it gives insights into the

role of sncRNAs in infection. Therefore, in our infection assay with Mtb H37Rv 6206 and BMDM, whether sncRNAs were produced independent of any infection was determined. For that, RNA was harvested from Mtb H37Rv 6206 auxotroph grown for 3 hours in the leu^+ pan^+ macrophage media in parallel to those used in the intracellular infection (Figure 11C). This analysis revealed that sncRNA-1 and sncRNA-6 were not detected in Mtb H37Rv 6206 independent of infection whereas sncRNA-8 had a detectable basal level expression. The expression of sncRNA-8 increased around 100-fold within 3 hours post-infection in macrophages. As a side note, the expression levels of these three sncRNAs were higher in BMDMs versus THP-1, which may indicate a regulation on their induction or processing in a cell type-based manner. Finally, in our infection assay with Mtb H37Rv 6206 and BMDM, the colony-forming units (CFU) of Mtb H37Rv 6206 were determined over the infection period in parallel to the q-RT-PCR assay (Figure 12). This analysis revealed that at day 6, CFU counts were stabilized indicating that the increase was observed in the sncRNA expression is not a direct result of an increase in the bacterial burden. To verify the q-RT-PCR with a complementary assay, northern blotting was used to detect the sncRNA. For these experiments, RNA was isolated from BMDM infected with Mtb H37Rv 6206 at MOI of 3:1. However, an appreciable signal for sncRNA-1 was not detected in the northern blots. Then, this assay was repeated with MOI of 10:1 and again a signal was not detected. Next, the same membrane was probed for 5S RNA overnight wherein a weak signal was observed only after an overnight exposure (Figure 13). These experiments implied that northern blotting may not be sensitive enough to detect the endogenous sncRNAs. It is likely that the detection of sncRNAs was hampered by the abundance of the eukaryotic RNAs. Therefore, for most of the later studies, the miRCURY^{TM} locked nucleic acid quantitative miRNA PCR system, which is usually referred to as the miRNA-specific q-RT-PCR, was employed for detection of the sncRNAs.

To expand on the *in vitro* findings with an *in vivo* model of tuberculosis, a screen was developed with Rhesus macaque primates infected with Mtb (CDC1551) for the presence of the sncRNAs. First, RNA from lung biopsies were sequenced from 6 primates; 2 uninfected controls, two with active pulmonary TB (PTB), and two latent tuberculosis infection (LTBI). sncRNA-1, sncRNA-5, sncRNA-8, sncRNA-19, and sncRNA-26 were variably detected in 3/6 infected primates, albeit at extremely low sequence reads with RPKM of 1-4. Given the sensitivity of the miRNA-specific q-RT-PCR assays in detecting the sncRNAs, the presence of sncRNA-1, sncRNA-6, and sncRNA-8 was assessed using a second cohort of macaque primates, 4 uninfected, 4 with PTBI, and 4 with LTB. The mycobacterial strains used included CDC1551 and the Erdman strain. To improve on the signal to noise ratio, sRNAs were first enriched for smaller RNAs. Of the 3 sncRNAs screened, sncRNA-8 was detected in 4 of the 8 infected primates (Figure 14). The expression of sncRNA-8 was independent of the mycobacterial strain used, unrelated to the CFU/g lung tissue levels and was observed in both PTB and LTBI groups of primates.

2.4 Discussion

Recent studies have shown that many bacteria including mycobacteria produce sRNAs in response to environmental stress. Since most of these studies have focused on the identification of sRNAs between 50-350 nts, smaller RNAs <50 nts are often overlooked. In this chapter, RNA-seq strategy was used to detect eukaryotic miRNAs and the smaller noncoding RNAs produced by Mtb H37Rv in the infected macrophage cultures. Using such techniques, 35 distinct Mtb-encoded sncRNAs were identified, all present within putative larger transcripts with predicted hairpin loop structures. While many of the sncRNAs could be byproducts of the degradation of the larger transcripts as some were implicated as a part of previously identified sRNAs, our attention was focused on 3, sncRNA-1, sncRNA-6 and sncRNA-8, as these predominated the sequence reads and had smallest

MFE suggesting the most stable RNA secondary structures. The expression of each increased significantly over the course of infection implying a role in Mtb pathogenesis. More intriguingly, sncRNA-1 and sncRNA-6 were solely expressed upon infection, whereas sncRNA-8 had a basal level expression that was elevated significantly in infected cells. These observations suggest an active transcriptional mechanism exists whereby certain sncRNAs are increased during mycobacterial infections in mammalian cells. sncRNA-8 having the highest expression level was detected *in vivo*. In lung biopsies of two of four primates that had either active or latent TB, sncRNA-8 was detectable (Figure 14). In one of the latently infected primates, CFUs were detected which may be an indication of a reactivating infection (Figure 14B, cdc-3). Detection of sncRNA-8 in infected primates independent of the CFU counts suggests that it can be used as a clinical diagnostic of TB. In future studies, it would be intriguing to analyze human sera isolated from patients with TB for sncRNA-8 expression. Overall, in this chapter, the identification of a new set of sRNAs named sncRNAs in Mtb that have been overlooked before has been presented. Having verified their expression in this chapter, next chapter will analyze their function in more detail.

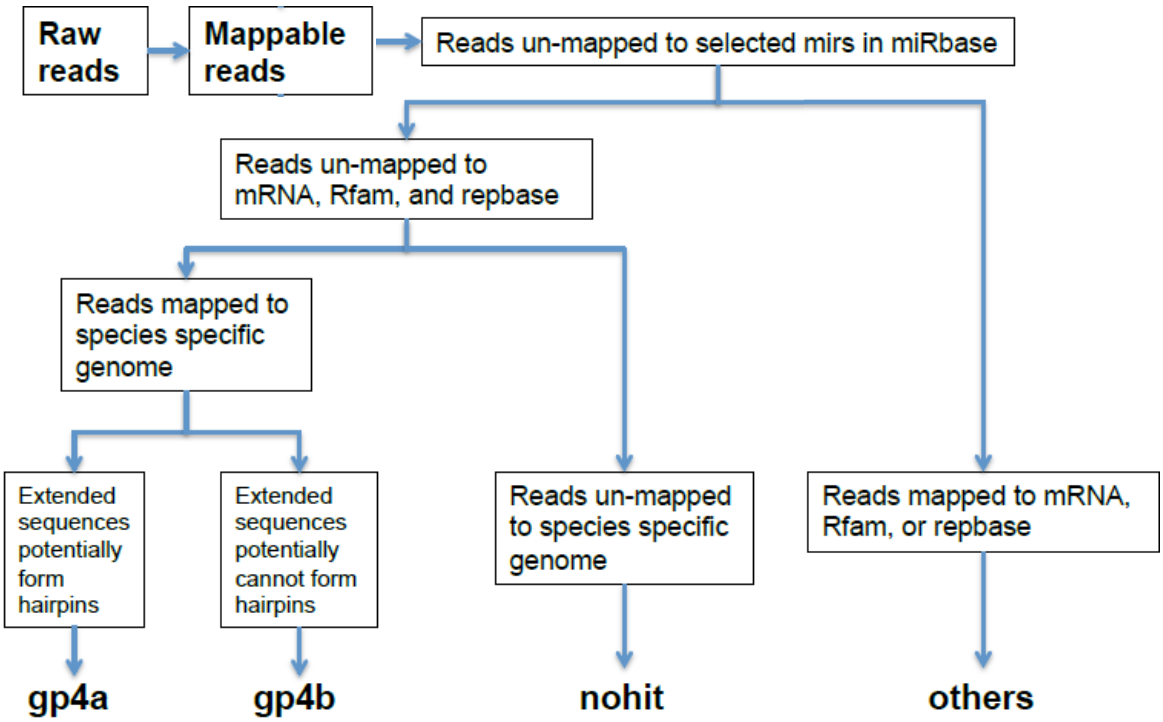


Figure 7. Data analysis pipeline for the discovery of Mtb-encoded sncRNAs. The data analysis was done by LC life Sciences. Novel sncRNAs identified in this study fall into gp4a.

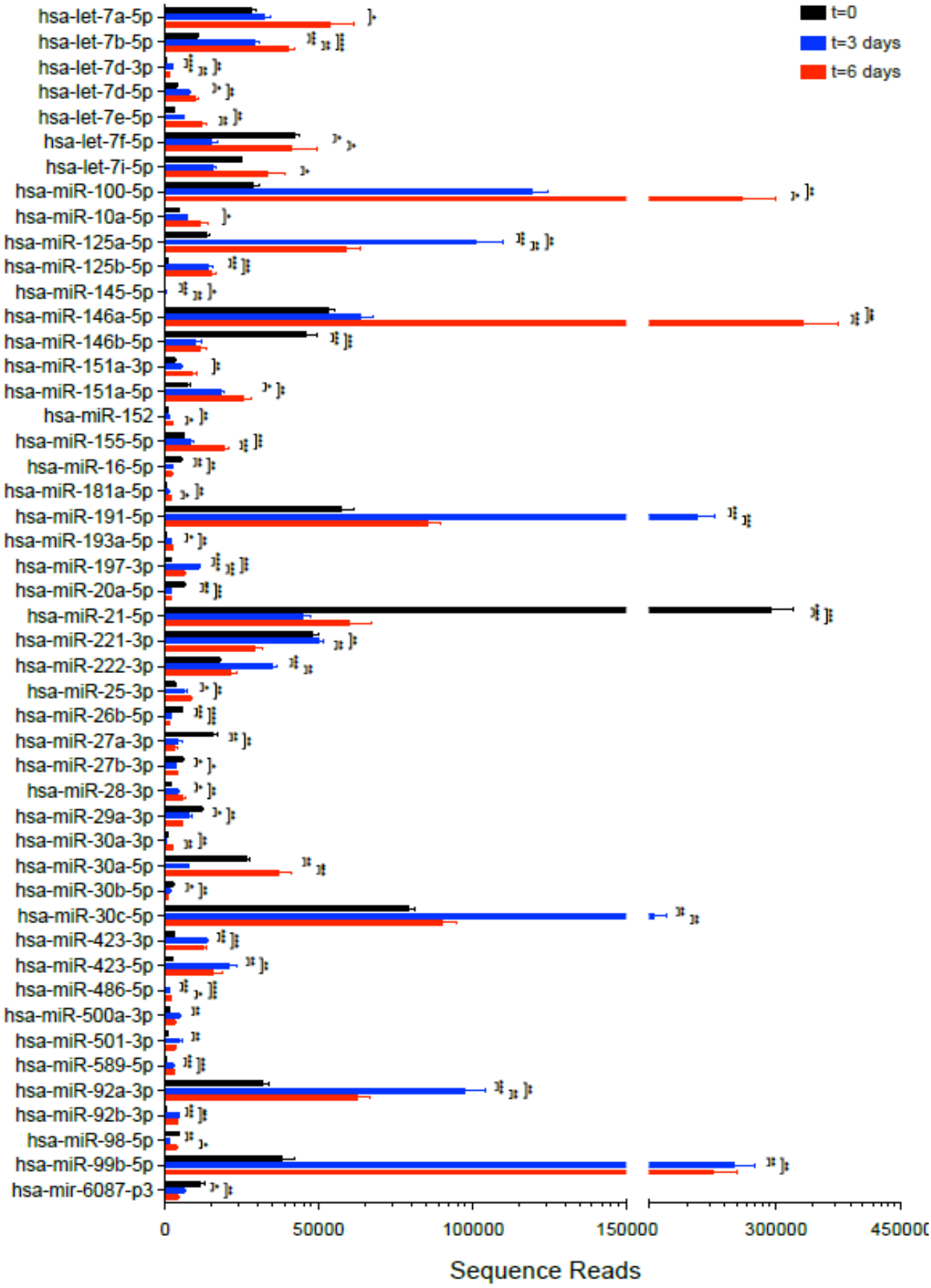

t=0
t=3 days
t=6 days
hsa-let-7a-5p
hsa-let-7b-5p
hsa-let-7d-3p
hsa-let-7d-5p
hsa-let-7e-5p
hsa-let-7f-5p
hsa-let-7i-5p
hsa-miR-100-5p
hsa-miR-10a-5p
hsa-miR-125a-5p
hsa-miR-125b-5p
hsa-miR-145-5p
hsa-miR-146a-5p
hsa-miR-146b-5p
hsa-miR-151a-3p
hsa-miR-151a-5p
hsa-miR-152
hsa-miR-155-5p
hsa-miR-16-5p
hsa-miR-181a-5p
hsa-miR-191-5p
hsa-miR-193a-5p
hsa-miR-197-3p
hsa-miR-20a-5p
hsa-miR-21-5p
hsa-miR-221-3p
hsa-miR-222-3p
hsa-miR-25-3p
hsa-miR-26b-5p
hsa-miR-27a-3p
hsa-miR-27b-3p
hsa-miR-28-3p
hsa-miR-29a-3p
hsa-miR-30a-3p
hsa-miR-30a-5p
hsa-miR-30b-5p
hsa-miR-30c-5p
hsa-miR-423-3p
hsa-miR-423-5p
hsa-miR-486-5p
hsa-miR-500a-3p
hsa-miR-501-3p
hsa-miR-589-5p
hsa-miR-92a-3p
hsa-miR-92b-3p
hsa-miR-98-5p
hsa-miR-99b-5p
hsa-mir-6087-p3
0
50000
100000
150000
300000
450000
Sequence Reads

Figure 8. *hsa*-miRNA expression profiles in Mtb-infected THP-1 macrophages reveals a set of differentially regulated miRNAs. The sequence reads at the beginning of the experiment (t=0) is shown in black, the sequence reads 3 days post infection (t=3days) is shown in blue and the sequence reads 6 days post infection (t=6days) is shown in red. Statistical analysis was done using one-way-ANOVA WITH *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

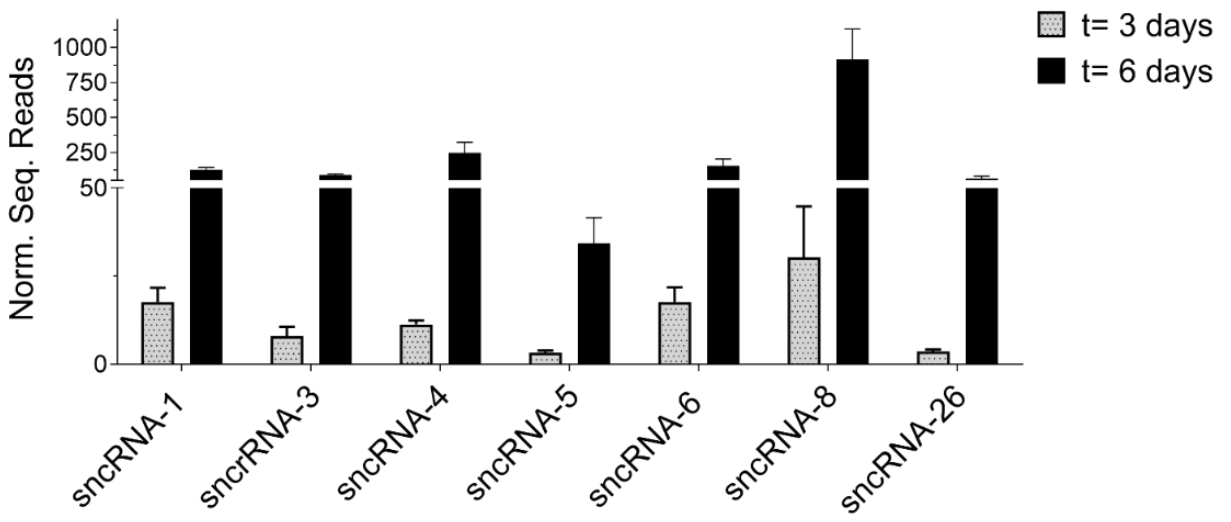


Figure 9. Normalized sequence reads for the seven most abundant sncRNAs 3- and 6-days post infection. RNA was isolated and sequenced from THP-1 macrophages infected with Mtb H37Rv. An algorithm selected for those that mapped only to the Mtb genome and resided with a predicted hairpin loop structure. This analysis revealed 35 such RNAs called sncRNAs, 7 most abundant of which are shown.

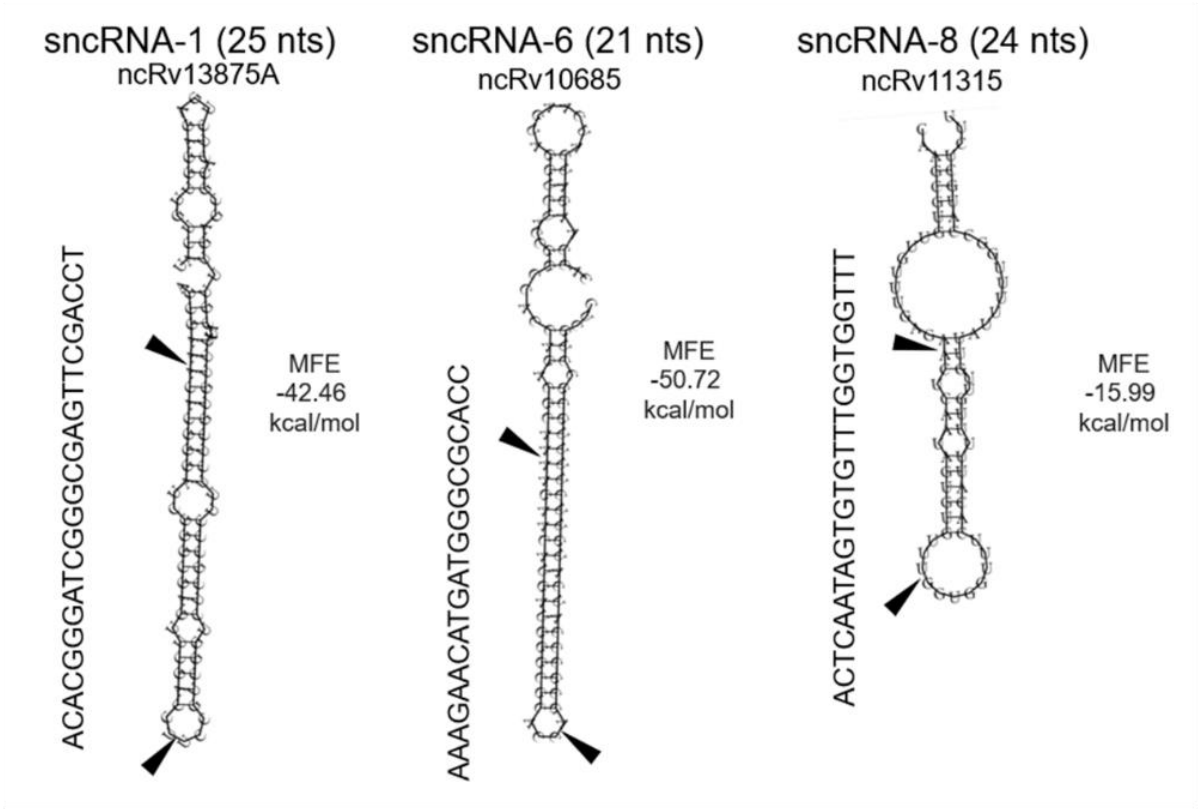
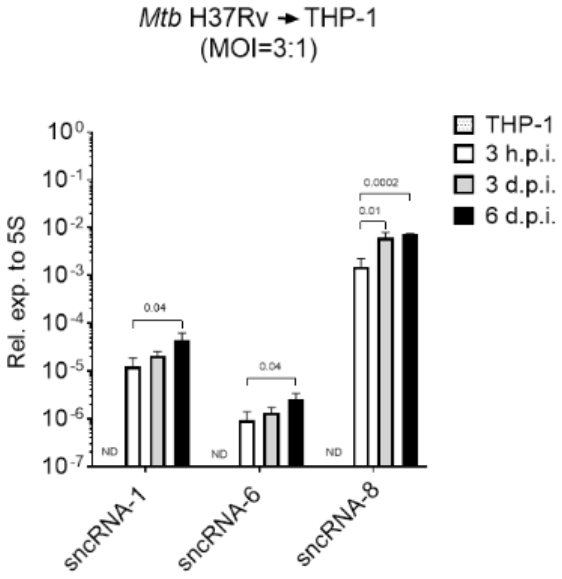
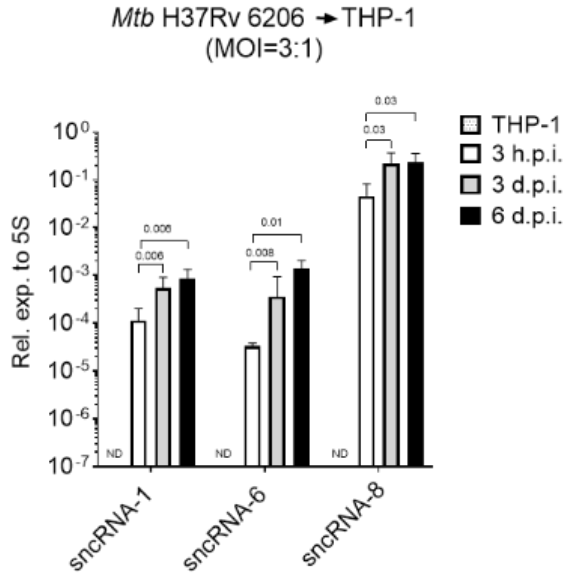


Figure 10. Predicted secondary structure of the precursor of sncRNA-1, sncRNA-6 and sncRNA-8 reveals hairpin loops. Precursors are expected to be cleaved at sites shown with arrows to yield the mature forms. The structures were predicted using RNAfold online tool. The sizes of the mature sncRNAs, which are indicated with arrows, are shown in parenthesis. The formal names of the sncRNAs are shown underneath of the sncRNA nicknames. The MFE values, which were calculated using the RNAfold tool are shown on the right side of each sncRNA.

A


Mtb H37Rv → THP-1
(MOI=3:1)
Rel. exp. to 5S
THP-1
3 h.p.i.
3 d.p.i.
6 d.p.i.
0.04
0.04
0.0002
0.01
ND
sncRNA-1
sncRNA-6
sncRNA-8

B


Mtb H37Rv 6206 → THP-1
(MOI=3:1)
Rel. exp. to 5S
THP-1
3 h.p.i.
3 d.p.i.
6 d.p.i.
0.006
0.006
0.01
0.008
0.03
0.03
ND
sncRNA-1
sncRNA-6
sncRNA-8

C

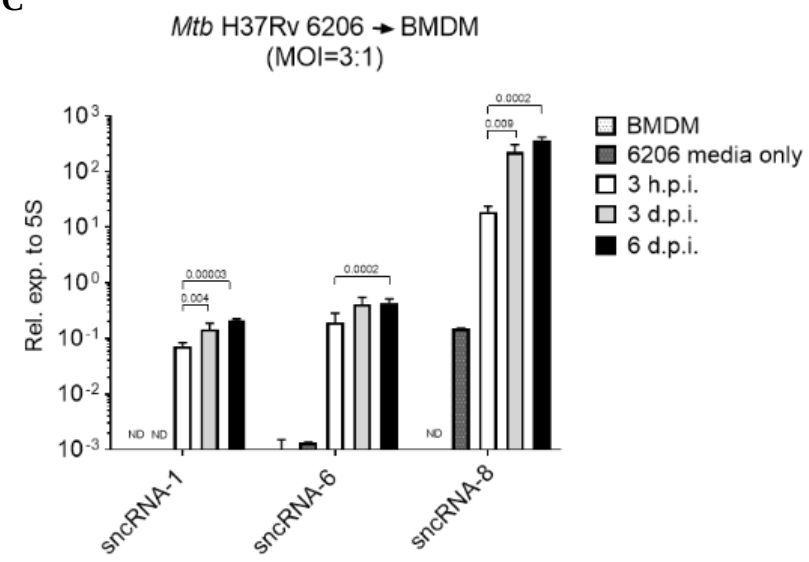

Mtb H37Rv 6206 → BMDM
(MOI=3:1)
Rel. exp. to 5S
BMDM
6206 media only
3 h.p.i.
3 d.p.i.
6 d.p.i.
0.00003
0.004
0.0002
0.0002
0.009
ND
sncRNA-1
sncRNA-6
sncRNA-8

Figure 11. sncRNA-1, sncRNA-6 and sncRNA-8 levels increase during the course of Mtb infections in macrophages as revealed by miRNA-specific q-RT-PCR assays. (A-C) The RNA was isolated from (A-B) THP-1 cells or (C) BMDMs, either uninfected or infected with either the (A) wild-type Mtb H37Rv or the (B-C) auxotroph mutant Mtb H37Rv 6206 at MOI of 3:1 at multiple time points matching the RNA-seq assay. The expression of sncRNAs were quantified using the miRNA-specific q-RT-PCR system. The expression of the sncRNAs were normalized to the mycobacterial 5S designed for a region different than the *hsa*-5S. Statistical analysis was done using one-way-ANOVA with *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$.

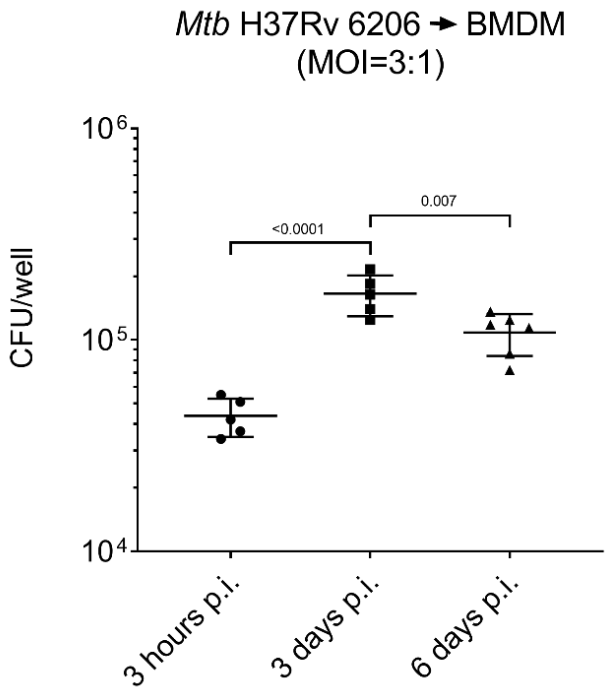


Figure 12. The CFU burden of Mtb in infected bone marrow derived macrophages initially increases and then decreases over 6-day infection period. CFU was determined over the infection period in parallel to the q-RT-PCR assay. The BMDM were infected with Mtb H37Rv 6206 for 3 hours and then either lysed in 0.05% SDS (3 hours p.i.) or washed for later time points (3 days p.i. and 6 days p.i.). The lysed mycobacteria were serially diluted, and the dilutions were plated onto supplemented 7H10 plates. The CFU was counted 3-4 weeks after the plating. Statistical analysis was done using one-way-ANOVA.

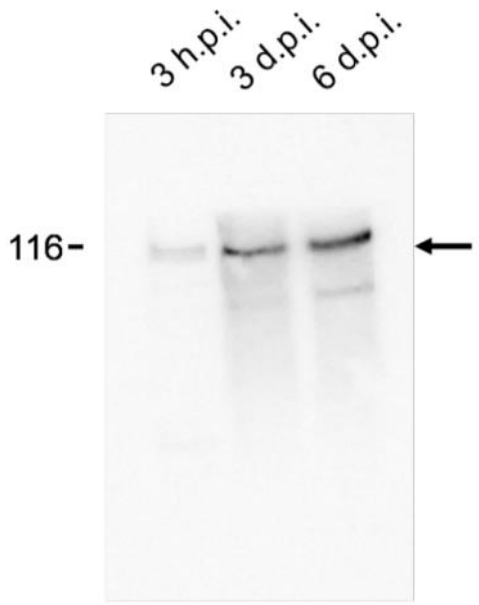


Figure 13. The levels of the mycobacterial 5S RNA are weakly detected in infected BMDM with northern blotting assays. RNA was isolated from BMDM infected with Mtb H37Rv 6206 at MOI of 10:1 after 3 hours, 3 days and 6 days of infection. They were run on 20% acrylamide gel and transferred to a Nylon+ membrane. The membrane was first probed for sncRNA-1, but no signal was detected after overnight exposure. Next, the same membrane was probed for 5S RNA and the signal detected after overnight exposure is presented.

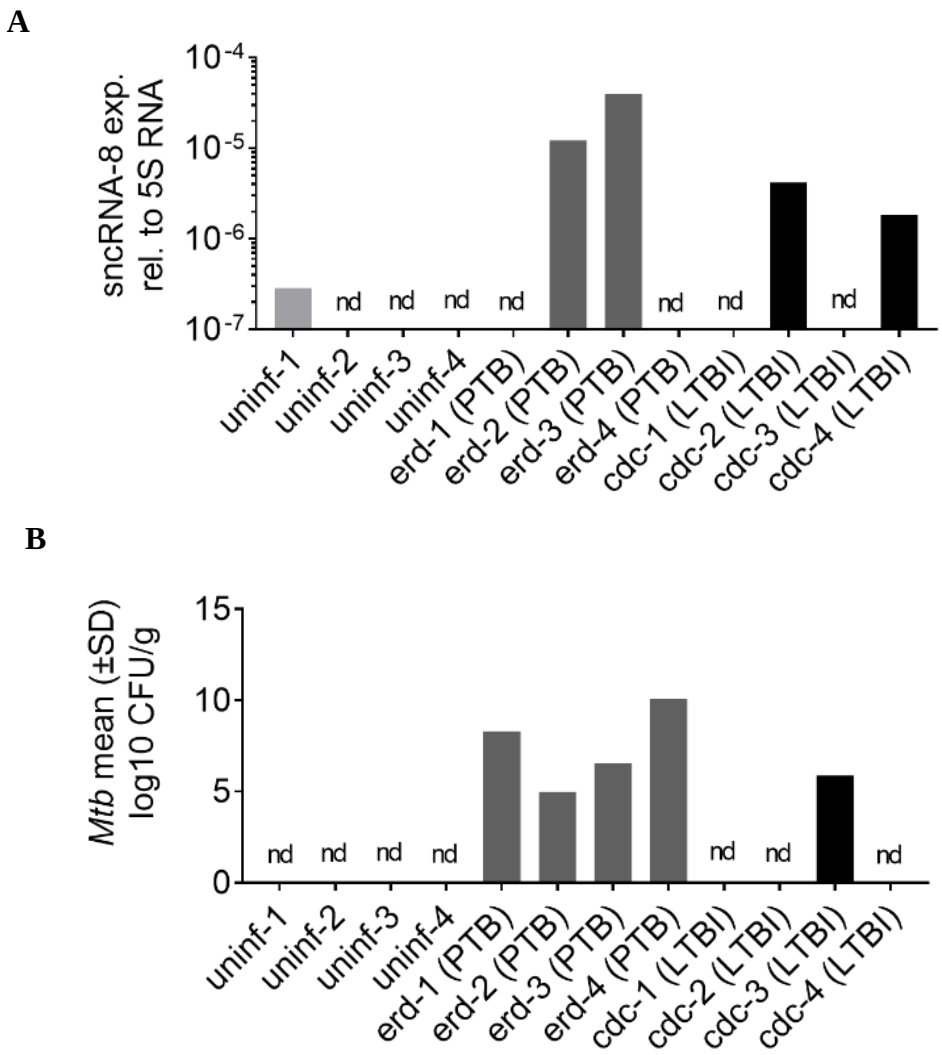


Figure 14. sncRNA-8 is detected from lung samples of Mtb-infected primates independent of the bacterial burden. RNA was isolated from a cohort of macaque primates, 4 uninfected, 4 with PTBI, and 4 with LTB. The mycobacterial strains used included CDC1551 and the Erdman strain. To improve on the signal to noise ratio, sRNAs were first enriched for smaller RNAs. (A) sncRNA-8 was detected in 4 of the 8 infected primates using the miRNA-specific q-RT-PCR assay. (B) The bacterial burden was quantified in these primates per gram of the lung tissue. Nd indicates not determined.

Table 2. Gene ontology analysis on predicted targets of *hsa*-miRNAs differentially expressed upon Mtb infection.

