ABSTRACT OF THE DISSERTATION

TOXIN-ANTITOXIN SYSTEMS IN MYCOBACTERIAL INFECTIONS

By Unnati Chauhan

Dissertation Director:
Nancy A. Woychik

With an increase in antibiotic therapy failure, there has been a surge of research to understand antibiotic persistence. The distinct subpopulation of cells that can survive antibiotic treatment are called “persisters”. The persister cells also account for group of cells that cause “persistent infection” by evading human immune system, for example Mycobacterium tuberculosis (Mtb) causing latent tuberculosis infection. The bacteria can use same mechanisms for both antibiotic persistence and immune evasion. Several studies have implicated toxin-antitoxin (TA) systems as an important player in cell persistence. TA systems are composed of two adjacent genes: one encoding a toxin that causes growth arrest, and second encoding a cognate antitoxin that inhibits the toxin activity. Upon stress, the toxin is activated through degradation of its corresponding antitoxin. Unlike exotoxins, toxins of TA systems are not secreted. Instead, they act upon the cell itself by inducing growth inhibition that can lead to a dormant state. The Mtb genome consists of an unusually high number of TA systems (>80) compared to other bacteria, with the Virulence-Associated Protein (Vap) BC family accounting for 50 of those TA systems. These VapC toxins are endoribonucleases that target unique single-stranded RNA sequences through recognition of a combination of sequence and structure
determinants. To date, the RNA targets of all VapC toxins within Mtb cell have not been identified and the mechanism these toxins use to increase persistence is not understood. This study aims to understand the role of three Mtb VapC toxins: VapC2, VapC4, and VapC21. To do so, a specialized 5’ RNA-seq method was used to accurately detect toxin-cleaved RNAs. This method utilizes distinct 5’ end left by the toxin upon cleavage. This method also allows to uncover any ribosome stalling, if present. We identified that VapC2 and VapC21 solely targets initiator tRNA cleaving at a single site within the anticodon stem-loop. Depletion of initiator tRNA causes inhibition in translation and eventually growth arrest. This decrease in growth might be necessary for Mtb to transition from active growth to latency during infection. We also identified that VapC4 targets tRNA^{Cys} at a single site within anticodon sequence, thus inactivating it. Depletion of the pool of tRNA^{Cys} causes ribosome stalling at Cys codons in actively translated transcripts. This ribosome stalling also allowed us to uncover several unannotated Cys-containing ORFs. VapC4 mimics a state of Cys starvation, which then activated Cys-attenuation at sORFs to globally redirect metabolism towards the synthesis of free Cys. The resulting newly enriched pool of Cys feeds into the synthesis of mycothiol, the glutathione counterpart in this pathogen that is responsible for maintaining cellular redox homeostasis during oxidative stress, as well as into a circumscribed subset of cellular pathways that enable cells to defend against oxidative and copper stresses characteristically endured by Mtb within macrophages. This study also designed to understand the role of VapC toxin in *Mycobacterium abscessus* (Mab) infections. Mab infections are on the rise, and it is the most notoriously difficult to treat due to its extreme resistance to antibiotics and disinfectants. Mab may enlist TA system to increase
antibiotic persistence. A total of 22 novel putative TA systems were identified in 128 Mab clinical strains available on NCBI. VapC5 toxin was one of the more abundant toxins present in these clinical strains. Unlike Mtb VapC toxins, VapC5 toxin cleaves multiple tRNAs at a single site within their anticodon resulting in reduction in translation. VapC5 also induces the expression of multiple genes that underlie persistence/intrinsic antibiotic resistance in Mab. Finally, we also demonstrated that VapC5 expression increased persister cell formation after treatment with two Mab antibiotics. Understanding how TA systems play a role in persister cell formation can help us design better antibiotics or use currently available antibiotics more effectively.
ACKNOWLEDGEMENTS AND DEDICATION

First and foremost, I would like to thank my mentor Dr. Nancy A. Woychik, without whom I would not be writing this. I haven’t met any mentor as caring as Nancy has been in both my professional and personal life. Nancy has not only helped me grow to become a better scientist, but also a better person. Nancy’s moto for us has been “Work hard and play hard” and we live by it. With lab beach trips, dance parties to “New Rules”, easter egg hunts, and watching crazy videos during lab meetings (yes, I still remember the time when we watched a video of an Indonesian woman found inside a Python and it still haunts my dreams), Nancy has made my PhD journey unforgettable.

To my current and former lab mates: Valdir, Heather, Julia P., Julia G., Eduardo, Mike D., Mike B., Parth, Mel, and Megan, thank you so much for making lab both fun and productive. A massive thank you to Valdir for training me on almost every technique and allowing me to annoy him with any (and many) questions I had about my projects. I would also like to give a massive thanks to Heather for listening to all of my rants, both science and life related.

To my committee members: Dr. Paul Copeland, Dr. Gary Brewer, and Dr. Sam Gu, thank you for all the guidance and asking me tough questions. Because of that, I became the scientist that I am today.

To the administrative staff: Diane Murano, Carolyn Ambrose, Caryn Caroll, and Letice Smith, Tina Marottoli, and Jamie Carr for handling all the “behind the scenes” paperwork.

To my lovely roommates: Sierra Swords, Meghan Arnolds, and Juliet Gardiner (honorary roommate), and my dearest friends: Rajasi, Aesha, Carolina, Caleb, Ilija, and
Russ thank you so much for always being there when I needed to talk and/or rant. You have always been there cheering me on and supporting me.

To my amazing family papa, mummy, and Jay, for your unconditional love and support, thank you. You guys have always been there for me no matter what, encouraging me, and providing me with an escape from the life of science when needed. Mummy and Papa, I cannot thank you enough for all the sacrifices you have made, including leaving your family in India and arriving to a completely new country, so Jay and I can have better education and better life. Without you, I would not be the strong person that I am today.

Last but not the least, to my love, Josh, thank you for my being rock and helping me through the darkest time. Thank you for always pushing me and believing in me, even when I was at my lowest. You have always been there for me even if it means driving me to the lab at midnight in the middle of the storm just so I would get there safe. Thank you for not getting mad at me while waiting in the car, because I said I would be done in 10 minutes but was still in the lab an hour later. Surely without you, I would have become the mad scientist.

Oh no, how can I forget?! Also, a super special thanks to my two cutest, bestest, kitties in the whole wide world: Momo & Taro.

This work was funded by a New Jersey Commission on Cancer Research Pre-Doctoral Grant and NIH R01 AI154464. Chapter 2 is in Press at *Proceedings of the National Academy of Sciences of the USA* (Barth et al., 2021).
# TABLE OF CONTENTS

**ABSTRACT OF THE DISSERTATION** .................................................................................. ii

**ACKNOWLEDGEMENTS AND DEDICATION** ................................................................. v

**LIST OF ILLUSTRATIONS** ........................................................................................... xi

**Introduction** ..................................................................................................................... 1

- Epidemiology of tuberculosis ........................................................................................ 1
- Pathology of tuberculosis .............................................................................................. 3
- Evasion of the Immune System ..................................................................................... 6
- Role of Toxin-Antitoxin systems in tuberculosis infection ........................................... 9
- VapBC TA systems in *Mycobacterium tuberculosis* ..................................................... 13
- Rationale ......................................................................................................................... 14

**Chapter 1: Mycobacterium tuberculosis VapC2 and VapC21 toxins cleave initiator tRNA inhibiting translation** ..................................................................................... 15

1.1 **Abstract** .................................................................................................................... 15

1.2 **Introduction** ............................................................................................................... 15

1.3 **Materials and Methods** ............................................................................................ 17

- Strains, Plasmids, and Reagents .................................................................................. 17
- Purification of VapC2 and VapC21 toxins .................................................................... 18
- In vitro synthesis of tRNA. .......................................................................................... 19
- In vitro tRNA cleavage assay ....................................................................................... 19
- RNA isolation ............................................................................................................... 20
- 5’ monophosphate RNA-seq ....................................................................................... 20
- Labeling of newly synthesized protein ........................................................................ 21
Expression of phylogenetically related VapC toxins, VapC2 and VapC21, leads to growth inhibition 22
VapC2 and VapC21 cleave multiple tRNAs in vitro ................................................................. 25
VapC2 and VapC21 cleave and inactivate only tRNA^fMet in vivo ............................................. 28
VapC2 and VapC21 inhibit protein synthesis .......................................................................... 30
The C-terminal WR residues of VapB2 and VapB20 antitoxins are necessary for toxin inhibition..... 31
VapC2 and VapC21 toxins are structurally similar, yet the VapB2 and VapB21 antitoxins are not
interchangeable ...................................................................................................................... 33

1.5 Discussion .......................................................................................................................... 34

Chapter 2: The VapC4 toxin helps Mycobacterium tuberculosis evade immune clearing. 37

2.1 Abstract ............................................................................................................................... 37