Biological Process	**Number**	**% gene hit against total number of genes**	**% gene hit against total number of process hits**
Metabolic process	308	58.3	19.30%
Cellular process	280	53.0	17.50
Cell communication	205	38.8	12.80
Apoptosis	80	15.2	
MAPKKK cascade	41	7.8	
Cytokine-mediated signaling pathway	32	6.1	
Transmembrane receptor protein tyrosine kinase signaling pathway	28	5.3	
G-protein coupled receptor protein signaling pathway	27	5.1	
I-kappaB kinase/NF-kappaB cascade	21	4.0	
Transmembrane receptor protein serine/ threonine kinase signaling pathway	17	3.2	
JAK-STAT cascade	12	2.3	
Developmental process	142	26.90	8.9
Immune system process	108	20.50	6.8
Macrophage activation	21	4.0	
B-cell mediated immunity	17	3.2	
Response to interferon-gamma	9	1.7	
Natural killer cell activation	8	1.5	
Cell cycle	107	20.30	6.70
Response to stimulus	83	15.70	5.20
System process	58	11.00	3.60

Table 3. RNA sequence reads of Mtb-encoded sncRNAs in infected macrophages

Small RNA	**Original RNA Seq**		**t =0**	**t= 3 d**			**t= 6 d**		
Designation[a]	**Identifier[b]**		Mean	SD	Mean	SD	Mean	SD	p-value[c]
sncRNA-1	PC-5p-55054_108	gp4a	0	0	17.6	4	126.6	16.4	1.30E-10
sncRNA-2	PC-5p-96794_33	gp4a	0	0	2.9	0.3	48.3	9.8	2.08E-10
sncRNA-3	PC-3p-128923_18	gp4a	0	0	8.1	2.5	89.2	7.7	4.32E-10
sncRNA-4	PC-3p-10118_1339	gp4a	0	0	11.4	1.2	247	75.9	1.04E-09
sncRNA-5	PC-5p-131326_17	gp4a	0	0	3.4	0.9	34.3	7.3	1.92E-09
sncRNA-6	PC-3p-27840_350	gp4a	0	0	17.9	4.6	154.2	48	4.15E-09
sncRNA-7	PC-3p-70005_66	gp4a	0	0	0.9	0.2	42.6	13.7	4.42E-09
sncRNA-8	PC-3p-9177_1491	gp4a	0	0	30	14.4	915.9	215.8	4.66E-09
sncRNA-9	PC-3p-71935_63	gp4a	0	0	4.5	1.7	38.6	11.1	1.40E-08
sncRNA-10	PC-3p-117155_22	gp4a	0	0	1.8	0.3	13.5	5.8	4.97E-08
sncRNA-11	PC-5p-74813_58	gp4a	0	0	3.6	1.3	36	14.1	5.10E-08
sncRNA-12	PC-5p-96661_33	gp4a	0	0	4.5	2.4	63.8	12.6	5.22E-08
sncRNA-13	PC-5p-298538_5	gp4a	0	0	4.4	2.1	16.8	5	1.07E-07
sncRNA-14	PC-5p-140336_15	gp4a	0	0	2.2	1.3	29.9	5.6	1.33E-07
sncRNA-15	PC-3p-88717_40	gp4a	0	0	2.2	1.3	27.5	6.2	1.53E-07
sncRNA-16	PC-5p-140583_15	gp4a	0	0	1.6	1.1	27	4	1.63E-07
sncRNA-17	PC-5p-137373_16	gp4a	0	0	1.5	1	16.1	3.9	2.46E-07
sncRNA-18	PC-5p-106320_27	gp4a	0	0	1.6	1.1	27.6	8.8	3.01E-07
sncRNA-19	PC-3p-139225_16	gp4a	0	0	3	1.7	40.3	13.4	4.23E-07
sncRNA-20	PC-3p-100291_30	gp4a	0	0	3.1	2.1	34.3	14.6	1.04E-06
sncRNA-21	PC-3p-399156_3	gp4a	0	0	0	0	5.6	5.6	4.93E-06

sncRNA-22	PC-5p-127118_19	gp4a	0	0	1	0.5	12.8	12.9	7.02E-06
sncRNA-23	PC-5p-280865_5	gp4a	0	0	0.2	0.4	8.9	2.2	5.73E-03
sncRNA-24	PC-3p-138156_16	gp4a	0	0	0.4	0.4	9.6	4	5.87E-03
sncRNA-25	PC-5p-73360_60	gp4a	0	0	1.3	1.5	40.4	18.6	6.11E-03
sncRNA-26	PC-3p-47073_145	gp4a	0.2	0.4	4	0.7	63.8	15.8	6.24E-03
sncRNA-27	PC-3p-118120_22	gp4a	0	0	0.5	0.5	10.5	4	6.51E-03
sncRNA-28	PC-3p-187786_9	gp4a	0	0	0.9	0.9	22.1	4.8	6.87E-03
sncRNA-29	PC-3p-265291_6	gp4a	0	0	0.3	0.5	9.9	6.2	7.67E-03
sncRNA-30	PC-5p-157943_12	gp4a	0	0	0.6	0.5	8.9	3.8	8.31E-03
sncRNA-31	PC-5p-131024_18	gp4a	0	0	0.3	0.5	7.7	5.2	8.43E-03
sncRNA-32	PC-5p-116466_22	gp4a	0	0	1.2	1.1	17.2	5.6	9.21E-03
sncRNA-33	PC-5p-118223_21	gp4a	0.2	0.4	4.1	3.3	29.9	3.2	9.52E-03
sncRNA-34	PC-5p-288803_5	gp4a	0	0	1.1	0.9	12.5	2.8	9.86E-03
sncRNA-35	PC-5p-110779_25	gp4a	0	0	1.9	1.8	10.6	1	1.44E-02

Table 4. The sequence and genomic location of 35 distinct Mtb sncRNAs

sncRNA designation [a]	**Nucleotide Sequence**	**Length**	**Genomic Location** [b]	**Nearby coding sequence** [c]	**Other designation**
sncRNA-1[d]	ACACGGGAUCGG GCGAGUUCGACCU	25	4352927-4352951	b/w[e] esxA (Rv3875) and Rv3876	ncRv13875
sncRNA-2	AAGACACACGGA UACCCUUUGC	22	3343151-3343172	b/w mutT (Rv2985) and hupB (Rv2986c)	ncRv12985c
sncRNA-3	ACGGGCAGACUA GAGUACUGCAGGGGAG AC	30	1472478-1472507	rrs (Rvnr01)	ncRv0001
sncRNA-4f	GAAAUGACGCAA UGACCUCU	20	3481440-3481459	Rv1199c and b/w Rv1200	ncRv11199A
sncRNA-5[d,f]	CGUUUCGAAGGA UCACGCGAUGACCGCCC	29	3481390-3481418	b/w Rv1199c and Rv1200	ncRv11199B
sncRNA-6f	AAAGAACAUGA UGGGCGCACC	21	786003-786083	b/w iron elongation factor (Rv0685) and (Rv0686)	ncRv10685
sncRNA-7	UGGUCAGAGAU GCAGUAACUC	21	4071678-4071698	b/w Rv3632 and Rv3633	ncRva13632c
sncRNA-8[d]	ACUCAAUAGUGU GUUUGGUGGUUU	24	1471701-1471724	b/w murA (Rv1315) and rrs (Rvnr01)	ncRv11315 (mcr3)
sncRNA-9	CGGCCGGUUCA UAUGACAC	19	1518217-1518237	b/w fabG (Rv1350) and Rv1351	ncRv11350
sncRNA-10	AAAGGAGGUGA CUCGAUGCU	20	51723-51742	b/w Rv0047c and b/w Rv0048c	ncRv10047

sncRNA-11	GAUCCCGAAG AUUCGCUCC	19	675996-676016	b/w PE_PGRS 7 (Rv0578c) and Rv0579	ncRv10578
sncRNA-12	CUACCGGCCAC CAGACGCAAAAGA	24	786014-786037	b/w tuf (Rv0685) and Rv0686	ncRv10685
sncRNA-13	CCGAGUCCUGAG GGCUGCAGUGACCCC	27	2096781-2096807	b/w Rv1846c and Rv1847	ncRv11846
sncRNA-14	CAACCGCGAAA UGACACUGAUGU	23	3640359-3640381	b/w whiB2 (Rv3260c) and fbiA (Rv3261)	ncRv13260
sncRNA-15	GCGUGAGUCGU UAAAGGUUGAUC	23	3621447-3621469	b/w Rv3241c and Rv3242c	ncRv13241
sncRNA-16	CAAGCGUGUUG UUUGAGAACU	21	1471683-1471703	b/w murA (Rv1315) and rrs (Rvnr01)	ncRv11315
sncRNA-17	CGAAAUCUUG UUUGACGCGGAC	22	1547634-1547655	b/w Rv1374c and Rv1375	ncRv11374 (MTS1082)
sncRNA-18	UCAGGGUCGG ACAUAUCGCCUCC	23	4071702-4071724	b/w Rv3632 and Rv3633	ncRv13632c
sncRNA-19	AGAGUCACG CCGGGUCUGCCC	21	1413108-1413128	b/w Rv1264 and Rv1265	ncRv11264c (mcr11)
sncRNA-20	UGUCUGCUCGC CGAGGCCUACC	22	456251-456272	b/w secE2 (Rv0379) and into Rv0380c	ncRv10379
sncRNA-21	CUUGCCCCAUC GUUCGUUGACUCUGCG UC	29	293713-293741	b/w fadA2 (Rv0243)	ncRv10243 (mcr14)

				and fadE5 (Rv0244c)	
sncRNA-22	CCAGUGAUCG GCGGUCUC	18	2626602-2626624	b/w PPE18 (Rv1196) and esxK (Rv1197)	ncRv11196
sncRNA-23	CGUGGUGGACG CAGAGUCAACGGA	24	3226334-3226357	b/w ffh (Rv2916c) and Rv2917	ncRv12916
sncRNA-24	UGAGUGGUCG GUAGUUGUC	19	4303389-4303411	b/w Rv2979c and Rv2980	ncRv12979
sncRNA-25	AACGAUCAGU AGUAGGCC	18	1196209-1196230	b/w echA9 (Rv1071c) and Rv1072	ncRv11071
sncRNA-26f	CGGCAACUGAA UACUGACC	19	4075666-4075684	b/w Rv3190c and Rv3191c	ncRv13190
sncRNA-27	CGCAACCUGGC CACCAAUUC	20	4318492-4318515	b/w PPE55 (Rv3347c) and Rv3348	ncRv13347
sncRNA-28	CGUUUGAGCGC ACUCUGAGAGGU	23	4352983-4353005	b/w esxA (Rv3875) and Rv3876	ncRv13875
sncRNA-29	UUUGGAGGUU CUGAGAUGAGUCC	23	611046-611069	b/w Rv0518 and Rv0519c	ncRv10518c
sncRNA-30	GAGCCGUCGGG ACCACACAGUCAUU	25	3862549-3862573	b/w rplM (Rv3443c) and esxT (Rv3444c)	ncRv13443
sncRNA-31	AGAUUGUUCA UCGGGGGCU	19	892259-892277	b/w cfp29 (Rv0798c) and Rv0799c	ncRv10798

sncRNA-32	CGGUAGCCAG UAGUUCAUCC	20	333375-333394	b/w Rv0277c and PE_PGRS 3 (Rv0278c)	ncRv10277
sncRNA-33	UGGUUUGUCG CGUUGUUCGU	20	1472422-1472441	rrs (Rvrn01)	ncRv0001B
sncRNA-34	CAAGAUGUCG AUUGCUCACC	20	3621389-3621408	b/w Rv3241c and Rv3242c	ncRv13241
sncRNA-35	CCUAGUUAUCU GCGCCGAGCGUGAAC	26	3987361-3987386	b/w Rv3547 into Rv3548c	ncRv13547c

[a]sncRNA name based on the p value, calculated from n=3/time point

[b]sncRNA location on the Mtb H37Rv genome

[c]Genes on which or near the sncRNA is found

[d]sncRNAs identified in TB infected primate lung tissue

[e]b/w = between

[f]in multiple locations

CHAPTER 3

Functional characterization of sncRNA-1, sncRNA-6, and sncRNA-8

3.1 Introduction

In the previous chapter, smaller noncoding RNAs (sncRNAs) produced by *Mycobacterium tuberculosis* (Mtb) in infected cells were sequenced and quantitated in a precise time course. The expression of three of these sncRNAs; sncRNA-1, sncRNA-6, and sncRNA-8 was verified with a miRNA-specific q-RT-PCR assay, confirming the increased expression of each in infected cells. In this chapter, the function of these was characterized employing a gain-of-function approach. Previous studies have used mycobacterial overexpression vector pKA-303 which has a strong constitutive promoter *rrnB* (K. B. Arnvig et al., 2011). To study the function of sncRNA-1, sncRNA-6 or sncRNA-8, longer DNA segments that contained the sncRNA sequences, which included the predicted hairpin loop structure, were cloned into pKA-303. This vector was used since the endogenous regulatory elements that control the expression of sncRNAs are not known. The plasmids were individually expressed in Mtb H37Rv 6230 (ΔRD1, Δ*panCD*), which lacks the region of difference (RD1) pathogenicity locus hence the sncRNA-1 sequence (Sambandamurthy et al., 2006). This strain has a doubling time of 3-4 hours, which is faster than the wild-type Mtb H37Rv that has a doubling time of 22-24 hours. The growth kinetics of Mtb H37Rv 6230 makes it a good model organism to study the wild-type Mtb H37Rv. Transcriptome of the Mtb H37Rv 6230 overexpressing any sncRNA was analyzed compared to the Mtb H37Rv 6230 with the control vector. Functional annotation of differentially expressed genes suggested different roles for each sncRNA, whereas sncRNA-1 was analyzed in more detail as it was in the RD1 locus.

3.2 Material and Methods

3.2.1 Bacterial strains and growth conditions

Mtb H37Rv 6230 (ΔRD1, Δ*panCD*) was grown in 7H9 media supplemented with 10% oleic albumin dextrose catalase (OADC) (Remel, Thermo-Fisher), 5% glycerol (G9012, Sigma-Aldrich), 0.05% tyloxapol (T0307, Sigma-Aldrich), 0.2% casamino acids (223050, BD Bacto™) and 24 g/ml pantothenate (D-Pantothenic acid hemicalcium salt, Sigma-Aldrich).

3.2.2 Cloning procedures

Mtb H37Rv genomic DNA was obtained from Colorado State University under a TB Vaccine Testing and Research Materials Contract with the National Institutes of Health (N01-AI-40091). pKA-303 was used as an expression vector for the sncRNAs. It originally had a kanamycin resistance cassette which was replaced with a hygromycin resistance cassette because pKA-303 with kanamycin resistance cassette could not be propagated in Mtb H37Rv 6230. Subcloning was done using NEBuilder® HiFi DNA Assembly Master Mix (E2621S, NEB Inc.) according to the manufacturer's instructions. Primer design was done using NEBuilder Assembly tool v2.2.5 (http://nebuilder.neb.com/#!/). DNA sequences encompassing the genomic locations of the individual sncRNAs were cloned into the modified pKA-303. Plasmid transformations were first done with chemically competent *Escherichia coli* (*E. coli*) (DH5alpha). DNA sequencing confirmed sequence insertion and orientation. Plasmids were then electroporated into Mtb H37Rv 6230 as described elsewhere. Transformants were selected in the presence of hygromycin, used at a final concentration of 100 µg/ml.

3.2.3 RNA isolation, sequencing, and data analysis

RNA was extracted from individual clones and sequenced at The Genomics and Microarray core at UT Southwestern Medical Center. Gene expression differences were determined with an associative t-test, as described (I. Dozmorov & Centola, 2003; I. Dozmorov & Lefkovits, 2009; I. M. Dozmorov et al., 2011). The data is available in the Gene Expression Omnibus (GEO)

repository (GSE146228). KEGG GO term analysis was done using DAVID 6.8 available online (https://david.ncifcrf.gov/). Remaining RNA was used to quantify mature and precursor sncRNA expression. sncRNA expression was quantified using the miRNA-specific q-RT-PCR assay as described in Chapter 2. Precursor sncRNA expression was quantified using high-capacity cDNA reverse transcription kit (ThermoFisher Scientific Inc.) and SYBR green PCR master Mix (ThermoFisher Scientific Inc.).

3.2.4 Northern blotting techniques for sncRNAs

Mtb H37Rv 6230 clones expressing the control vector or sncRNA-1 vector were grown in supplemented 7H9 for 1 week. RNA was extracted as described in Chapter 2. Five µg RNA was resolved on a 20% acrylamide gel (29:1) for 2-3 hours in a 1X Tris-cl, borate, EDTA buffer (TBE). The RNA bands were then transferred to Nylon N^+ membranes at 80 volts for 1 hour in 0.5X TBE buffer. RNA was UV crosslinked and the membrane was prehybridized at 39ºC for 1-2 hours in rapid-hybridization buffer (GE/Amersham). These conditions were fully described elsewhere (Belkaya et al., 2011). It was hybridized overnight at 39ºC with Starfire probe specific for sncRNA-1 or 5S rRNA (Starfire™, Idt DNA technologies, Inc.). The 5S probe detects a band near 116 nts. The membrane was washed twice in buffer (2x SSC buffer with 0.1% SDS), wrapped in saran wrap, and imaged with a phosphoimager.

3.3 Results

3.3.1 sncRNA-1, sncRNA-6, or sncRNA-8 overexpression in Mtb H37Rv 6230

In the previous chapter, it was shown that sncRNAs were upregulated in infected cells, implying that they may have functions related to Mtb survival within macrophages. In order to gain insights into the function of sncRNA-1, sncRNA-6 and sncRNA-8, these sncRNAs were individually expressed in an auxotrophic mutant of Mtb called Mtb H37Rv 6230 (ΔRD1, Δ*panCD*) in an

exogenous expression plasmid. Advantages of Mtb H37Rv 6230 are that it has a faster doubling time (3-4 hours) than Mtb H37Rv, which makes it a good model for genetic studies. Second, it can be used in BSL2 Plus conditions, obviating the need for BLS3 facilities (Sambandamurthy et al., 2006; Sampson et al., 2004; Volpe et al., 2006). We cloned the genomic regions of the sncRNAs into pKA-303, an RNA expression vector that uses the *rrnB* promoter (Figure 15) (K. B. Arnvig et al., 2011). The longer DNA segments containing sncRNA-1, sncRNA-6 and sncRNA-8 were 208 nts, 333 nts, and 321 nts, respectively, within each were the appropriate hairpin loops evident in the predicted RNA secondary structures of the transcripts. The long DNA segments were used in the pKA vector because the endogenous regulatory elements that control the expression of the different sncRNAs remain unknown. The plasmids were electroporated into Mtb H37Rv 6230, and three different transformants/construct were picked and expanded in the selection media. RNA was extracted with growing cultures and the sncRNA expression was quantified using miRCURY^{TM} locked nucleic acid quantitative miRNA PCR system (Figure 16).

In clones bearing the sncRNA-1 expression vector, a 20,000-fold increase in expression was detected. Such an extremely high level is not surprising since Mtb H37Rv 6230 lacks the RD1 region which encompasses the sncRNA-1 genomic location. Therefore, the sncRNA-1 level detected in the control vector is consider zero (Figure 16A). sncRNA-6 was barely detectable in clones with the control vector but was highly expressed in the clones containing the sncRNA-6 expression vector (Figure 16A). In contrast, SncRNA-8 had a high basal level expression. Overexpression of the precursor sncRNA-8 transcript in Mtb H37Rv 6230 failed to augment sncRNA-8 expression (Figure 16A). To determine if longer precursor transcripts containing the sncRNAs were induced, the 200-300 nts precursor transcripts were quantified with primers designed to detect the longer RNA species. A presumed precursor of sncRNA-8 was highly

expressed in the clones that had sncRNA-8 expression vector. Contrasting sncRNA-8, precursor of sncRNA-1 and sncRNA-6 were barely detectable in the transformants that with the respective sncRNA expression cassettes (Figure 16B). These experiments suggest the very high levels of the premature sncRNA-8 limit the ability of putative processing enzymes from generating the mature sncRNA-8 sRNA. Alternatively, the levels of mature sncRNA-8 may have reached a saturation point. Note also that sncRNA-1 and sncRNA-6 are more readily detected in infected cells (Figure 11).

3.3.2 Transcriptomic comparisons among Mycobacterial clones overexpressing sncRNA-1, sncRNA-6, or sncRNA-8

The elevated expression of the sncRNA-1 and sncRNA-6 in Mtb H37Rv 6230 clones enabled us to determine their effects on the Mtb transcriptome along with sncRNA-8. RNA was extracted from individual clones and sequenced, with the differentially expressed genes log2 >1.5 fold (DEGs) identified. This analysis revealed a unique transcriptome in Mtb when each sncRNA was individually overexpressed (Figure 17).

Subsequently, a KEGG GO term analysis was performed on the DEGs using DAVID 6.8, which revealed over/under-represented KEGG terms. Overexpression of sncRNA-1 resulted in under-representation of aromatic compound catabolic process and over-representation of fatty acid metabolic processes (Figure 18A). These data suggest sncRNA-1 regulates Mtb metabolism. Overexpression of sncRNA-6 resulted in under-representation of genes coupled to growth and stress response including antibiotic and heat response (Figure 18B). Lastly, expression of the sncRNA-8 vector resulted in under-representation of genes coupled to transcription and over-representation of genes coupled to plasma membrane transport (Figure 18C).

3.3.3 In-depth functional characterization of sncRNA-1 implicates it in the regulation of Mtb fatty acid metabolism

Of the 3 sncRNAs that differentially impacted the Mtb transcriptome, sncRNA-1 was subsequently studied for several reasons. First, sncRNA-1 has never been identified whereas the primary transcripts for sncRNA-6 nor sncRNA-8 were previously detected (DeJesus et al. 2017 and DiChiara et al., 2010). Second, sncRNA-1 is in the RD1 region, which is the major pathogenicity site in Mtb, implying that sncRNA-1 may contribute to Mtb pathogenicity. Third, the RD1 region is lacking in the auxotrophic mutant Mtb H37Rv 6230 (ΔRD1, Δ*panCD*), which allows for a true gain-of-function assay in the absence of endogenous sncRNA-1. To confirm overexpression of sncRNA-1, northern blotting was performed. All 3 clones contained the complete processed 25-nt form of sncRNA-1 (Figure 19 and Figure 20). Some weakly expressed longer transcripts RNAs (25-116 nts) were also detected. Probing for 5S RNA indicated comparable RNA loading (Figure 19).

Moreover, in this assay the specificity of the LNA^{TM} quantitative miRNA PCR probes was tested using primers designed to detect the complement sequence of sncRNA-1 (sncRNA-1 comp) and a 51-nt presumed precursor sequence containing sncRNA-1 and its complement (sncRNA-1 pre) (Figure 20). q-RT-PCR performed with the primers specific to sncRNA-1 comp revealed a lower detection than the probes specific for sncRNA-1 (Figure 20). This could implicate a mechanism for maturation of sncRNA-1 where the precursor is cleaved to release the mature sncRNA-1 and its complement, which is less efficiently protected from RNases compared to sncRNA-1. Moreover, the miRNA-specific q-RT-PCR primers designed to detect sncRNA-1 pre yielded much lower expression of the sncRNA-1 precursor than either sncRNA-1 and its complement (Figure

20). Overall, these tests imply a distinct stabilization of the mature sncRNA-1 vs its complement sequence or a presumed precursor sequence.

For the analysis of sncRNA-1, the DEGs in Mtb H37Rv 6230 were compared between three vector control and 3 sncRNA-1 overexpressing clones using a cut-off of log2 1.5-fold (Figure 17, first heat map). A subset of these genes is shown in Figure 21. The KEGG GO term analysis was done on the DEGs (Figure 18A). This analysis showed enrichment for GO category coupled to fatty acid metabolic activity.

Mycobrowser, a database with genomic information specific to mycobacteria, was used to peruse the functions of diverse DEGs involved in fatty acid metabolism (Table 5). The data reveal that many of the genes upregulated by sncRNA-1 were coupled to fatty acid biogenesis. The results for 2 such genes, *desA2* (Rv1094) and *fabG4* (Rv0242c), was verified using q-RT-PCR (Figure 21). Rv1094 catalyzes the conversion of saturated fatty acids to unsaturated fatty acids whereas Rv0242c is involved in the first reduction step of oleic acid biogenesis. Overall, the data suggested that sncRNA-1 positively regulates fatty acid biogenesis in Mtb.

3.4 Discussion

In this chapter, a gain-of-function approach was used to study the function of sncRNA-1, sncRNA-6 and sncRNA-8. The overexpression of sncRNA-1 and sncRNA-6 clearly affected the mycobacterial transcriptome. KEGG GO term analysis on differentially expressed genes expressing sncRNA-1, sncRNA-6 or sncRNA-8 expression vector suggested that they may regulate fatty acid metabolism, stress response, and transcription, respectively. A heat map comparing 3 independent clones/construct revealed that sncRNA-1 and sncRNA-6 upregulated 9 and 13 mRNAs, respectively, compared to the control vector (Figure 17). Fifteen and 64 mRNAs were down-regulated in the presence of either sncRNA-1 and sncRNA-6 when compared to the

controls (Figure 17). In clones overexpressing sncRNA-6, the number of down-regulated genes was much higher than the number of upregulated genes, which may have indicated that it conferred a negative regulatory function. However, many of the DEGs in these clones encode hypothetical proteins of unknown function, which have complicated further functional analysis. In clones that contained sncRNA-8 expression vector, 38 genes were down-regulated, whereas 17 genes were upregulated (Figure 17). However, since the overexpression of the precursor sncRNA-8 transcript in Mtb failed to augment the mature sncRNA-8 expression (Figure 16A), DEGs in this group cannot be attributed to the overexpression of sncRNA-8. Therefore, difficulty in overexpressing sncRNA-8 in Mtb have limited the functional studies on this sncRNA.

More in-depth functional analyses have focused on sncRNA-1 since it was not identified in any prior transcriptome studies and it exists within a key pathogenicity locus, RD1. 25-nt mature sncRNA-1 was successfully overexpressed as verified with northern blotting and q-RT-PCR with the detection of weakly expressed unprocessed form. The use of LNA^{TM} miRNA-specific q-RT-PCR system was tested to detect sncRNA-1, which revealed that the probes used in this assay were specific for sncRNA-1. RNA-seq of the fast-growing auxotroph mutant of Mtb H37Rv 6230 overexpressing sncRNA-1 suggested a novel role for this sncRNA in the positive regulation of fatty acid metabolism. Many genes that were upregulated were involved in fatty acid biosynthesis, except for FadA2 and FadE23, which are involved in fatty acid degradation (Table 5). Noteworthy, DesA2 (Rv1094), which catalyzes the conversion of saturated fatty acids to unsaturated fatty acids, and FabG4 (Rv0242c), which is involved in the first reduction step of fatty acid biosynthesis, were upregulated. Paralleling this, many of the genes that were downregulated were involved in fatty acid degradation (Table 5), again implying that sncRNA-1 positively regulated components of the

fatty acid biogenesis pathway. Next chapter have focused on in-depth analysis of the role of sncRNA-1 in fatty acid metabolism.

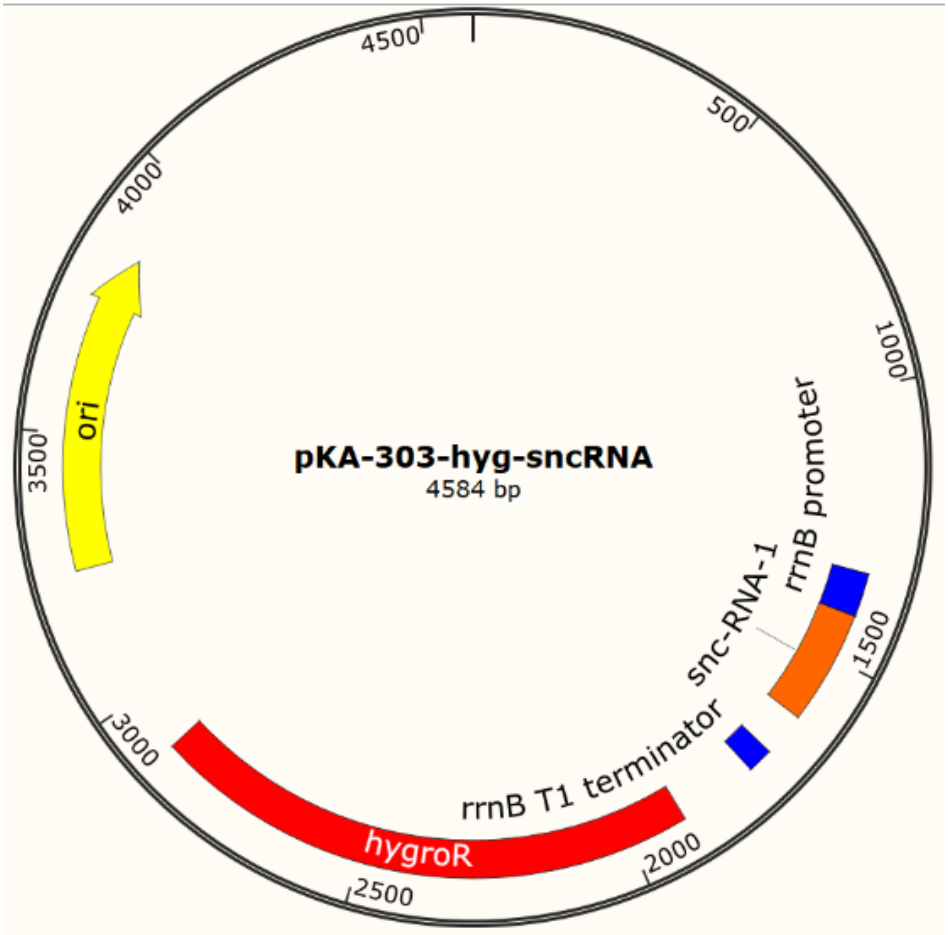


Figure 15. Plasmid map for the pKA-303 sncRNA expression vector. The plasmid was used to constitutively express the larger transcripts bearing sncRNA-1, sncRNA-6, or sncRNA-8 under the strong *rrnB* promoter. The DNA segments containing sncRNA-1, sncRNA-6 and sncRNA-8 were 208 nts, 333 nts, and 321 nts, respectively. Hygromycin was used for selection of the positive clones. The plasmid map shown is for sncRNA-1.

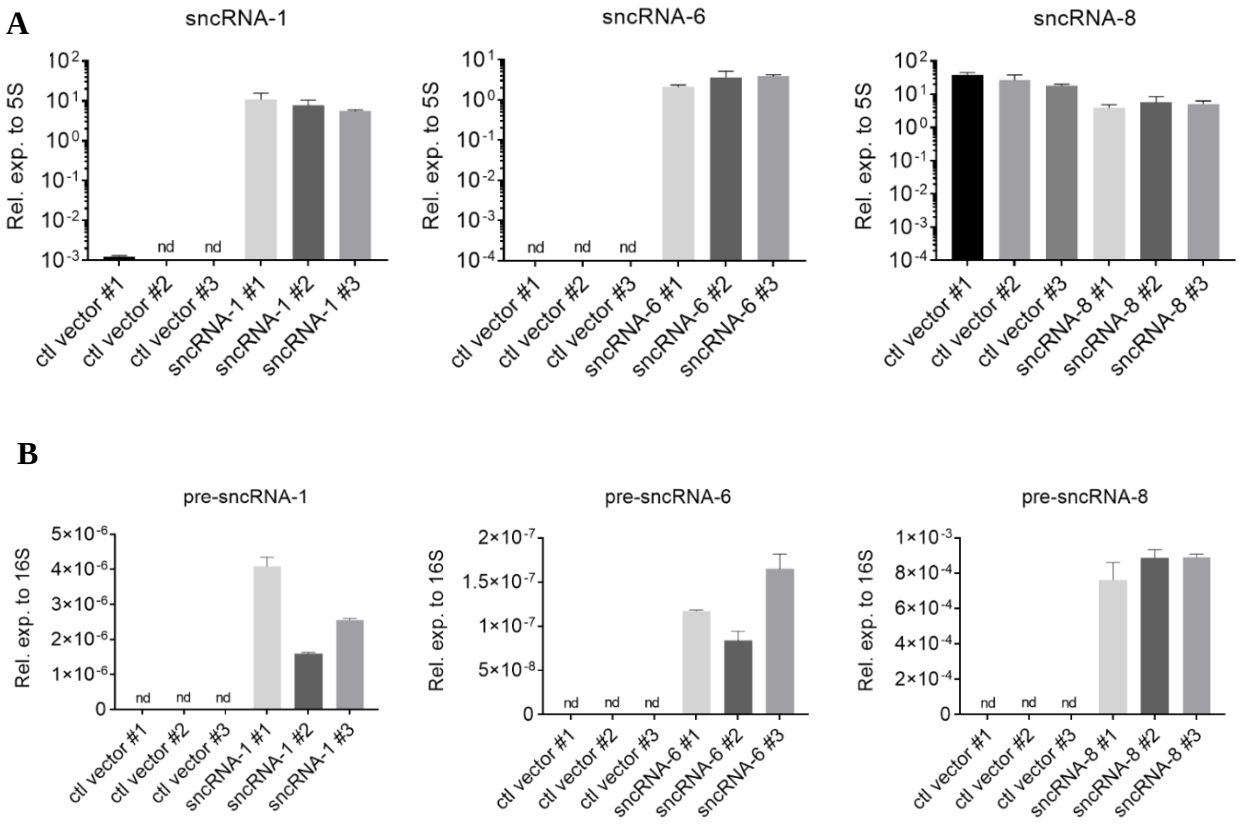


Figure 16. A q-RT-PCR assay establishes the successful overexpression of sncRNA-1 and sncRNA-6 in independently isolated clones of Mtb HR37Rv 6230. RNA was isolated from cultures grown in selective media at the stationary phase. (A) Exogenous sncRNA expression was quantified using the miRNA-specific q-RT-PCR assay relative to mycobacterial 5S RNA. (B) Precursor RNAs were quantified relative to 16S RNA using conventional q-RT-PCR assay. Nd indicates not detected.

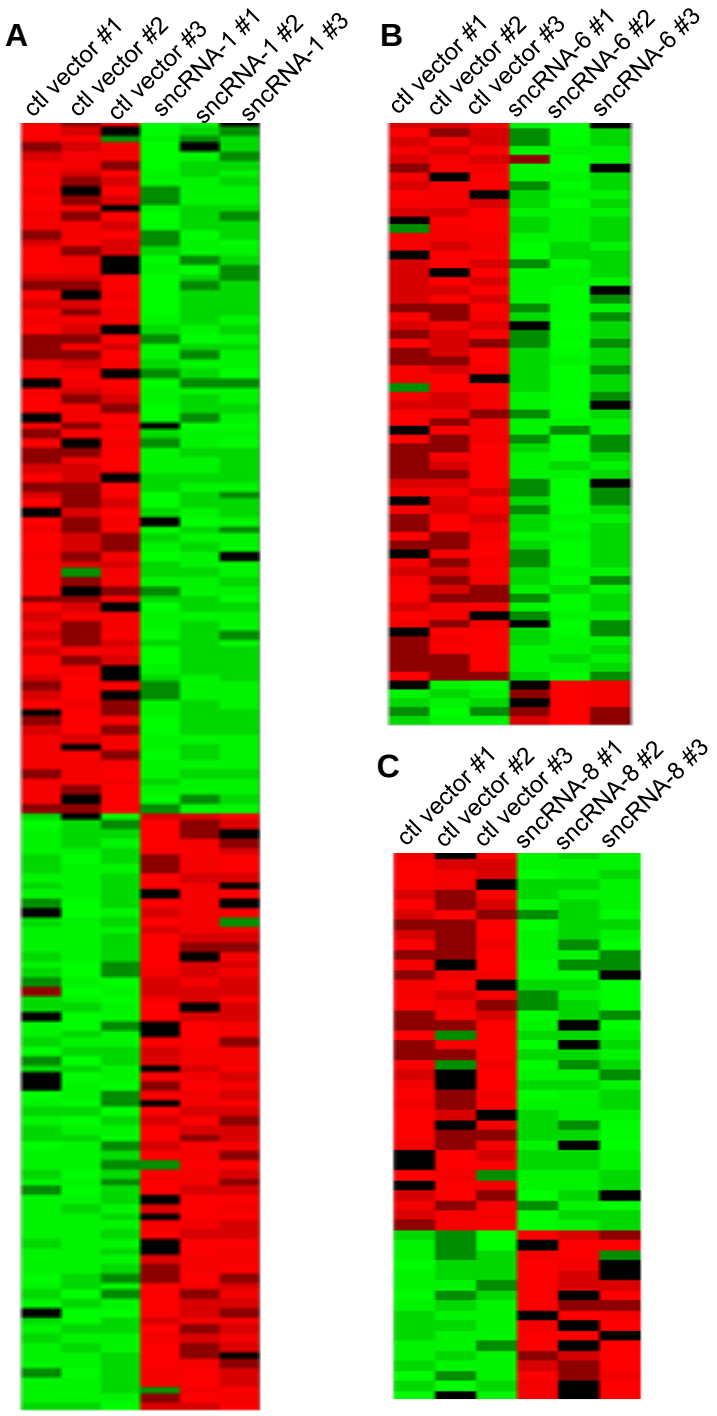

A
ctl vector #1
ctl vector #2
ctl vector #3
sncRNA-1 #1
sncRNA-1 #2
sncRNA-1 #3
B
ctl vector #1
ctl vector #2
ctl vector #3
sncRNA-6 #1
sncRNA-6 #2
sncRNA-6 #3
C
ctl vector #1
ctl vector #2
ctl vector #3
sncRNA-8 #1
sncRNA-8 #2
sncRNA-8 #3

Figure 17. Transcriptome analysis of Mtb H37Rv 6230 clones overexpressing sncRNA-1, sncRNA-6 or sncRNA-8 reveals distinct transcript patterns dependent on the sncRNA. RNA was isolated from 3 independent clones for each of the three sncRNAs that were over-expressed along with 3 independent clones for the vector control. RNA was isolated from each and sequenced. Genes differentially expressed, based on a log2 >1.5 fold difference, were compared along with those whose expression does not change between groups. Transcriptome analysis on differentially expressed genes revealed that different sets of genes were affected when (A) sncRNA-1, (B) sncRNA-6 or (C) sncRNA-8 was overexpressed compared to the control vector in Mtb H37Rv 6230.

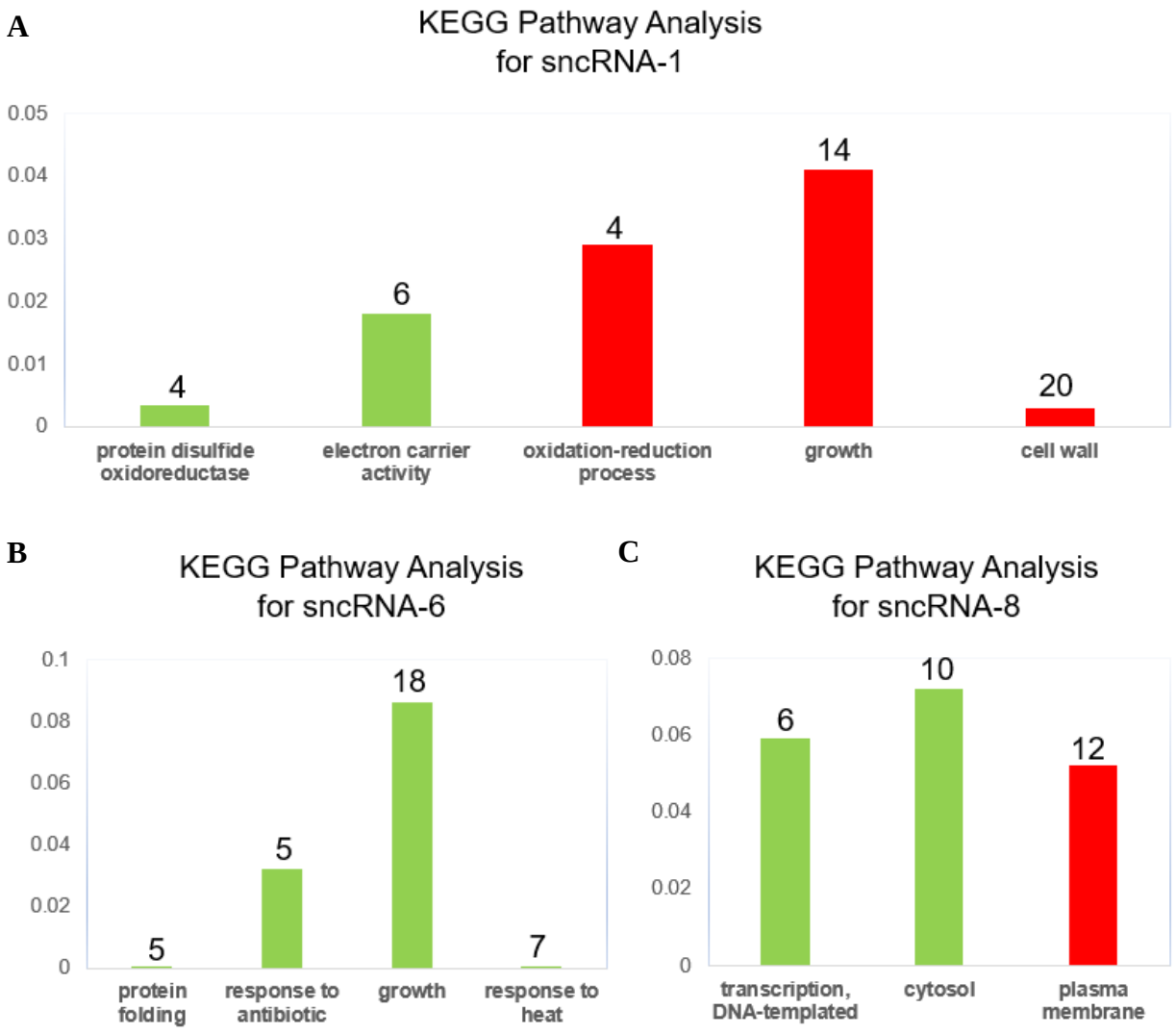


Figure 18. KEGG Pathway analysis on DEGs in Mtb H37Rv 6230 with sncRNA-1, sncRNA-6 or sncRNA-8 indicates that distinct pathways are affected by specific sncRNAs. KEGG pathway analysis was done using DAVID 6.8 online tool. Number of genes in each pathway is shown on top the bars. Red bars show over-representation, whereas green bars show underrepresentation.