2.2 Introduction .......................................................................................................................... 38

2.3 Materials and Methods ....................................................................................................... 40

2.4 Results ................................................................................................................................. 47
VapC4 cleaves many Mtb tRNAs within their anticodon stem loop in vitro .................................................. 47
VapC4 does not completely inhibit protein synthesis .................................................................................. 49
VapC4 only targets tRNA\textsubscript{Cys} in Mtb cells .................................................................................. 51
tRNA\textsubscript{Cys-GCA} cleavage leads to ribosome stalling at Cys codons ..................................................... 55
An Mtb sORF is the ortholog of a Cys-responsive attenuator in \textit{M. smegmatis} ............................................. 59
VapC4-triggered ribosome stalling at leaderless sORFs identifies additional Cys attenuation-responsive regulons .................................................................................................................. 62
VapC4 regulates growth through targeted manipulation of discrete pathways .......................................... 63
VapC4 upregulates sulfur metabolism and oxidative stress pathways ..................................................... 65
VapC4 upregulates the CsoR and RicR copper resistance pathways .......................................................... 70
Many copper response genes are metabolic enzymes involved in sulfate assimilation and oxidative stress .................................................................................................................................................. 73

2.5 Discussion ................................................................................................................................. 75

\textbf{Chapter 3:} \textit{Toxin-mediated enhancement of intrinsic antibiotic resistance in} \textit{Mycobacterium abscessus} ............................................................................................................. 81

3.1 Abstract ........................................................................................................................................ 81

3.2 Introduction ................................................................................................................................... 82

3.3 Materials and Methods .................................................................................................................. 85

Toxin-antitoxin Identification .................................................................................................................. 85
Strains, Plasmids and Reagents .................................................................................................................. 86
RNA isolation ........................................................................................................................................... 86
Northern analysis of tRNA levels ............................................................................................................ 87
Labeling of newly synthesized Mab proteins ........................................................................................... 88
5' RNA-seq ............................................................................................................................................. 89
RNA-Seq ............................................................................................................................................ 90
Proteomics ........................................................................................................................................ 91
Antibiotic MICs ................................................................................................................................. 91
Persister Cell Assay ......................................................................................................................... 92

3.4 Results ........................................................................................................................................... 92
Many Mab clinical strains harbor TA systems. ..................................................................................... 92
VapC5 expression leads Mab cell growth arrest ................................................................................... 93
VapC5 targets six tRNAs, including tRNAmet, for cleavage within their anticodons ......................... 95
VapC5 tRNA targets share sequence and structural features. ............................................................ 98
VapC5 induces WhiB7 and multiple genes that underlie intrinsic antibiotic resistance in Mab. ..... 100
VapC5 expression reprograms the proteome to exclusively stockpile ribosomes and translation factors. .............................................................................................................................................. 104
VapC5 expression enhances persister formation. .............................................................................. 109

3.5 Discussion ..................................................................................................................................... 110

Conclusions ....................................................................................................................................... 116

References ......................................................................................................................................... 118
LIST OF ILLUSTRATIONS

Figure 1: Tuberculosis global incidence. ........................................................................................................ 2

Figure 2: Latent (a) and active (b) stages of tuberculosis infection .............................................................. 4

Figure 3: Schematic representation of Mycobacterium tuberculosis cell wall and how it modulates
macrophage recruitment. .......................................................................................................................... 7

Figure 4: The Four Main Types of TA Modules .......................................................................................... 10

Figure 5: Chromosomal map of mtb H37Rv TA systems .......................................................................... 12

Figure 6: Phylogenetic tree of the 50 Mtb VapC toxins. ............................................................................. 24

Figure 7: Expression of VapC2 and VapC21 leads to growth inhibition. ...................................................... 25

Figure 8: VapC2 and VapC21 cleave multiple tRNAs in vitro to some extent but only tRNA^{fMet} to
completion. .............................................................................................................................................. 27

Figure 9: VapC2 and VapC21 cleave and inactivate only tRNA^{fMet} in vivo. ........................................... 29

Figure 10: VapC2 and VapC21 inhibit protein synthesis. ........................................................................... 30

Figure 11: The C-terminal WR residues of VapB2 and VapB20 antitoxins are necessary for inhibition of
toxin. ......................................................................................................................................................... 32

Figure 12: VapC2 and VapC21 toxins are structurally similar, yet the VapB2 and VapB21 antitoxins are not
interchangeable. ......................................................................................................................................... 33

Figure 13: VapC4 cleaves multiple tRNAs in vitro. ..................................................................................... 48

Figure 14: VapC4 inhibits growth but does not completely inhibit translation ........................................... 50

Figure 15: tRNA^{Cys-GCA} is the sole VapC4 target in vivo ......................................................................... 53

Figure 16: VapC4 cleavage generates 5’ monophosphate ends and lacks specificity in in vitro assays.... 54

Figure 17: 5’ OH RNA-seq reveals ribosome stalling at Cys codons with single nucleotide resolution.... 56

Figure 18: Mtb expressing VapC4 preferentially downregulates translation of Cys-rich proteins............. 58

Figure 19: Downstream genes of identified Cys-containing sORFs are generally upregulated............... 61

Figure 20: DAVID analysis of transcripts down- and upregulated by VapC4 .............................................. 65

Figure 21: DAVID analysis of transcripts down- and upregulated by VapC4 ............................................. 69
Figure 22: VapC4 mimics Cys starvation to activate the oxidative and Cu stress responses .......................... 72

Figure 23: VapC4 upregulates key components of Cu homeostasis ....................................................... 75

Figure 24: VapC4 reprograms Mtb physiology to protect against oxidative and Cu stress......................... 79

Figure 25: Many M. abscessus clinical strains contain an M. tuberculosis VapC5 ortholog whose expression leads to cell growth arrest......................................................................................................... 94

Figure 26: VapC5 targets multiple tRNAs within their anticodon in vivo .................................................. 97

Figure 27: VapC5 cleavage is sequence and structure specific ................................................................. 99

Figure 28: VapC5 induces the expression of WhiB7, ribosomal proteins and multiple genes that underlie intrinsic antibiotic resistance in Mab ................................................................................................. 103

Figure 29: VapC5 predominantly upregulates the synthesis of ribosomal proteins ................................. 108

Figure 30: VapC5 potentiates intrinsic resistance in Mab ........................................................................ 110

Figure 31: Mab VapC5 toxin subverts the efficacy of first line Mab antibiotics by activating multiple resistance pathways .................................................................................................................................. 114
INTRODUCTION

Epidemiology of tuberculosis

Tuberculosis (TB) has been dated back to 2400 B.C. in Egyptian mummies, making it one of the oldest infectious diseases affecting mankind. TB is a leading cause of death from a single infectious agent, ranking above HIV/AIDS since 2007, and only recently was surpassed by COVID-19 (WHO, 2020). Many consider TB to be the disease of the past, especially in developed countries such as the United States where the disease is almost eradicated (Figure 1A). However, TB continues to be a strain on the health resources for many developing countries (Figure 1B). Two thirds of the global TB cases originate in eight developing nations: India (26%), Indonesia (8.5%), China (8.4%), the Philippines (6.0%), Pakistan (5.7%), Nigeria (4.4%), Bangladesh (3.6%), and South Africa (3.6%) (WHO, 2020). HIV-positive individuals are also commonly coinfect with TB, accounting for 30% of HIV related deaths (WHO, 2021). An estimated 10 million (range, 8.9—11.0 million) were infected with TB globally in 2019, killing 1.4 million (range, 1.3—1.5 million) people (WHO, 2020). This number has been declining consistently in recent years (Figure 1C, D). However, the COVID-19 pandemic has threatened this progress, impacting TB burden and vulnerability through a variety of ways. The COVID-19 pandemic has caused disruption in TB health services which resulted in the decline in TB diagnosis and treatment (McQuaid et al., 2021). There are many reasons for this decline: equipment shortages, lockdowns making attaining health care difficult (which also increased transmission rates within households) and reallocation of TB resources. Dr. Lucica Ditiu, Executive Director of the Stop TB
Partnership, stated: “Twelve years of impressive gains in the fight against TB, including in reducing the number of people who were missing from TB care, have been tragically reversed by another virulent respiratory infection.” Therefore, there is a need to call for strategies to fight COVID-19 and TB together.