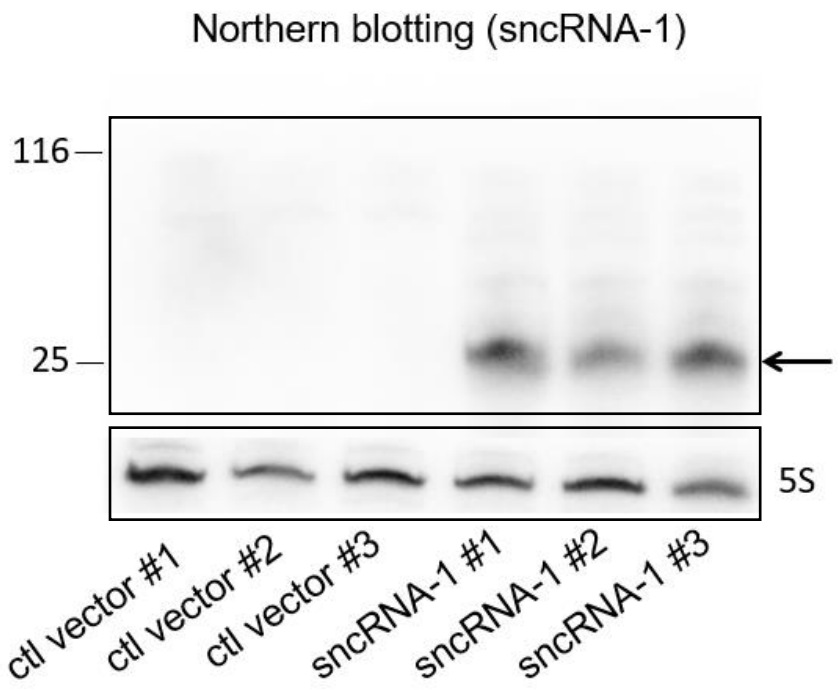


Figure 19. Northern blot for Mtb H37Rv 6230 clones expressing sncRNA-1 vector reveals 25-nt sncRNA-1 expression. sncRNA-1 expression vector the ctl vector was electroporated to Mtb H37Rv 6230 and selected clones were expanded. RNA was isolated from cultures at the stationary phase and the RNA was ran for a northern blotting. The membrane was probed for first sncRNA-1 later for 5S RNA. sncRNA-1 is marked with the arrow.

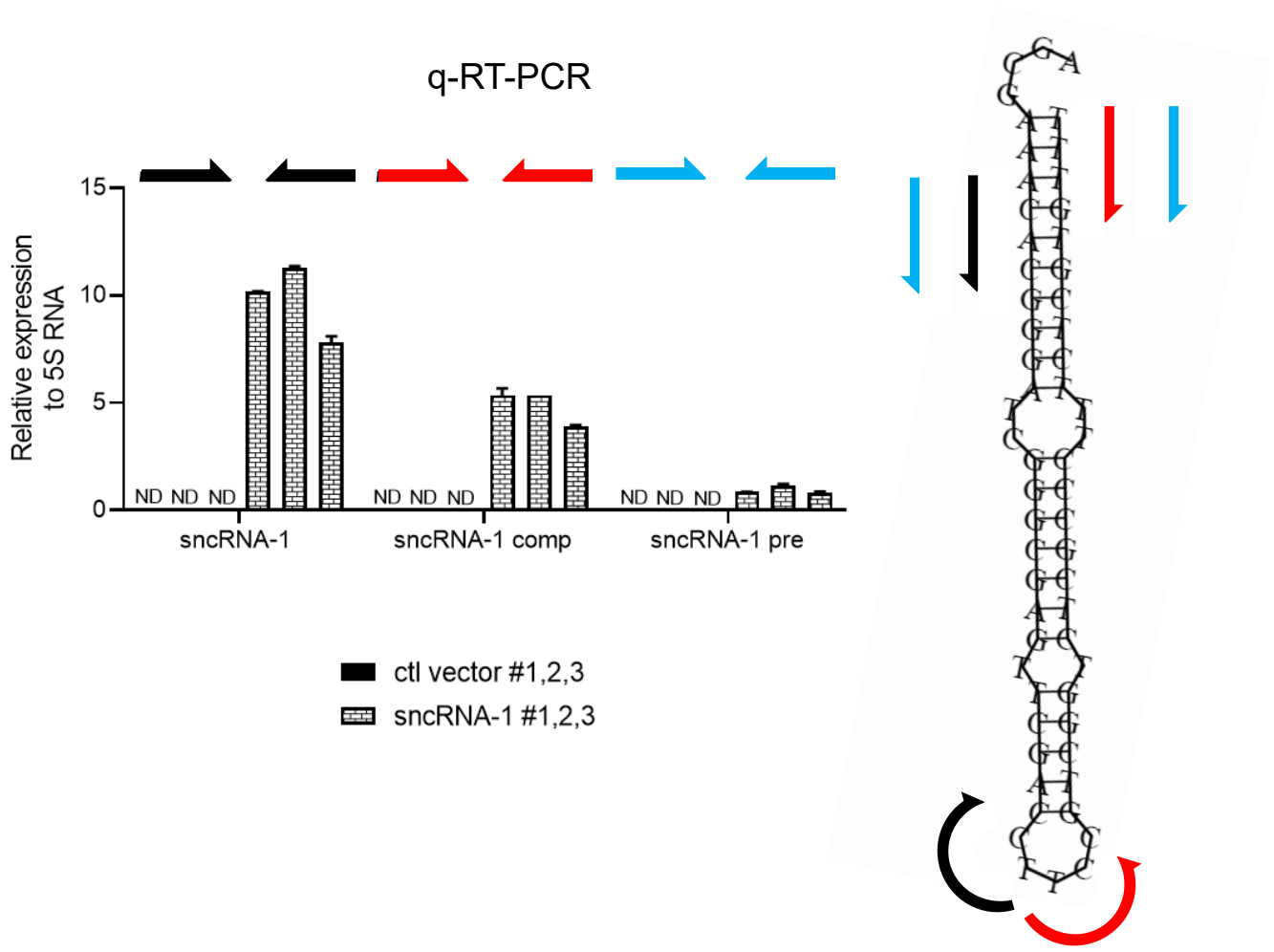


Figure 20. Quantification of sncRNA-1 using the LNA™ PCR probes demonstrates assay specificity. In three clones of Mtb H37Rv 6230 expressing either the ctl vector or the sncRNA-1 expression vector, sncRNA-1, its complement and a putative precursor was quantified using primers specific for these regions. The sncRNA-1 primers were designed for region shown between black arrows on the right. The primers for the sncRNA-1 complement (shown as sncRNA-1 comp on the graph) were designed for the region shown with the red arrows on the right. The primers for the putative sncRNA-1 precursor were designed for the entire region residing between blue arrows as shown on the right. The expression in each case was relative to the expression of the mycobacterial 5S RNA. Nd indicates not determined.

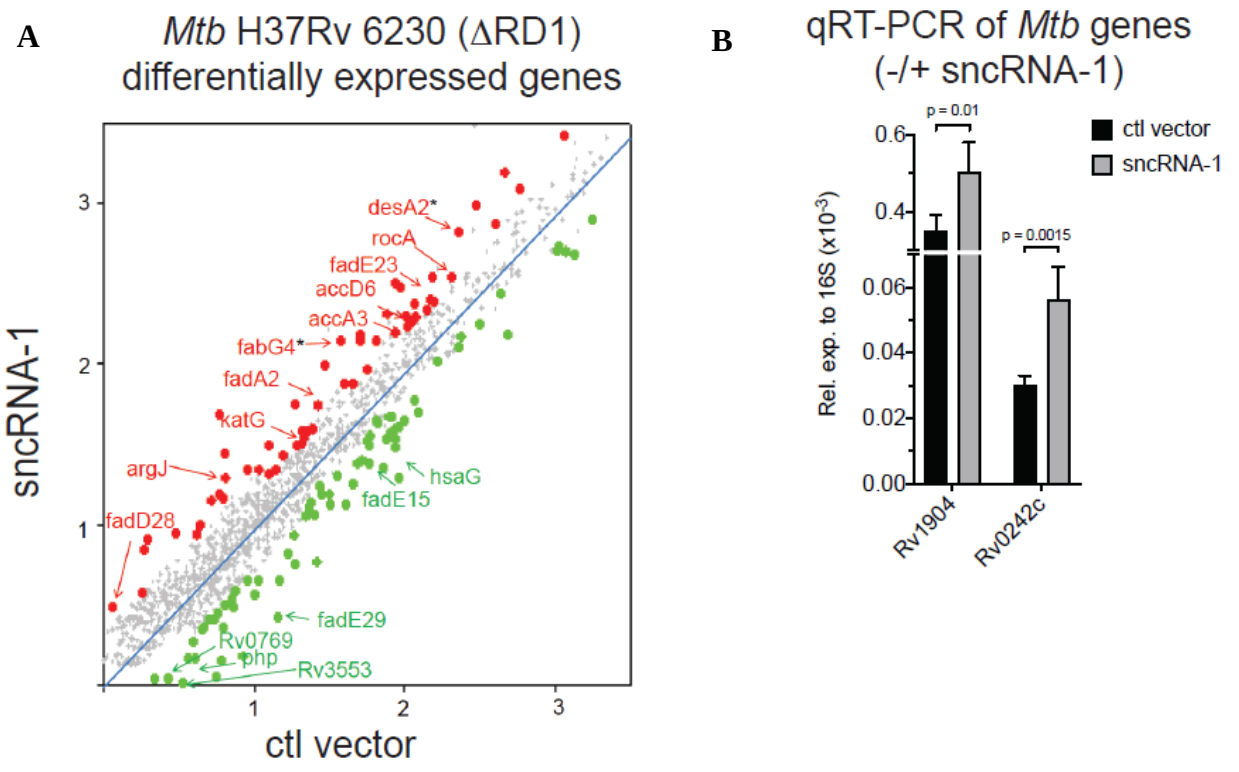


Figure 21. Overexpression of sncRNA-1 in Mtb H37Rv 6230 enhances Rv0242c and Rv1094 expression. RNA was isolated from Mtb H37Rv 6230 expressing either the ctl vector or the sncRNA-1 expression vector and then sequenced. The transcriptome of Mtb H37Rv 6230 with the sncRNA-1 vector was analyzed compared to the Mtb H37Rv 6230 with the ctl vector. The fold changes in the differentially expressed genes with $p<0.05$ was plotted. (A) The red indicated overexpressed genes, whereas the green indicated underrepresented gene in clones that had the sncRNA-1 expression vector compared to the ctl vector. (B) The expression of two of these genes, Rv1904 and Rv0242c, were verified with conventional q-RT-PCR. Statistical analysis was done with the Student's T-test. Two of the DEGs Rv1094 and Rv0242c were quantified with q-RT-PCR.

Table 5. Putative functions of the DEGs involved in fatty acid metabolism in Mtb 6230 overexpressing sncRNA-1

Status	Gene name	Function
up	desA2 (Rv1094)*	Thought to catalyze the principal **conversion of saturated fatty acids to unsaturated fatty acids**.
	accD6	Involved in **fatty acid biosynthesis** (mycolic acids synthesis)
	fabG4 (Rv0242c)*	Involved in the **fatty acid biosynthesis** pathway (first reduction step)
	fadD28	Involved in **phthiocerol dimycocerosate biosynthesis**. Thought to be involved in the release and transfer of mycoserosic acid from mas onto the DIOLS.
	fadA2	Function unknown, but involved in **lipid degradation**
	accA3	Involved in **long-chain fatty acid synthesis** (at the first step).
	fadE23	Function unknown but involved in **lipid degradation**.
	rocA	Involved in the **production of acetyl-CoA** in the beta-ketoadipate pathway (at the third step)
down	fadE15	Function unknown, but involvement in **lipid degradation**.
	echA19	Enoyl-CoA hydratase or crotonase is an enzyme that hydrates the double bond between the second and third carbons on 2-trans/cis-enoyl-CoA: ECH is essential to **metabolizing fatty acids in beta oxidation** to produce both acetyl CoA and energy in the form of ATP.

	hsaG	Supposed involvement in one, or several, **catabolic pathways** [catalytic activity: acetaldehyde + CoA + NAD(+) = acetyl-CoA + NADH]. Predicted to be involved in lipid catabolism.
	fadE29	Function unknown but involved in **lipid degradation**.
	php	Probable phosphotriesterase Php, involved in **catabolic reactions**

CHAPTER 4

sncRNA-1 positively regulates genes coupled to oleic acid production

4.1 Introduction

Bacteria including mycobacteria often express small RNAs (sRNAs) in response to stress. Historically, these sRNAs have been identified by sequencing cloned cDNA libraries and more often now, RNA-sequencing to obtain the entire mycobacterial transcriptome. However, both techniques have bias towards RNAs >50 nts and overlook sRNAs are <50 nucleotides (nts) since this population needs to be enriched prior to sequencing. In the previous chapter, a transcriptome analysis was done followed by a gene ontology analysis in *Mycobacterium tuberculosis* (Mtb) overexpressing sncRNA-1 vs control vector. This analysis revealed that sncRNA-1 may regulate fatty acid metabolism. In the current chapter, the role of sncRNA-1 is analyzed in more detail. Whether sncRNA-1 provides growth advantage to Mtb in cultures in the absence of fatty acids specifically oleic acid was tested. The focus on oleic acid was because a putative sncRNA-1 binding site was identified in the 5' UTR of the two genes (Rv0242c and Rv1094) associated with oleic acid biogenesis. Such seed-sequence complementary interactions are common for sRNAs. Second, the mechanism by which sncRNA-1 regulates these genes is investigated. There are multiple mechanisms by which a sRNA can regulate the expression of the target gene. One possibility is a direct interaction mediated by Watson-Crick base pairing between the sRNA and its target. Site-directed mutagenesis of putative binding sites between sncRNA-1 and Rv0242c was performed, which revealed that a direct interaction was required for such a regulation. Finally, the necessity and sufficiency of sncRNA-1 in infection was tested. My findings have shown that sncRNA-1 supported mycobacterial growth, partly via the positive regulation of enzymes

important for oleic acid synthesis. Overall, this chapter explores the role of sncRNA-1 in oleic acid production and its contribution to a successful infection.

4.2 Materials and Methods

4.2.1 Bacterial strains and growth conditions

Mtb H37Rv 6230 (ΔRD1, Δ*panCD*) was grown in the 7H9 supplemented with 10% oleic albumin dextrose catalase (OADC) (Remel, Thermo-Fisher) and 5% glycerol (G9012, Sigma- Aldrich), 0.05% tyloxapol (T0307, Sigma-Aldrich), 0.2% casamino acids (223050, BD Bacto™) and 24 µg/ml pantothenate (D-Pantothenic acid hemicalcium salt, Sigma-Aldrich). The media for Mtb H37Rv 6206 (ΔleuC, Δ*panCD*) was similar to that used for 6230 with an additional supplement of 80 ug/ml leucine (Sigma-Aldrich). Bone marrow derived macrophages (BMDM) were used for the infection assays. BMDM were obtained from the femur, tibia, and spinal column of the C57BL/6 mice. Macrophages were differentiated using macrophage media, 30% of L929 supernatant (ATCC CCL-1^{TM}) mixed with 70% of RPMI containing 10% FCS (Weischenfeldt & Porse, 2008). Briefly, red blood cells were lysed, and the remaining cells cultured in macrophage media overnight. Non-adherent cells were collected and cultured for four days in macrophage media. On day 4, cells were detached from the culture plates and stained with an antibody detecting CD11b (Catalog # ab24874, Abcam). After confirming positive staining for the CD11b marker, the BMDMs were plated for infection experiments in RPMI media containing 10% FCS.

4.2.1 Cloning strategies

Rv0242c was cloned into a protein expression vector called pSUM-Kan-MCS1-gfp (Eitson et al., 2012) (Figure 22). The kanamycin cassette was replaced with a zeocin resistance cassette while GFP was replaced with the Rv0242c gene along with its 5’ UTR. Subcloning was done using NEBuilder® HiFi DNA Assembly Master Mix. Plasmids expressing control vector or sncRNA-1

and Rv0242c were electroporated into Mtb H37Rv 6230 in combination. Transformants were selected in the presence of 100 µg/ml hygromycin and 100 µg/ml zeocin. RNA was isolated and Rv0242c expression was quantified with a q-RT-PCR assay, using the 16S rRNA as a reference gene.

4.2.2 Site-directed mutagenesis strategies to mutate precursor sncRNAs

For site-directed mutagenesis, primers were designed using NEBasechanger™ tool. The proposed binding site in the 5' UTR of Rv0242c in the expression vector (pSUM-Zeo-Rv0242) was PCR amplified with primers designed for site directed mutagenesis. PCR product was treated with KDL mixture (NEB cat #0554) for 30 mins 37ºC and transformed into *Escherichia coli* (*E. coli*) DH5alpha. Clones were randomly selected and propagated. Plasmids were isolated with ZymoPURE™ Plasmid Miniprep Kit (Cat #D4209, Zymo Corp.) and sequenced by the McDermott Sanger Sequencing Core at UT Southwestern Medical Center. The plasmids with the desired mutations were electroporated to Mtb H37Rv 6230. The clones that were selected were confirmed to contain the proper sequences. At least 3 clones/construct were selected for further study. RNA was extracted and Rv0242c expression was quantified using a q-RT-PCR assay relative to 16S rRNA.

4.2.3 Mtb growth assays

Mtb H37Rv 6230 clones containing the control vector, or the sncRNA-1 vectors were grown in 7H9 supplemented with 100 µg/ml hygromycin and 5% glycerol or with 5% glycerol and 10% OADC. Optical density at 600 nm was measured in 96 well plates in triplicates every 24 hours for 7 days. In some experiments, Mtb H37Rv 6230 clones expressing the control vector or sncRNA-1 vector were grown in 7H9 supplemented with 100 µg/ml hygromycin, 5% glycerol and 10% ADC or 10% OADC.

4.2.4 sncRNA loss-of-function assays

A locked nucleic acid power inhibitor (LNA-PI) targeting Mtb-encoded sncRNA-1 was designed and synthesized by Exiqon Inc., based on their proprietary nucleotide modifications (Exiqon Inc., Woborn, MA, now part of Qiagen Corp.). A control LNA that does not target any known eukaryotic gene was included as a control (5'-acgtctatacgccca-3'). A fluorescent tag was covalently attached at the 5' end of the inhibitor to monitor uptake. LNA-PI against sncRNA-1 (350 nM final concentration) was used to stain Mtb H37Rv 6230 or 6206 by incubating the cells for 16 hours or 3 days, respectively. The mycobacteria were subsequently fixed in 2% PFA and fluorescence was detected by flow cytometry. For CFU experiments, BMDMs were infected at a MOI of 3:1. Infected BMDMs were lysed in 0.05% SDS at different time points post infection and Mtb was plated on 7H10 agar plates containing 10% OADC, along with other supplements needed for the survival of the auxotroph mutants. The number of bacterial counts was measured to compare the total bacterial burden.

4.2.5 Mtb infection assays with macrophages

BMDMs were infected with Mtb H37Rv 6206 or Mtb H37Rv 6230 clones expressing either the ctl vector or sncRNA-1 vector at MOI of 3:1. Three hrs. post-infection BMDMs were washed and cultured for indicated time courses. For CFU enumeration, the infected cells were lysed in PBS containing SDS (0.05% v/v). Ten-fold serial dilutions were made in the same buffer, and samples were spread on 7H10 agar plates containing 10% OADC and other supplements required for the auxotroph mutants to determine the total bacterial burden.

4.3 Results

4.3.1 Elevated levels of sncRNA-1 support Mtb growth in oleic acid deficient growth media.

The transcriptome analysis on Mtb H37Rv 6230 clones overexpressing sncRNA-1 versus the control vector suggested that sncRNA-1 regulated fatty acid metabolism through the genes involved in oleic acid biogenesis in Mtb. This would imply that sncRNA-1 could provide growth advantage to Mtb H37Rv 6230 if grown in OADC deficient media. OADC refers to Oleic acid, along with Albumin, Dextrose, and Catalase, which is a part of the complete growth media used to maintain Mtb in the laboratory. To test the role of sncRNA-1 on the growth in OADC deficient media, Mtb H37Rv 6230 clones with either the sncRNA-1 vector or control vector were grown in the presence of OADC (7H9/OADC^{+}), or in the absence of OADC (7H9/OADC^{-}). Mtb H37Rv 6230 was used as it has a short doubling time, and it lacks the region of difference (RD1) region where sncRNA-1 resides. As shown in Figure 23, clones expressing ctl vector or sncRNA-1 vector had similar growth kinetics in the presence of OADC. However, clones expressing the ctl vector had growth defect in OADC deficient media, whereas clones expressing sncRNA-1 vector did not (Figure 23). This experiment revealed that sncRNA-1 positively regulates Mtb growth in media lacking OADC.

4.3.2 sncRNA-1 positively regulates the expression of Rv0242c, a gene required for oleic acid biogenesis

sRNAs can modulate gene expression through post transcriptional gene regulation. The sRNAs transcribed from intergenic regions often regulate the expression of targets using a 6-7 nts long seed sequence that has partial complementarity (K. Arnvig & Young, 2012). Since sncRNA-1 is an intergenic RNA, we reasoned it may have similar targeting properties of a sRNA. A homology search was done to identify genes that could be targeted by the seed sequence of sncRNA-1. A sequence among different mycobacteria species containing sncRNA-1 were compared, revealing a highly conserved 8-nt region. Using this sequence as a probe for the sRNA target prediction

program TargetRNA2, 39 putative sncRNA-1 targets were identified. Two targets included Rv1094 (*desA2*) and Rv0242c (*fabG4*), both having a putative sncRNA-1 binding site in their 5' UTR. Both with two other genes form a pathway involved in the oleic acid production in Mtb (available online: (https://www.genome.jp/kegg-bin/show_pathway?map00061) (Figure 24). Moreover, both genes were upregulated in clones overexpressing sncRNA-1 (Figure 22). Overall, these data suggest that sncRNA-1 may be targeting them using a seed sequence. This possibility is supported by the growth advantage observed in OADC- deficient media as sncRNA-1 upregulates genes involved in de novo oleic acid production (Figure 23).

To test this hypothesis, Mtb H37Rv 6230 clones expressing the control vector, or the sncRNA-1 vector were grown in 7H9 media containing Albumin, Dextrose, and Catalase in the presence of Oleic acid (7H9-ADC/oleic acid^{+}) or in the absence of Oleic acid (7H9-ADC/oleic acid^{-}). Consistent with previous findings, clones with control vector or sncRNA-1 had comparable growth kinetics in the presence of oleic acid (Figure 25). However, clones expressing sncRNA-1 had a growth advantage compared to clones expressing just the control vector in the absence of oleic acid (Figure 25). Of note, there was some growth in the clones with control vector in the absence of oleic acid. This is probably due to the presence of the remaining OADC supplements in the media, Albumin, Dextrose, and Catalase. In summary, the data support our conclusion that sncRNA-1 positively regulates the oleic acid biogenesis in Mtb, which promotes Mtb growth in media lacking oleic acid.

4.3.3 sncRNA-1 regulates Rv0242c transcript levels using a 7-nt seed sequence interacting with the 5' untranslated region of the gene

After identifying putative targets, additional experiments were undertaken to confirm that sncRNA-1 directly regulated the 5' UTR segments of the target genes. First, sncRNA-1 was

expressed with one of the targets, Rv0242c, in Mtb H37Rv 6230. A 5' UTR upstream region and the coding segment of Rv0242c were exogenously expressed along with mutations in the 5'UTR that eliminate the sncRNA-1 binding site. As negative controls, the control vector and a construct overexpressing an unrelated sncRNA, sncRNA-6 were used. The clones were grown in double selection media and the expression of Rv0242c was quantified with RNA isolated from stationary phase cultures. Expression of Rv0242c was induced 3-fold in the presence of sncRNA-1 but not the sncRNA-6 nor the control vector (Figure 26). This experiment suggests sncRNA-1 positively regulates the expression of Rv0242c by targeting the 5' UTR of this gene. To test if a direct interaction between the sncRNA-1 seed sequence and its binding site in the 5' UTR of Rv0242c, site-directed mutagenesis experiments were designed to disrupt this interaction. First, the sncRNA-1 binding site in the 5' UTR of Rv0242c was mutated and two mutants were created and named as Rv0242c_M1 and Rv0242c_M2 (Figure 27). In Rv0242c_M1, all 8 nts were mutated while maintaining the overall predicted RNA secondary structure when compared to the wild type. When expressed along with the sncRNA-1, the expression of Rv0242c in Rv0242c_M1 was no longer induced, establishing that the 5'UTR sequence was critical for the positive induction of this gene (Figure 26). In Rv0242c_M2, 4 of 8 nucleotides were mutated, which caused a bending of the 5' UTR of Rv0242c as suggested by RNA secondary sequence predictions. Interestingly, the exogenous expression of Rv0242c in Rv0242c_M2 was eliminated completely, regardless of the presence or absence of sncRNA-1. This finding established that the organization of the 5'UTR has an essential role in the expression of Rv0242c. The normal 5'UTR sequence has a predicted ribosomal binding site (AAGG), and our data suggest that sncRNA-1 interacts with this region to improve the expression of Rv0242c.

To restore the upregulation of the Rv0242c with the 5'UTR mutated at the 8 nts (Rv0242c_M1), the seed sequence in sncRNA-1 was re-engineered to enable direct base pairing with the new sequence in Rv0242c_M1 (Figure 28). Of note, several nucleotides in the pre-sncRNA-1 were also mutated to preserve the predicted hairpin loop structure so that this sncRNA would be generated as a 25-nt species. Named sncRNA-1_mut, this 25-nt new sRNA was confirmed to be overexpressed as a mature sncRNA, as verified using the LNA^{TM} quantitative miRNA PCR system with primers designed for sncRNA-1_mut (Figure 29). Experimentally, sncRNA-1_mut or the wild-type sncRNA-1 was co-expressed with Rv0242c_M1. Rv0242c_M1 was induced in the presence of sncRNA-1_mut but not the wild-type sncRNA-1 (Figure 30). These experiments support the notion that a direct interaction between the sncRNA-1 seed sequence and the 5' UTR of Rv0242c is required for the upregulation of the transcript.

To further support this conclusion, a novel reporter assay was developed. The 5'UTR promoter region of the Rv0242c was cloned upstream of the *GFP* gene (R_5'U_GFP). It was co-transformed in Mtb H37Rv 6230 with either with the control vector or sncRNA-1 vector (Figure 31). When quantified, medium fluorescent intensity was 1.5-fold higher in clones having the sncRNA-1 vector compared to the control vector further showing that sncRNA-1 induces the expression of the target through the 5' UTR region.

4.3.4 The 25-nt sncRNA-1 directly targets Rv0242c

An alternate explanation to the findings is that a larger RNA transcript containing sncRNA-1 could have upregulated Rv0242c. To address this, the predicted hairpin loop structure of pre-sncRNA-1 was mutated, creating two distinct mutants termed sncRNA-1_L1 and sncRNA-1_L2. The mutations were designed to change the predicted RNA secondary structure to prevent processing of sncRNA-1. When plotted with the RNAfold web server, sncRNA-1_L1 and sncRNA-1_L2

were predicted to have larger loops (Figure 32). To test whether these mutants would impact the expression of Rv0242c, each was independently expressed, with several clones/construct studied. RNA was isolated from stationary phase cultures. sncRNA-1 expression levels were at least 100-fold less with either sncRNA-1_L1 or sncRNA-1_L2 compared to the wild-type positive control (Figure 33A). This was verified by northern blotting with a sncRNA-1 specific probe (Figure 33B). The enhanced expression of Rv0242c with sncRNA-1 was no longer observed in the presence sncRNA-1_L1 or sncRNA-1_L2 (Figure 34). These data strongly suggest that mature sncRNA-1 is responsible for the induction of Rv0242c, However, we cannot rule out the potential contribution of the unprocessed RNA transcript containing the sncRNA-1 sequence. Moreover, these findings also suggest the presence of a nucleotide sequence/stem loop-dependent RNA processing system in Mtb.

4.3.5 Overexpression of sncRNA-1 enables Mtb survival in macrophages

Oleic acid is an unsaturated fatty acid that can be produced by Mtb in infected cells to support mycobacterial survival, suggesting that the increased levels of sncRNA-1 could contribute to Mtb pathogenicity. To assess this, various infection assays were set up with Mtb H37Rv auxotrophic mutants overexpressing sncRNA-1. First, BMDM was infected with Mtb H37Rv 6230 (ΔRD1, Δ*panCD*) clones expressing either the control vector or sncRNA-1 expression vector at MOI of 3:1. The bacterial burden was quantified at 3 hours and 3 days post infection. Although initial CFU counts were comparable, Mtb H37Rv 6230 clones expressing the control vector had a severe survival defect due to the lack of the RD1 region at 3 days post infection (Figure 35). This was partially rescued by the expression of sncRNA-1, revealing a role for this sncRNA in mycobacterial survival.

The assay was repeated with Mtb H37Rv 6206 auxotroph (Δ*panCD*, Δ*leuC*). This auxotroph retains the RD1 region hence making it more comparable to the wild-type Mtb H37Rv in terms of the growth kinetics (slow growing mutant) in the supplemented media. Therefore, BMDM were infected with Mtb H37Rv 6206 expressing either the control vector or the sncRNA-1 vector at MOI of 3:1. The bacterial burden was quantified 3 hours or 6 days post infection. Like Mtb H37Rv 6230, the bacterial burden was comparable at 3 hours post infection. However, at 6 days post infection, Mtb H37Rv 6206 expressing the control vector had a small survival defect, with the overexpression of sncRNA-1 providing some growth survival advantage, although the sample size did not enable significance (Figure 35). Together, these two assays revealed that sncRNA-1 supports Mtb survival in infected cells, with a more prominent effect noted in the absence of the RD1 region.

4.3.6 Locked nucleic acid power inhibitors antagonizing sncRNA-1 result in diminished Mtb survival in infected cells

To complement the overexpression studies, a loss-of-function assay was designed for sncRNA-1. To do that, locked nucleic acid power inhibitors (LNA-PIs) were used. LNA-PIs are 14-16 nts antisense RNAs that antagonize the miRNAs (Jepsen & Wengel, 2004). Their backbone is modified which makes them nuclease resistant and transfection independent. This process is known as gymnosis and the molecular processes governing it are unknown (Stein et al., 2010). In addition to the LNA targeting sncRNA-1 (LNA-PI-1), a negative LNA-PI not targeting any known gene in Mtb or humans (LNA-Ctl) was also tested. These LNA-PIs were screened in Mtb H37Rv 6206, which retains the sncRNA-1 or Mtb H37Rv 6230 overexpressing sncRNA-1 (Mtb H37Rv 6230 OE). First, the uptake of the LNA-PIs was verified with flow cytometry. Mtb H37Rv 6206 and Mtb H37Rv 6230 OE were incubated with LNA-PI-1 for three days or overnight, respectively

and then analyzed with flow cytometry. Mtb incorporated LNA-PI-1 without any transfection requirements (Figure 36).

BMDM was infected with Mtb H37Rv 6206 or Mtb H37Rv 6230 OE, previously treated with LNA-Ctl or LNA-PI-1. Three- and six-days post-infection, RNA was isolated to determine the targeting efficiency against sncRNA-1. miRNA-specific qRT-PCR revealed that the levels of sncRNA-1, but not sncRNA-8 were reduced in Mtb H37Rv 6206 and Mtb H37Rv 6230 OE treated with LNA-PI-1 compared to LNA-Ctl (Figure 37). To determine if this reduction was due to the degradation of sncRNA-1 or whether sncRNA-1 was simply hybridized to LNA-PI-1, northern blotting was performed. sncRNA-1 levels detected under conditions where the LNA-PI-1 or LNA control were used were comparable, revealing that the LNAs do not promote sRNA degradation (Figure 38A). The reduced detection of sncRNA-1 by LNA-PI-1 could also be explained by the LNA-PIs interfering with the q-RT-PCR where LNA-based primers are used to detect the miRNAs. To test this, RNA was isolated from Mtb H37Rv 6230 overexpressing sncRNA-1 and subsequently incubated with increasing amounts of LNA-PI-1 prior to cDNA synthesis and quantification. Detection of sncRNA-1 was reduced as the amount of LNA-PI-1 increased in the reaction (Figure 38B). These analyses revealed that LNA-PI-1 reduced the detection of sncRNA-1 in the q-RT-PCR without inducing its degradation.

Given the potential of LNA-PI-1 to functionally antagonize sncRNA-1, we examined the consequences during infections. BMDM were infected with Mtb H37Rv 6206 or Mtb H37Rv 6230 OE that had been pre-cultured with LNA-Ctl or LNA-PI-1. The bacterial burden was quantified at 3 hours for both strains and 6 days or 3 days post infection, for Mtb H37Rv 6206 or Mtb H37Rv 6230 OE, respectively. Initial bacterial burden, determined at 3 hours post infection, was comparable between LNA-Ctl vs LNA-PI-1 in either Mtb H37Rv 6206 or Mtb H37Rv 6230 OE

(Figure 39). At 6-days post-infection, Mtb H37Rv 6206 cultures incubated with the LNA-PI-1 had a more than 2-fold decrease in CFU counts compared to the LNA-Ctl.

We next examined the impact of antagonizing sncRNA-1 on the expression of Rv0242c transcript. Mtb H37Rv 6230 that had a Rv0242c expression cassette was incubated with either the LNA-PI-1 or the LNA-Ctl (Figure 40). RNA was isolated and the expression of Rv0242c was quantified. The expression of Rv0242c was reduced compared to the H37Rv 6230 OE treated with LNA-Ctl (Figure 40). This finding further confirms that sncRNA-1 positively regulates Rv0242c.

4.4 Discussion

My findings indicate that sncRNA-1 directly and positively regulates the expression of Rv0242c transcript. Rv0242c encodes a noncanonical 3-oxoacyl-thioester reductase, which transfers a hydride group from NADH to reduce the β-oxoacyl to β-hydroxyacyl (Dutta, 2018). This is the first reduction step in the production of oleic acid in Mtb. Oleic acid is an unsaturated fatty acid found in the Mtb membrane which contributes to maintenance of the optimal fluidity (Schaefer & Lewis, 1965; Supply et al., 2013). Moreover, oleic acid is required to produce mycolic acid which is the primary fatty acid in the Mtb membrane (Srivastava et al., 2015). Inhibition of the oleic acid production results in Mtb death (DeJesus et al., 2017). Therefore, antibiotics suppressing enzymes involved in oleic acid production would be promising drugs. Our studies show that sncRNA-1 positively regulates genes involved in oleic acid production via a direct interaction. Thus, drugs targeting sncRNA-1 production could provide a novel approach at targeting Mtb in infected patients.

sncRNA-1 uses its seed sequence to target the 5'UTR of Rv0242c. Reciprocal mutagenesis of these regions abrogated the regulation on Rv0242c. In bacteria, transcription is initiated by the recognition of a -10 element which has a specific sequence in the promoter region by a specific

sigma factor bound to the RNA Polymerase (Kumar, 1981). Then, the sigma factor is released which allows the elongation of the transcription. This process yields an mRNA with a short 5' UTR. The 5'UTR has important elements for translation as it encompasses the ribosomal binding site (RBS) (Alifano, Bruni, & Carlomagno, 1994; Van Assche, Van Puyvelde, Vanderleyden, & Steenackers, 2015). RNAfold was used to predict the secondary structure of the 5' UTR of Rv0242c and the mutants that were generated. While the 5' UTR of Rv0242c_M1 was comparable to the 5' UTR of the wild-type Rv0242c, a bending in the 5' UTR of Rv0242c_M2 was predicted. Therefore, it is possible that the excess of the ribosomal machinery is blocked in Rv0242c_M2, which is why the exogenous expression of the Rv0242c transcript is abrogated in this mutant. I hypothesize that sncRNA-1 stabilizes the structure of this region, enhancing expression of Rv0242c.

In this chapter the role of sncRNA-1 in infection was explored with gain- and loss- of function assays. First, sncRNA-1 was overexpressed in either Mtb H37Rv 6206 (Δ*panCD*, Δ*leuC*) or Mtb H37Rv 6230 (ΔRD1, Δ*panCD*). This increased the CFUs/well in infected BMDM, which was more prominent in Mtb H37Rv 6230 (3 days p.i.) compared to Mtb H37Rv 6206 (6 days p.i.). Mtb H37Rv 6230 is a fast-growing version of Mtb for unknown reasons, and has a 3-4 hours of division cycle, whereas Mtb H37Rv 6206 is slow growing and has 22-24 hours of division cycle. Therefore, 3 days at which CFUs/well was determined for Mtb H37Rv 6230 corresponds to around 18 division cycle whereas 6 days at which CFUs/well was determined for Mtb H37Rv 6206 corresponds to around 6 division cycle. This difference in their growth kinetics may explain the difference in the significance of the CFUs/well for these two strains. Another explanation is related to their genetic differences. Mtb H37Rv 6230 is lacking the RD1 region which is important for Mtb virulence and is more susceptible to killing by the macrophages. Genes in the RD1 region

include ESAT-6 and CFP-10 which are implicated in the pore formation within phagolysosomes (Gao et al., 2004; Smith et al., 2008). These pores are thought be required for the secretion of the virulence factors, but it is possible that they may be involved in the import of certain nutrients such as fatty acids. Mtb H37Rv 6230 overexpressing sncRNA-1, the lack of these import function may be rescued by the upregulation of the genes involved in oleic acid biogenesis. This is an interesting concept because it could allow macrophages to impose nutritional immunity as inside macrophages, Mtb relies on the host fatty acids. It can be speculated that the upregulation of sncRNA-1 may oppose such a mechanism. The kinetics of oleic acid metabolism and its import inside macrophages is a topic that needs to be studied in more detail.

To complement the overexpression studies, the sncRNA-1 was also antagonized before infecting the BMDM in either Mtb H37Rv 6206 (Δ*panCD*, Δ*leuC*) or Mtb H37Rv 6230 (ΔRD1, Δ*panCD*) overexpressing sncRNA-1 as this strain is lacking the endogenous sncRNA-1. This significantly reduced the bacterial burden on macrophages. In Mtb H37Rv 6230 overexpressing sncRNA-1, there was a less dramatic decrease in the CFU counts compared to Mtb H37Rv 6206. This may be explained by the fact that Mtb H37Rv 6230 has a much faster doubling time, which might have caused a dilution in LNA-PIs. It is also possible that the survival phenotype that was observed in these assays is not solely related to the oleic acid metabolism and there may be additional targets of sncRNA-1. Overall, the data show that sncRNA-1, a novel smaller noncoding RNA contributes to Mtb pathogenicity.

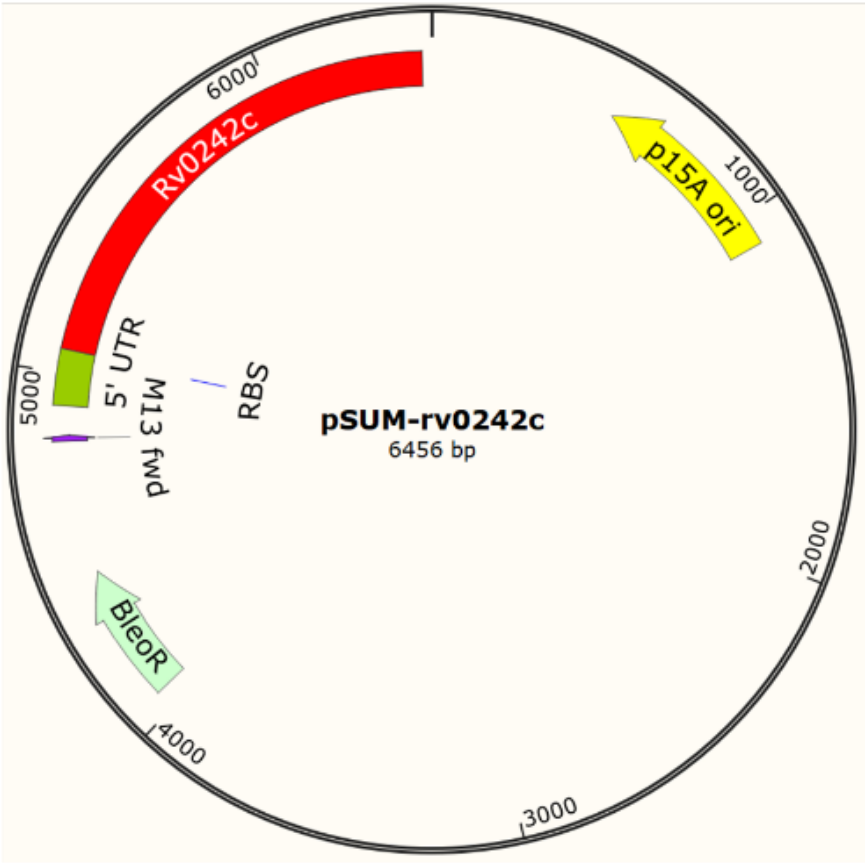


Figure 22. Plasmid map for the protein expression vector pSUM-zeo-rv0242c used to express Rv0242c in Mtb H37Rv 6230.

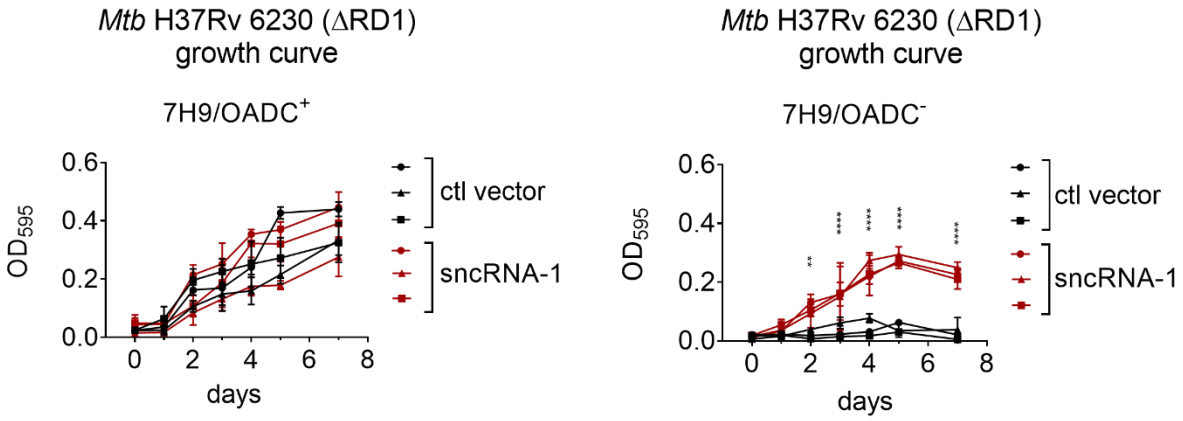


Figure 23. Overexpression of sncRNA-1 confers a growth advantage to Mtb H37Rv 6230 grown in OADC deficient media. sncRNA-1 expression vector or the ctl vector was electroporated to Mtb H37Rv 6230 and cultures selected on antibiotics were expanded. Cultures in the stationary phase were washed with media lacking OADC and diluted to expand in the presence or absence of OADC. The growth of the bacteria was monitored for 7 days by measuring OD at 595 nm. The statistical significance was calculated using paired T-test. *P<0.05, **p<0.01, ***p<0.001, ****p<0.0001

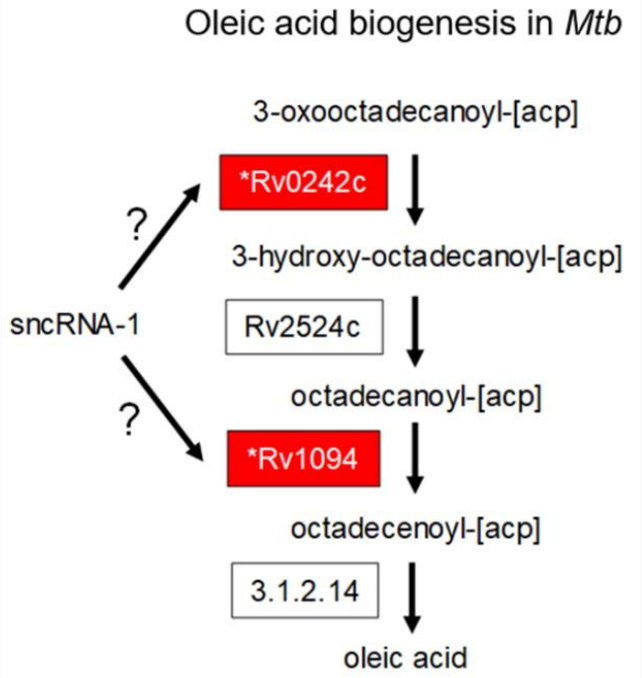


Figure 24. Oleic acid biogenesis pathway in Mtb based on the KEGG mycobacterial fatty acid biogenesis pathway. Two putative targets of sncRNA-1 which are Rv0242c and Rv1904 are shown in red.

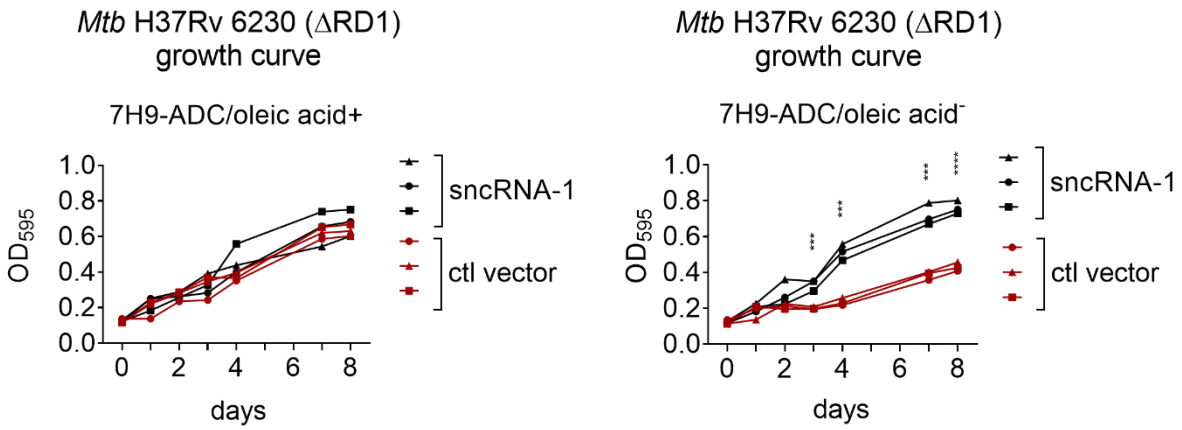


Figure 25. Overexpression of sncRNA-1 confers growth advantage to Mtb H27Rv 6230 grown in the absence of oleic acid. sncRNA-1 expression vector or the ctl vector was electroporated to Mtb H37Rv 6230 and cultures selected on antibiotics were expanded. Cultures in the stationary phase were washed with media lacking OADC and diluted to expand in the presence or absence of oleic acid. The growth of the bacteria was monitored for 8 days by measuring OD at 595 nm. The statistical significance was calculated using paired T-test. *P<0.05, **p<0.01, ***p<0.001, ****p<0.0001

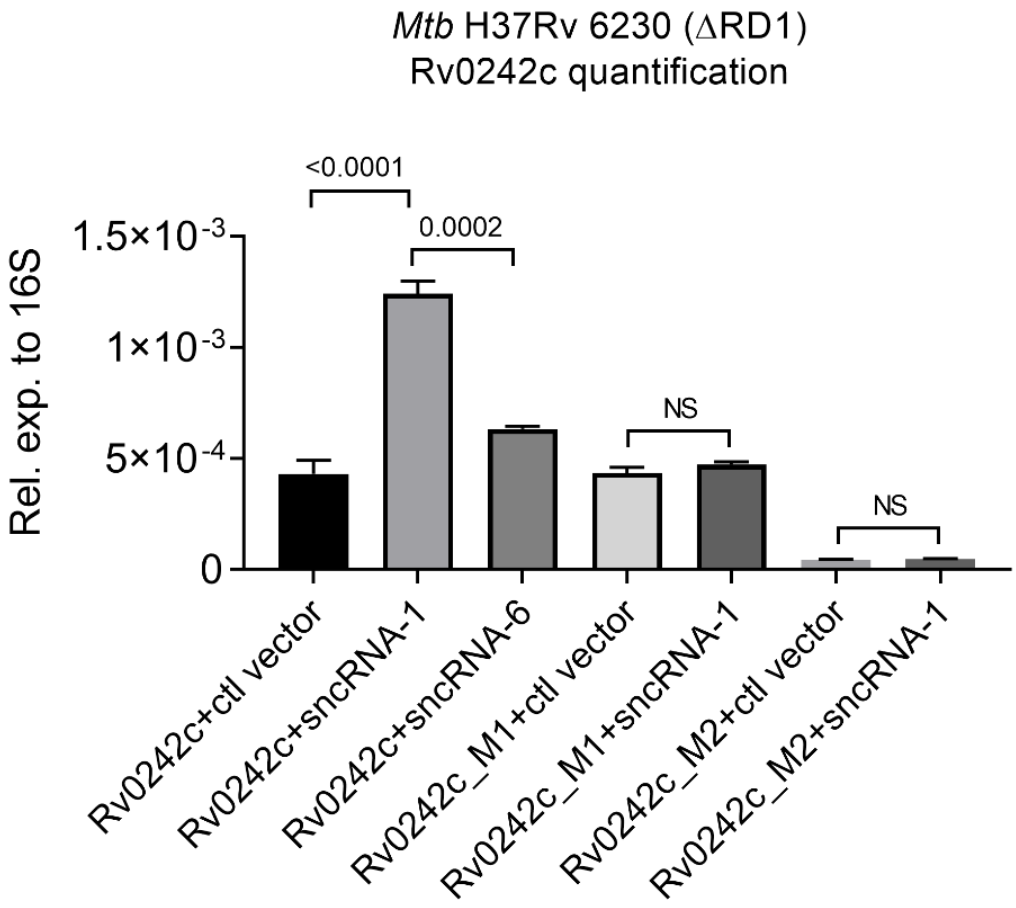


Figure 26. sncRNA-1 directly induces the expression of Rv0242c transcript. One of the putative targets of sncRNA-1, Rv0242c including its 5' UTR, was expressed with either the ctl vector, sncRNA-1 or un unrelated sncRNA, sncRNA-6. The expression of Rv0242c was determined with conventional q-RT-PCR. Mutations were introduced to the proposed binding site of sncRNA-1 in the 5' UTR of Rv0242c to make 2 different mutants. For mutant 1 (Rv0242c_M1), expressing sncRNA-1 with this mutant could not enhance the expression of Rv0242c transcript. For mutant 2 (Rv0242c_M2), the exogenous expression was abrogated independent of the presence or absence of sncRNA-1. The statistical significance was calculated using T-test.

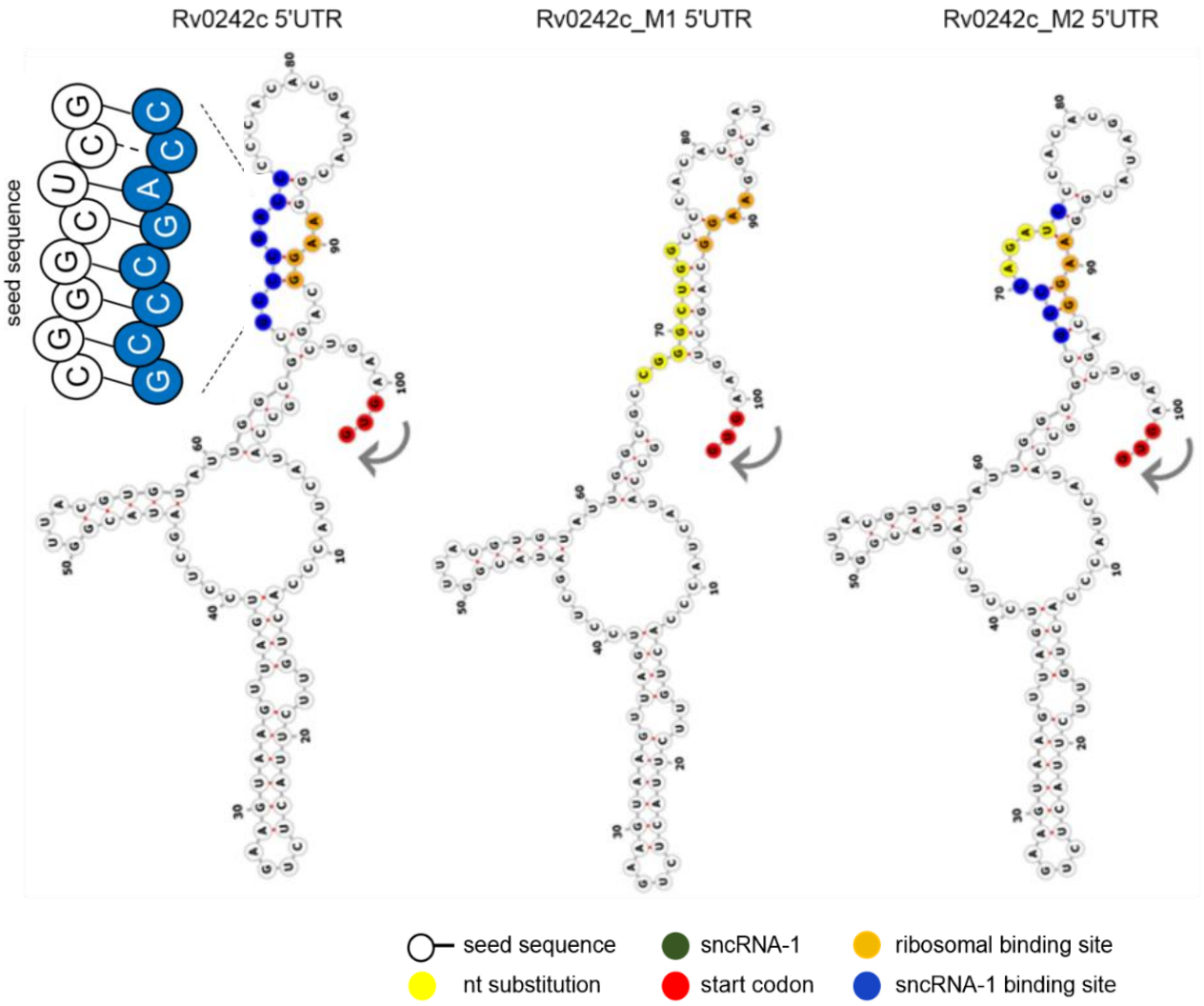


Figure 27. Putative secondary structures of the 5' UTR of Rv0242c, Rv0242c_M1, Rv0242c_M2. The structures were predicted using RNAfold online tool. The sncRNA-1 seed sequence is shown in circles. Putative sncRNA-1 binding site is shown in blue. Start codon is shown in red. Introduced mutations are shown in yellow. Predicted ribosomal binding site is shown in orange.

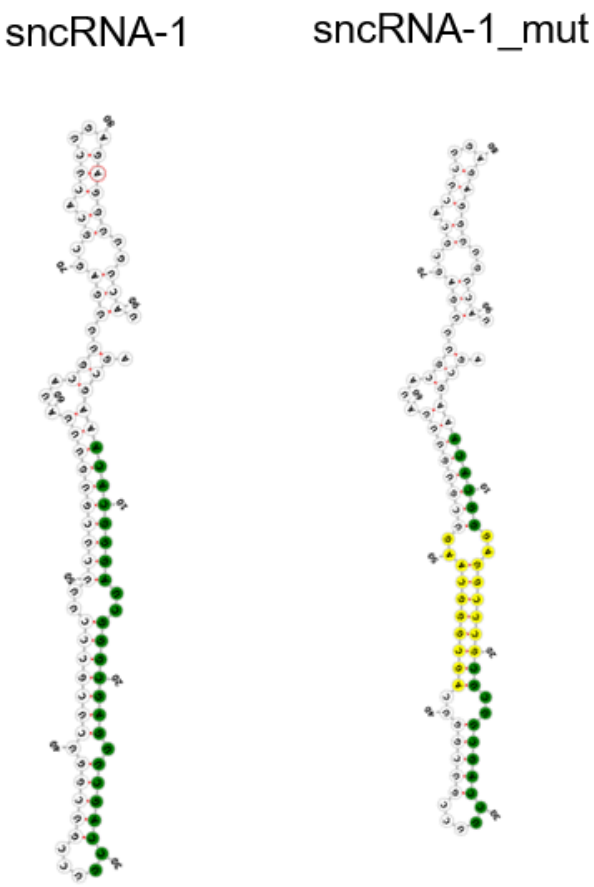


Figure 28. Predicted RNA secondary structure of transcripts containing control or sncRNA-1 and sncRNA-1 with selectively mutated nucleotides. sncRNA-1 sequence is shown in green, and the introduced mutations are shown in yellow.

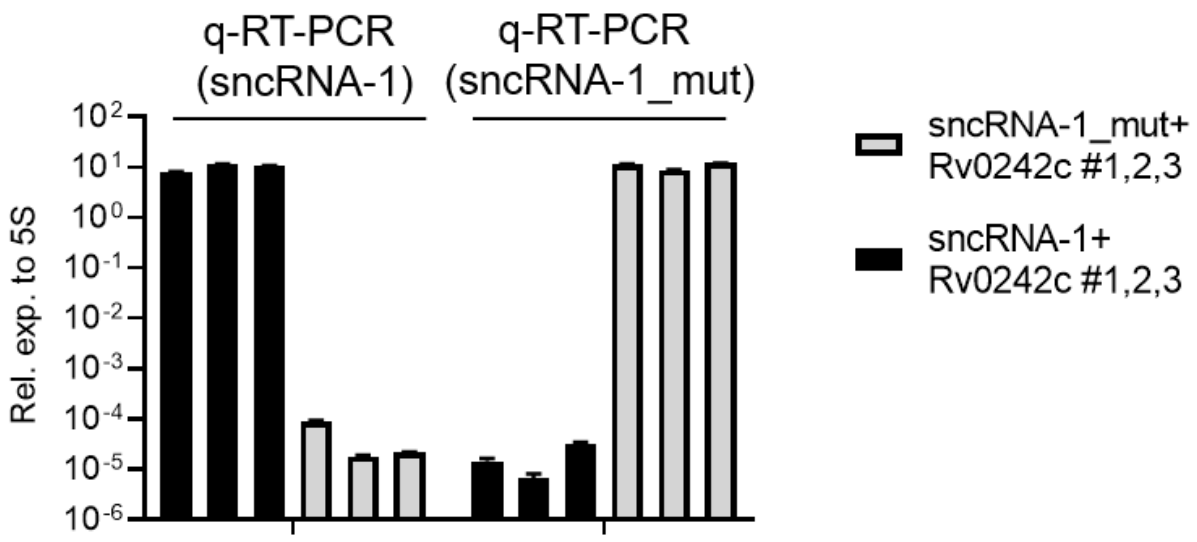


Figure 29. sncRNA-1 and sncRNA-1_mut are distinguished with miRNA-specific q-RT-PCR probes. Wild-type sncRNA-1 or the mutant sncRNA-1 (sncRNA-1_mut) was expressed in Mtb H37Rv 6230 along with the mutant Rv0242c (Rv0242c_M1) expression vector. RNA was isolated from the three different cultures at the stationary phase and the expression of sncRNA-1 or sncRNA-1_mut was quantified with a miRNA-specific q-RT-PCR assay using primers specific for each construct. sncRNA-1 expression that was detected in clones that had the sncRNA-1_mut construct and vice versa are considered background.

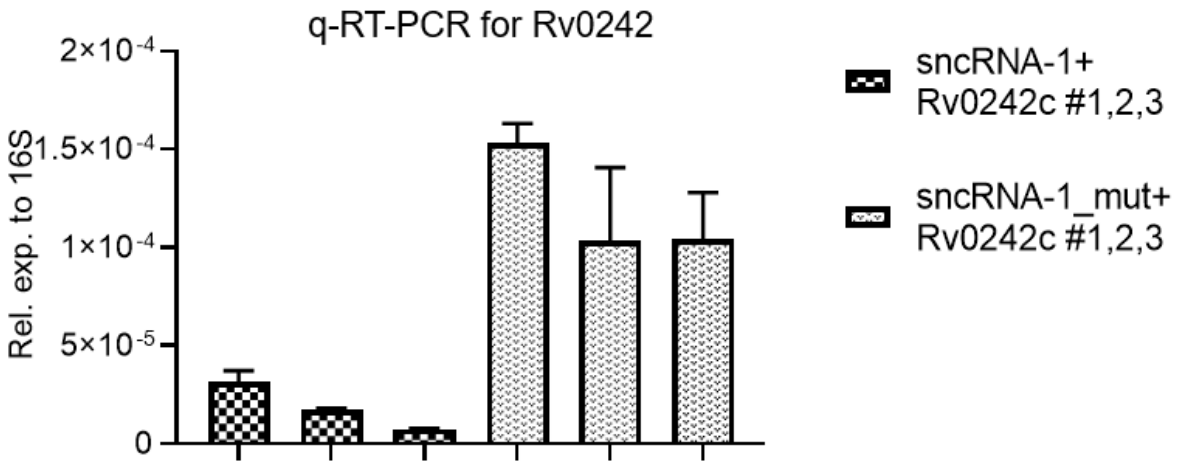


Figure 30. The expression of a mutated version of Rv0242c is enhanced when the corresponding Watson base-pair is present in sncRNA-1 (sncRNA-1_mut). Wild-type sncRNA-1 or the mutant sncRNA-1 (sncRNA-1_mut) was expressed in Mtb H37Rv 6230 along with the mutant Rv0242c (Rv0242c_M1) expression vector. RNA was isolated from the three different cultures at the stationary phase and the expression of Rv0242c was quantified with the conventional RT-PCR assay.

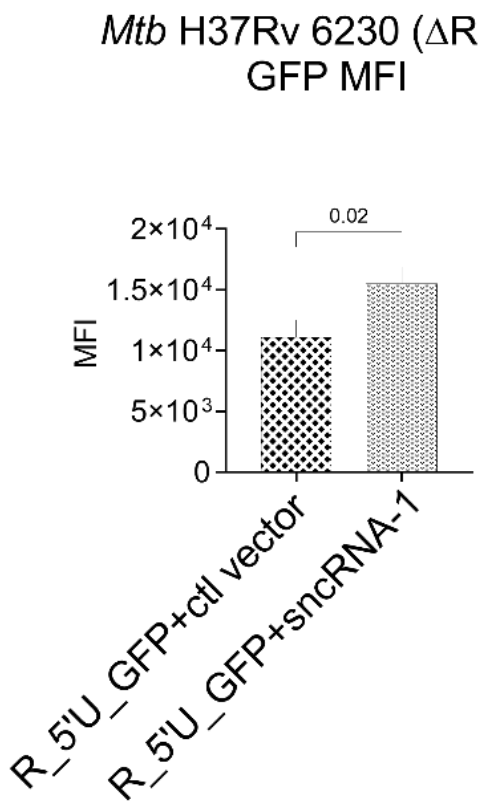


Figure 31. sncRNA-1 enhances GFP expression when the 5' UTR of Rv0242c is cloned 5' to the transcriptional start site of GFP. The 5' UTR of Rv0242c was cloned upstream of a GFP expression gene cassette. This reporter construct was expressed in Mtb H37Rv 6230 along with either the ctl vector or the sncRNA-1 expression vector. After reaching to the stationary phase the cells were transferred to a plate and the GFP intensity was measured on a plate reader.

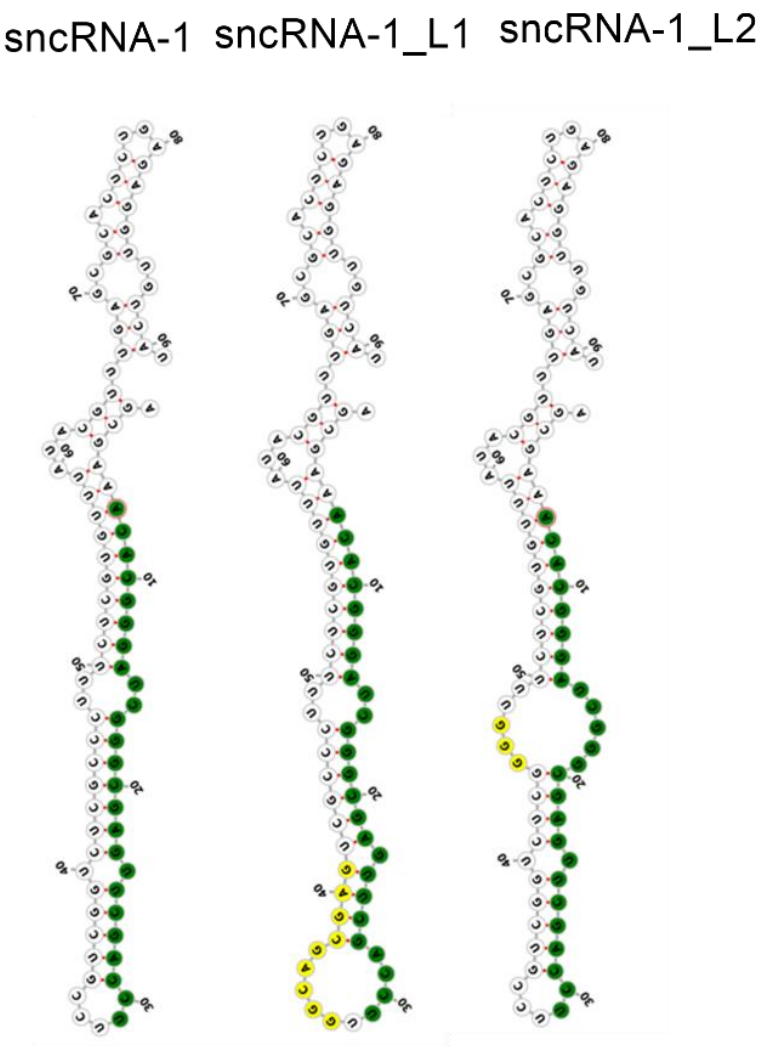


Figure 32. Putative secondary structures of sncRNA-1, sncRNA-1_L1 and sncRNA-1_L2. The structures were predicted using RNAfold online tool. Mature sncRNA-1 sequence is shown in green, and the introduced mutations are shown in yellow.

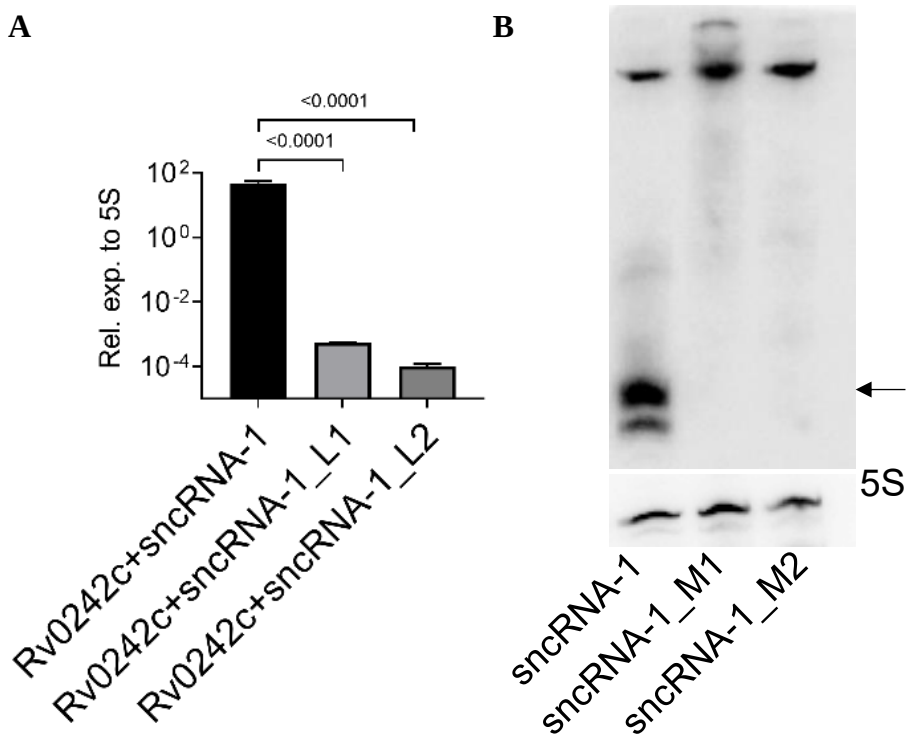


Figure 33. Mutations in the predicted hairpin loop structure of the precursor sncRNA-1 sequence reduces the expression of sncRNA-1. 9 nts mutations were introduced to the loop in sncRNA-1_M1 and 3 nts mutations were inserted to the hairpin in sncRNA-1_M2. The mutants or the wildtype sncRNA-1 was expressed in Mtb H37Rv 6230 along with the wild-type Rv0242c. The RNA was isolated from stationary phase cultures and the expression of sncRNA-1 was quantified with a miRNA-specific q-RT-PCR. The remaining RNA was run on a gel for a northern blotting. The membrane was first probed for sncRNA-1 and then for the 5S RNA. sncRNA-1 is shown with an arrow. The statistical significance was calculated using one-way-ANOVA.

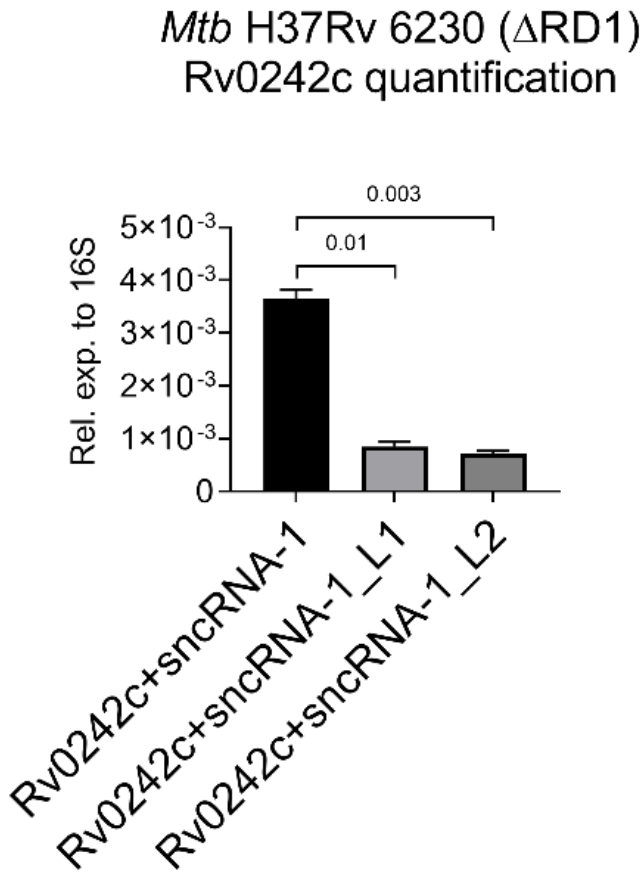


Figure 34. Rv0242c is not upregulated in Mtb H37Rv 6230 expressing the sncRNA-1_M1 or sncRNA-1_M2 mutant relative to wild-type sncRNA-1. The mutants of sncRNA-1 or the wildtype sncRNA-1 was expressed in Mtb H37Rv 6230 along with the wild-type Rv0242c. The RNA was isolated from stationary phase cultures and the expression of Rv0242c was quantified with a conventional q-RT-PCR. The statistical significance was calculated using one-way-ANOVA.