**Figure 1: Tuberculosis global incidence.** (A) Countries (in blue) that had an estimated TB incidence rate of less than 10 per 100,000 population in 2019. (B) Countries that had at least 100,000 incident cases of TB in 2019. (C) Global trend in the estimated TB incidence rate (blue), 2000–2019. (D) Global trend in the estimated number of TB deaths, 2000–2019. The blue shaded area is the uncertainty interval. Source: World Health Organization, 2020.
**Pathology of tuberculosis**

TB is an airborne infectious disease caused by *Mycobacterium tuberculosis* (Mtb), which typically infects lungs; however, it can also infect other parts of the body as well. When an infected person who has pulmonary or laryngeal TB disease coughs, sneezes, shouts, or sings, it causes droplets containing bacteria to be released into the air and then transmission can occur when an uninfected person inhales that infectious droplet. TB is a lower respiratory tract disease and small droplets containing few cells (1-10 Mtb cells) are the ones that initiate lung infection (Wells et al., 1948). Small droplets allow Mtb to evade the mucosal immune response that is present in the nasopharynx and establish an infection of the lower respiratory system. Exposure to Mtb can result in two broad outcomes: 1) the body can eradicate the infection by a combination of innate and adaptive immunity or 2) if the immune system is not able to eliminate the pathogen, tuberculosis disease will develop.

Tuberculosis infection can present itself in two forms: active and latent (Figure 2). When Mtb enter human lungs, it will encounter the resident alveolar macrophages, which can internalize Mtb. Macrophages serve as a major host cell type for Mtb infection. Mtb-infected phagocytes can then migrate into lung interstitium and invade through infecting either alveolar epithelium or alveolar macrophages migrating into lung parenchyma. During this stage, dendritic cells and monocytes can carry Mtb to the nearby lymph node to prime T-cells. This will trigger cell-mediated immunity and cause other immune cells (including T-cells and B-cells) to migrate to the site of infection and form an organized structure called a granuloma (Figure 2a). Formation of granulomas will contain Mtb; however, not all Mtb cells are killed. These surviving cells within the granuloma represent latent tuberculosis infection. If the immune system is compromised or the bacterial load
within the granuloma becomes too great, the granuloma will fail to contain the infection and the bacteria can break through and infect the rest of the lungs or even other organs causing the active disease (Figure 2b).

Figure 2: Latent (a) and active (b) stages of tuberculosis infection. Source: (Pai et al., 2016)

In an active tuberculosis infection, the mycobacterium is proliferating, leading to a variety of symptoms such as fever, lack of appetite, weight loss, fatigue, and hemoptysis. During this stage, the disease is transmissible and, if not treated, can lead to death. Individuals with an active form of the disease will usually test positive for tuberculin skin test (TST) or interferon-γ release assay (IGRA). These individuals will also benefit from recommended treatment regimen for active TB disease. The first-line anti-TB drugs used
in the treatment regimen are isoniazid, rifampin, ethambutol, and pyrazinamide. The regimens for TB disease typically consist of an intensive phase, which consist of four drugs and last for 2 months, followed by a continuation phase that consist of two drugs and last from 4-7 months. Full treatment for TB can take six to nine months.

In contrast to active disease, latent tuberculosis infection is asymptomatic and non-transmissible. About 90-95% of all TB infection will result in latent infection and it is estimated that one-quarter of the world’s population is infected with the latent form of tuberculosis (Houben & Dodd, 2016). In the latent form of TB, although the bacteria are in a dormant state, it can be reactivated if the immune system of the infected individual is immunocompromised. Mtb can persist in human lungs for decades in the dormant form while tolerating the host immune response and antibiotics; although recent studies suggest that TB incubation period is much shorter than previously thought (Behr et al., 2018). Overall, if not treated, ~5-10% of people with latent TB will develop the active disease at some point in their lives. Treatment for TB is extensive, long, and with side effects, which can discourage people with latent TB from seeking treatment. In recent years, the Center for Disease Control and the National Tuberculosis Controllers Association have started recommending shorter rifamycin-based treatment (3-4 months) for latent TB over a 6- or 9-month isoniazid therapy. The persister cells play a major role in latent tuberculosis infection as well as in the emergence of drug-resistant strains; however, the molecular mechanism underlying persistence is not fully understood. Defining factors that contribute to this complex growth control mechanism and persistence in mycobacteria will provide a better understanding of the latent infection as well as help create more efficient drugs to treat latent disease.


**Evasion of the Immune System**

Mtb is a highly evolved pathogen with its innate resistance to many antimicrobial drugs. This antimicrobial resistance can be credited to Mtb’s unique but complex cell wall that has low permeability and several efflux pumps. The cell wall mainly consists of peptidoglycan (PG), mycolic acid (MA), and arabinogalactan (AG) (Figure 3). These three layers are covalently attached to each other creating MA-AG-PG complex (MAPc), also referred to as the cell wall core (Brennan, 2003). On top of the Mtb core cell wall, there are free lipid layers consisting of phthiocerol dimycocerosate (PDIM) and phenolic glycolipids (PGL). PGLs and PDIMs. Although their function is not fully understood, there is evidence suggesting that they play an important role in Mtb virulence and evasion of the host immune system.

When a common bacterial pathogen enters the human body, the first line of the innate immune response’s defense is induced through recognition of pathogen associated molecular patterns (PAMPs). PAMPs are typically present on the surface of the bacteria and will trigger toll-like receptor pathways, thus recruiting bactericidal macrophages (iNOS positive) to the site of infection. Like other pathogens, Mtb also contains PAMPs on its surfaces; however, PDIMs coats these PAMPs thus preventing recruitment of bactericidal macrophages. Mtb also contains PGs on its surface that allows recruitment of growth-permissive macrophages (iNOS negative). Therefore, with the help of PDIMs and PGLs, mtb can regulate macrophage recruitment to the site of infection and end up infecting growth-permissive macrophages, the preferred host cell.
Macrophages act as the first line of defense against pathogens by engulfing them and placing the bacilli into a vacuole. This pathogen containing vacuole is called phagosome and can undergo a maturation process. The maturation process will lead to fusion of phagosome with lysosomes: an acidic vacuole that carries a set of hydrolytic enzymes that allow breakdown of the pathogen. Mtb, a successful macrophage pathogen, has developed a variety of strategies to evade macrophage defenses. Mtb halts phagosome maturation by preventing mycobacterial vacuole from fusing with lysosomes and stabilizes Figure 3: Schematic representation of Mycobacterium tuberculosis cell wall and how it modulates macrophage recruitment. Source: (Siegrist & Bertozzi, 2014).
the phagosomal pH to be between 6.2 and 6.5 (Queval et al., 2017). This blockade of acidification is necessary for mtb to enter the macrophage cytosol (Simeone et al., 2015).

Phagosomal rupture, as well as getting access to the macrophage cytosol, is central to Mtb’s success as a pathogen. Once inside the phagosome, with the help of PDIMs and the ESX-1 type VII secretion (ESX/T7S) system, Mtb triggers phagosomal rupture and provides cytosolic access. ESX/T7S is the most critical virulence factor that determines Mtb fate within the macrophage. Upon deletion of ESX-1 locus, Mtb loses its virulence and results in an attenuated strain (Simeone et al., 2012). The attenuated strain used in Bacillus Calmette-Guérin (BCG) vaccine, the only TB vaccine currently available, has the ESX-1 locus deleted (Hsu et al., 2003). Phagosomal rupture and access to cytosol is a critical step for Mtb survival and can ultimately lead to the host cell death.

Control of the host cell-death is important for Mtb pathogenesis. The host-cell death continues to be a controversial topic as to whether the host cell’s response to induce death, either through apoptosis or necrosis, is beneficial or damaging to Mtb infection. Previously, it was thought that apoptosis (typically triggered by an attenuated strain of Mtb) is a defense the host cell uses to prevent the spread of bacilli, while necrosis (typically triggered by a virulent strain of Mtb) allows bacilli to spread across other macrophages (Behar et al., 2010). However, recent studies have suggested that apoptosis can also be triggered by virulent strains of Mtb (Aguilo et al., 2013). It is speculated that apoptosis is useful for dissemination of bacilli while necrosis is used to enhance bacterial replication (Queval et
Therefore, both apoptosis and necrosis could be beneficial to Mtb, and it can trigger one or the other based on the stage of infection.

Mtb and humans have evolved together, resulting in a very sophisticated host-pathogen interaction, a battle that still continues. Mtb has developed amazing strategies to survive under harsh conditions presented by host-cells; however, it is not fully understood how Mtb is able to achieve these tasks and it remains a highly investigative area. Latency is the key for Mtb being such a successful pathogen. But, how Mtb infection is reactivated from latency, or how exactly does the bacilli survive the first attack from immune system in the first place, remains unanswered. Toxin-antitoxin genes have been proposed to play a role in these processes and might be the missing ingredient in answering these questions.

**Role of Toxin-Antitoxin systems in tuberculosis infection**

Toxin-antitoxin (TA) systems are almost exclusively present in free living bacteria and archaea and, at a minimum, characteristically act as growth regulators. TA systems are widely studied due to their implications in cellular processes such as bacterial persistence, stress response, biofilm formation, antiphage defense, genomic stability, plasmid maintenance, and programmed cell death (Magnuson, 2007; Page & Peti, 2016; Short et al., 2018; Wang & Wood, 2011). The toxin in a TA system is typically a protein that causes growth inhibition by interfering with vital cellular process. The antitoxin in a TA system can be either a non-coding RNA or a low molecular weight protein that neutralizes toxin activity. TA systems can be classified into four main types (Figure 4). Type I antitoxins are RNA molecules that bind to the toxin RNA to inhibit its translation (Figure 4a). Type III antitoxins are also RNA molecules; however, they bind directly to the toxin and inhibit
its activity (Figure 4c). Type II TA systems consist of a bi-cistronic operon, where antitoxins are proteins that directly bind to the toxin, neutralizing its activity (Figure 4b).