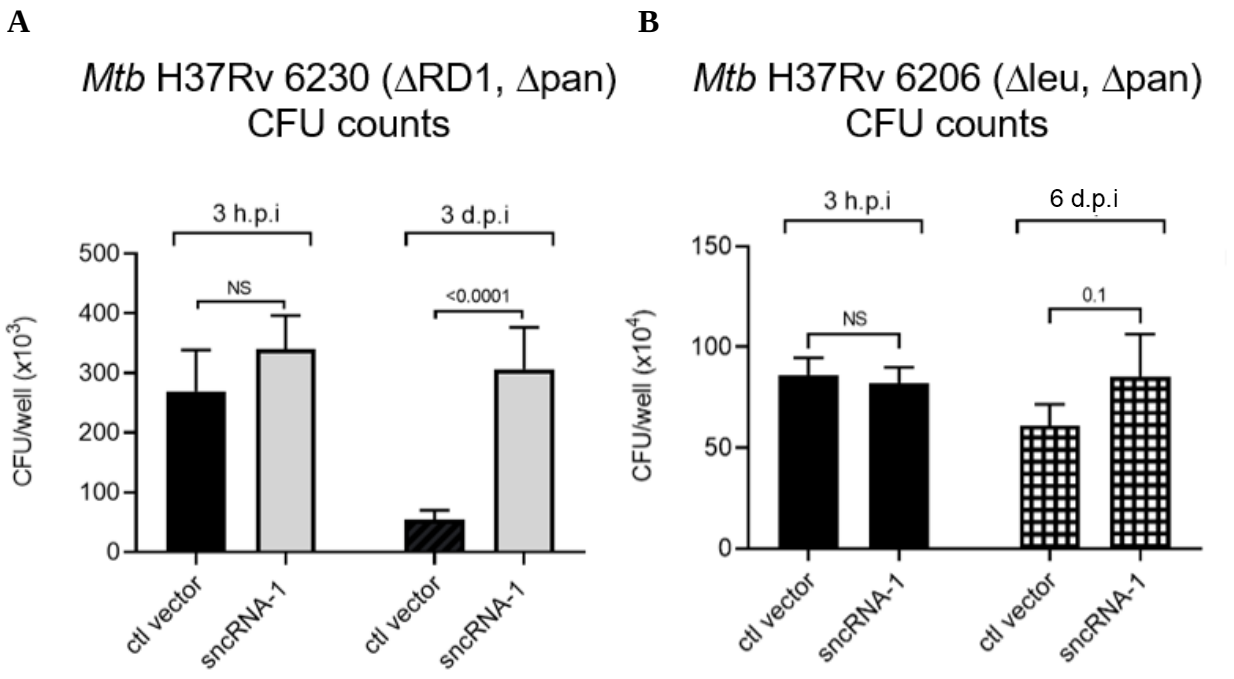


Figure 35. sncRNA1 overexpression in Mtb H37Rv 6230 and Mtb H37Rv 6206 enhances bacterial survival in Mtb-infected BMDM. Ctl vector or sncRNA-1 vector was expressed in Mtb H37Rv 6230 or Mtb H37Rv 6206. BMDM was infected with (A) Mtb H37Rv 6230 or (B) Mtb H37Rv 6206 expressing either the ctl vector or the sncRNA-1 vector. Cells were washed after 3 hours post infection. Half of the samples were lysed and plated to determine the initial bacterial burden. The remaining half were lysed on 3-days post infection for MtbH37Rv 6230 and 6-days post infection for Mtb H37Rv 6206 and plated. The CFU counts were determined 1 week later for Mtb H37Rv 6230 and 3 weeks later for Mtb H37Rv 6206. Statistical significance was determined using student's T-test.

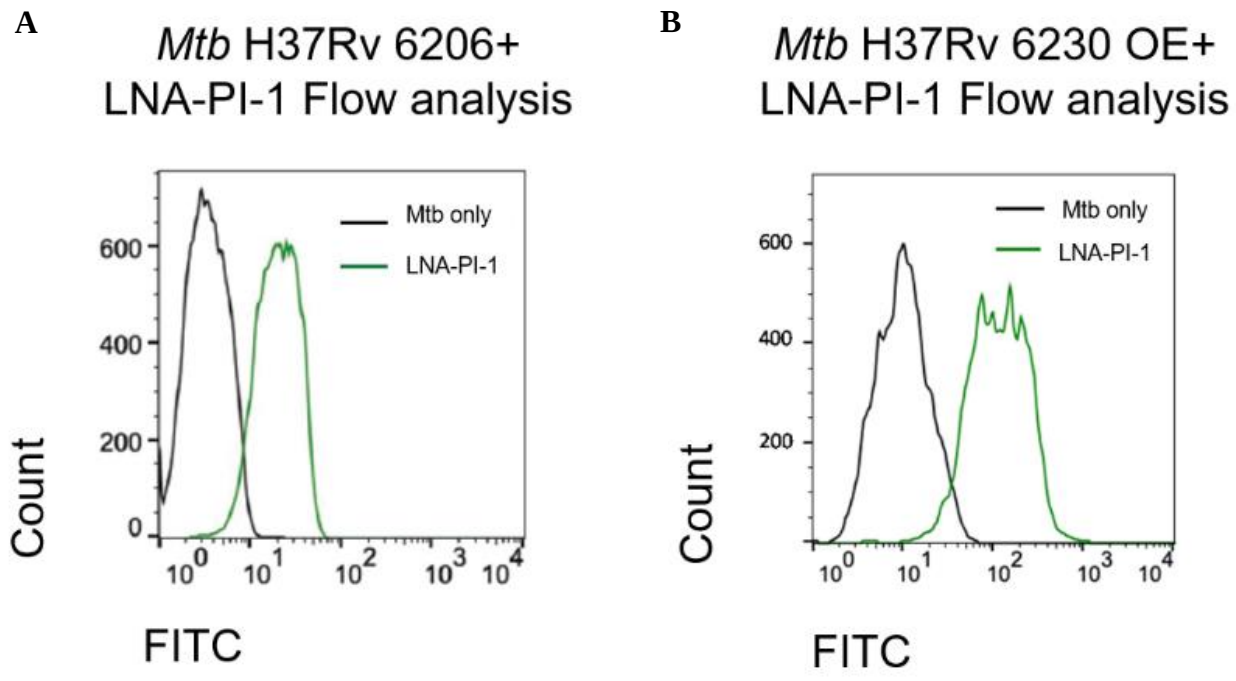


Figure 36. Mtb H37Rv 6206 or Mtb H37Rv 6230 are detected by flow cytometry using fluorochrome labeled LNA-PIs. (A) Mtb H37Rv 6206 or (B) Mtb H37Rv 6230 expressing sncRNA-1 vector were incubated with LNA-PI-1 for 3 days or overnight, respectively. The cells were fixed and analyzed relative to the Mtb only settings with the flow cytometry on the FITC channel. The data was analyzed using FlowJo.

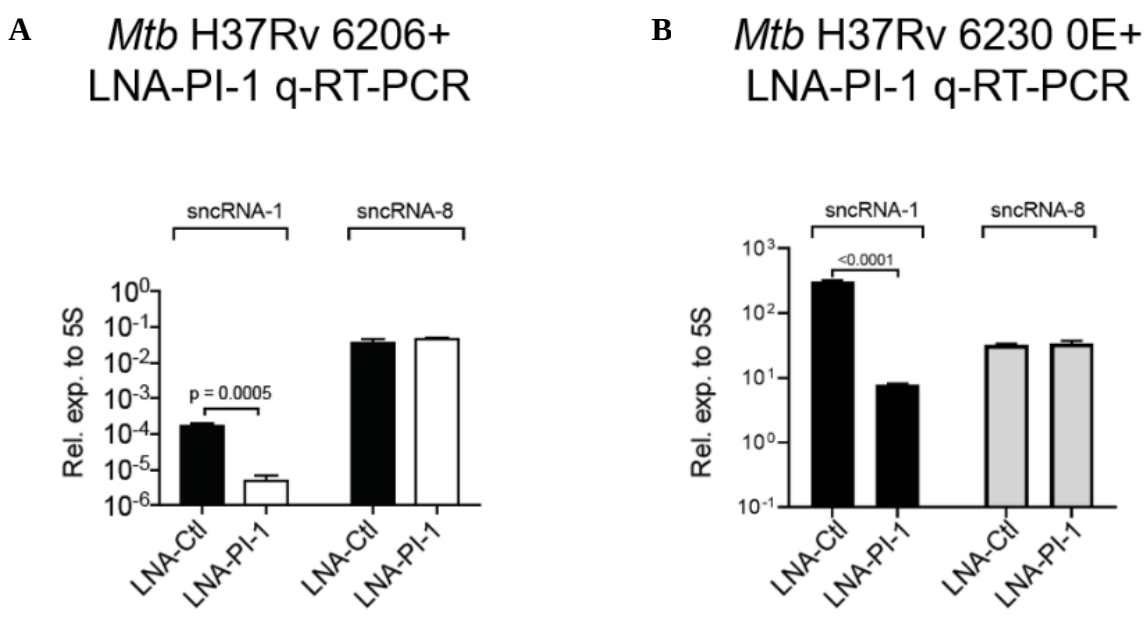


Figure 37. LNA-PI-1 specifically reduces the detection of sncRNA-1 but not sncRNA-8 in Mtb H37Rv 6206 and in Mtb H37Rv 6230. RNA was isolated from (A) Mtb H37Rv 6206 or (B) Mtb H37Rv 6230 overexpressing sncRNA-1 (Mtb H37Rv 6230 OE) treated with either LNA-Ctl or LNA-PI-1. Expression of sncRNA-1 or an unrelated sncRNA, sncRNA-8 was quantified using miRNA-specific q-RT-PCR. The statistical significance was determined using student's T-test.

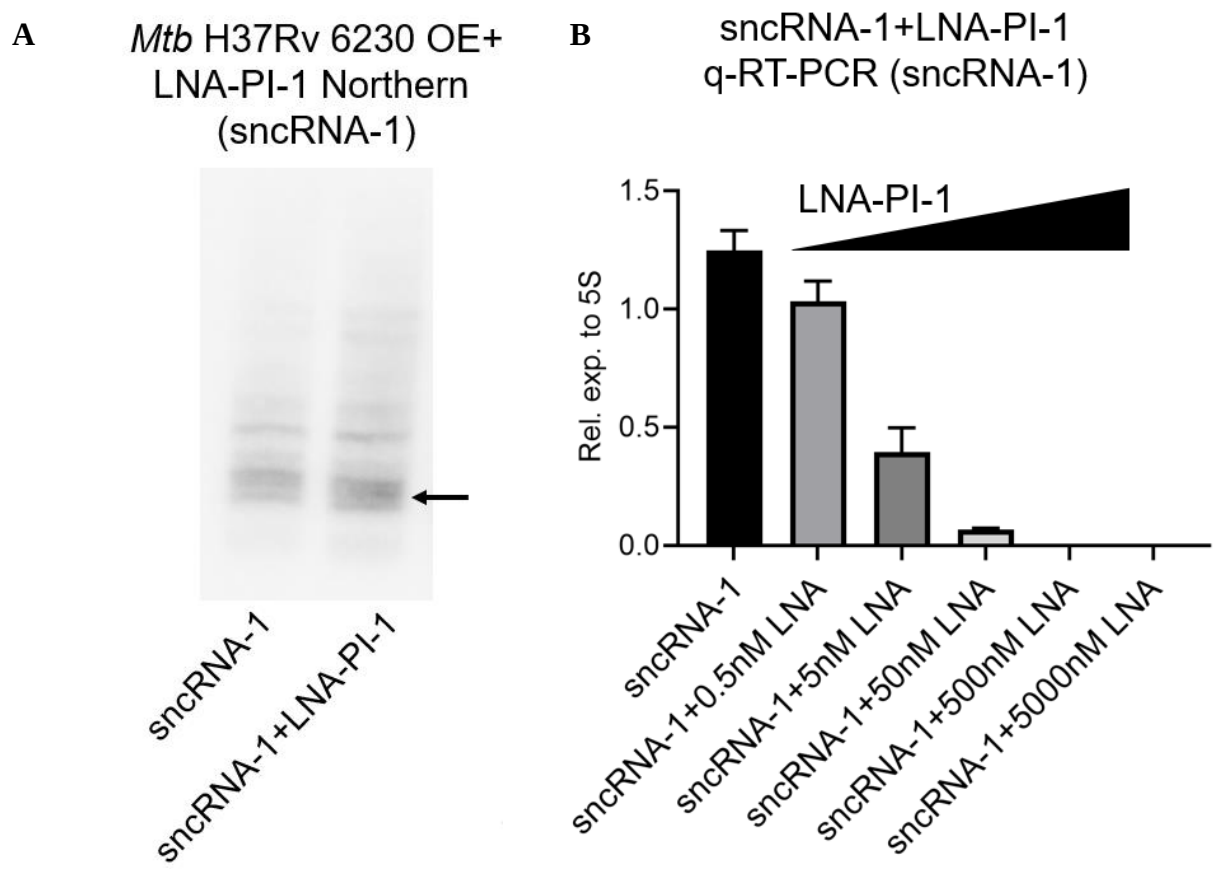


Figure 38. sncRNA-1 is detected in LNA-PI-1 treated mycobacterial cultures by both northern blotting and miRNA-specific q-RT-PCR. (A) RNA was isolated from Mtb H37Rv 6230 overexpressing sncRNA-1 treated with LNA-PI-1 and run on a gel to do a northern blotting. The membrane was probed for sncRNA-1 which is shown with the arrow. (B) RNA isolated from Mtb H37Rv 6230 overexpression sncRNA-1 was incubated with increasing amounts of LNA-PI-1 just before the cDNA synthesis. sncRNA-1 expression was then quantified using specific primers.

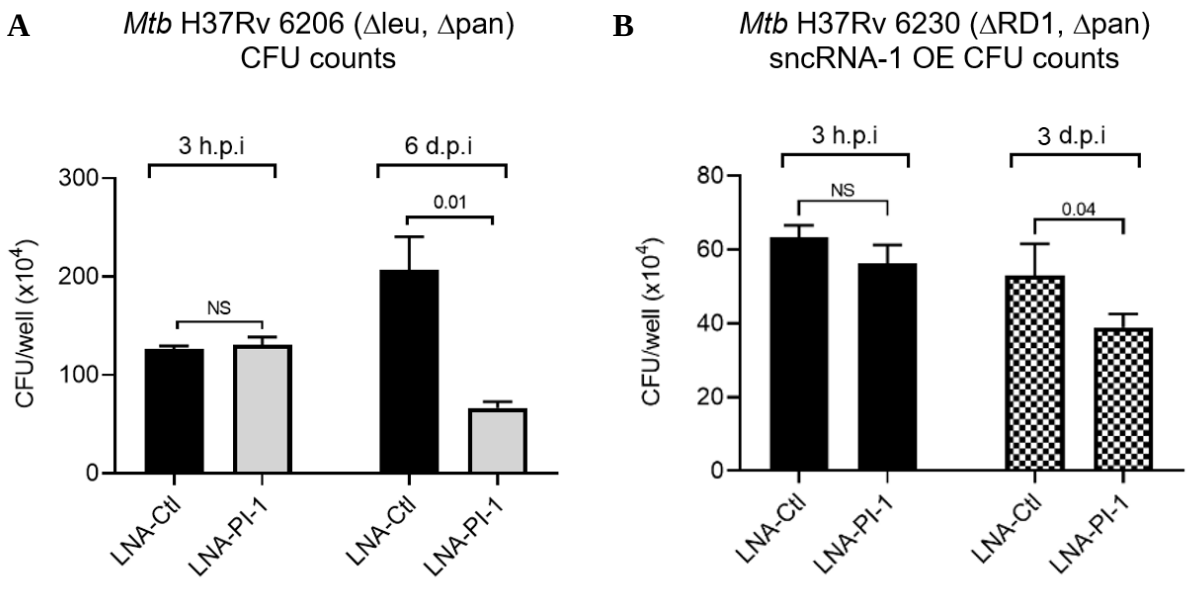


Figure 39. Antagonizing sncRNA-1 functions with LNA-PI-1 in Mtb H37Rv 6206 or in Mtb H37Rv 6230 reduces bacterial survival in infected BMDM. (A) BMDM were infected with Mtb H37Rv 6206 or (B) Mtb H37Rv 6230 overexpressing sncRNA-1 incubated with either LNA-Ctl or LNA-PI-1 for 3 days or 1 day, respectively. Cells were washed after 3 hours post infection. Half of the samples were lysed and plated to determine the initial bacterial burden. The remaining samples were lysed on 3 days- or 6-days post infection for Mtb H37Rv 6230 or Mtb H37Rv 6206, respectively. The CFU counts were determined 1 week later for Mtb H37Rv 6230 and 3 weeks later for Mtb H37Rv 6206. Statistical significance was determined using Student's T-test.

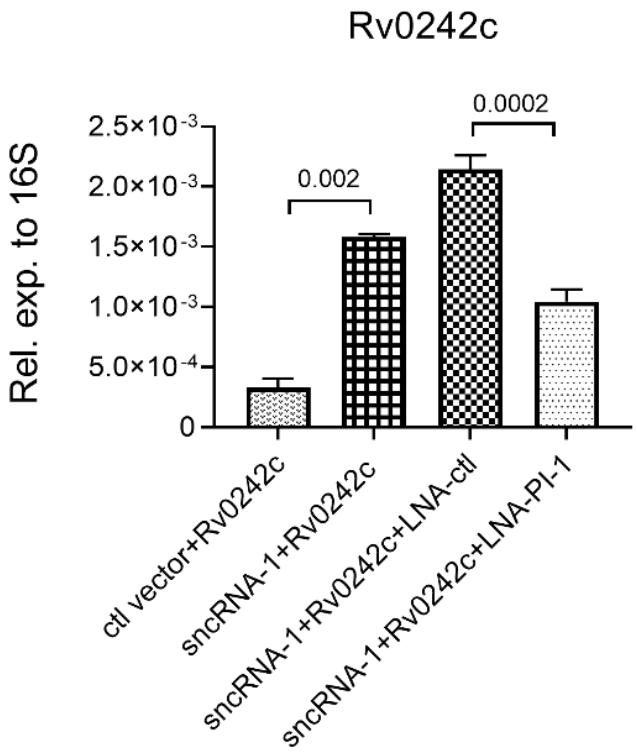


Figure 40. The elevated expression Rv0242c by sncRNA-1 is suppressed by the addition of specific LNA inhibitors. Rv0242c was expressed in Mtb H37Rv 6230 in the presence of ctl vector or sncRNA-1 vector. The cultures were treated with LNA-PI-1 or LNA-Ctl and the RNA was isolated. The expression of Rv0242c was quantified using regular q-RT-PCR. Statistical significance was determined using one-way-ANOVA.

CHAPTER 5

The role of the Mtb degradosome in the processing of precursor sncRNAs into mature sncRNA species

5.1 Introduction

Multiple mycobacterial small RNAs (sRNAs) undergo a poorly characterized RNA processing after their transcription (Girardin & McDonough, 2020; Moores, Riesco, Schwenk, & Arnvig, 2017; Sikova et al., 2019). For example, MTS2823 exists as a 300-nt form in exponentially growing mycobacterial cultures with a second 250-nt transcript generated from this longer transcript when the cells reach stationary phase (K. B. Arnvig et al., 2011). MTS1338 exists as a >400 nts sRNA that is cleaved to a mature 108 nts sequence (Moores et al., 2017). MTS0997 undergoes 3' end trimming of between 3-14 nts (Girardin & McDonough, 2020).

Several RNA processing enzymes have been identified in *Mycobacterium tuberculosis* (Mtb). The Mycobacterial core RNA degradosome includes enzymes involved in the clearance of excess RNA such as ribonuclease E (RNase E), polynucleotide phosphorylase (PNPase or GpsI), ribonuclease J (RNase J) and the ATP-dependent RNA helicase RhlE (Figure 41) (Plocinski et al., 2019). Except for RhlE, the other enzymes are essential for *in vitro* mycobacterial growth, determined with transposon-based mutagenesis screens for essential genes (DeJesus et al., 2017). RNase E is an endonuclease cleaving at A/U rich single stranded RNAs (ssRNA) after recognizing the proximal 5' phosphate (Baek et al., 2019). PNPase is a 3'→5' exonuclease, whereas RNase J is a 5'→3' exonuclease, cleaving ssRNA until a double-stranded RNA (dsRNA) section is reached (Jester, Romby, & Lioliou, 2012).

Sikova *et al.* investigated the role of these enzymes in the processing of the sRNA, MTS2823. Using the CRISPR interference platform developed for mycobacteria, the selective knock down

of each RNase reveals that the level of *M. smegmatis* MTS2823, called Ms1, increases 30% when PNPase is downregulated (Plocinski et al., 2019; Rock et al., 2017; Sikova et al., 2019). Contrasting this, Plocinski *et al.* have shown that the sRNA is completely lost when PNPase was knocked down in Mtb (Plocinski et al., 2019). These species-dependent seemingly contadicting findings demonstrate the need for more studies to address the role of RNA prcoessing enzymes in Mtb.

sncRNA-1 positively regulates Rv0242c, controlling the oleic acid production in Mtb (Chapter 4). SncRNA-1 is a 25-nt RNA formed from a precursor transcript >115 nts that contains a hairpin loop structure (pre-sncRNA-1) (Figure 28). Nucleotide substitutions disrupting the predicted hairpin loop structure and a presumed cleavage site of the precursor transcript eliminate the formation of sncRNA-1. In addition, my findings reveal that the complementary strand of sncRNA-1 is present at levels less than sncRNA-1 (Figure 20). These experiments support the presence of an RNA processing pathway that is both sequence and structure dependent. Moreover, the ability to overexpress sncRNA-1 with an RNA expression vector in the auxotrophic mutants of Mtb, Mtb H37Rv 6206 (Mtb 6206) or Mtb H37Rv 6230 (Mtb 6230) independent of any eukaryotic infection model confirm the existence of a prokaryotic RNA processing or degradation system for sRNAs.

Based on our findings along with that reported elsewhere, the sRNAs produced in Mtb are processed by poorly defined RNA enzymes. This concept is comparable to the maturation of eukaryotic miRNAs, which are 22 nts post transcriptional regulatory RNAs. In eukaryotes, a miRNA is synthesized as a primary miRNA (reviewed by (Suzuki & Miyazono, 2011)). Briefly, a primary miRNA (pri-miRNA) >1000 nts is cleaved by an enzyme called Drosha, producing a precursor miRNA (pre-miRNA) between 60-120 nts bearing a hairpin loop structure. This pre-

miRNA is further processed by an RNase III type enzyme called Dicer, yielding a short double-stranded RNA. Most often, one of the strands, the mature miRNA, is bound by an Argonaute type of protein to form the RNA-induced silencing complex. The complex uses a "seed" sequence in the RNA to selectively induce the degradation of a target gene, most often by hybridizing to a sequence within the 3' UTR. In Mtb, there is an RNase III type enzyme called *rnc* (Rv2925c) (Akey & Berger, 2005). Similar to its eukaryotic version, *rnc* also digests dsRNA by recognizing distal (2nts) and proximal (3nts) boxes making it a putative sncRNA processing enzyme (Figure 42) (K. Zhang & Nicholson, 1997). Based on these findings and interpretations, the potential contribution of mycobacterial RNase III and the core RNA degradosome enzymes RNase E, RNase J and PNPase on sncRNA processing was investigated.

5.2 Materials and Methods

5.2.1 Bacterial strains and growth conditions

Mycobacterium smegmatis (*M. smegmatis*) and *Mycobacterium avium* (*M. avium*) was grown in 7H9 supplemented with 10% oleic albumin dextrose catalase (OADC) (Remel, Thermo-Fisher) and 5% glycerol (G9012, Sigma- Aldrich), 0.05% tyloxapol (T0307, Sigma-Aldrich), 0.2% casamino acids (223050, BD Bacto™). For Mtb H37Rv 6230 (ΔRD1, Δ*panCD*), the media was supplemented with 24 µg/ml pantothenate (D-Pantothenic acid hemicalcium salt, Sigma-Aldrich) and for Mtb H37Rv 6206 (Δ*leuC*, Δ*panCD*), it was similar to that used for 6230 with an additional supplement of 80 mg/ml leucine (Sigma-Aldrich).

5.2.2 Site directed mutagenesis of sncRNA-6

For site directed mutagenesis, primers were designed using NEBasechanger™ tool. sncRNA-6 expression vector (pKA-KAN-sncRNA-6) was PCR amplified with primers designed for site directed mutagenesis. PCR product was treated with KDL mixture (NEB cat #0554) for 30 mins

37ºC and transformed into *Escherichia coli* (*E. coli*) DH5alpha. Clones were randomly selected and propagated. Plasmids were isolated with ZymoPURE™ Plasmid Miniprep Kit (Cat #D4209, Zymo Corp.) and sequenced by the McDermott Sanger Sequencing Core at UT Southwestern Medical Center.

5.2.3 Overexpression of sncRNAs in mycobacteria

Wild-type and mutant sncRNA-6 plasmids were expressed in *M. avium*, whereas sncRNA-8 was expressed in *M. smegmatis*. The electroporation was done as described in Chapter 4. Briefly, cells were washed 2 times in 10% glycerol, and 100 ul of bacteria were mixed with 4-6 ug of plasmids. They were briefly pulsed with 1.3 kV of power and transferred to 400 ul of media without antibiotics. They were subsequently plated with the corresponding antibiotic for selection. After colonies grew, 1 week for *M. smegmatis* and 3-4 weeks for *M. avium*, individual clones were cultured in liquid media with the antibiotics.

5.2.4 RNA extraction and northern blotting

Total RNA was extracted with the Qiagen miRNeasy kit following the manufacturers' instructions (Qiagen, Valencia, CA). Samples were first treated with DNase (Ambion-Thermo-Fisher). Where required, the RNA was concentrated with an RNA Pure and Concentrate kit (Zymo Research, CA). RNA quality and quantity were determined using a NanoDrop 2000/2000c spectrophotometer and/or Bioanalyzer (Thermo Fisher Scientific Inc.). Five µg RNA was resolved on a 20% acrylamide gel (29:1) for 2-3 hours in a 1X Tris-cl, borate, EDTA buffer (TBE). The RNA bands were then transferred to Nylon N^+ membranes at 80 volts for 1 hour in 0.5X TBE buffer. RNA was UV crosslinked and the membrane was prehybridized at 39ºC for 1-2 hours in rapid-hybridization buffer (GE/Amersham). These conditions were fully described elsewhere (Belkaya et al., 2011). It was hybridized overnight at 39ºC with Starfire probe specific for sncRNA-1 or 5S rRNA

(Starfire^{TM}, Idt DNA technologies, Inc.). The 5S probe detects a band near 116 nts. The membrane was washed twice in buffer (2x SSC buffer with 0.1% SDS), wrapped in saran wrap, and imaged with a Phosphoimager.

5.2.5 CRISPR interference assays

Non-targeting CRISPR interference (CRISPRi) plasmids and those targeting RNase E, PNPase and RNase J that were designed for *M. smegmatis* were kindly provided by Dr. Jarmila Hnilicova (Complex of Biomedical Institute at Krc, Czech Republic) (Sikova et al., 2019). CRISPRi plasmids targeting RNase E, PNPase and RNase III for Mtb were designed as described earlier (Rock et al., 2017). Briefly, pJR965 was mutated using NEB site directed mutagenesis kit and the sequences were verified with Sanger sequencing. CRISPRi plasmids were electroporated to mycobacteria as described in previous chapters and clones were selected on plates with antibiotics. Liquid cultures were started using single clones. sgRNA expression was induced using 100 ng/ml anhydrous tetracycline (Atc). To determine the knock down efficiency RNA was extracted as described before and q-RT-PCR was done using high-capacity cDNA reverse transcription kit (ThermoFisher Scientific Inc.) and SYBR green PCR master Mix (ThermoFisher Scientific Inc.). In certain cases, northern blotting was done to complement qPCR.

5.2.6 *In vitro* cleavage assay

Protein expression plasmids designed for the purification of RNase E, RNase J, and PNPase were kind gifts from Dr. Przemyslaw Plocinski (University of Lodz, Poland). Recombinant protein purification was done as described before (Plocinski et al., 2019). Briefly, the plasmids were transformed into *E. coli* Bl21 strains and single colonies were selected. 5 ml cultures were grown overnight at 37 centigrade degrees. Next day, the cultures were expanded to 200 ml and they were grown until the OD reached 0.6. The induction was done with 1 mM Isopropyl β-D-1-

thiogalactopyranoside (IPTG) for 16 hours at 20 ^{o}C. Next day, the cultures were spined down at 5000 rpm for 15 minutes. The cells were resuspended in the lysis buffer (50 mM Na2HPO4, 300 mM NaCl, 10 mM imidazole at pH 8). They were lysed with lysozyme at final concentration of 100 ug/ml for 30 minutes on ice. The lysis buffer was supplemented with protease inhibitors (leupeptin, aprotinin, benzamidine, PMSF, pepstatin A). After 30 minutes, the cells were sonicated on ice 3 times for 30 seconds. They were spined down at 20,000 rpm for 10 minutes at 4 ^{o}C. Meanwhile, the Ni-NTA Agarose beads (Qiagen #30210) were equilibrized with the lysis buffer. The lysate was mixed with the beads and they were incubated on a shaker for 1 hour at 4 ^{o}C. After the incubation, the beads were spined down at 500 g for 1 minute and they were washed with the wash buffer (50 mM Na2HPO4, 300 mM NaCl, 50 mM imidazole at pH 8) 3 times. Purified protein was eluted with the elution buffer (50 mM Na2HPO4, 300 mM NaCl, 500 mM imidazole at pH 8) twice. Eluate was passed through Pierce Protein Concentrator PES (Thermo Scientific #88539) according to the manufacturer's instructions. Concentrated eluate was washed with the storage buffer (10 mM Tris-HCl, pH 8.0, 25 mM NH_4Cl, 100 μM EDTA, 10 mM $MgCl_2$, 30 μM DTT) to remove imidazole. The concentration of the washed eluate was determined using Nanodrop at 280 nm and then it was mixed with equal amount of glycerol. The isolated protein was used in the *in vitro* RNA cleavage assays. The assays were done as described before (Sikova et al., 2019). Briefly, 400 ng synthetic RNA was mixed with 10 mM $MgCl_2$, 10 mM Na_2HPO_4, 20 mM Tris-HCl pH 7.6 and 1-10 μM RNase enzyme. They were kept at 37 °C for 30 minutes. The reaction mixture was mixed with 2X RNA loading buffer (95% formamide, 20 mM EDTA pH 8.0) and boiled at 70 °C for 5 minutes. They were run on 7 M urea 20% polyacrylamide gels. The fluorescence was visualized using Typhoon FLA 9500 laser scanner. Later the gels were stained with GelRed (LabMark) for 30 minutes and scanned again.

5.2.7 Overexpression of RNases in mycobacteria

Mycobacterial expression plasmids designed for the overexpression of RNase E, RNase J, and PNPase were kind gifts from Dr. Przemyslaw Plocinski (University of Lodz, Poland) (Plocinski et al., 2019). They were electroporated to Mtb 6230 along with sncRNA-8. Clones expressing both plasmids were selected using double selection media. RNA was extracted and qPCR for RNases and northern for sncRNA-8 was done as described previously.

5.3 Results

5.3.1 sncRNA-6 is produced by structure-dependent processing from a longer precursor RNA transcript

sncRNA-1 and sncRNA-6 are mature sRNAs of 25 and 21 nts, respectively. sncRNA-1 is processed from a precursor sncRNA in a nucleotide sequence and RNA secondary structure dependent manner. Thus, targeted nucleotide substitutions with the precursor form of sncRNA-1 (pre-sncRNA-1) result in the loss of the mature form with a concomitant appearance of a larger transcript >116 nts (Figure 33). To investigate whether the mature sncRNA-6 depends on similar sequence-dependent processing steps, specific nucleotide substitutions were introduced into the precursor sequence of sncRNA-6 (pre-sncRNA-6) (Figure 43). Mutations were introduced at either the predicted hairpin loop structure or at the putative cleavage site that creates the mature sncRNA-6 (Figure 43). The mature 21-nt form of sncRNA-6 was not formed when either 3 nts in the putative cleavage site or 4 nts within the hairpin loop were mutated (Figure 43). Only larger RNA transcripts containing precursor sequences of were present (Figure 43). Taken together, these findings support the existence of a small RNA processing system in mycobacteria.

5.3.2 Endogenous levels of sncRNA-8 are evident in multiple mycobacterial strains in culture

To expand on the processing requirements for the sncRNAs, sncRNA-8 was characterized, in part because this is the only sncRNA that expresses a basal level of expression in mycobacteria without the need for infections in eukaryotic cells (Figure 44). SncRNA-8 was detected as a 24-nt sncRNA in stationary *M. smegmatis* cultures, with comparable levels of endogenous expression noted in in Mtb 6206 and Mtb 6206 (Figure 44). Of note, attempts at overexpressing sncRNA-8 in *M. smegmatis* caused a reduction in the levels of the mature sncRNA (Figure 45). Yet, the levels of the precursor form of sncRNA-8 (pre-sncRNA-8) were increased, with a predominant >100-nt transcript evident by northern blotting with sncRNA-8 specific probes (Figure 45). This finding suggests that overexpressing pre-sncRNA-8 may reduce the effectiveness of sncRNA processing by overwhelming the RNA processing components present in Mtb.

5.3.3 Overexpression of RNase E, RNase J or PNPase in mycobacteria degrades sncRNA-8

Several putative RNA processing enzymes have been identified in Mtb and these form part of the degradosome complex. To address whether these enzymes were coupled to sncRNA processing, 3 candidates, RNase E, RNase J and PNPase were overexpressed to assess their impact on sncRNA formation. In *M. smegmatis*, no clones were recovered following attempts to electroporate each of the 3 overexpression cassettes, suggesting that high levels of the RNases were lethal. Subsequent experiments were done with the Mtb 6230 as several clones per construct were successfully propagated with the expression of RNase E and PNPase increased by more than 100-fold, with RNase J elevated 3-fold (Figure 46).

With these overexpressing clones, the levels of precursor of sncRNA-8 (pre-sncRNA-8) were compared with the levels of mature sncRNA-8 with q-RT-PCR (Figure 47A) and miRNA-specific RT-PCR (Figure 47B), respectively. To quantitate the effectiveness of the RNA processing, the

ratio of sncRNA to precursor sncRNA was calculated. Overexpression of RNase E and RNase J resulted in decreased levels of pre-sncRNA-8 while the mature sncRNA-8 at 24-nts remained similar except for PNPase where the level of 24-nt sncRNA-8 increased about 3-fold (Figure 47C). Regardless of which RNases were overexpressed, in all cases the levels of the 24-nt sncRNA-8 and the precursor form were less than in control clones (Figure 47). If these RNases were selectively involved in sncRNA-8 processing, we would have expected the band intensity at lower sizes to be higher as the band intensity at larger sizes diminished. Therefore, it is more likely that these enzymes are involved in the nonspecific clearance of excess RNAs as they make the core RNA degradosome. Thus, our approach using an overexpression system was inconclusive as to whether the core RNA degradosome components are involved in sncRNA-8 processing. Consequently, a knock down-based approach was used as described next.

5.3.4 Targeting of PNPase by CRISPR interference assays reduces sncRNA-8 degradation

In a second approach to define the potential contributions of enzymes that form part of the core RNA degradosome in sncRNA processing, CRISPR interference (CRISPRi) was used to knock down the RNases. This technology was successfully developed for mycobacteria (Rock et al., 2017). An inducible expression cassette with dead Cas9 (dCas9) and a single stranded guide RNA targeting several distinct RNases was performed using previously developed strategies used with *M. smegmatis* (Sikova et al., 2019). In the published reports, Sikova *et al.* characterized a sRNA in *M. smegmatis* called Ms1, which sequesters RNA polymerase (RNAP) during stationary phase. As cultures enter exponential phase, Ms1 is degraded, releasing RNAP to promote transcription. They hypothesized that the core RNA degradosome is involved in Ms1 decay. Knock down of RNase E, RNase J and PNPase was performed and the amount of Ms1 quantitated. They found

that the levels of Ms1 were increased when PNPase was knocked down, implying that this RNase is involved in Ms1 decay.

Given the success of their approach, we selectively targeted RNase E, PNPase and RNase J. Each enzyme was selectively knocked down in *M. smegmatis* by 20-150-fold, determined using 3 independently isolated clones (Figures 48). The levels of mature sncRNA-8 were visualized by northern blotting (Figure 49). Expression of the sgRNA revealed an alter processing of sncRNA-8 with the appearance of multiple bands larger than 25 nts (Figure 44 and Figure 49). However, compared to the ctl sgRNA, the most prominent change in the banding pattern of sncRNA-8 was observed with the knock down of PNPase where a smearing between 25 nts and 42 nts appeared. RNase E and J gave inconsistent findings. Moreover, staining the gel of the RNA isolated from the *M. smegmatis* clones treated with the CRISPRi vectors revealed a smearing for low molecular weight RNAs in clones deficient of PNPase (Figure 50). This suggested that PNPase may be specifically contributing to the clearance of low molecular weight RNAs including sncRNAs. Taken together, our findings suggest that the targeted disruption of PNPase has the most pronounced effect on the mature sncRNA-8, wherein a larger sized intermediate is evident. To further test these findings, this assay was repeated in Mtb 6206, where the expression of sgRNA does not alter the processing of sncRNA-8 (Figure 53).

To better understand how PNPase contribute to the processing of pre-sncRNA-8, this enzyme was knocked down with CRISPRi in Mtb 6206. In addition, we also targeted RNase III, which is reported to recognize and cleave dsRNA, leaving 2-3 nts overhangs (Court et al., 2013). Defects following knock down would be consistent for RNase III processing of the precursor sncRNAs. Plasmids encoding dCas9 and sgRNAs targeting PNPase or RNase III were electroporated into Mtb 6206. Two independent clones/construct were obtained and the knock down efficiency of

each enzyme was assessed with q-RT-PCR (Figure 51). Expression of a control sgRNA not targeting any Mtb genes was used to ensure that the CRISPRi did not influence endogenous levels of each RNases. RNase III transcript expression was reduced 20-fold relative to the control, whereas the level of PNPase transcript level was reduced by 74-fold (Figure 51). With these clones, the amounts of the mature 24-nt sncRNA-8 and its precursor form were determined by RT-PCR. As a readout of processing of the 24-nt sncRNA-8 from the precursor form, the ratio of sncRNA-8 to pre-sncRNA-8 was calculated in Mtb 6206 when RNases were knocked down. Both sncRNA-8 and its precursor form accumulated in clones wherein PNPase was specifically targeted (Figure 52A-B). The ratio of sncRNA-8 to pre-sncRNA-8 was significantly enhanced in only one of the clones deficient of PNPase (Figure 52C). In contrast, the targeting of RNase III did not enhance the levels of sncRNA-8 and its precursor suggesting that RNase III is not involved in sncRNA-8 processing (Figure 52). The q-RT-PCR results were confirmed by northern blotting. The sncRNA-8 levels in Mtb 6206 using the ctl sgRNA was comparable to those where RNase III (RNC) was targeted, again suggesting that this enzyme is not involved in sncRNA-8 processing (Figure 53A). More importantly, both the mature sncRNA-8 levels and the precursor form were significantly increased in all the clones where PNPase was knocked down (Figure 53A). Northern blotting for a 5S RNA revealed a comparable RNA loading (Figure 53B). Taken together, these results support the notion that PNPase degrades small RNAs and is not coupled to the processing of the pre-sncRNA-8 into 24-nt sncRNA-8.

5.3.5 Purified PNPase degrades sncRNA-8

To complement the knock down and overexpression assays, a biochemical enzymatic assay was used to determine if PNPase and/or the RNases could directly process the small RNAs. For these biochemical assays, a synthetic RNA was designed to match a 42-nt precursor form of sncRNA-8

(syn-sncRNA-8). This 42-nt synthetic RNA had 6-Carboxyfluorescein (FAM) attached to its 5' end to enable direct visualization of the various RNA species in low molecular RNA gels (Figure 54A). PNPase and RNase J were independently expressed and purified using pET expression vectors and HIS-tagged constructs (Figure 54B). RNase J was used as a negative control as previous data suggested it is not involved in sncRNA-8 degradation. Second, this enzyme is a 5'→3' exonuclease and will not digest the synthetic RNA since the 5'end is covalently coupled to FAM. Syn-sncRNA-8 was incubated with 1 uM recombinant PNPase or RNase J at 37 centigrade degrees for 30 minutes. This RNA degradation assay revealed that PNPase, but not RNase J was able to degrade syn-sncRNA-8 under these conditions (Figure 55). Next, the assay was repeated with the titration of PNPase between 0.1-1 uM. PNPase <0.4 uM was no longer able to degrade the 42-nt syn-sncRNA-8 (Figure 56). No degradation of the syn-sncRNA-8 was observed with the high doses of RNase J.

Since the experiments suggested PNPase is the key enzyme involved in sncRNA processing/degradation, we used small molecular weight inhibitors of this enzyme in the assays. Previous studies have shown that citrate and ATP can block PNPase activity in *E. coli* (Del Favero et al., 2008; Stone et al., 2017). We found that concentrations >5 mM ATP blocked the mycobacterial PNPase, whereas 10 mM citrate was unable to do so (Figure 57A). To further support that, higher doses of citrate was used which revealed that 250 mM citrate was able to partially block PNPase activity (Figure 57B). These experiments confirmed that the degradation of syn-sncRNA-8 was specifically mediated by PNPase. Overall, these data showed that PNPase degrades sncRNA-8 *in vivo* and *in vitro*.

5.3.6 PNPase degrades sncRNA-1 in Mtb 6206

We had initially hypothesized that PNPase is involved in the processing of sncRNAs, but our data indicated that it degrades sncRNA-8. Next, we hypothesized that this observation can be extended to other sncRNAs including sncRNA-1. To test whether PNPase degrades sncRNA-1, RNA was isolated from Mtb 6206 expressing either the control sgRNA or sgRNA targeting PNPase. In these clones, expression of the mature 25-nt sncRNA-1 and its precursor form (pre-sncRNA-1) were determined by RT-PCR. As a readout of processing of the 25-nt sncRNA-1 from the precursor, the ratio of sncRNA-1 to pre-sncRNA-1 was calculated. Like sncRNA-8, both sncRNA-1 and its precursor form accumulated in clones wherein PNPase was specifically targeted (Figure 52 and Figure 58). The ratio of sncRNA-1 to pre-sncRNA-1 was not significantly enhanced indicating that PNPase is not involved in sncRNA-1 processing (Figure 58). Overall, these data support that PNPase degrades low molecular RNAs including sncRNAs.

5.4 Discussion

Many studies including ours show that certain small RNAs in mycobacteria undergo processing. In Chapter 4, it was shown that sncRNA-1 is processed in a secondary structure dependent manner. In this chapter, these observations have been expanded to sncRNA-6 as structural changes in the precursor of sncRNA-6 abrogated its processing. To identify the sncRNA processing enzymes, RNase III and the core RNA degradosome including RNase E, RNase J and PNPase were studied for their contribution to the processing of sncRNA-8. First, a gain-of-function approach was employed. However, this approach was incompetent to address the question. Next, a loss-of-function approach was utilized, which revealed that sncRNA-8 accumulated in the deficiency of PNPase but not the other enzymes tested. Further confirming these were the *in vitro* RNA degradation assays with a synthetic putative sncRNA-8 precursor and recombinant PNPase.

Finally, it was shown that sncRNA-1 levels also enhanced in clones deficient of PNPase. Overall, our findings revealed the importance of PNPase in degrading sncRNA-8 and sncRNA-1.

PNPase is a 3'→5' exonuclease and it was implicated in the decay of small RNAs in mycobacteria and others including *E. coli* (Andrade, Pobre, Matos, & Arraiano, 2012; Sikova et al., 2019). When we knocked down PNPase in Mtb, the levels of sncRNA-1 and sncRNA-8 were enhanced, but more interestingly the levels of their precursors also increased. This suggested that PNPase degrades both the sncRNAs and their precursors. Therefore, we cannot rule out PNPase as a sncRNA processing enzyme as the rate of PNPase activity on the precursor vs the mature form may determine the stability of the sncRNAs. Future studies with RNA decay assays of sncRNAs in PNPase deficient Mtb may elucidate the role of PNPase on sncRNA processing vs degradation. Overall, my data implicates the role of PNPase in Mtb RNA regulation, but other enzymes involved in sncRNA processing remain unidentified. Future studies will focus on the identification of sncRNA processing enzymes using unbiased approaches.

There are many questions that came up with the preliminary observations in this chapter. For instance, it is not clear how PNPase recognizes small RNAs including sncRNAs. It would be interesting to investigate whether this is determined by structural features of the sRNA, or if post-transcriptional modifications are involved. Another question to explore is whether the PNPase-mediated decay of sncRNAs has a regulatory function. I previously established that Mtb 6230 (lacking endogenous sncRNA-1 expression) has a growth defect when cultured in the absence of oleic acid. Moreover, this growth impairment could be overcome when sncRNA-1 was overexpressed. If PNPase is involved in the degradation of sncRNA-1, the manipulating the levels of this enzyme would be predicted to affect sncRNA-1 expression and mycobacterial growth. This may be critical for Mtb growth in certain growth conditions including in the absence of oleic acid.

However, to better understand the role of PNPase in different conditions where sncRNAs may be involved, we need to better characterize the roles of sncRNAs such as sncRNA-8. Further studies will focus on those whose function remains unexplored. Better characterization of sncRNA functions and their regulatory elements will determine if they are suitable drug targets to help the eradication of TB.

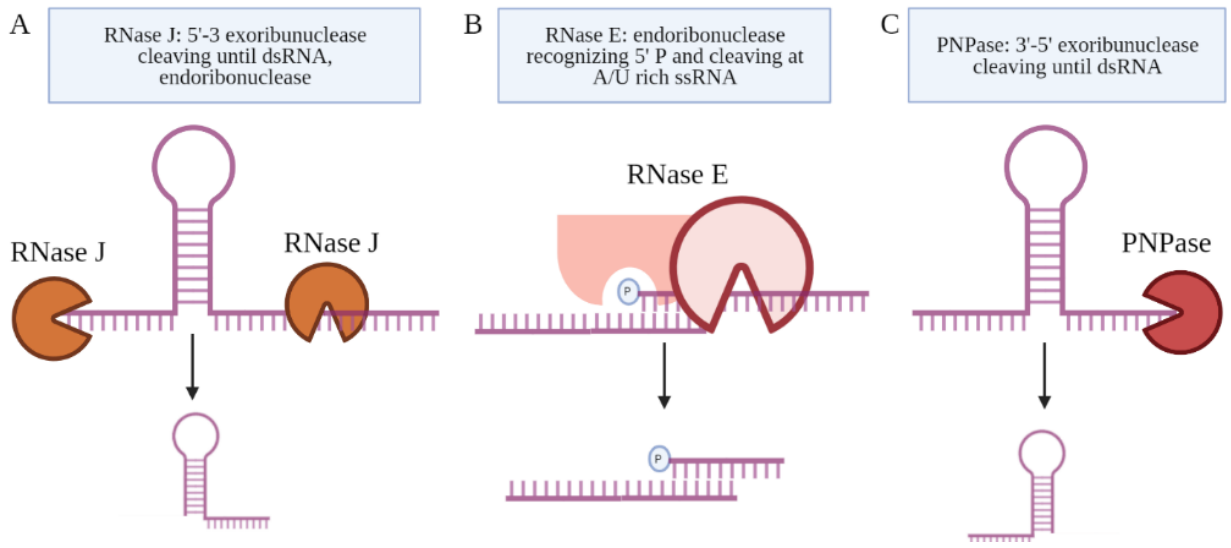


Figure 41. The enzymes of the core Mtb RNA degradosome, RNase J, RNase E and PNPase have different selectivity for RNA. RNase J is a 5'→3' exonuclease having also endonuclease activity. RNase E is an endonuclease cleaving at A/U rich sites after recognizing proximal 5' phosphate. PNPase is a 3'→5' exonuclease. Its activity is blocked by the presence of a double-stranded RNA.

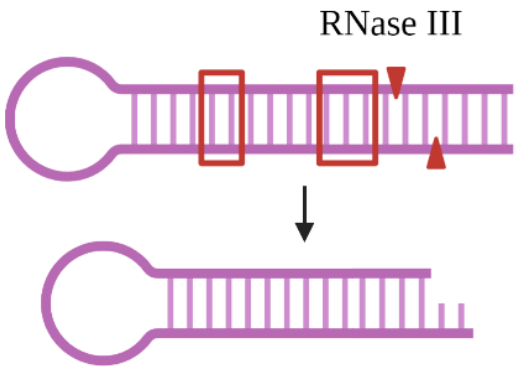


Figure 42. RNase III digests double-stranded RNA by recognizing distal and proximal boxes. This cleavage creates a 2-3 nts overhang. The RNase III (RNC) recognition sites are shown with red boxes and the RNase III cleavage sites are shown with red arrows.

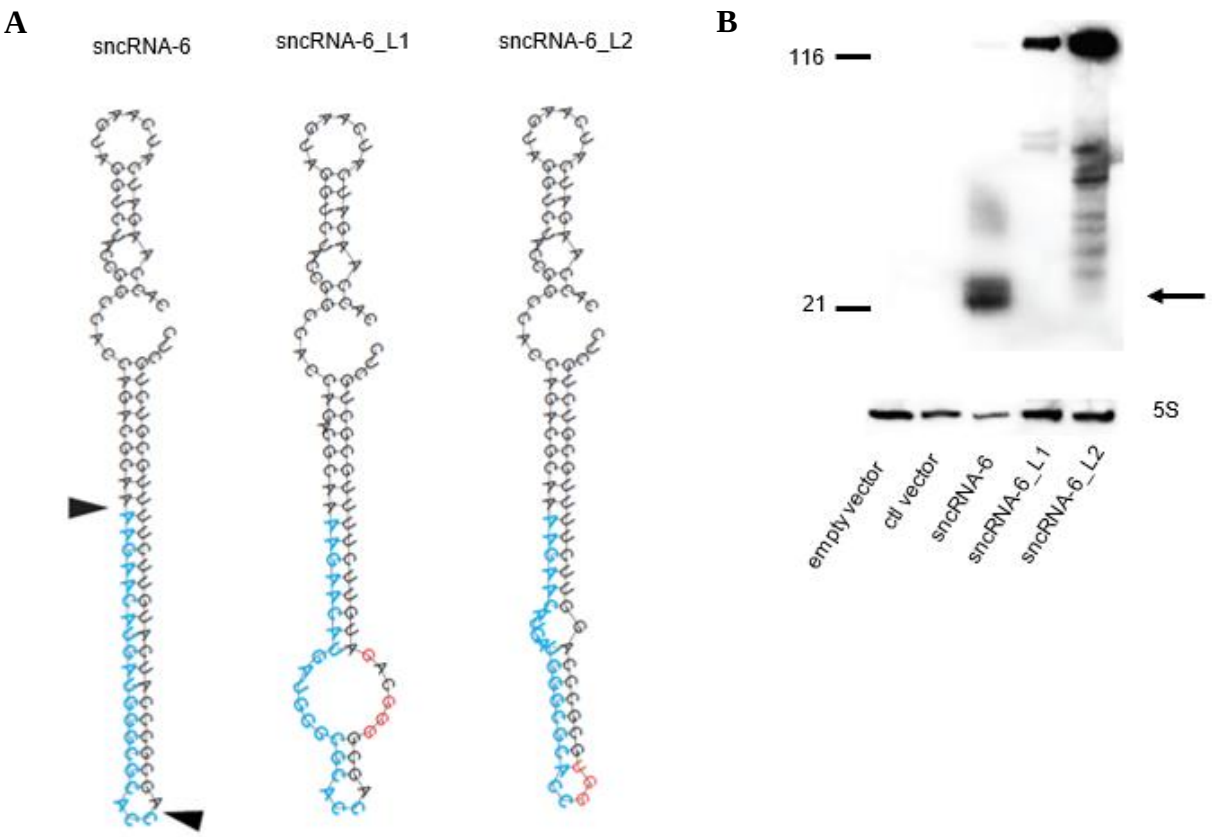


Figure 43. The precursor sncRNA-6 undergoes a sequence dependent processing into sncRNA-6 in *M. avium*. (A) Wild-type sncRNA-6 sequence is shown between arrows. Four nucleotide substitutions in the predicted hairpin structure and three nucleotide substitutions in the putative cleavage sites were introduced in sncRNA-6_L1 and sncRNA-6_L2, respectively. These are shown in red. (B) The wild-type and the mutant constructs were expressed in *M. avium*. RNA was isolated from stationary phase cultures and run on a gel for a northern blotting. The membrane was probed first for sncRNA-6 later for 5S RNA. The sncRNA-6 is shown with an arrow. Unprocessed forms are detected in clones bearing the mutant constructs.

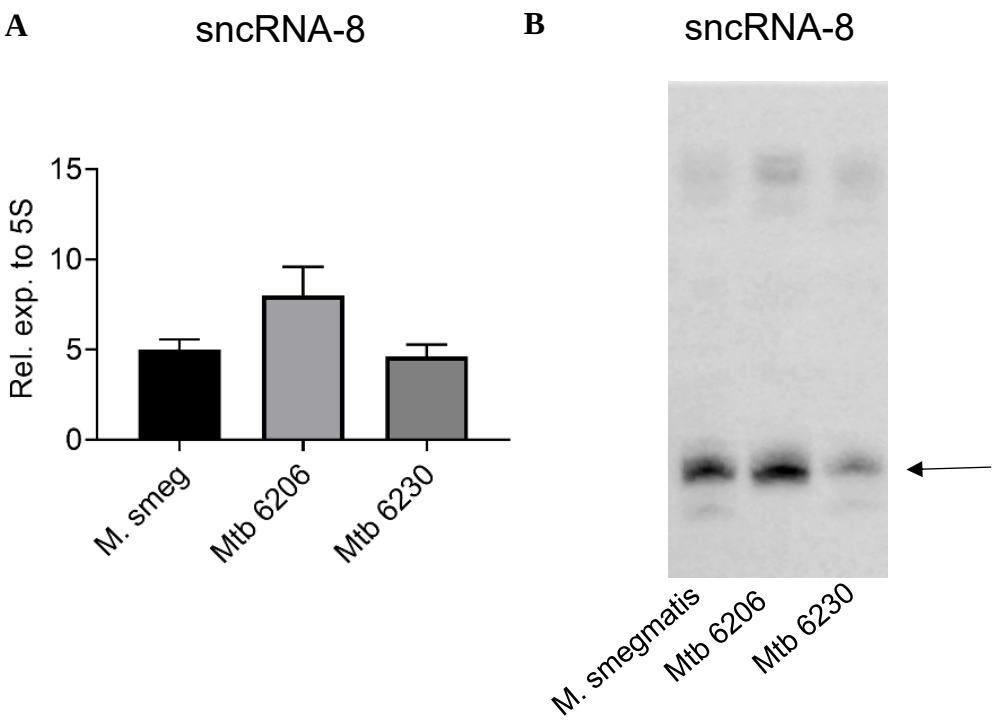


Figure 44. The mature sncRNA-8 is detected in diverse mycobacterial species including *M. smegmatis*, Mtb 6206 and Mtb 6230 with both miRNA-specific q-RT-PCR and northern blotting. RNA was isolated from *M. smegmatis*, Mtb 6206 and Mtb 6230 cultures at the stationary phase. (A) The expression of sncRNA-8 was quantified relative to 5S RNA using a miRNA-specific q-RT-PCR assay. (B) Equal amount from the remaining RNA was run on a gel for northern blotting where the membrane was probed for sncRNA-8. sncRNA-8 is shown with the arrow.

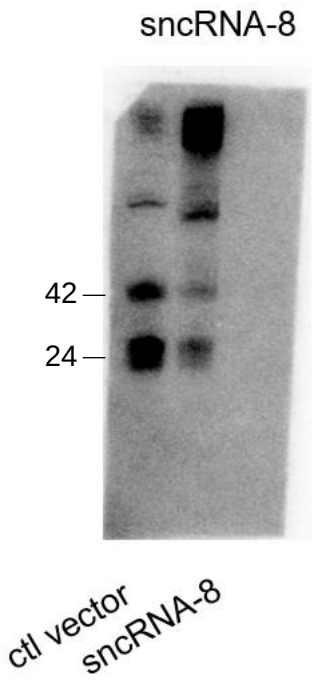


Figure 45. *M. smegmatis* does not process the precursor sncRNA-8 into the mature form when it is overexpressed. sncRNA-8 or the ctl vector was expressed in *M. smegmatis*. RNA was isolated to do a northern blotting. Probing the membrane for sncRNA-8 revealed that once overexpressed, most of sncRNA-8 is detected at >100 nts.

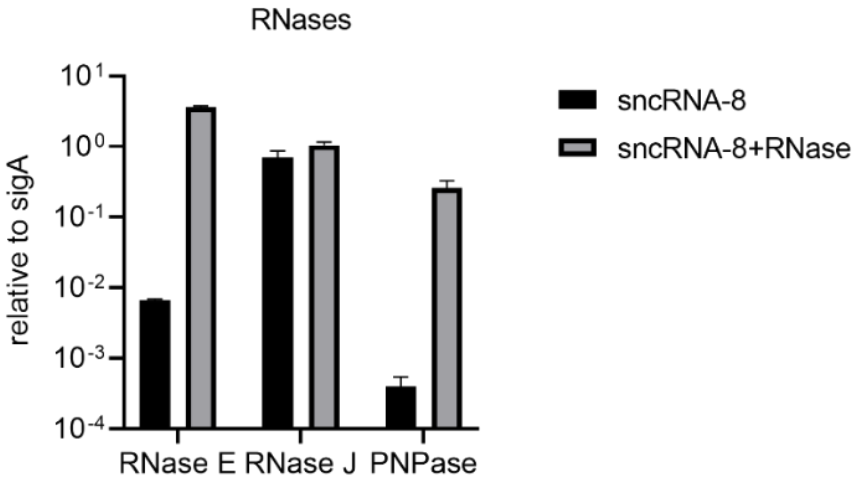


Figure 46. Diverse RNases are overexpressed Mtb 6230. RNase E, RNase J, or PNPase was electroporated to Mtb 6230 along with sncRNA-8 expression vector. Clones were selected on double selection media and propagated. RNA was isolated from cultures at the stationary phase and the corresponding RNase expression in the clones was quantified using conventional q-RT-PCR relative to *sigA* expression. Shown is the representative of the two experiments.

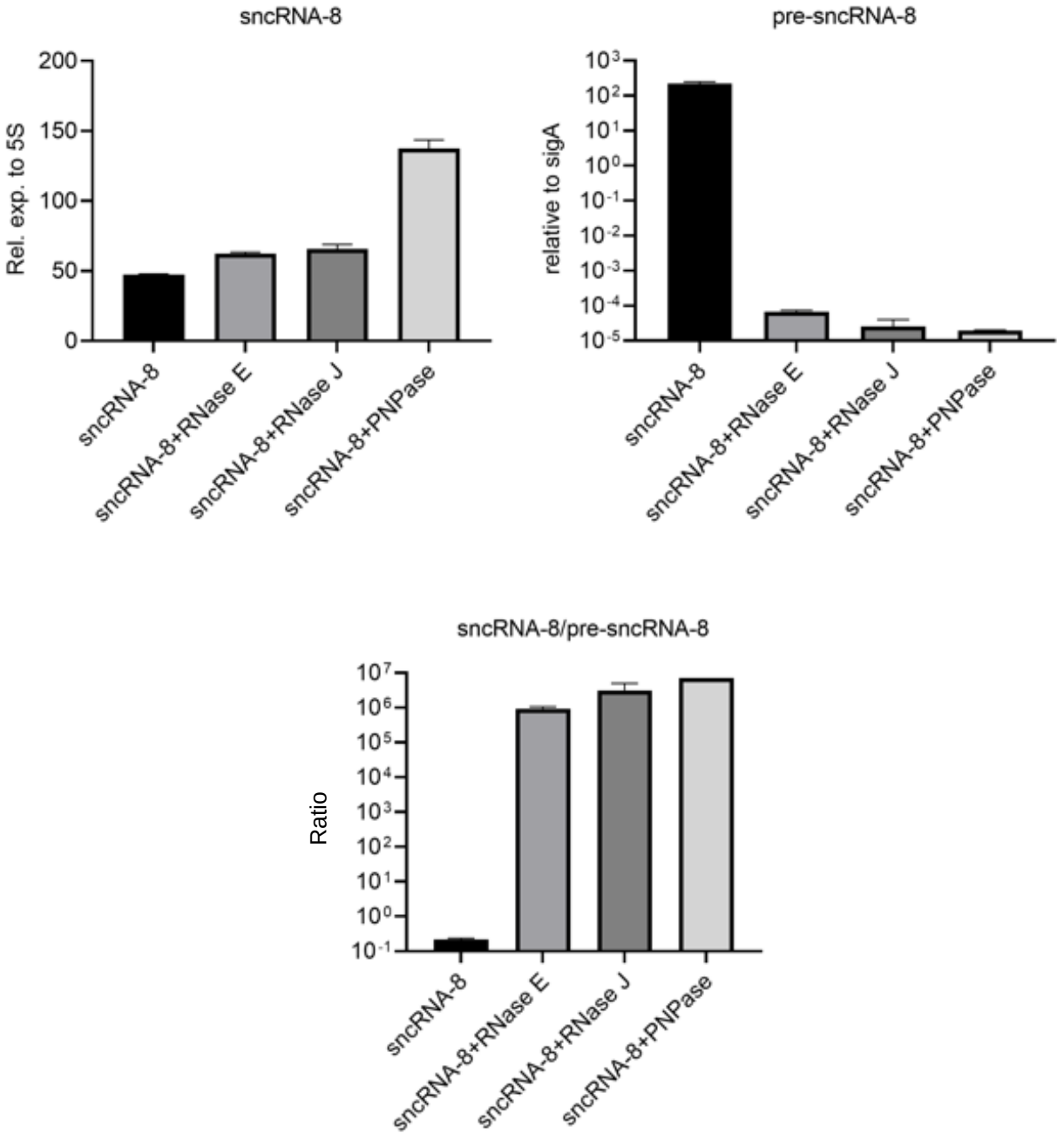

sncRNA-8
Rel. exp. to 5S
200
150
100
50
0
sncRNA-8
sncRNA-8+RNase E
sncRNA-8+RNase J
sncRNA-8+PNPase
pre-sncRNA-8
relative to sigA
10^3
10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
sncRNA-8
sncRNA-8+RNase E
sncRNA-8+RNase J
sncRNA-8+PNPase
sncRNA-8/pre-sncRNA-8
Ratio
10^7
10^6
10^5
10^4
10^3
10^2
10^1
10^0
10^{-1}
sncRNA-8
sncRNA-8+RNase E
sncRNA-8+RNase J
sncRNA-8+PNPase

Figure 47. RNases in the mycobacterial core RNA degradosome can rescue the sncRNA-8 processing defect in Mtb 6230 overexpressing sncRNA-8. RNase E, RNase J, or PNPase was electroporated to Mtb 6230 along with sncRNA-8 expression vector. Clones were selected on double selection media and propagated. RNA was isolated from cultures at the stationary phase. The expression of the pre-sncRNA-8 and the expression of sncRNA-8 was determined with a regular and a miRNA-specific q-RT-PCR, respectively (A-B). The ratio of sncRNA-8 to pre-sncRNA-8 was calculated by dividing the ΔΔCt value of sncRNA-8 to the ΔΔCt value of its precursor for each clone ©.

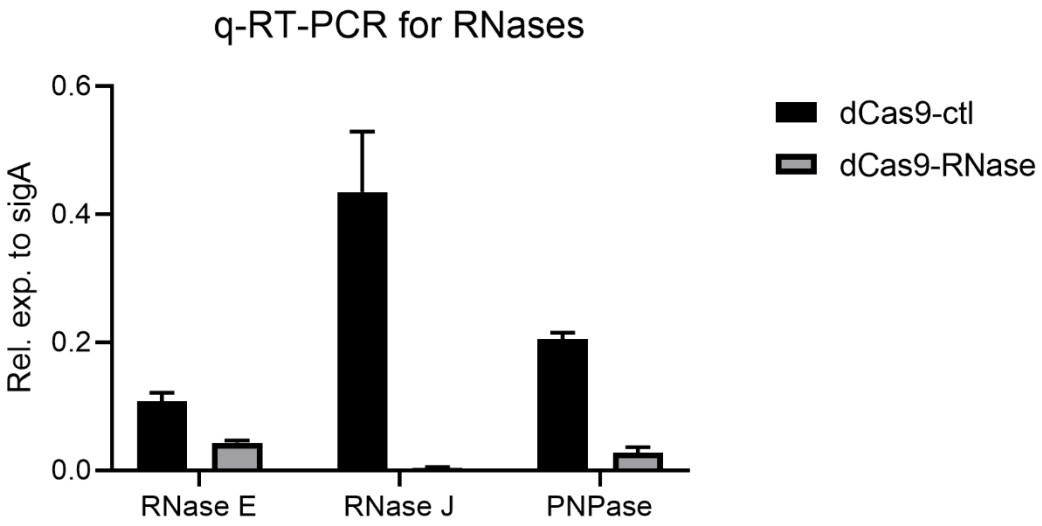


Figure 48. RNase E, RNase J and PNPase can be selectively knocked down in *M. smegmatis*. Non-targeting CRISPRi plasmid or the CRISPRi plasmids to knock down RNase E, RNase J or PNPase was electroporated to *M. smegmatis* and the clones were expanded. Once they reached the stationary phase, they were diluted with fresh media and the knock down was induced using ATc. Next day RNA was isolated, and the expression of RNases was quantified using conventional q-RT-PCR relative to *sigA* expression.

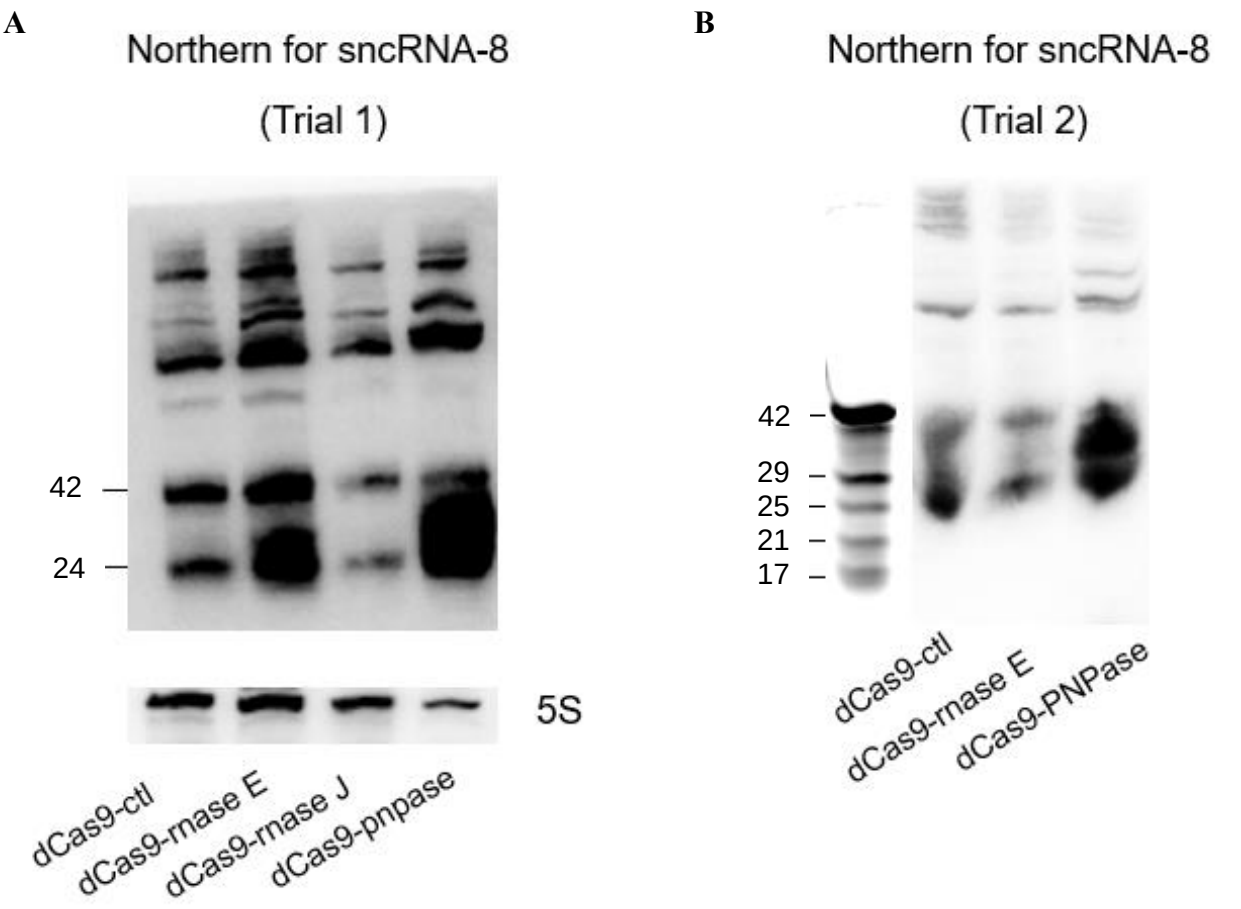


Figure 49. Endogenous sncRNA-8 processing is altered when RNase E and PNPase is knocked down in *M. smegmatis*. In trial 1, non-targeting CRISPRi plasmid or the CRISPRi plasmids to knock down RNase E, RNase J or PNPase was electroporated to *M. smegmatis* and the clones were expanded. Once they reached the stationary phase, they were diluted with fresh media and the knock down was induced using ATc. Next day RNA was isolated and run on a gel. (A) Northern blotting was performed using sncRNA-8 specific probes. (B) The membrane was later probed for 5S RNA. In trial 2, knock down of RNase J was omitted.

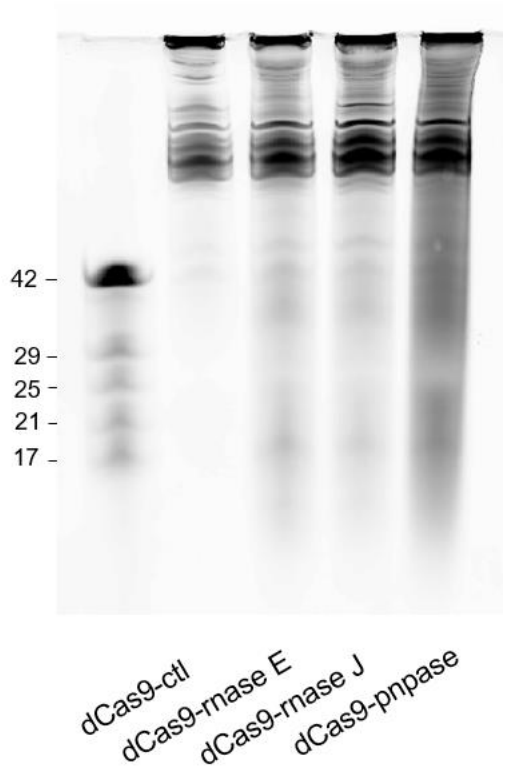


Figure 50. Low molecular weight RNA degradation is disrupted in *M. smegmatis* when PNPase is knocked down. RNA was isolated from *M. smegmatis* clones that had either the ctl CRISPRi vector or the CRISPRi vector targeting RNase E, RNase J or PNPase. The RNA was run on a gel and stained with GelRed for 30 mins. It was visualized with Typhoon 9500 Fluorescence filter.

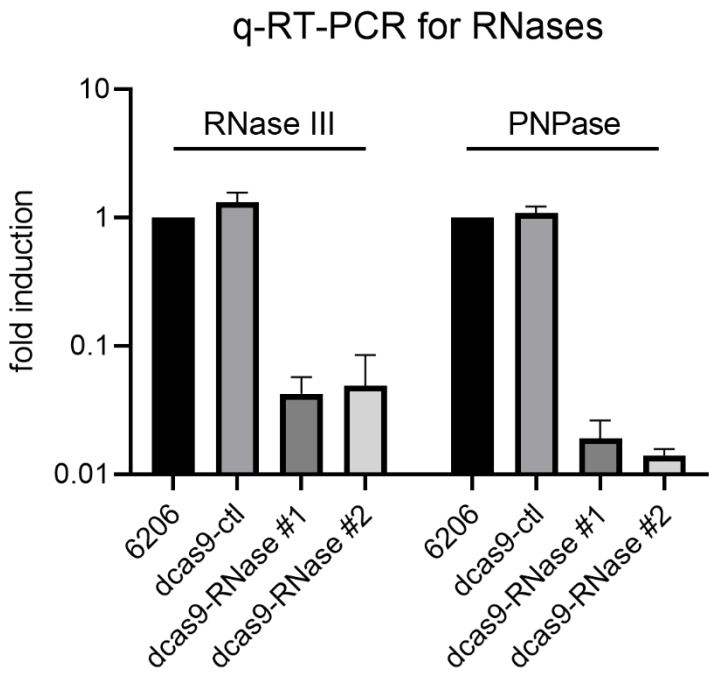


Figure 51. RNase III, RNase E and PNPase can be selectively knocked down in Mtb 6206. Ctl CRISPRi plasmid or the CRISPRi plasmids targeting RNase III, RNase E or PNPase were electroporated to Mtb 6206 and the clones were selected on plates. The cultures were expanded until they reached the stationary phase. Then, they were diluted to induce the knock down of RNases using ATc. Three days later, the RNA was isolated to quantify the expression of corresponding RNases with regular q-RT-PCR. The expression of RNase III (RNC) and PNPase was determined relative to *sigA* and then the level of gene expression in each sample was normalized to the level of gene expression in Mtb 6206 with no treatment. Shown are the data for 2 clones of Mtb 6206 with the targeting CRISPRi vectors.