![Figure 4: The Four Main Types of TA Modules](https://example.com/figure4)

Figure 4: The Four Main Types of TA Modules. Source: (Harms et al., 2018)

Lastly, type IV TA systems have antitoxins that are proteins; however, these antitoxins counteract or reverse the effects of toxin without direct interaction (Figure 4d). Two single instances have been identified where toxin-antitoxin systems cannot be categorized into any of the four types discussed above, leading to the generation of type V and type VI TA systems.

Type II TA systems are extensively studied and most abundantly present in the Mtb genome. Type II TA systems contain a stable toxin and labile antitoxin. Under normal conditions, the antitoxin can interact with the toxin, neutralizing its activity and forming a stable complex. However, under stress, the antitoxin is degraded, freeing toxin to perform its growth inhibitory activity, such as hampering DNA replication or protein translation. TA systems have been suggested to play a central role in bacterial persistence. The term bacterial persistence is used to define distinct subpopulation of cells that can survive antibiotic treatment. Unlike resistance where the bacteria undergo a genetic mutation that is heritable, persistence results from a phenotypic switch where the bacteria enter a dormant
state, slowing down its metabolic activity to survive bactericidal antibiotics. The first time TA systems were linked to persistence was in the *Escherichia coli* gene *HipA* (high persister protein A). *HipA* is a kinase that inactivates glutamyl-tRNA GltX and mutations in the *hipA* gene led to ~100- to ~1000-fold increase in persistence formation (Kaspy et al., 2013; Moyed & Bertrand, 1983). Further studies showed that HipA activity causes ppGpp-mediated stringent response that activates type II RNase toxins, whose activities has also been linked to persistence (Germain et al., 2015). Several type II RNase toxins are found to be upregulated in persister cells and direct role of these toxins in persister cells formation has also been established by ectopic expression experiments as well as with deletion experiments (Helaine et al., 2014; Keren et al., 2004; Shah et al., 2006; Vazquez-Laslop et al., 2006). Although there have been studies to establish a role of TA systems with persistence, the mechanism by which they increase persistence is still not fully understood.

TA systems are stress modules that decrease physiological activity to allow cells to enter a persistent state where the cells’ metabolic activity is shut down and are tolerant to antibiotic stress. *Mtb*’s nonreplicating persistent state is analogous to this state where the cell has low metabolic activity and is insensitive to antibiotic treatments. *Mtb* has an unusually high number of TA systems (over 80) spread throughout the genome (Figure 5), while its non-pathogenic counterpart *Mycobacterium smegmatis* only has four (Sala et al., 2014). The majority of *Mtb* TA systems belong to five well-studied type II TA families: VapBC (50 systems), MazEF (10 systems), HigBE (two systems), RelBA (two systems), and ParDE (two systems) (Sala et al., 2014). These TA systems are triggered by stress relevant to latent tuberculosis infection, such as antibiotic exposure, immune response, or hypoxia (Albrethsen et al., 2013; Keren et al., 2011; Korch et al., 2015; Ramage et al.,
Defining factors that contribute to this complex growth control mechanism and persistence in mycobacteria will provide a better understanding of the latent infection as well as help create more efficient drugs to treat latent forms of the disease.

Figure 5: Chromosomal map of mtb H37Rv TA systems. Source: (Sala et al., 2014)
VapBC TA systems in Mycobacterium tuberculosis

The Virulence-Associated Protein (Vap) TA family accounts for more than half of TA systems present in Mtb. The toxin, VapC, contains a PIN (PilT N-terminal) domain, which comprises the catalytic site for the toxin endoribonuclease activity and the VapB antitoxin neutralizes the toxin. The PIN domain is a highly conserved 130 amino acid motif that can be found across all three domains of life. It contains five conserved residues (4 acidic and 1 serine/threonine) that coordinate Mg$^{2+}$ or Mn$^{2+}$ at its catalytic site. These five residues are positioned in a very similar manner as of PIN Domain of RNase H, suggesting that VapC and RNase H share their catalytic mechanism (Arcus et al., 2011). Although previously mischaracterized as an mRNA cleaver, VapC toxins most often possess tRNA cleavage activity (Ramage et al., 2009).

To better understand the mechanism of these toxins, VapCs have been studied across multiple bacteria. The first identification of a tRNA target for a VapC toxin was for VapCs from Shigella flexnerii and Salmonella enterica serovar Typhimurium LT2 (Winther & Gerdes, 2011). It was determined that these toxins are site-specific tRNases that cleave initiator tRNA between the anticodon stem and loop. Cleaving of the initiator tRNA shuts down translation and results in growth arrest. Later studies found several of Mtb VapC toxins are also site-specific tRNase cleaving tRNA$^{Cys}$, tRNA$^{Leu}$, tRNA$^{Gln}$, tRNA$^{Ser}$ and tRNA$^{Trp}$ (Cruz et al., 2015; Winther et al., 2016). Some of these tRNAs were redundantly cleaved by multiple toxins, including: VapC25, VapC29, VapC33, VapC37, and VapC39 cleave tRNA$^{Trp}$, VapC28 and VapC30 cleave tRNA$^{Ser}$, and VapC11, VapC15, and VapC32 cleave tRNA$^{Leu}$ (Winther et al., 2016).
Interestingly, not all VapC toxins cleave tRNA; VapC20 and VapC26 in Mtb was found to cleave the sarcin-raicin loop of 23S rRNA (Winther et al., 2016; Winther et al., 2013). The majority of experiments studying Mtb toxins were performed using either in vitro assay or ectopic expression of the toxin in mycobacterium model organism, *Mycobacterium smegmatis*. However, in order to determine the accurate toxin target, it is important to study the toxin in its native host (Cintrón et al., 2019). Several of Mtb VapC toxins target have been identified; however, the majority of them remain to be characterized.

**Rationale**

Tuberculosis has been around for at least 6000 years, yet it remains to be a major cause of mortality worldwide. Mtb’s success as a pathogen is based on its ability to establish a latent infection. The involvement of TA systems in persistence and stress related to latency (nutrient starvation, oxidative stress, hypoxia, and antimicrobial) has been well established but the molecular mechanism of how TA systems trigger this response remain to be determined. Since Mtb contains a vast number of TA systems, finding targets of these toxins will help us understand how these TA systems are involved in persistence, thus providing a framework that can be used to design new therapeutic drugs to treat latent tuberculosis infection and/or prevent further Mtb progression.
Chapter 1: Mycobacterium tuberculosis VapC2 and VapC21 toxins cleave initiator tRNA inhibiting translation

1.1 Abstract

The Mycobacterium tuberculosis (Mtb) genome contains an abundance of toxin-antitoxin (TA) systems, 50 of which belong to the VapBC family. Here we used a specialized RNA-seq approach, 5’ RNA-seq, to accurately identify the in vivo RNA target of two phylogenetically related VapC toxins, VapC2 and VapC21. Both toxins exclusively cleave and inactivate initiator tRNA\(^{f\text{Met}}\) at a single, identical site within their anticodon loop. Consistent with the essential role and global requirement for initiator tRNA\(^{f\text{Met}}\) in bacteria, expression of each VapC toxin led to growth arrest and translation inhibition. Finally, we identified two conserved, contiguous amino acids (WR) at the carboxy terminus of the antitoxins that are essential for interaction with their cognate toxins. This short, discrete site of VapB-VapC interaction can be a drug target that enables constitutive inactivation of tRNA\(^{f\text{Met}}\) by both VapC toxins. More specifically, identification of small molecules that bind to and block the antitoxin WR domain can be exploited for development of highly specific antituberculars, particularly for latent tuberculosis (TB) infections.

1.2 Introduction

Among the ~90 Type II TA systems in Mtb, most belong to the 50-member VapBC family of RNases (designated VapC1 through VapC-50). All VapC toxins possess a PIN (PilT amino-terminal) domain containing a conserved quartet of acidic
residues and a fifth invariant serine or threonine residue, all of which are responsible for coordinating divalent cation(s) in the catalytic center (Arcus et al., 2011). It is not understood why the Mtb genome has the highest number of VapC TA systems among free-living bacteria; most bacteria have just one (Pandey & Gerdes, 2005). Curiously, the handful of Mtb VapC toxins studied in detail have different targets that lead to dramatically different endpoints. The Mtb VapC20 and 26 toxins cleave at 23S rRNA a single site within the highly conserved sarcin-ricin loop that inactivates the ribosome and leads to widespread translation inhibition (Winther et al., 2016; Winther et al., 2013). By contrast, the Mtb VapC4 and VapC11 are isoacceptor-specific tRNases (Barth et al., 2021; Cintrón et al., 2019). The VapC4 toxin exclusively inactivates the sole tRNA^{Cys} through cleavage at a single site within the anticodon sequence. Reduction in the levels of tRNA^{Cys} mimics a state of Cys starvation that leads to the reprogramming of Mtb metabolism to specifically defend against oxidative and copper stresses (Barth et al., 2021).