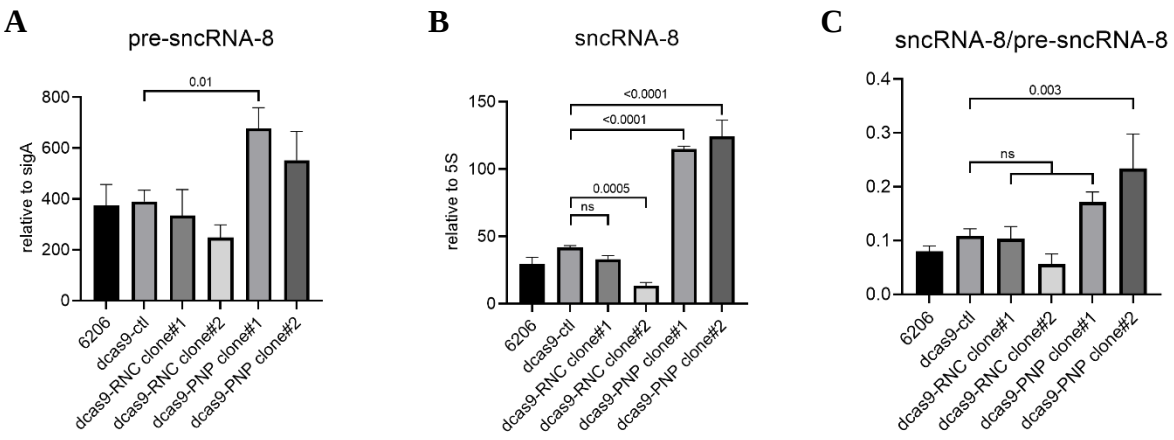


Figure 52. sncRNA-8 and its precursor accumulate when PNPase is knocked down in Mtb 6206. Ctl CRISPRi plasmid or the CRISPRi plasmids targeting RNase III, RNase E or PNPase were electroporated to Mtb 6206 and the clones were selected on plates. The cultures were expanded until they reached to the stationary phase. Then, they were diluted to induce the knock down of RNases using ATc. Three days later, the RNA was isolated. (A) Expression of the pre-sncRNA-8 was determined with regular q-RT-PCR and the (B) expression of 24-nt sncRNA-8 was assessed with a miRNA-specific RT-PCR. (C) The ratio of sncRNA-8 to its precursor was determined by dividing the level of sncRNA-8 to the level of pre-sncRNA-8 in each sample. Shown are the data for 2 clones of Mtb 6206 with the targeting CRISPRi vectors.

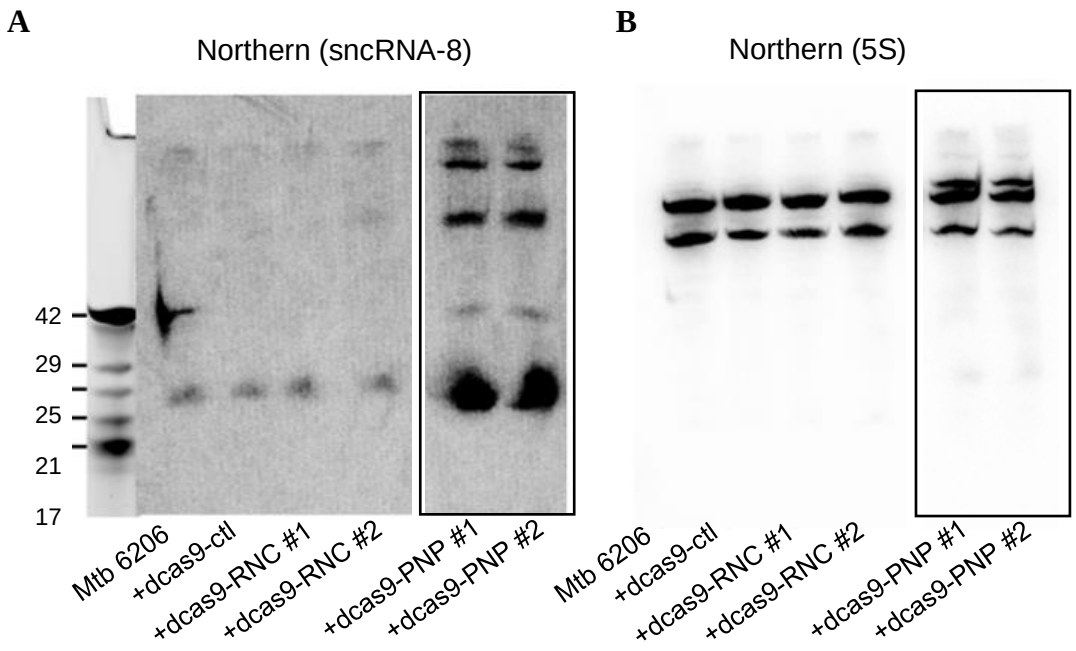


Figure 53. Northern blot confirms that sncRNA-8 and pre-sncRNA-8 accumulate when PNPase is knocked down in Mtb 6206. (A) Ctl CRISPRi plasmid or the CRISPRi plasmids targeting RNase III (RNC), RNase E (RNE) or PNPase (PNP) were electroporated to Mtb 6206 and the clones were selected on plates. The cultures were expanded until they reached the stationary phase. Then, they were diluted to induce the knock down of RNases using ATc. Three days later, the RNA was isolated and ran on a gel for northern blotting with a sncRNA-8 probe. (B) The membrane was stripped and probed again for the 5S RNA. Shown are the data for 2 clones of Mtb 6206 with the targeting CRISPRi vectors.

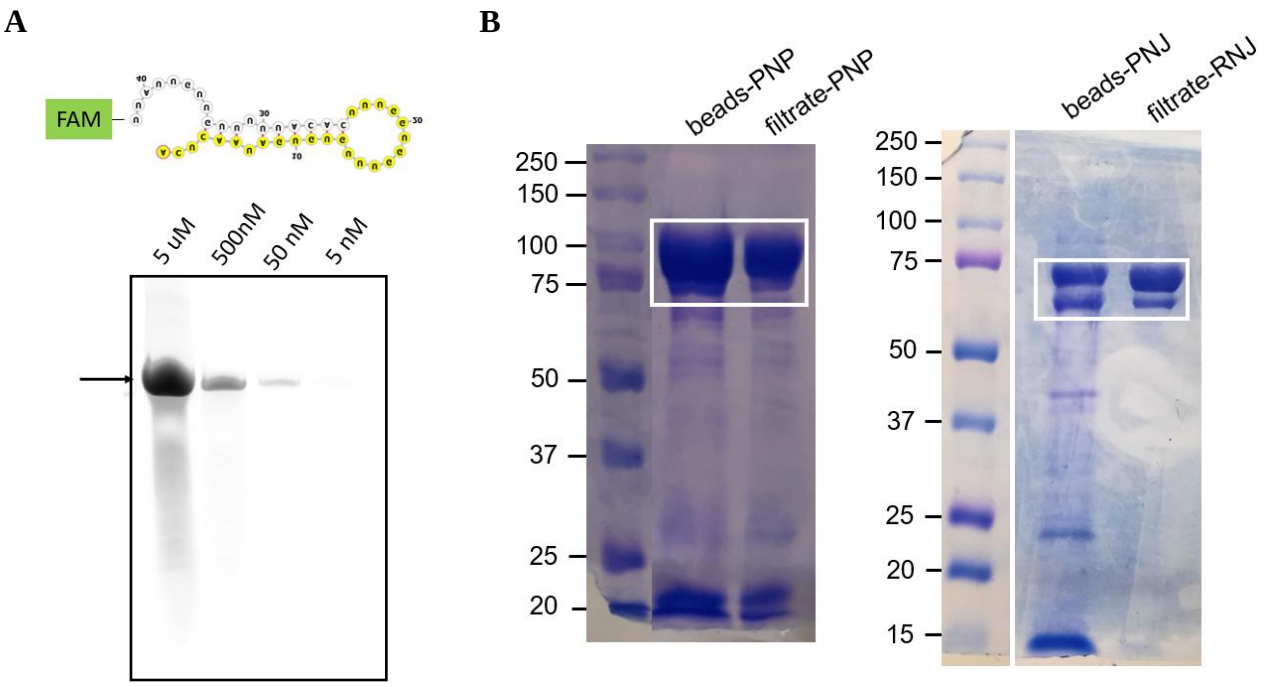


Figure 54. A schematic of the synthetic RNA containing sncRNA-8 sequence used as a substrate to test the catalytic activity of the recombinant PNPase. (A) An arbitrarily selected larger transcript containing sncRNA-8 sequence was synthesized. It was attached to FAM on the 5' end. This construct was diluted between 5 uM and 5 nM and run on a gel. The gel was visualized on Typhoon 9500 with the fluorescence filter with no staining which revealed dose dependent detection of the synthetic RNA. (B) PNPase and RNase J were recombinantly expressed in *E. coli* and purified with Ni-NTA agarose beads. The eluates were passed through desalting columns to get the filtrates. They were run on SDS gel and stained with Coomassie staining.

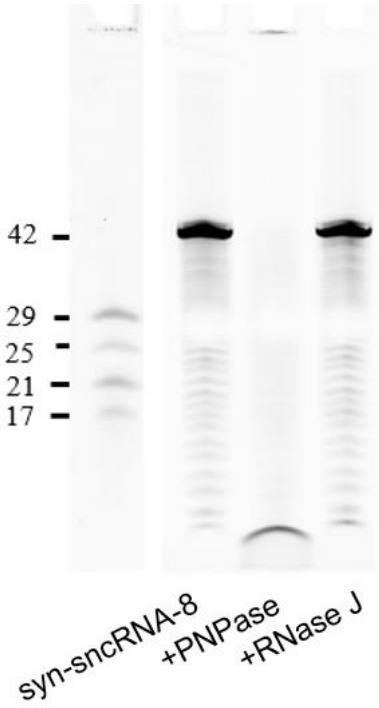


Figure 55. PNPase can degrade the putative pre-sncRNA-8 *in vitro*. The arbitrarily selected larger transcript containing sncRNA-8 sequence (syn-sncRNA-8) was incubated with either the recombinant PNPase or RNase J in the RNA degradation buffer at 37 ^{o}C for 30 mins. The samples were heated to 70 ^{o}C for 5 minutes and loaded onto the gel with the RNA running buffer. The gel was visualized using the Typhoon 9500 with the fluorescence filter. Then, it was stained with GelRed to visualize the marker.

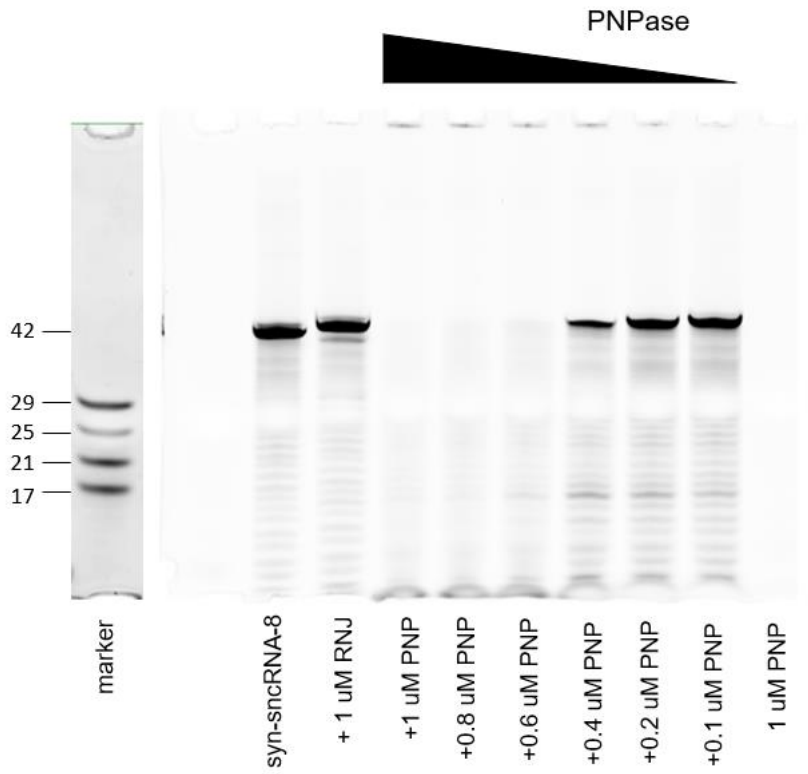


Figure 56. At least 0.4 uM of recombinant PNPase is necessary to degrade syn-sncRNA-8. Syn-sncRNA-8 was incubated with decreasing amounts of PNPase between 1 uM and 0.1 uM at 37 ^{o}C for 30 mins. The samples were heated to 70 ^{o}C for 5 minutes and loaded onto the gel with the RNA running buffer. RNase J was used as a negative ctl. Another ctl sample with the PNPase and the reaction buffer (no RNA) was loaded to the last lane. The gel was visualized using the Typhoon 9500 with the fluorescence filter. Then, it was stained with GelRed to visualize the marker.

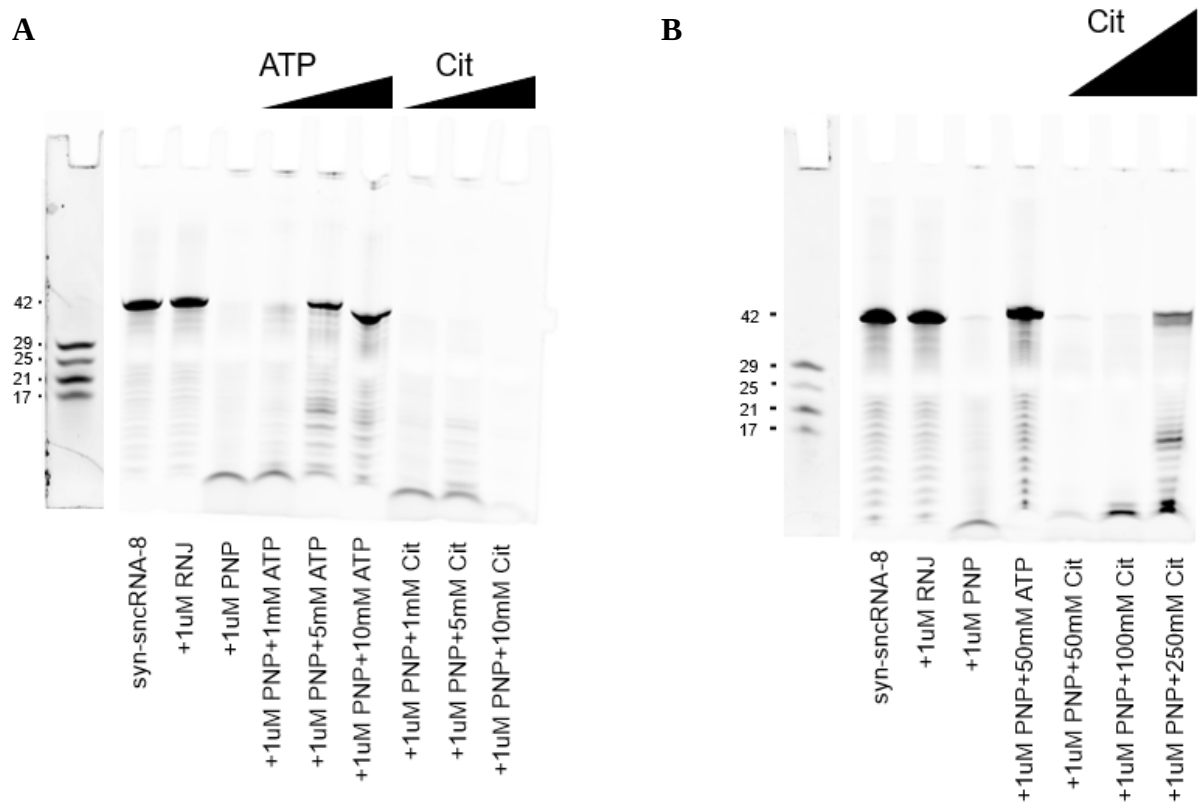


Figure 57. Low doses of ATP can block PNPase mediated degradation of syn-sncRNA-8, whereas high doses of citrate are required to inhibit PNPase activity. (A) Syn-sncRNA-8 was incubated with increasing amounts of ATP (1-10 Mm) and Citrate (1-10 mM) at 37 ^{o}C for 30 mins. (B) Higher doses of Citrate (50-250 mM) were tested to inhibit PNPase activity with 50 mM ATP being a positive ctl. The gels were visualized using the Typhoon 9500 with the fluorescence filter. Then, they were stained with GelRed to visualize the marker.

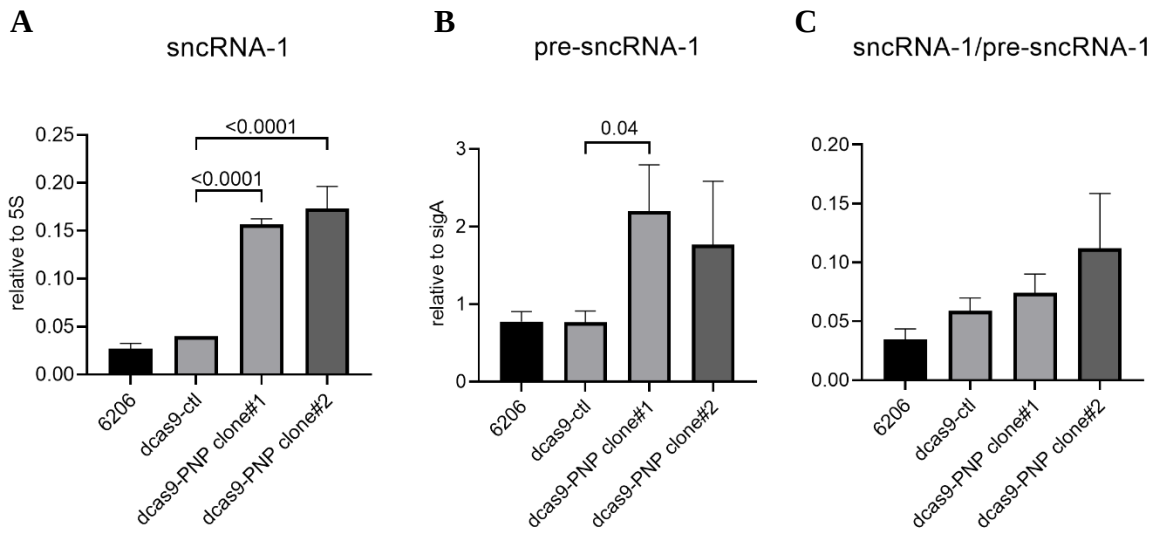


Figure 58. The role PNPase on sncRNA-8 degradation can be extended to sncRNA-1. Ctl CRISPRi plasmid or the CRISPRi plasmids targeting RNase III was electroporated to Mtb 6206 and the clones were selected on plates. The cultures were expanded until they reached the stationary phase. Then, they were diluted to induce the knock down of RNases using ATc. Three days later, the RNA was isolated to quantify the expression of sncRNA-1 and pre-sncRNA-1. (A) The expression of 25-nt sncRNA-1 was assessed with a miRNA-specific q-RT-PCR and the (B) expression of the pre-sncRNA-1 was determined with regular q-RT-PCR. (C) The ratio of sncRNA-1 to its precursor was determined by dividing the level of sncRNA-1 to the level of pre-sncRNA-1 in each sample. Shown are the data for 2 clones of Mtb 6206 with the targeting CRISPRi vectors.

CHAPTER 6

6.3 Overall Conclusions

Mycobacteria produce many distinct small RNAs, some of which are induced upon stress that can be due to environmental conditions, infections, hypoxia, and growth phases (K. B. Arnvig & Young, 2009; Coskun et al., 2020; Gerrick et al., 2018; Gottesman & Storz, 2011; Haning, Cho, & Contreras, 2014; Miotto et al., 2012). Most of the sRNAs range in size from 21-350 nts and were identified with conventional RNA-sequencing technologies. My work focused on RNAs <50 nts reported for many microbial species except *Mycobacterium tuberculosis* (Mtb). In my thesis, RNA-sequencing was used with Mtb-infected macrophage cultures to identify RNAs <50 nts produced by both eukaryotic and the prokaryotic cells, i.e. Mtb. The analysis detailed in Chapter 2 revealed both mammalian miRNAs, and previously unreported, a specific population of 35-Mtb-encoded RNAs that I named as smaller noncoding RNAs (sncRNA).

sncRNA designation was partly based on the predicted secondary structure of their precursor RNA intermediates with evidence of hairpin loops and other features more often defined for plant and animal sRNA. They ranged in size from 18-30 nucleotides and had an average GC content of 50%. They were distributed throughout the Mtb genome, with most contained in intergenic regions, and several replicated in multiple locations. Each was present in longer RNA sequences with predicted RNA secondary structures including stem loops. sncRNA-1, sncRNA-28, sncRNA-22 and sncRNA-30, were near the genes encoding virulence factors, including early secretory antigen target 6 (ESAT-6), ESAT-6-like proteins, culture fluid protein-10 (CFP-10), and Rv3876, suggesting an infection-dependent co-transcription. sncRNA-3, sncRNA-8, sncRNA-16 and sncRNA-33 clustered between Rvnr01 and Rv1315, ending at the coding sequences of a ribosomal RNA, which may indicate a role in the regulation of transcription. sncRNA-4, sncRNA-5,

sncRNA-6 and sncRNA-26 were present in multiple locations and often flanked by transposases, likely that these were being duplicated by transposition and perhaps important for transposition. Sequence similarities to validated herpeviridae and mammalian miRNAs were also noted. For example, sncRNA-5 had a similar stem-loop sequence as the γ-2 herpesvirus Rhesus primate rhadinovirus (RRV) rrv-miR-rR1-4, while sncRNA-34 has sequence similarity to rrv-miR-rR1-7-3p (O'Connor & Kedes, 2007). sncRNA-2 had a 12-base match with the miR-1185 family, while sncRNA-21 had similarity to miR-181a, a miR-181 family member that regulates thymopoiesis and Natural Killer T cell functions (Belkaya & van Oers, 2014; Henao-Mejia et al., 2013). Whether such similarities are reflected in the function of sncRNAs would be interesting to address.

sncRNA-1, sncRNA-6, and sncRNA-8 were further characterized, based in part on their surrounding RNA sequences with predicted stem loops with the strongest similarity to mammalian and plant pre-mature miRNAs. First, the expression of three of these, sncRNA-1, sncRNA-6, and sncRNA-8 were confirmed with a quantitative microRNA-based RT-PCR assay never used for such mycobacterial sncRNAs before. This analysis revealed that sncRNA-1 and sncRNA-6 are induced in infected cells implying a role in pathogenesis. sncRNA-8 levels were detectable independent of an infection yet enhanced in infected cells. This would suggest that sncRNA-8 has a function in free floating cells. Moreover, sncRNA-8 was detected in the lung biopsies of Rhesus macaque primates infected with Mtb albeit at low numbers. The scarcity of these noncoding RNAs from *in vivo* lung isolates could be due to the predominance of mammalian RNAs relative to the Mtb-encoded sRNAs in the lung biopsies, potential secretion of the sncRNAs in exosomes, and/or RNA degradation in the isolates. The low read count of the sncRNAs compared to mammalian miRNAs also suggests that the Mtb sncRNAs primarily impact mycobacterial gene expression, which in turn, would affect the eukaryotic transcriptome.

In Chapter 3, I functionally characterized 3 of the sncRNAs, sncRNA-1, sncRNA-6 and sncRNA-8 using a gain-of-function approach. The overexpression assays were done with Mtb H37Rv 6230 (ΔRD1, Δ*panCD*), an auxotrophic mutant of Mtb. This mutant belongs to fast-growing mycobacteria hence easier to handle for genetic studies. Plus, it can be used in BSL2 conditions (Pym et al., 2002). While sncRNA-1 and sncRNA-6 were successfully overexpressed in Mtb H37Rv 6230, the endogenous levels of sncRNA-8 prevented the overexpression of this sncRNA and hampering the functional analysis. Transcriptome analysis of the sncRNA-1 and sncRNA-6 overexpression clones suggested a different function for these sncRNAs. However, most of the genes differentially expressed in clones overexpressing sncRNA-6 belong to hypothetical protein category complicating further analysis. In depth functional studies have focused on sncRNA-1. sncRNA-1 was in the region of difference 1 (RD1) region of the Mtb genome, which contains critical genes involved in Mtb pathogenicity. The RD1 region is deleted in Mtb H37Rv 6230 mutant, enabling me to use a mycobacterial strain without sncRNA-1. KEGG GO term analysis of the differentially expressed genes in clones that had sncRNA-1 expression vector compared to the control vector revealed it regulates genes involved in fatty acid metabolism, which was unique to this sncRNA.

In Chapter 4, I studied the impact of sncRNA-1 on two of the putative targets, Rv0242c and Rv1094. Both genes were upregulated in clones overexpressing sncRNA-1 and they are components of oleic acid biogenesis pathway in Mtb. Each had a putative sncRNA-1 binding site in their respective 5’ UTR, which is a common regulatory site for small RNAs. The sncRNA-1 regulation of the 5’UTR was confirmed for Rv0242c by introducing reciprocal mutations in the putative regulatory site within sncRNA-1 and/or the 5’ UTR of Rv0242c. Interestingly, Mtb lacking sncRNA-1 had growth defects in oleic acid deficient media, whereas this was rescued by

sncRNA-1 expression. Moreover, clones lacking sncRNA-1 had survival defects in infected macrophages, whereas clones overexpressing sncRNA-1 had survival benefit. In infected cells, Mtb primarily relies on host lipids in macrophages, which is a rich source of oleic acid (Agarwal et al., 2020; Nazarova et al., 2019). While the upregulation of the Mtb-encoded genes that enable oleic acid synthesis might be counterintuitive, I speculate that sncRNA-1 upregulation at late stages of infection may be an anticipatory response to the upcoming fatty acid starvation. The knock down of sncRNA-1 using locked nucleic acid power inhibitors (LNA-PI) further supported a role for this sncRNA in mycobacterial survival. Specifically, antagonizing sncRNA-1 reduced the numbers of CFUs/well in macrophages. This suggests that sncRNA-1 may target additional genes that are able to modulate mycobacterial infectivity. Overall, these data show that sncRNA-1 contributes to Mtb pathogenicity.

In Chapter 5, I initiated a search to identify putative processing enzymes responsible for the formation of the mature sncRNAs. I discovered that mutations in the predicted hairpin loop structure of both sncRNA-1 and sncRNA-6 eliminates formation of the mature sncRNA. Using the detection of mature sncRNAs relative to their precursor as a readout of processing, I examined the role of several recently characterized core RNA degradosome components on sncRNA processing. CRISPR interference-based knock down of RNase E, RNase J and PNPase, which make the core RNA degradosome in Mtb and another enzyme RNase III, which digests dsRNA with hairpin loop revealed that RNase E, RNase J and RNase III are not involved the specific cleavage of the precursor of the sncRNA into the respective mature form. Of significance, the selective knock down of PNPase led to an accumulation of both the mature and precursor forms of sncRNA-8 and sncRNA-1. This result suggested that PNPase directly degrades all small RNAs in the nucleotides *in vivo*. This was confirmed for the *in vitro* RNA degradation assays, in which

PNPase cleaves and degrades sncRNAs. Thus, it remains an open question about the identity of the RNA processing enzymes that mediate site-specific cleavages of larger RNA transcripts into the mature sncRNAs. More studies are needed to elucidate the regulatory role of PNPase in Mtb.

6.4 Future directions

In this thesis, Mtb-encoded sncRNAs have been identified and for three of these, sncRNA-1, sncRNA-6 and sncRNA-8, functional studies have been performed (Figure 59). More studies are needed to elucidate the function of sncRNA-6 and sncRNA-8 as the gain-of-function approaches have been incompetent. Loss-of-function studies using LNA-PIs may allow us to better characterize the function of these sncRNAs in the future. However, the role of sncRNA-1 has been characterized in detail and it was shown that it regulates Rv0242c. Rv0242c encodes FabG4, 3-oxoacyl-thioester reductase, which is essential for the *in vitro* Mtb growth and for a successful infection (DeJesus et al., 2017; Dutta, 2018; Sassetti & Rubin, 2003). Rv0242c is induced in Mtb exposed to streptomycin implicating a role for it in mycobacterial resistance to this antibiotic (Sharma et al., 2010). Therefore, sncRNA-1 mediated regulation on Rv0242c suggests a role for sncRNA-1 in antibiotic resistance, which needs to be addressed in the future.

FabG4 mediates the production of oleic acid, which is an unsaturated fatty acid in Mtb (Dutta, 2018). Oleic acid alters the membrane fluidity in Mtb due to the presence of double bonds (Bloch & Segal, 1956). Plus, changes in the ratio of long chain to short chain fatty acids also contributes to membrane fluidity (Murinova & Dercova, 2014). Such mechanisms enable Mtb to adopt to extreme environmental conditions (Kang et al., 2013). Our data indicated that sncRNA-1 increases the expression of Rv0242c transcript. However, the contribution of sncRNA-1-mediated upregulation of Rv0242c to membrane fluidity hence bacterial adaptation to cell lysis remains

unexplored. Future studies might focus on the role of sncRNA-1 in changing membrane composition of Mtb through the regulation of Rv0242c.

Mtb primarily depends on the host lipids in infected macrophages, which are rich in oleic acid (Agarwal et al., 2020; Hidalgo, Carretta, & Burgos, 2021; Nazarova et al., 2019). Our data suggested that Mtb also enhances intrinsic oleic acid production by increasing the expression of genes involved in oleic acid production through sncRNA-1. Therefore, it is likely that there is an interplay between intrinsic oleic acid production and the uptake of oleic acid from the macrophages. It can be speculated that Mtb depends on oleic acid imported from the host in the early phases of the infection, whereas the upregulation of sncRNA-1 may be an anticipatory response to an upcoming fatty acid starvation. It is well established that macrophages limit nutrients to starve bacteria upon infection, which is commonly referred to as nutritional immunity (Cumming, Addicott, Adamson, & Steyn, 2018). Future studies will elucidate the kinetics of the oleic acid metabolism in Mtb in infected macrophages to better understand this interplay.

The mechanisms by which sncRNAs get induced in the first place also remains unexplored. Two potential mechanisms include the production of sncRNAs in operons and their independent transcription. In the first hypothesis, sncRNAs are co-transcribed with their neighboring genes which get cleaved to release sncRNAs or sncRNA precursors. In the second hypothesis, sncRNA production requires the involvement of cis- and trans-regulatory elements that mediate sncRNA transcription. Both hypotheses are plausible currently, and they need to be tested in the future. Many questions remain unanswered regarding the regulatory elements that control the expression of the sncRNAs.

Despite our attempts to identify the sncRNA processing enzymes in Mtb, how sncRNAs are maturated remains unknown. I have shown that RNase III and the core RNA degradosome

including RNase J and RNase E are not involved in sncRNA processing. My data indicated that PNPase has a role in sncRNA degradation. However, the role of other RNA processing enzymes that mediate the specific cleavage and site-directed modification of the precursor sncRNAs remain unexplored. In fact, these processing enzymes may be growth phase or stress-dependent, altering the function of the RNA in selected physiological conditions, which also needs to be addressed. The mycobacterial proteins related to RNA processing are not well annotated. Using Mycobrowser, candidate genes involved in RNA processing in Mtb are shown (Table 6). However, more studies are needed to investigate the role of these putative RNA binding/processing proteins since most have been identified via computational predictions. As summarized, the last decade has identified many distinct Mtb-encoded RNAs. The next decade will likely characterize RNA processing enzymes in Mtb. Identification of new sRNAs/sncRNAs involved in pathogenesis and their regulatory mechanisms will enhance our understanding of tools that Mtb utilizes to escape macrophage killing, which will eventually help eradicate the TB.

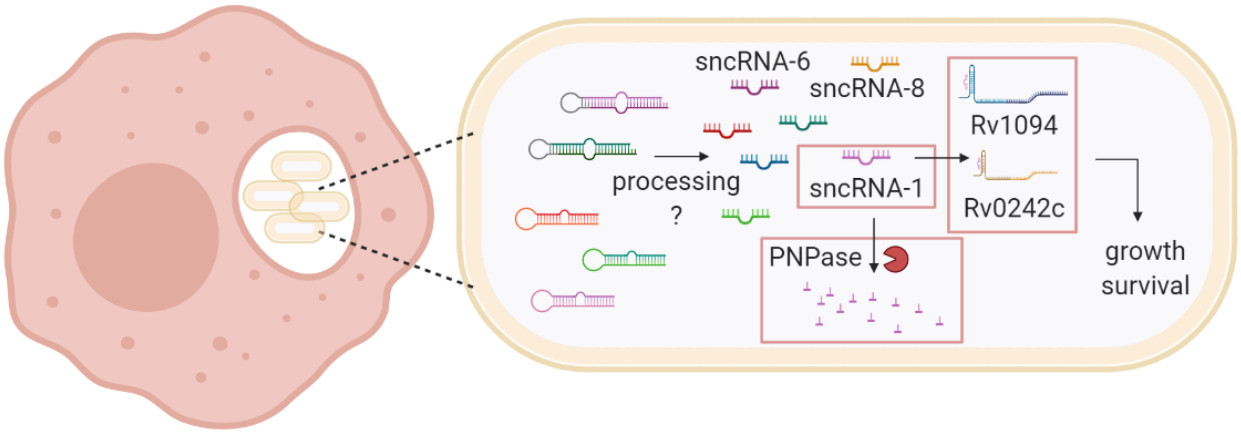


Figure 59. Identification of smaller noncoding RNAs produced by *Mycobacterium tuberculosis* in infected macrophages that regulate Mtb growth and survival is summarized. Mtb produces a set of sncRNAs in infected macrophages, three of which are sncRNA-1, sncRNA-6 and sncRNA-8. SncRNA-1 enhances the expression of Rv0242c and Rv1094 transcript, which induces growth and survival of Mtb. PNPase, a component of the core RNA degradosome, degrades the sncRNAs. The induction and processing requirements for sncRNAs remain unknown.

Table 6. Mtb-encoded RNA processing enzymes or components of the degradosome implicated in small RNA formation identified with Mycobrowser.

Gene ID	Name	Species	Putative Function
MTB000026	rnpB	*M. bovis, Mtb, M. haemophilum*	RNA component of RNase P: RNase P catalyzes the removal of the 5'-leader sequence from pre-tRNA to produce the mature 5' terminus.
Rv1340	rphA	*M. marinum, M. leprae, M. bovis, Mtb*	Probable ribonuclease RphA (RNase PH)
Rv2092c	helY	*M. marinum, M. leprae, M. bovis, Mtb, M. abscessus*	DNA helicase activity
Rv2179c	rnt	*Mtb, M. smegmatis, M. leprae, M. marinum*	Conserved hypothetical protein
Rv2228c	Rv2228c	*N/A*	Multifunctional protein. Has RNase H, alpha-ribazole phosphatase, and acid phosphatase activities.
Rv2407	rnz	*Mtb, M. smegmatis, M. leprae, M. marinum, M. bovis*	Endonucleolytic cleavage of RNA, removing extra 3' nucleotides from tRNA precursor, generating 3' termini of tRNAs.
Rv2444c	rne	*M. bovis, Mtb, M. leprae, M. marinum, M. smegmatis*	Putative RNase E. Plays central role in the maturation of 5S and 16S rRNAs and the majority of tRNAs.

			Also involved in the degradation of most mRNAs.
Rv2511	orn	*Mtb, M. smegmatis, M. leprae, M. marinum, M. bovis*	Involved in RNA degradation: 3'-to-5' exoribonuclease specific for small oligoribonucleotides.
Rv2681	rnd	*Mtb, M. smegmatis, M. leprae, M. marinum, M. bovis*	Conserved hypothetical protein
Rv2752c	rnj	*M. bovis, Mtb, M. leprae, M. marinum, M. smegmatis*	Conserved hypothetical protein
Rv2783c	gpsI (pnp)	*M. marinum, M. leprae, M. bovis, Mtb, M. smegmatis, M. abscessus*	Involved in mRNA degradation. Hydrolyses single-stranded polyribonucleotides processively in the 3' to 5' direction.
Rv2902c	rnhB	*M. marinum, M. leprae, M.bovis, Mtb, M. abscessus*	Probable ribonuclease HII protein RnhB
Rv2907c	rimM	*M. marinum, M. leprae, M. bovis, Mtb, M. smegmatis, M. abscessus*	Essential for efficient processing of 16S rRNA. Probably part of the 30S subunit prior to or during the final step in the processing of 16S free 30S ribosomal subunits. It could be some accessory protein needed for efficient assembly of the 30S subunit.

Rv2925c	rnc	*M. marinum, M. leprae, M. bovis, Mtb, M. smegmatis, M. abscessus*	Digests double-stranded RNA. Involved in the processing of ribosomal RNA precursors and of some mRNAs
Rv3853	rraA	*M. leprae, M. bovis, Mtb*	Regulator of RNase E activity a RraA
Rv3923c	rnpA	*M. marinum, M. leprae, M. bovis, Mtb, M. smegmatis, M. abscessus*	Ribonuclease P protein component RnpA

REFERENCES

Afonyushkin, T., Vecerek, B., Moll, I., Blasi, U., & Kaberdin, V. R. (2005). Both RNase E and RNase III control the stability of sodB mRNA upon translational inhibition by the small regulatory RNA RyhB. *Nucleic Acids Res, 33*(5), 1678-1689. doi:10.1093/nar/gki313

Agarwal, P., Combes, T. W., Shojaee-Moradie, F., Fielding, B., Gordon, S., Mizrahi, V., & Martinez, F. O. (2020). Foam Cells Control Mycobacterium tuberculosis Infection. *Front Microbiol, 11*, 1394. doi:10.3389/fmicb.2020.01394

Agrawal, N., Dasaradhi, P. V., Mohmmed, A., Malhotra, P., Bhatnagar, R. K., & Mukherjee, S. K. (2003). RNA interference: biology, mechanism, and applications. *Microbiol Mol Biol Rev, 67*(4), 657-685. doi:10.1128/MMBR.67.4.657-685.2003

Akey, D. L., & Berger, J. M. (2005). Structure of the nuclease domain of ribonuclease III from M. tuberculosis at 2.1 A. *Protein Sci, 14*(10), 2744-2750. doi:10.1110/ps.051665905

Ali Syeda, Z., Langden, S. S. S., Munkhzul, C., Lee, M., & Song, S. J. (2020). Regulatory Mechanism of MicroRNA Expression in Cancer. *Int J Mol Sci, 21*(5). doi:10.3390/ijms21051723

Alifano, P., Bruni, C. B., & Carlomagno, M. S. (1994). Control of mRNA processing and decay in prokaryotes. *Genetica, 94*(2-3), 157-172. doi:10.1007/BF01443430

Andrade, J. M., Pobre, V., Matos, A. M., & Arraiano, C. M. (2012). The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA, 18*(4), 844-855. doi:10.1261/rna.029413.111

Arnvig, K., & Young, D. (2012). Non-coding RNA and its potential role in Mycobacterium tuberculosis pathogenesis. *RNA Biol, 9*(4), 427-436. doi:10.4161/rna.20105

Arnvig, K. B., Comas, I., Thomson, N. R., Houghton, J., Boshoff, H. I., Croucher, N. J., . . . Young, D. B. (2011). Sequence-based analysis uncovers an abundance of non-coding RNA in the total transcriptome of Mycobacterium tuberculosis. *PLoS Pathog, 7*(11), e1002342. doi:10.1371/journal.ppat.1002342

Arnvig, K. B., & Young, D. B. (2009). Identification of small RNAs in Mycobacterium tuberculosis. *Mol Microbiol, 73*(3), 397-408. doi:10.1111/j.1365-2958.2009.06777.x

Baek, Y. M., Jang, K. J., Lee, H., Yoon, S., Baek, A., Lee, K., & Kim, D. E. (2019). The bacterial endoribonuclease RNase E can cleave RNA in the absence of the RNA chaperone Hfq. *J Biol Chem, 294*(44), 16465-16478. doi:10.1074/jbc.RA119.010105

Barnhill, E. C., Crucello, A., Houserova, D., King, V. M., Amin, S. V., Roberts, J. T., . . . Borchert, G. M. (2019). Characterization of novel small RNAs (sRNAs) contributing to the desiccation response of Salmonella enterica serovar Typhimurium. *RNA Biol, 16*(11), 1643-1657. doi:10.1080/15476286.2019.1653680

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell, 136*(2), 215-233. doi:10.1016/j.cell.2009.01.002

Behr, M. A., Edelstein, P. H., & Ramakrishnan, L. (2018). Revisiting the timetable of tuberculosis. *BMJ, 362*, k2738. doi:10.1136/bmj.k2738

Belkaya, S., Silge, R. L., Hoover, A. R., Medeiros, J. J., Eitson, J. L., Becker, A. M., . . . van Oers, N. S. (2011). Dynamic modulation of thymic microRNAs in response to stress. *PLoS One, 6*(11), e27580. doi:10.1371/journal.pone.0027580

Belkaya, S., & van Oers, N. S. (2014). Transgenic expression of microRNA-181d augments the stress-sensitivity of CD4(+)CD8(+) thymocytes. *PLoS One, 9*(1), e85274. doi:10.1371/journal.pone.0085274

Bloch, H., & Segal, W. (1956). Biochemical differentiation of Mycobacterium tuberculosis grown in vivo and in vitro. *J Bacteriol, 72*(2), 132-141. doi:10.1128/jb.72.2.132-141.1956

Blondal, T., Jensby Nielsen, S., Baker, A., Andreasen, D., Mouritzen, P., Wrang Teilum, M., & Dahlsveen, I. K. (2013). Assessing sample and miRNA profile quality in serum and plasma or other biofluids. *Methods, 59*(1), S1-6. doi:10.1016/j.ymeth.2012.09.015

Cambau, E., & Drancourt, M. (2014). Steps towards the discovery of Mycobacterium tuberculosis by Robert Koch, 1882. *Clin Microbiol Infect, 20*(3), 196-201. doi:10.1111/1469-0691.12555

Cameron, T. A., Matz, L. M., & De Lay, N. R. (2018). Polynucleotide phosphorylase: Not merely an RNase but a pivotal post-transcriptional regulator. *PLoS Genet, 14*(10), e1007654. doi:10.1371/journal.pgen.1007654

Carrier, M. C., Lalaouna, D., & Masse, E. (2018). Broadening the Definition of Bacterial Small RNAs: Characteristics and Mechanisms of Action. *Annu Rev Microbiol, 72*, 141-161. doi:10.1146/annurev-micro-090817-062607

Cavanagh, A. T., & Wassarman, K. M. (2014). 6S RNA, a global regulator of transcription in Escherichia coli, Bacillus subtilis, and beyond. *Annu Rev Microbiol, 68*, 45-60. doi:10.1146/annurev-micro-092611-150135

Chakrabarty, S., Kumar, A., Raviprasad, K., Mallya, S., Satyamoorthy, K., & Chawla, K. (2019). Host and MTB genome encoded miRNA markers for diagnosis of tuberculosis. *Tuberculosis (Edinb), 116*, 37-43. doi:10.1016/j.tube.2019.04.002

Chen, J., Morita, T., & Gottesman, S. (2019). Regulation of Transcription Termination of Small RNAs and by Small RNAs: Molecular Mechanisms and Biological Functions. *Front Cell Infect Microbiol, 9*, 201. doi:10.3389/fcimb.2019.00201

Chousalkar, K. K., Cheetham, B. F., & Roberts, J. R. (2009). LNA probe-based real-time RT-PCR for the detection of infectious bronchitis virus from the oviduct of unvaccinated and vaccinated laying hens. *J Virol Methods, 155*(1), 67-71. doi:10.1016/j.jviromet.2008.09.028

Cohen, S. B., Gern, B. H., Delahaye, J. L., Adams, K. N., Plumlee, C. R., Winkler, J. K., . . . Urdahl, K. B. (2018). Alveolar Macrophages Provide an Early Mycobacterium tuberculosis Niche and Initiate Dissemination. *Cell Host Microbe, 24*(3), 439-446 e434. doi:10.1016/j.chom.2018.08.001

Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., . . . Barrell, B. G. (1998). Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature, 393*(6685), 537-544. doi:10.1038/31159

Conradie, F., Diacon, A. H., Ngubane, N., Howell, P., Everitt, D., Crook, A. M., . . . Nix, T. B. T. T. (2020). Treatment of Highly Drug-Resistant Pulmonary Tuberculosis. *N Engl J Med, 382*(10), 893-902. doi:10.1056/NEJMoa1901814

Coskun, F. S., Srivastava, S., Raj, P., Dozmorov, I., Belkaya, S., Mehra, S., . . . van Oers, N. S. C. (2020). sncRNA-1 Is a Small Noncoding RNA Produced by Mycobacterium tuberculosis in Infected Cells That Positively Regulates Genes Coupled to Oleic Acid Biosynthesis. *Front Microbiol, 11*, 1631. doi:10.3389/fmicb.2020.01631

Court, D. L., Gan, J., Liang, Y. H., Shaw, G. X., Tropea, J. E., Costantino, N., . . . Ji, X. (2013). RNase III: Genetics and function; structure and mechanism. *Annu Rev Genet, 47*, 405-431. doi:10.1146/annurev-genet-110711-155618

Cumming, B. M., Addicott, K. W., Adamson, J. H., & Steyn, A. J. (2018). Mycobacterium tuberculosis induces decelerated bioenergetic metabolism in human macrophages. *Elife, 7*. doi:10.7554/eLife.39169

Dang, T. H. Y., Tyagi, S., D'Cunha, G., Bhave, M., Crawford, R., & Ivanova, E. P. (2019). Computational prediction of microRNAs in marine bacteria of the genus Thalassospira. *PLoS One, 14*(3), e0212996. doi:10.1371/journal.pone.0212996

DeJesus, M. A., Gerrick, E. R., Xu, W., Park, S. W., Long, J. E., Boutte, C. C., . . . Ioerger, T. R. (2017). Comprehensive Essentiality Analysis of the Mycobacterium tuberculosis Genome via Saturating Transposon Mutagenesis. *mBio, 8*(1). doi:10.1128/mBio.02133-16

Del Favero, M., Mazzantini, E., Briani, F., Zangrossi, S., Tortora, P., & Deho, G. (2008). Regulation of Escherichia coli polynucleotide phosphorylase by ATP. *J Biol Chem, 283*(41), 27355-27359. doi:10.1074/jbc.C800113200

Dheda, K., Barry, C. E., 3rd, & Maartens, G. (2016). Tuberculosis. *Lancet, 387*(10024), 1211-1226. doi:10.1016/S0140-6736(15)00151-8

DiChiara, J. M., Contreras-Martinez, L. M., Livny, J., Smith, D., McDonough, K. A., & Belfort, M. (2010). Multiple small RNAs identified in Mycobacterium bovis BCG are also expressed in Mycobacterium tuberculosis and Mycobacterium smegmatis. *Nucleic Acids Res, 38*(12), 4067-4078. doi:10.1093/nar/gkq101

Doench, J. G., & Sharp, P. A. (2004). Specificity of microRNA target selection in translational repression. *Genes Dev, 18*(5), 504-511. doi:10.1101/gad.1184404

Dozmorov, I., & Centola, M. (2003). An associative analysis of gene expression array data. *Bioinformatics, 19*(2), 204-211. doi:10.1093/bioinformatics/19.2.204

Dozmorov, I., & Lefkovits, I. (2009). Internal standard-based analysis of microarray data. Part 1: analysis of differential gene expressions. *Nucleic Acids Res, 37*(19), 6323-6339. doi:10.1093/nar/gkp706

Dozmorov, I. M., Jarvis, J., Saban, R., Benbrook, D. M., Wakeland, E., Aksentijevich, I., . . . Lefkovits, I. (2011). Internal standard-based analysis of microarray data2--analysis of functional associations between HVE-genes. *Nucleic Acids Res, 39*(18), 7881-7899. doi:10.1093/nar/gkr503

Dutta, D. (2018). Advance in Research on Mycobacterium tuberculosis FabG4 and Its Inhibitor. *Front Microbiol, 9*, 1184. doi:10.3389/fmicb.2018.01184

Ehrt, S., & Schnappinger, D. (2009). Mycobacterial survival strategies in the phagosome: defence against host stresses. *Cell Microbiol, 11*(8), 1170-1178. doi:10.1111/j.1462-5822.2009.01335.x

Eitson, J. L., Medeiros, J. J., Hoover, A. R., Srivastava, S., Roybal, K. T., Ainsa, J. A., . . . van Oers, N. S. (2012). Mycobacterial shuttle vectors designed for high-level protein expression in infected macrophages. *Appl Environ Microbiol, 78*(19), 6829-6837. doi:10.1128/AEM.01674-12

Esteller, M. (2011). Non-coding RNAs in human disease. *Nat Rev Genet, 12*(12), 861-874. doi:10.1038/nrg3074

Feng, Z. S., Zhao, L., Wang, J., Qiu, F. Z., Zhao, M. C., Wang, L., . . . Ma, X. J. (2018). A multiplex one-tube nested real time RT-PCR assay for simultaneous detection of respiratory syncytial virus, human rhinovirus and human metapneumovirus. *Virol J, 15*(1), 167. doi:10.1186/s12985-018-1061-0

Fu, Y., Li, W., Wu, Z., Tao, Y., Wang, X., Wei, J., . . . Zhang, F. (2018). Detection of mycobacterial small RNA in the bacterial culture supernatant and plasma of patients with

active tuberculosis. *Biochem Biophys Res Commun, 503*(2), 490-494. doi:10.1016/j.bbrc.2018.04.165

Furuse, Y., Finethy, R., Saka, H. A., Xet-Mull, A. M., Sisk, D. M., Smith, K. L., . . . Cullen, B. R. (2014). Search for microRNAs expressed by intracellular bacterial pathogens in infected mammalian cells. *PLoS One, 9*(9), e106434. doi:10.1371/journal.pone.0106434

Gao, L. Y., Guo, S., McLaughlin, B., Morisaki, H., Engel, J. N., & Brown, E. J. (2004). A mycobacterial virulence gene cluster extending RD1 is required for cytolysis, bacterial spreading and ESAT-6 secretion. *Mol Microbiol, 53*(6), 1677-1693. doi:10.1111/j.1365-2958.2004.04261.x

Gerrick, E. R., Barbier, T., Chase, M. R., Xu, R., Francois, J., Lin, V. H., . . . Fortune, S. M. (2018). Small RNA profiling in Mycobacterium tuberculosis identifies MrsI as necessary for an anticipatory iron sparing response. *Proc Natl Acad Sci U S A, 115*(25), 6464-6469. doi:10.1073/pnas.1718003115

Gimpel, M., & Brantl, S. (2017). Dual-function small regulatory RNAs in bacteria. *Mol Microbiol, 103*(3), 387-397. doi:10.1111/mmi.13558

Girardin, R. C., & McDonough, K. A. (2020). Small RNA Mcr11 requires the transcription factor AbmR for stable expression and regulates genes involved in the central metabolism of Mycobacterium tuberculosis. *Mol Microbiol, 113*(2), 504-520. doi:10.1111/mmi.14436

Gottesman, S. (2004). The small RNA regulators of Escherichia coli: roles and mechanisms*. *Annu Rev Microbiol, 58*, 303-328. doi:10.1146/annurev.micro.58.030603.123841

Gottesman, S., & Storz, G. (2011). Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harb Perspect Biol, 3*(12). doi:10.1101/cshperspect.a003798

Gu, H., Zhao, C., Zhang, T., Liang, H., Wang, X. M., Pan, Y., . . . Zen, K. (2017). Salmonella produce microRNA-like RNA fragment Sal-1 in the infected cells to facilitate intracellular survival. *Sci Rep, 7*(1), 2392. doi:10.1038/s41598-017-02669-1

Guirado, E., & Schlesinger, L. S. (2013). Modeling the Mycobacterium tuberculosis Granuloma - the Critical Battlefield in Host Immunity and Disease. *Front Immunol, 4*, 98. doi:10.3389/fimmu.2013.00098

Guirado, E., Schlesinger, L. S., & Kaplan, G. (2013). Macrophages in tuberculosis: friend or foe. *Semin Immunopathol, 35*(5), 563-583. doi:10.1007/s00281-013-0388-2

Guo, H., Ingolia, N. T., Weissman, J. S., & Bartel, D. P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature, 466*(7308), 835-840. doi:10.1038/nature09267

Haning, K., Cho, S. H., & Contreras, L. M. (2014). Small RNAs in mycobacteria: an unfolding story. *Front Cell Infect Microbiol, 4*, 96. doi:10.3389/fcimb.2014.00096

Henao-Mejia, J., Williams, A., Goff, L. A., Staron, M., Licona-Limon, P., Kaech, S. M., . . . Flavell, R. A. (2013). The microRNA miR-181 is a critical cellular metabolic rheostat essential for NKT cell ontogenesis and lymphocyte development and homeostasis. *Immunity, 38*(5), 984-997. doi:10.1016/j.immuni.2013.02.021

Hidalgo, M. A., Carretta, M. D., & Burgos, R. A. (2021). Long Chain Fatty Acids as Modulators of Immune Cells Function: Contribution of FFA1 and FFA4 Receptors. *Front Physiol, 12*, 668330. doi:10.3389/fphys.2021.668330

Hnilicova, J., Jirat Matejckova, J., Sikova, M., Pospisil, J., Halada, P., Panek, J., & Krasny, L. (2014). Ms1, a novel sRNA interacting with the RNA polymerase core in mycobacteria. *Nucleic Acids Res, 42*(18), 11763-11776. doi:10.1093/nar/gku793

Houben, D., Demangel, C., van Ingen, J., Perez, J., Baldeon, L., Abdallah, A. M., . . . Peters, P. J. (2012). ESX-1-mediated translocation to the cytosol controls virulence of mycobacteria. *Cell Microbiol, 14*(8), 1287-1298. doi:10.1111/j.1462-5822.2012.01799.x

Hu, W., Jain, A., Gao, Y., Dozmorov, I. M., Mandraju, R., Wakeland, E. K., & Pasare, C. (2015). Differential outcome of TRIF-mediated signaling in TLR4 and TLR3 induced DC maturation. *Proc Natl Acad Sci U S A, 112*(45), 13994-13999. doi:10.1073/pnas.1510760112

Ignatov, D. V., Salina, E. G., Fursov, M. V., Skvortsov, T. A., Azhikina, T. L., & Kaprelyants, A. S. (2015). Dormant non-culturable Mycobacterium tuberculosis retains stable low-abundant mRNA. *BMC Genomics, 16*, 954. doi:10.1186/s12864-015-2197-6

Jepsen, J. S., & Wengel, J. (2004). LNA-antisense rivals siRNA for gene silencing. *Curr Opin Drug Discov Devel, 7*(2), 188-194. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/15603252

Jester, B. C., Romby, P., & Lioliou, E. (2012). When ribonucleases come into play in pathogens: a survey of gram-positive bacteria. *Int J Microbiol, 2012*, 592196. doi:10.1155/2012/592196

Jousselin, A., Metzinger, L., & Felden, B. (2009). On the facultative requirement of the bacterial RNA chaperone, Hfq. *Trends Microbiol, 17*(9), 399-405. doi:10.1016/j.tim.2009.06.003

Kajitani, M., & Ishihama, A. (1991). Identification and sequence determination of the host factor gene for bacteriophage Q beta. *Nucleic Acids Res, 19*(5), 1063-1066. doi:10.1093/nar/19.5.1063

Kajitani, M., Kato, A., Wada, A., Inokuchi, Y., & Ishihama, A. (1994). Regulation of the Escherichia coli hfq gene encoding the host factor for phage Q beta. *J Bacteriol, 176*(2), 531-534. doi:10.1128/jb.176.2.531-534.1994

Kalayinia, S., Arjmand, F., Maleki, M., Malakootian, M., & Singh, C. P. (2021). MicroRNAs: roles in cardiovascular development and disease. *Cardiovasc Pathol, 50*, 107296. doi:10.1016/j.carpath.2020.107296

Kang, S. M., Choi, J. W., Lee, Y., Hong, S. H., & Lee, H. J. (2013). Identification of microRNA-size, small RNAs in Escherichia coli. *Curr Microbiol, 67*(5), 609-613. doi:10.1007/s00284-013-0411-9

Kehl, T., Backes, C., Kern, F., Fehlmann, T., Ludwig, N., Meese, E., . . . Keller, A. (2017). About miRNAs, miRNA seeds, target genes and target pathways. *Oncotarget, 8*(63), 107167-107175. doi:10.18632/oncotarget.22363

Kumar, S. A. (1981). The structure and mechanism of action of bacterial DNA-dependent RNA polymerase. *Prog Biophys Mol Biol, 38*(3), 165-210. doi:10.1016/0079-6107(81)90013-4

Kusunose, E., Ichihara, K., Noda, Y., & Kusunose, M. (1976). Superoxide dismutase from Mycobacterium tuberculosis. *J Biochem, 80*(6), 1343-1352. doi:10.1093/oxfordjournals.jbchem.a131407

Lamichhane, G., Arnvig, K. B., & McDonough, K. A. (2013). Definition and annotation of (myco)bacterial non-coding RNA. *Tuberculosis (Edinb), 93*(1), 26-29. doi:10.1016/j.tube.2012.11.010

Lee, Y. S., & Dutta, A. (2009). MicroRNAs in cancer. *Annu Rev Pathol, 4*, 199-227. doi:10.1146/annurev.pathol.4.110807.092222

Li, M., Xia, Y., Gu, Y., Zhang, K., Lang, Q., Chen, L., . . . Li, X. (2010). MicroRNAome of porcine pre- and postnatal development. *PLoS One, 5*(7), e11541. doi:10.1371/journal.pone.0011541

Liao, D., Fan, Q., & Bao, L. (2013). The role of superoxide dismutase in the survival of Mycobacterium tuberculosis in macrophages. *Jpn J Infect Dis, 66*(6), 480-488. doi:10.7883/yoken.66.480

Mahairas, G. G., Sabo, P. J., Hickey, M. J., Singh, D. C., & Stover, C. K. (1996). Molecular analysis of genetic differences between Mycobacterium bovis BCG and virulent M. bovis. *J Bacteriol, 178*(5), 1274-1282. doi:10.1128/jb.178.5.1274-1282.1996

Mai, J., Rao, C., Watt, J., Sun, X., Lin, C., Zhang, L., & Liu, J. (2019). Mycobacterium tuberculosis 6C sRNA binds multiple mRNA targets via C-rich loops independent of RNA chaperones. *Nucleic Acids Res, 47*(8), 4292-4307. doi:10.1093/nar/gkz149

Mendell, J. T., & Olson, E. N. (2012). MicroRNAs in stress signaling and human disease. *Cell, 148*(6), 1172-1187. doi:10.1016/j.cell.2012.02.005

Meyer, C., Grey, F., Kreklywich, C. N., Andoh, T. F., Tirabassi, R. S., Orloff, S. L., & Streblow, D. N. (2011). Cytomegalovirus microRNA expression is tissue specific and is associated with persistence. *J Virol, 85*(1), 378-389. doi:10.1128/JVI.01900-10

Mika, F., & Hengge, R. (2013). Small Regulatory RNAs in the Control of Motility and Biofilm Formation in E. coli and Salmonella. *Int J Mol Sci, 14*(3), 4560-4579. doi:10.3390/ijms14034560

Miotto, P., Forti, F., Ambrosi, A., Pellin, D., Veiga, D. F., Balazsi, G., . . . Cirillo, D. M. (2012). Genome-wide discovery of small RNAs in Mycobacterium tuberculosis. *PLoS One, 7*(12), e51950. doi:10.1371/journal.pone.0051950

Moores, A., Riesco, A. B., Schwenk, S., & Arnvig, K. B. (2017). Expression, maturation and turnover of DrrS, an unusually stable, DosR regulated small RNA in Mycobacterium tuberculosis. *PLoS One, 12*(3), e0174079. doi:10.1371/journal.pone.0174079

Morita, T., Maki, K., & Aiba, H. (2005). RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev, 19*(18), 2176-2186. doi:10.1101/gad.1330405

Murinova, S., & Dercova, K. (2014). Response mechanisms of bacterial degraders to environmental contaminants on the level of cell walls and cytoplasmic membrane. *Int J Microbiol, 2014*, 873081. doi:10.1155/2014/873081

Nathan, C., & Shiloh, M. U. (2000). Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. *Proc Natl Acad Sci U S A, 97*(16), 8841-8848. doi:10.1073/pnas.97.16.8841

Nazarova, E. V., Montague, C. R., Huang, L., La, T., Russell, D., & VanderVen, B. C. (2019). The genetic requirements of fatty acid import by Mycobacterium tuberculosis within macrophages. *Elife, 8*. doi:10.7554/eLife.43621

Nejman-Falenczyk, B., Bloch, S., Licznerska, K., Dydecka, A., Felczykowska, A., Topka, G., . . . Wegrzyn, G. (2015). A small, microRNA-size, ribonucleic acid regulating gene expression and development of Shiga toxin-converting bacteriophage Phi24Beta. *Sci Rep, 5*, 10080. doi:10.1038/srep10080

Nurmohamed, S., Vincent, H. A., Titman, C. M., Chandran, V., Pears, M. R., Du, D., . . . Luisi, B. F. (2011). Polynucleotide phosphorylase activity may be modulated by metabolites in Escherichia coli. *J Biol Chem, 286*(16), 14315-14323. doi:10.1074/jbc.M110.200741

O'Brien, J., Hayder, H., Zayed, Y., & Peng, C. (2018). Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Front Endocrinol (Lausanne), 9*, 402. doi:10.3389/fendo.2018.00402

O'Connor, C. M., & Kedes, D. H. (2007). Rhesus monkey rhadinovirus: a model for the study of KSHV. *Curr Top Microbiol Immunol, 312*, 43-69. doi:10.1007/978-3-540-34344-8_2

Pellin, D., Miotto, P., Ambrosi, A., Cirillo, D. M., & Di Serio, C. (2012). A genome-wide identification analysis of small regulatory RNAs in Mycobacterium tuberculosis by RNA-Seq and conservation analysis. *PLoS One, 7*(3), e32723. doi:10.1371/journal.pone.0032723

Petri, R., Malmevik, J., Fasching, L., Akerblom, M., & Jakobsson, J. (2014). miRNAs in brain development. *Exp Cell Res, 321*(1), 84-89. doi:10.1016/j.yexcr.2013.09.022

Pichon, C., & Felden, B. (2005). Small RNA genes expressed from Staphylococcus aureus genomic and pathogenicity islands with specific expression among pathogenic strains. *Proc Natl Acad Sci U S A, 102*(40), 14249-14254. doi:10.1073/pnas.0503838102

Plocinski, P., Macios, M., Houghton, J., Niemiec, E., Plocinska, R., Brzostek, A., . . . Dziembowski, A. (2019). Proteomic and transcriptomic experiments reveal an essential role of RNA degradosome complexes in shaping the transcriptome of Mycobacterium tuberculosis. *Nucleic Acids Res, 47*(11), 5892-5905. doi:10.1093/nar/gkz251

Pym, A. S., Brodin, P., Brosch, R., Huerre, M., & Cole, S. T. (2002). Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines Mycobacterium bovis BCG and Mycobacterium microti. *Mol Microbiol, 46*(3), 709-717. doi:10.1046/j.1365-2958.2002.03237.x

Queval, C. J., Brosch, R., & Simeone, R. (2017). The Macrophage: A Disputed Fortress in the Battle against Mycobacterium tuberculosis. *Front Microbiol, 8*, 2284. doi:10.3389/fmicb.2017.02284

Roca, F. J., Whitworth, L. J., Redmond, S., Jones, A. A., & Ramakrishnan, L. (2019). TNF Induces Pathogenic Programmed Macrophage Necrosis in Tuberculosis through a Mitochondrial-Lysosomal-Endoplasmic Reticulum Circuit. *Cell, 178*(6), 1344-1361 e1311. doi:10.1016/j.cell.2019.08.004

Rock, J. M., Hopkins, F. F., Chavez, A., Diallo, M., Chase, M. R., Gerrick, E. R., . . . Fortune, S. M. (2017). Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. *Nat Microbiol, 2*, 16274. doi:10.1038/nmicrobiol.2016.274

Rojas, M., Olivier, M., Gros, P., Barrera, L. F., & Garcia, L. F. (1999). TNF-alpha and IL-10 modulate the induction of apoptosis by virulent Mycobacterium tuberculosis in murine macrophages. *J Immunol, 162*(10), 6122-6131. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10229855

Russell, D. G. (2013). Trp'ing tuberculosis. *Cell, 155*(6), 1209-1210. doi:10.1016/j.cell.2013.11.015

Sambandamurthy, V. K., Derrick, S. C., Hsu, T., Chen, B., Larsen, M. H., Jalapathy, K. V., . . . Jacobs, W. R., Jr. (2006). Mycobacterium tuberculosis DeltaRD1 DeltapanCD: a safe and limited replicating mutant strain that protects immunocompetent and immunocompromised mice against experimental tuberculosis. *Vaccine, 24*(37-39), 6309-6320. doi:10.1016/j.vaccine.2006.05.097

Sampson, S. L., Dascher, C. C., Sambandamurthy, V. K., Russell, R. G., Jacobs, W. R., Jr., Bloom, B. R., & Hondalus, M. K. (2004). Protection elicited by a double leucine and pantothenate auxotroph of Mycobacterium tuberculosis in guinea pigs. *Infect Immun, 72*(5), 3031-3037. doi:10.1128/IAI.72.5.3031-3037.2004

Sassetti, C. M., & Rubin, E. J. (2003). Genetic requirements for mycobacterial survival during infection. *Proc Natl Acad Sci U S A, 100*(22), 12989-12994. doi:10.1073/pnas.2134250100

Schaefer, W. B., & Lewis, C. W., Jr. (1965). Effect of oleic acid on growth and cell structure of mycobacteria. *J Bacteriol, 90*(5), 1438-1447. doi:10.1128/jb.90.5.1438-1447.1965

Shah, N. S., Auld, S. C., Brust, J. C., Mathema, B., Ismail, N., Moodley, P., . . . Gandhi, N. R. (2017). Transmission of Extensively Drug-Resistant Tuberculosis in South Africa. *N Engl J Med, 376*(3), 243-253. doi:10.1056/NEJMoa1604544

Sharma, P., Kumar, B., Singhal, N., Katoch, V. M., Venkatesan, K., Chauhan, D. S., & Bisht, D. (2010). Streptomycin induced protein expression analysis in Mycobacterium tuberculosis by two-dimensional gel electrophoresis & mass spectrometry. *Indian J Med Res, 132*, 400-408. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/20966518

Shmaryahu, A., Carrasco, M., & Valenzuela, P. D. (2014). Prediction of bacterial microRNAs and possible targets in human cell transcriptome. *J Microbiol, 52*(6), 482-489. doi:10.1007/s12275-014-3658-3

Sikova, M., Janouskova, M., Ramaniuk, O., Palenikova, P., Pospisil, J., Bartl, P., . . . Krasny, L. (2019). Ms1 RNA increases the amount of RNA polymerase in Mycobacterium smegmatis. *Mol Microbiol, 111*(2), 354-372. doi:10.1111/mmi.14159

Simonsen, L., & Viboud, C. (2021). A comprehensive look at the COVID-19 pandemic death toll. *Elife, 10*. doi:10.7554/eLife.71974

Smith, J., Manoranjan, J., Pan, M., Bohsali, A., Xu, J., Liu, J., . . . Gao, L. Y. (2008). Evidence for pore formation in host cell membranes by ESX-1-secreted ESAT-6 and its role in Mycobacterium marinum escape from the vacuole. *Infect Immun, 76*(12), 5478-5487. doi:10.1128/IAI.00614-08

Srivastava, S., Chaudhary, S., Thukral, L., Shi, C., Gupta, R. D., Gupta, R., . . . Gokhale, R. S. (2015). Unsaturated Lipid Assimilation by Mycobacteria Requires Auxiliary cis-trans Enoyl CoA Isomerase. *Chem Biol, 22*(12), 1577-1587. doi:10.1016/j.chembiol.2015.10.009

Stein, C. A., Hansen, J. B., Lai, J., Wu, S., Voskresenskiy, A., Hog, A., . . . Koch, T. (2010). Efficient gene silencing by delivery of locked nucleic acid antisense oligonucleotides, unassisted by transfection reagents. *Nucleic Acids Res, 38*(1), e3. doi:10.1093/nar/gkp841

Stone, C. M., Butt, L. E., Bufton, J. C., Lourenco, D. C., Gowers, D. M., Pickford, A. R., . . . Callaghan, A. J. (2017). Inhibition of homologous phosphorolytic ribonucleases by citrate may represent an evolutionarily conserved communicative link between RNA degradation and central metabolism. *Nucleic Acids Res, 45*(8), 4655-4666. doi:10.1093/nar/gkx114

Supply, P., Marceau, M., Mangenot, S., Roche, D., Rouanet, C., Khanna, V., . . . Brosch, R. (2013). Genomic analysis of smooth tubercle bacilli provides insights into ancestry and pathoadaptation of Mycobacterium tuberculosis. *Nat Genet, 45*(2), 172-179. doi:10.1038/ng.2517

Suzuki, H. I., & Miyazono, K. (2011). Emerging complexity of microRNA generation cascades. *J Biochem, 149*(1), 15-25. doi:10.1093/jb/mvq113

Tornheim, J. A., & Dooley, K. E. (2019). The Global Landscape of Tuberculosis Therapeutics. *Annu Rev Med, 70*, 105-120. doi:10.1146/annurev-med-040717-051150

Touati, D. (2000). Iron and oxidative stress in bacteria. *Arch Biochem Biophys, 373*(1), 1-6. doi:10.1006/abbi.1999.1518

Tsui, H. C., Leung, H. C., & Winkler, M. E. (1994). Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in Escherichia coli K-12. *Mol Microbiol, 13*(1), 35-49. doi:10.1111/j.1365-2958.1994.tb00400.x

Van Assche, E., Van Puyvelde, S., Vanderleyden, J., & Steenackers, H. P. (2015). RNA-binding proteins involved in post-transcriptional regulation in bacteria. *Front Microbiol, 6*, 141. doi:10.3389/fmicb.2015.00141

Vogel, J., & Luisi, B. F. (2011). Hfq and its constellation of RNA. *Nat Rev Microbiol, 9*(8), 578-589. doi:10.1038/nrmicro2615

Volpe, E., Cappelli, G., Grassi, M., Martino, A., Serafino, A., Colizzi, V., . . . Mariani, F. (2006). Gene expression profiling of human macrophages at late time of infection with Mycobacterium tuberculosis. *Immunology, 118*(4), 449-460. doi:10.1111/j.1365-2567.2006.02378.x

Voskuil, M. I., Bartek, I. L., Visconti, K., & Schoolnik, G. K. (2011). The response of mycobacterium tuberculosis to reactive oxygen and nitrogen species. *Front Microbiol, 2*, 105. doi:10.3389/fmicb.2011.00105

Wang, B. (2013). Base Composition Characteristics of Mammalian miRNAs. *J Nucleic Acids, 2013*, 951570. doi:10.1155/2013/951570

Wang, C. H., & Kuo, H. P. (2001). Nitric oxide modulates interleukin-1beta and tumour necrosis factor-alpha synthesis, and disease regression by alveolar macrophages in pulmonary tuberculosis. *Respirology, 6*(1), 79-84. doi:10.1046/j.1440-1843.2001.00302.x

Wei, Z., Liu, X., Feng, T., & Chang, Y. (2011). Novel and conserved micrornas in Dalian purple urchin (Strongylocentrotus nudus) identified by next generation sequencing. *Int J Biol Sci, 7*(2), 180-192. doi:10.7150/ijbs.7.180

Weinberg, Z., Barrick, J. E., Yao, Z., Roth, A., Kim, J. N., Gore, J., . . . Breaker, R. R. (2007). Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. *Nucleic Acids Res, 35*(14), 4809-4819. doi:10.1093/nar/gkm487

Weischenfeldt, J., & Porse, B. (2008). Bone Marrow-Derived Macrophages (BMM): Isolation and Applications. *CSH Protoc, 2008*, pdb prot5080. doi:10.1101/pdb.prot5080

Yang, R., Xi, C., Sita, D. R., Sakai, S., Tsuchiya, K., Hara, H., . . . Kawamura, I. (2014). The RD1 locus in the Mycobacterium tuberculosis genome contributes to the maturation and secretion of IL-1alpha from infected macrophages through the elevation of cytoplasmic calcium levels and calpain activation. *Pathog Dis, 70*(1), 51-60. doi:10.1111/2049-632X.12075

Zapf, R. L., Wiemels, R. E., Keogh, R. A., Holzschu, D. L., Howell, K. M., Trzeciak, E., . . . Carroll, R. K. (2019). The Small RNA Teg41 Regulates Expression of the Alpha Phenol-Soluble Modulins and Is Required for Virulence in Staphylococcus aureus. *mBio, 10*(1). doi:10.1128/mBio.02484-18

Zhang, K., & Nicholson, A. W. (1997). Regulation of ribonuclease III processing by double-helical sequence antideterminants. *Proc Natl Acad Sci U S A, 94*(25), 13437-13441. doi:10.1073/pnas.94.25.13437

Zhang, Y. J., Reddy, M. C., Ioerger, T. R., Rothchild, A. C., Dartois, V., Schuster, B. M., . . . Rubin, E. J. (2013). Tryptophan biosynthesis protects mycobacteria from CD4 T-cell-mediated killing. *Cell, 155*(6), 1296-1308. doi:10.1016/j.cell.2013.10.045

Zhao, C., Zhou, Z., Zhang, T., Liu, F., Zhang, C. Y., Zen, K., & Gu, H. (2017). Salmonella small RNA fragment Sal-1 facilitates bacterial survival in infected cells via suppressing iNOS induction in a microRNA manner. *Sci Rep, 7*(1), 16979. doi:10.1038/s41598-017-17205-4