Type II TA systems are composed of adjacent genes encoding two small (~10 kDa) proteins, a toxin and its cognate antitoxin that inhibits toxin activity through formation of a stable TA protein-protein complex. In the prevailing model, stress conditions are proposed to play a major role in regulating the toxin-antitoxin balance by selectively activating antitoxin degradation (Harms et al., 2018; Masuda & Inouye, 2017). Thus, specific stress conditions lead to a preponderance of free toxin, which exerts its growth-regulating and/or other functions within the bacterial cells (Harms et al., 2018; Masuda & Inouye, 2017). To obtain a broader understanding of Mtb VapC toxin functions and how they might collaborate to survive the spectrum of stresses associated
with the Mtb infection cycle, we are systematically identifying the RNA targets of each toxin and studying how they impact Mtb growth and physiology. Here we identified tRNA<sup>fMet</sup> as a new target for the Mtb VapC family. We demonstrate that expression of either of the two toxins with this activity, VapC2 and VapC21, leads to strong inhibition of protein synthesis and cell growth. Finally, we highlight how their properties could be harnessed for development of shorter, more effective treatment for latent tuberculosis infections.

1.3 **Materials and Methods**

**Strains, Plasmids, and Reagents**

VapC2 (Rv0301) and VapC21 (Rv2757c) were amplified by PCR from Mtb H37Rv genomic DNA. The amplified gene was cloned adjacent to anhydrotetracycline (ATC)-inducible promoter in the pMC1s plasmid or pristinamycin (PTC)-inducible promoter in the pNW2046 (pMYT769-derived) plasmid. Induction of VapC toxins was obtained by adding ATC to the media to a final concentration of 200 ng/ml, and replenishing it every 48 hours or adding PTC to the media to a final concentration of 1 μg/ml.

All Mtb experiments were performed using Mtb mc<sup>2</sup> 6206 (ApanCD AleuCD, generously provided by William Jacobs laboratory, Albert Einstein College of Medicine). Mtb cells were grown under constant shaking at 200 rpm at 37° C in 7H9 Middlebrook media containing 1x OADC supplement (Sigma), 0.05% of tyloxapol, and appropriate antibiotic (kanamycin at 25 μg/ml or spectinomycin at 100μg/ml) for plasmid selection.
The media was also supplemented with 50 µg/ml of pantothenic acid and 100 µg/ml of leucine for growth of the attenuated strain Mtb mc² 6206.

For mutation studies, VapB2 (Rv0300) and VapB21 (Rv2758c) were PCR amplified from Mtb H37Rv genomic DNA and cloned into an isovaleronitrile (IVN)-inducible pNIT plasmid. Mutations were performed using PCR mutagenesis and sequence confirmed. *M. smegmatis* mc²155 used for mutation studies were grown in 7H9 Middlebrook media containing 1x ADN supplement (0.5% bovine albumin, 0.2% Dextrose, 0.085% NaCl), 0.05% Tween80, and appropriate antibiotic (kanamycin at 25 µg/ml). M. smegmatis cells were grown under constant shaking at 200rpm at 37°C. For co-expression, cells were induced at 0.1 OD₆₀₀ by adding ATC and IVN to a final concentration of 50 ng/mL and 1 µM, respectively.

**Purification of VapC2 and VapC21 toxins**

*E. coli* BL21(DE3) cells transformed with either pET28a-vapC2 or pET28a-vapC21 were grown in M9 minimal medium supplemented with 0.1% glycerol and 50 µg/ml kanamycin at 37°C to exponential phase. Induction of the protein was achieved by adding 1 mM isopropyl 1-thio-D-galactopyranoside (IPTG). After 3 h, cells were harvested by centrifugation and resuspended in lysis buffer (50 mM NaH₂PO₄ (pH 8.0), 500 mM NaCl, 20 mM imidazole, 10 mM β-mercaptoethanol, 1 mM PMSF, 1 mg/ml lysozyme). Cell pellets were then lysed by sonication and lysates applied to a nickel-nitrilotriacetic acid (Ni-NTA) resin (Qiagen) to purify the protein as described previously (Sharp et al., 2012).
**In vitro synthesis of tRNA.**

M. tuberculosis tRNAs were synthesized in vitro fusing a synthetic DNA oligonucleotide containing the T7 RNA polymerase promoter and the 5’ end of the tRNA gene of interest and a second oligonucleotide corresponding to the 3’ end of the tRNA gene. Oligonucleotides were design to have a region of overlap, thus providing a template for PCR. The annealed oligonucleotides were then extended using Taq DNA polymerase to create dsDNA containing the entire tRNA gene preceded by the T7 promoter (tDNA). The product was then resolved on a 2% agarose gel to confirm its size and purified using the QIAquick Gel Extraction Kit (Qiagen). Two-hundred ng of the tDNA was then transcribed in vitro using the RiboMAX Large Scale RNA Production System (Promega) as recommended by the manufacturer. The transcription reaction was separated on a 9% polyacrylamide, 7 M urea gel and visualized by staining with SYBR Gold to confirm the size and purity of the transcribed tRNA. The tRNA transcript was then excised from the gel and incubated for 18 h at 37°C in elution buffer (1 mM EDTA, 0.5 M ammonium acetate, 10 mM magnesium acetate, 0.1% SDS). The eluate was ethanol precipitated and resuspended in nuclease-free water.

**In vitro tRNA cleavage assay**

VapC2 and VapC21 cleavage reactions were performed as described previously (Cruz et al., 2015) with few modifications. Briefly, ~50 pmol of purified VapC toxin was individually incubated with 2 pmol of each of the 45 in vitro synthesized tRNAs from the Mtb H37Rv genome for 3 hours at 37°C in 10mM HEPES (pH 7.5), 15mM KCl, 3mM MgCl₂, and 10% glycerol. Cleavage reactions were run on a 15% Urea-PAGE gel,
stained with SYBR Gold and visualized in a UV transilluminator.

**RNA isolation**

Mycobacterial cells were grown in the absence or presence of inducer for 72 hours. Cells were centrifuged at 2000 × g for 10 min at 4 °C, and supernatants were removed. The cell pellets were resuspended in 1 ml of Trizol and transferred to lysing tubes (Bertin Corp.) containing 0.1 mm glass beads. Cells were lysed in 3 cycles of 30 s at 10,000 rpm using Precellys Evolution homogenizer (Bertin Corp.) with 2 minutes cool down periods on ice in between each cycle. The lysate was centrifuged at 14000 rpm at 4°C and the supernatant was used to extract total RNA using the Direct-zol™ RNA MiniPrep Plus kit (Zymo Research). RNA was treated with an additional step of genomic DNA removal in a TURBO DNase (Thermo Fisher) digestion reaction for 30 min at 37°C, and re-purified using the Zymo RNA Clean and Concentrator kit. The extracted RNA was quantified by spectrophotometry using a µCuvette in a BioSpectrometer (Eppendorf).

**5’ monophosphate RNA-seq**

Preparation of 5’ monophosphate libraries were performed as described in Barth et al (Barth et al., 2021). In brief, The Illumina small RNA 5’ adapter (5’-GUUCAGAGUUCUACAGUCCGACGAUNNNN-3’) was ligated to 3 mg of total RNA using T4 RNA ligase 1 (New England Biolabs) at 4 °C overnight. Ligated reaction was electrophoresed on 6% (wt/vol) polyacrylamide 7 M urea gel and excessive non-ligated adaptors were removed by gel excision and purification. The purified RNA was
used in a Superscript IV reverse transcription reaction (ThermoFisher) using the primer

\[ 5' - GCCTGGCACCAGAAGAA \\
TTCCANNNNNNNN-3' \] and the resulting cDNA was electrophoresed on 10% (wt/vol) polyacrylamide 7 M urea gel and size selected for fragments from 80 to 500 nts. The cDNA libraries were amplified in 12 cycles of PCR using Phusion HF DNA polymerase (ThermoFisher). The oligonucleotides used for PCR amplification were RP1 (5’-AATGATA

\[ CGGCGACCACCGAGATCTACGTTCCGGTGGCTACAGTCCGA-3' \]) and RPIX (5’-CAA

\[ GCAGAAGACGGCATACGAGATNNNNNGTGACTGGAGTTCCTTGGCACCCGAAGAATTCCA-3' \]), where the N’s represent the individual Illumina barcodes for each library. After electrophoresis on a 10% (wt/vol) polyacrylamide gel, amplified DNA between the sizes 150 and 450 bp was isolated by gel excision and purified. The libraries were sequenced in an Illumina HiSeq 4000 platform at New York University’s Genome Technology Center. Data analyses was performed as described in Barth et al. (Barth et al., 2019). Reads that had at least 1 read per million of mapped reads (rpm) for mRNAs and 5 rpm for tRNAs in the induced sample and a fold change of at least 20 were considered.

**Labeling of newly synthesized protein**

The effect of VapC toxins on protein synthesis was determined using a methionine-mimetic, azidohomoalanine (AHA). Mtb cells harboring VapC toxin were grown to 0.1 OD\textsubscript{600} and the toxin was induced by adding PTC. To assess the levels of
newly synthesized proteins, cultures were labeled with a methionine-mimetic, azidohomoalanine (AHA). AHA was added to the sample cultures at 1, 3, and 5 days to the final concentration 50 µM and incubated for 24 hours. Cells were pelleted, resuspended in lysis buffer (2% CHAPS, 8M Urea) and lysed using Precellys Evolution homogenizer as described earlier in RNA isolation. The lysate was centrifuged, and the proteins from the supernatant were linked to an alkyne-containing fluorophore (TAMRA) using the Click-IT Protein Reaction Buffer kit (ThermoFisher) while following manufacturer’s protocol. Ten microgram of protein from each sample were resolved on a 9% SDS-PAGE gel and imaged with a Typhoon FLA 9500 (GE Healthcare) image system.

1.4 Results

Expression of phylogenetically related VapC toxins, VapC2 and VapC21, leads to growth inhibition

We created a phylogenetic tree with all 50 Mtb VapC toxins (Figure 6). We identified two toxins, VapC2 and VapC21, together on one branch of the tree and posited that these two toxins likely have similar functions. Therefore, going forward we studied these two toxins in parallel.

Following convention in the field, we individually expressed the two toxins of interest from an inducible expression plasmid in *M. smegmatis* or Mtb. Tightly regulated inducible plasmids are used because the physiological trigger(s) of nearly all bacterial TA systems are not known. Stress triggers for VapC2 and VapC21 have also not been rigorously established. Expression of either toxin leads to strong growth arrest in both
mycobacterial model organisms (Figure 7). Growth begins to flatline approximately 5 hrs post induction in *M. smegmatis* (Figure 7A,B) and approximately 5 days post induction in Mtb (Figure 7C,D). This phenotype is in alignment with type II TA systems, where toxin expression characteristically leads to some degree of cell growth arrest regardless of toxin’s mechanism of action.
Figure 6: Phylogenetic tree of the 50 Mtb VapC toxins. VapC2 and VapC21 highlighted in red.
Figure 7: Expression of VapC2 and VapC21 leads to growth inhibition. (A,B)

Growth profile of *M. smegmatis* cells harboring anhydrotetracycline (ATC)-inducible plasmid expressing either VapC2 or VapC21 with (+ATC) or without (-ATC) the inducer. (C,D) Growth profile of Mtb cells harboring the pristinamycin (PTM)-inducible plasmid expressing either VapC2 or VapC21 with (+PTM) or without (-PTM) the inducer.

**VapC2 and VapC21 cleave multiple tRNAs in vitro**

Knowing that Mtb VapC toxins have been reported to cleave either tRNA or 23S rRNA, we first prepared recombinant VapC2 and VapC21 toxins and performed in vitro cleavage assays with a complete set of Mtb tRNAs. Only a few tRNAs among the 45
were either partially cleaved or completely degraded when incubated with each toxin (Figure 8A,B *shaded gray*). However, only tRNA\textsuperscript{fMet} (designated as “Met30” in Figure 8A,B) was efficiently cleaved and generated stable cleavage products corresponding to two tRNA halves Figure 8A,B *shaded black*). Although we know that this in vitro cleavage assay typically suggests more tRNA targets than the true target(s) we identify by 5' RNA-seq in vivo (Barth et al., 2021; Cintrón et al., 2019), it is useful to definitively identify the class of RNA favored by an RNase toxin. Therefore, VapC2 and VapC21 are tRNases.
Figure 8: VapC2 and VapC21 cleave multiple tRNAs in vitro to some extent but only tRNA^{Met} to completion. Cleavage assays with each of the 45 in vitro synthesized Mtb tRNAs. Individual tRNAs were cleaved by either VapC2 (A) or VapC21 (B) at 37°C for 3 hours. In vitro assay with tRNAs that showed some degree of cleavage were repeated to confirm the results shown. Enhanced exposure was required to view Arg17 in panel A. tRNA numbering from the Lowe lab genomic tRNA database http://gtrnadb.ucsc.edu (Chan & Lowe, 2016).
**VapC2 and VapC21 cleave and inactivate only tRNA^fMet in vivo**

We have demonstrated previously that endoribonuclease toxins must be expressed in their native host for reliable identification of their true in vivo target(s) (Barth et al., 2021; Cintrón et al., 2019). Therefore, we harvested total RNA from Mtb cells expressing VapC2 or VapC21 vs. controls and performed 5’ RNA-seq to identify the tRNA(s) cleaved by each toxin in vivo. 5’ RNA-seq enables genome-wide RNA target identification coupled with high resolution, single nt cleavage site mapping (Schifano et al., 2014). For VapC toxins, 5’ RNA-seq amplifies RNAs carrying a 5’-phosphate (5’-P) moiety generated upon cleavage by VapC. As suggested by in vitro cleavage assays, 5’ RNA-seq unequivocally identified tRNA^fMet as the sole target of each toxin (Figure 9A,B,C). Both toxins cleave only Mtb tRNA^fMet within the anticodon loop between nts 39 and 40 (Figure 9D); they do not cleave the other two Mtb elongator tRNA^Met species (metT, metV) shown near tRNA^fMet (metU) in Figure 9A,B heatmaps. The position of cleavage is identical to that of the only VapC toxin in Salmonella/Shigella that also exhibits specificity for initiator tRNA^fMet (Winther & Gerdes, 2011) and is in proximity to the G-C pair at the base of the stem that is crucial for initiation factor (IF)-3 discrimination of initiator tRNA versus elongator tRNA^fMet (Barraud et al., 2008) (Figure 9D).
Figure 9: VapC2 and VapC21 cleave and inactivate only tRNA<sup>fMet</sup> in vivo. (A) Heatmap representing fold change (induced/control) at each position for all 45 Mtb tRNA was generated from 5’-P RNA-seq libraries after induction of either VapC2 (A) or VapC21 (B) in Mtb. (C) Histogram representing the ratio of cleavage by VapC2 and VapC21 toxin identified using 5’ RNA-seq at each nucleotide within MetU gene (tRNA<sup>fMet<sub>30</sub></sup>) in Mtb after toxin induction. Genomic position and the negative strand sequence are shown. (D) Representation of VapC2 and VapC21 target, tRNA<sup>fMet</sup>. 
Anticodon sequence (CAU) is highlighted in white; toxin cleavage site is denoted with green arrow.

**VapC2 and VapC21 inhibit protein synthesis**

Since these two toxins disable a tRNA critical for translation of all bacterial mRNAs, we next tested if its role is to comprehensively shut down translation. We used a click chemistry approach for metabolic labeling to monitor *de novo* protein synthesis in *Mtb* cells with and without VapC2 or VapC21 toxin expression. Incorporation of the azide-containing Met mimetic azidohomoalanine (AHA) enabled fluorescent visualization of AHA-containing proteins upon coupling to the alkyne tetramethylrhodamine (TAMRA). As anticipated, translation was severely inhibited at all intervals tested (day 1, 2 and 3 following toxin induction; Figure 10A,B).

![Figure 10: VapC2 and VapC21 inhibit protein synthesis.](image)

Newly synthesized proteins were labeled with the methionine mimetic AHA in induced and uninduced samples 1, 3,
and 5 days post-induction of either VapC2 (A) or VapC21 (B) and visualized upon alkyne-TAMRA conjugation. Each lane was normalized to 10 μg of total protein.

The C-terminal WR residues of VapB2 and VapB20 antitoxins are necessary for toxin inhibition

The structure of the Mtb VapBC2 TA complex revealed 23% identity and 33% overall similarity. The carboxy-terminal VapB2 WR residues of the antitoxin are important for inhibition of VapC2 toxin enzymatic activity (Min et al., 2012). VapB2 W72 embeds into the crevice formed between the VapC toxin dimers; R73 disrupts the catalytic PIN domain of VapC2 by contacting all three aspartate residues that are essential for Mg coordination of these PIN domain (Min et al., 2012). Alignment of VapB2 and VapB21 indicated that the WR amino acid pair was also conserved in VapB21, but it was located 10 amino acids from the carboxy end (Figure 11A). We mutated the WR→AA for each antitoxin; the mutated antitoxins were no longer able to inhibit the activity of their cognate toxin using failure to arrest cell growth as the phenotypic endpoint (Figure 11B,C). Thus, these conserved WR residues are excellent candidates for inhibitors that block the interaction of the antitoxin with the toxin.
Figure 11: The C-terminal WR residues of VapB2 and VapB20 antitoxins are necessary for inhibition of toxin. (A) VapB2 and VapB21 sequence alignments highlighting the position of the WR residues generated using Clustal Omega v. 1.2.4 (Sievers & Higgins, 2018) (B) Growth profiles of *M. smegmatis* expressing VapC2 toxin from the ATC-inducible pMC1s plasmid plus either the empty IVN-inducible pNIT plasmid (control, red circle), wild type VapB2 antitoxin in pNIT (black square), or VapB2 W72A/R72A mutant antitoxin in pNIT (WR→AA, grey triangle). (C) Growth profiles of *M. smegmatis* expressing VapC21 from the ATC-inducible pMC1s plasmid plus either the empty isovaleronitrile (IVN)-inducible pNIT plasmid (control, red circle), wild type VapB21 antitoxin in pNIT (black square), or VapB21 (W79A/R80A) mutant antitoxin in pNIT (grey triangle). Growth curves performed in triplicate; standard deviation indicated.
VapC2 and VapC21 toxins are structurally similar, yet the VapB2 and VapB21 antitoxins are not interchangeable

Finally, we aligned the two toxin sequences to highlight their similarity (40% identity and 52% overall similarity, Figure 12A) and modeled their tertiary structures (Figure 12B).

Yet, the noncognate antitoxin-pairs were not functional, i.e. the VapB2-VapC21 and VapB21-VapC2 pairs still exhibited toxicity as assessed by growth arrest compared to the cognate pairs. Therefore, although there are reports of crosstalk between noncognate toxins and antitoxins in Mtb (Zhu et al., 2010), this was not the case for VapBC2 and VapBC21.

Figure 12: VapC2 and VapC21 toxins are structurally similar, yet the VapB2 and VapB21 antitoxins are not interchangeable. (A) VapC2 and VapC21 sequence alignments was generated using Clustal Omega v. 1.2.4 (Sievers & Higgins, 2018). The five conserved residues comprising the PIN domain are boxed in red. (B) Crystal
structure of Mtb VapC2 (PDB 3H87) and VapC21 (PDB 5SV2) were superimposed using PyMOL. (C) Growth profiles of M. smegmatis carrying VapC2 toxin in pMC1s plasmid with either control pNIT vector (red circle), VapB2 antitoxin (black square), or VapB21 antitoxin (grey triangle). Growth curves performed in triplicate; standard deviation indicated.

1.5 Discussion

To date, Mtb VapC toxins are now known to specifically cleave and inactivate two classes of RNAs (tRNAs and 23S rRNA), leading to two distinct outcomes. tRNA-cleaving toxins encompass those that target specific isoacceptors for mRNA codons and those that target tRNA\(^{f\text{Met}}\). Those that target tRNA isoacceptors used during elongation of newly synthesized proteins have a more subtle effect, they remodel the Mtb transcriptome and proteome toward an outcome beneficial to this pathogen. By contrast, bacterial toxins that cleave and inactivate initiator tRNA\(^{f\text{Met}}\) (Walling & Butler, 2018; Winther & Gerdes, 2011) or the 23S rRNA sarcin-ricin loop (Winther et al., 2016; Winther et al., 2013) are deadly. However, their deadly effects are not necessarily universal. Inactivation of ribosomes through cleavage at the extremely conserved sarcin-ricin loop is deadly to all organisms with ribosomes, from the simplest bacteria to humans. Inactivation of the initiator tRNA\(^{f\text{Met}}\), leading to global inhibition of translation initiation bacteria, is only deadly for bacteria but not higher organisms. This distinction enables suggest that the use of the VapBC2 and VapBC21 TA systems as novel targets for development of better antituberculars.
The drugs currently used for TB treatment were developed >50 years ago by the pharmaceutical industry (Williams & Duncan, 2007). At that time, the prescribed drug regimens led to a significant decline in TB in the industrialized world, so there was little incentive to continue their drug discovery and development programs. This gap in development has left us with a limited arsenal for treatment options with very few drugs that are also effective against Mtb persister cells present during latent TB. Though difficult to prove formally, it is believed that most bacilli in latent Mtb infection are not actively replicating but remain viable. Thus, there is an urgent need for the development of new TB drugs that 1) clear Mtb infection faster because they also kill persisters, 2) are effective against multi drug resistant (MDR)- and extreme drug resistant (XDR)-TB, 3) are less susceptible to being disarmed through the development of cellular resistance mechanisms, and 4) are not contraindicated in combination with HIV antiretroviral therapy.

The ability of VapC2 and VapC21 to specifically target tRNA\textsuperscript{fMet}, and no other bacterial tRNAs, suggest that the development of small molecule inhibitors or peptide inhibitors (I. G. Lee et al., 2015) of antitoxin association with its cognate toxin for the VapBC2 and VapBC21 would be a worthy path to pursue in follow-on studies. This approach should result in an excess of free toxin with specificity to initiator tRNA\textsuperscript{fMet}, leading to widespread translation inhibition and cell death.

Inhibition of protein-protein interactions through small molecule inhibitors enables the manipulation of key intracellular processes that dictate bacterial growth, cancer progression and viral replication. Although viewed as a more challenging strategy relative to small molecule inhibition of the enzymatic function of a single protein, the
number of success stories continues to grow. Some of the highest visibility examples include disruption of p53/MDM2 to induce apoptosis and inhibit tumor growth (Vassilev et al., 2004), disruption of ToxT transcriptional activator dimerization to inhibit transcription of the cholera toxin gene (Hung et al., 2005) and cancer treatment via inhibition of the antiapoptotic activity of Bcl family proteins by blocking their interaction with SH3 domain-containing antiapoptotic proteins (Oltersdorf et al., 2005).

Finally, this approach should overcome another Achilles' Heel of conventional antibiotic therapy—cells must be actively growing to be killed. Because all strains of pathogenic Mtb carry these two TA modules, this approach would kill not only actively growing Mtb, including multi-drug resistant strains, but Mtb persisters as well. This class of novel antituberculars has the potential to dramatically reduce the duration of treatment required for eradication of all viable Mtb cells from infected individuals. The attributable benefits of the development of efficacious, novel VapBC2- and VapBC21-based antibiotics from test tube to the clinic are far reaching and with the potential to save millions of lives worldwide.
Chapter 2: The VapC4 toxin helps *Mycobacterium tuberculosis* evade immune clearing

2.1 Abstract

The *Mycobacterium tuberculosis* (Mt) VapBC4 toxin-antitoxin system is essential for the establishment of Mtb infection. Using a multi-tier, systems-level approach, we uncovered the sequential molecular events triggered by the VapC4 toxin that activate a circumscribed set of critical stress survival pathways which undoubtedly underlie Mtb virulence. VapC4 exclusively inactivated the sole tRNA$^{\text{Cys}}$ through cleavage at a single site within the anticodon sequence. Depletion of the pool of tRNA$^{\text{Cys}}$ led to ribosome stalling at Cys codons within actively translating mRNAs. Genome mapping of these Cys-stalled ribosomes unexpectedly uncovered several unannotated Cys-containing ORFs. Four of these are small ORFs (sORFs) encoding Cys-rich proteins of less than 50 amino acids that function as Cys-responsive attenuators that engage ribosome stalling at tracts of Cys codons to control translation of downstream genes. Thus, VapC4 mimics a state of Cys starvation, which then activates Cys-attenuation at sORFs to globally redirect metabolism toward the synthesis of free Cys. The resulting newly enriched pool of Cys feeds into the synthesis of mycothiol, the glutathione counterpart in this pathogen that is responsible for maintaining cellular redox homeostasis during oxidative stress, as well as into a circumscribed subset of cellular pathways that enable cells to defend against oxidative and copper stresses characteristically endured by Mtb within macrophages. Our ability to pinpoint activation or downregulation of pathways that collectively align with Mtb virulence-associated
stress responses and the nonreplicating persistent state brings to light a novel, direct and
vital role for the VapC4 toxin in mediating these critical pathways.

2.2 Introduction

The vast majority of people infected with Mtb do not develop active tuberculosis
(TB). Instead, in most infected individuals this bacterial pathogen reprograms its
physiology to evade immune clearance and establish a latent TB infection. The Mtb
genome harbors ~90 Type II toxin-antitoxin (TA) systems (Ramage et al., 2009; Sala et
al., 2014), the majority belonging to the 50-member virulence associated protein VapBC
family (VapB - antitoxin, VapC - toxin) unique to pathogenic bacteria. In general, Type
II TA systems comprise adjacent genes encoding two small (~10 kDa) proteins, a toxin
and its cognate antitoxin that inhibits toxin activity through formation of a stable TA
protein-protein complex. In the prevailing model, stress conditions alter the toxin-
antitoxin balance, resulting in a preponderance of free toxin, which exerts its growth-
regulating and/or other functions within the bacterial cells (Harms et al., 2018; Masuda &
Inouye, 2017). Because the activity of individual Mtb toxins typically leads to growth
inhibition, they have been implicated in the establishment of latent TB infection (Arcus et
al., 2011; Arcus et al., 2005; Harms et al., 2018; Ramage et al., 2009; Sala et al., 2014).
Consistent with this, a subset of the 50 VapBC TA systems are differentially expressed
when subjected to stresses relevant to Mtb infection and non-replicating persistence TB
(Agarwal et al., 2018). More importantly, TA module deletion strains used for guinea pig
and mouse infection models demonstrated that VapBC3, VapBC4 and VapBC22 are
essential to establish Mtb infections (Agarwal et al., 2020; Agarwal et al., 2018). VapC4,
VapB5, VapC26 and VapC45 are required for Mtb survival in macrophages and lungs of mice and nonhuman primates (Slayden et al., 2018).

VapC toxins are structure and sequence-specific endoribonucleases (Cruz & Woychik, 2016). Only a few of the 50 Mtb VapC toxins have been studied in detail. In vitro cleavage studies demonstrated that VapC4 (aka VapC-mt4, Rv0595c) and VapC11 (aka VapCmt-11, Rv1561) require a cleavage consensus sequence and specifically target tRNA (Cintron et al., 2019; Cruz et al., 2015; Sharp et al., 2012). Other Mtb VapC toxins with apparent tRNase activity were identified using an RNA-VapC interaction screen in an M. smegmatis host (Winther et al., 2016). However, only recently did we learn that accurate tRNA target identification requires in vivo cleavage in the matched mycobacterial host cell (Cintrón et al., 2019).

Here we applied physiologically relevant, low level toxin expression studies in tandem with powerful genome-scale approaches to elucidate the mechanism underlying VapBC4 virulence. VapC4 cleaves and inactivates a single tRNA\textsubscript{Cys} in vivo. Its activity leads to widespread ribosome stalling at Cys codons, which enabled the discovery of unnannotated Cys codon-containing genes. We demonstrate that some are orthologous to sORFs that act as Cys-responsive attenuators, as recently described in M. smegmatis (Canestrari et al., 2020). Ribosome stalling also unleashes a cascade of downstream events that precisely manipulates the transcriptome based on codon-usage to downregulate growth and redirect metabolism toward synthesis of Cys, which in turn selectively activates genes involved in the oxidative stress and copper detoxification responses, endowing Mtb with the molecular defenses needed to withstand the deadly assaults of the host innate immune response.
2.3 Materials and Methods

Strains, plasmids and reagents

All experiments were performed using either Mtb strain H37Rv (ATCC 25618), Mtb mc2 6206 (ΔpanCD ΔleuCD, generously provided by William Jacobs laboratory, Albert Einstein College of Medicine) or Mtb mc2 6206 ΔVapBC4. VapC4 (Rv0595c locus) was amplified by PCR from Mtb H37Rv genomic DNA. The amplified gene was cloned under the control of an anhydrotetracycline (ATC)-inducible promoter in the pMC1s plasmid. Induction of VapC4 was obtained by adding ATC to the media to a final concentration of 200 ng/ml, and replenishing it every 48 hours. The average induction levels of the VapC4 transcript from the pMC1s plasmid expressed in H37Rv is 6.5 fold (Dataset S3).

Mtb cells were grown under constant shaking at 200 rpm at 37° C in 7H9 Middlebrook media containing 1x OADC supplement (Sigma), 0.05% of tyloxapol, and kanamycin at 25 μg/ml for plasmid selection. The media was supplemented to 50 μg/ml of pantothenic acid and 100 μg/ml of leucine for growth of the attenuated strain Mtb mc2 6206.

Construction of vapBC4 deletion strain

To accurately assess whether VapC4 preferentially downregulates translation of Cys-rich Mtb proteins (Figure 18), we created a strain with the VapBC4 (ΔVapBC4 or ΔRv0596c-Rv0595c) module deleted from the Mtb mc2 6206 parental strain using the ORBIT recombineering method (Murphy et al., 2018). The Bxb1 attP-containing
oligonucleotide was designed to contain the first and last 60 base pairs of the Rv0596c-Rv0595c toxin-antitoxin module (NWO3093, 5’-GAT CGT GAC CGC AGA CAT CGA GAC GAG GCG GCG CGA CAG CAG CTC CGC TAC ACT CTG AGC GGT TTG TCT GGT CAA CCA CCG CGG TCT CAG TGG TGT ACG GTA CAA ACC TGC GAT CCT GAC CCA AGA CAA CGA CTA CGC CGC CAT GCC CGA CGT CGA GGT CAT AAC GAT CTG ACG GTT G-3’). To confirm successful deletion, we performed PCR using oligonucleotides targeting the module flanking regions (NWO3126, 5’- ACA AAT CAC GGC ACT TCG GC – 3’ and NWO3127, 5’- GAG CGG CAA AGT CGT AGC AC – 3’) in combination with internal pKM464 plasmid oligonucleotides (NWO3012, 5’ - CAG GTA TCC GGT AAG CGG CA - 3’ and NWO3013, 5’ - CAC CGA TCC GGA GGA ACT GG - 3’).

RNA isolation

Mycobacterial cells were grown in the absence or presence of inducer for 24 or 72 hours. Cells were centrifuged at 2000 × g for 10 min at 4 °C, and supernatants were removed. The cell pellets were resuspended in 1 ml of Trizol and transferred to lysing tubes (Bertin Corp.) containing 0.1 mm glass beads. Cells were lysed in 4 cycles of 30 s at 9000 rpm using Precellys Evolution homogenizer (Bertin Corp.) with 1 min cool down periods on ice in between each cycle. The lysate was centrifuged at 14000 rpm at 4°C and the supernatant was used to extract total RNA using the Direct-zol™ RNA MiniPrep Plus kit (Zymo Research). RNA was treated with an additional step of genomic DNA removal in a TURBO DNase (Thermo Fisher) digestion reaction for 30 min at 37°C, and re-
purified using the Zymo RNA Clean and Concentrator kit. The extracted RNA was quantified by spectrophotometry using a μCuvette in a BioSpectrometer (Eppendorf).

**In vitro cleavage assays with purified VapC4**

VapC4 purification and cleavage reactions were performed as described in Sharp et al. and Cruz et al., respectively (Cruz et al., 2015; Sharp et al., 2012). Briefly, 10 picomoles of purified VapC4 was individually incubated with 2 pmol of each of the 45 in vitro synthesized tRNAs from the Mtb H37Rv genome or 2 µg of total RNA extracted from Mtb mc² 6206 for 3 hours at 37° C. Cleavage reactions with in vitro synthesized tRNAs were run on a 15% Urea-PAGE gel, stained with SYBR Gold and visualized in a UV transilluminator. Cleavage reactions using total RNA were purified using the Zymo RNA Clean and Concentrator kit following the manufacturer’s protocol and were used for constructing 5’ RNA seq libraries.

**5’ RNA-seq**

Preparation and analysis of 5’OH libraries was performed as described in Barth et al. (Barth et al., 2019). For 5’ monophosphate libraries, the same procedure was followed except for the steps of digestion with Terminator 5’-Phosphate-Dependent Exonuclease (Epicentre) and phosphorylation with T4 PNK (New England Biolabs). The libraries were sequenced in an Illumina HiSeq 2500/4000 platform at Genewiz Corp or New York University’s Genome Technology Center.

For data analysis, we only considered reads that had at least 1 read per million of mapped reads (rpm) for mRNAs and 5 rpm for tRNAs in the induced sample and a fold
change of at least 20. Frequency logos were generated with kpLogo (Wu & Bartel, 2017) or weblogo (Crooks et al., 2004).

To identify unannotated Cys codon-containing ORFs, we searched for unannotated reads with either of the two Cys codons (UGC and UGU) ∼15 nt from the RNA cleavage site (corresponding to stalled ribosome at the “hungry” codon) in the 5′OH RNA-seq dataset. We then determined if the stalled Cys codon was in-frame and fit within our parameters for a potential ORF: at least 10 amino acids with a start (ATG, GTG, or TTG) and a stop (TGA, TAA, or TAG) codon. We selected the longest possible ORF since transcription start sites for these ORFs are not known.

**Total RNA-seq**

In order to remove 16S/23S ribosomal RNA from the total RNA, the samples were treated using the NEBNext bacterial rRNA depletion kit (New England Biolabs). Approximately 100 ng of rRNA-depleted RNA was used to generate the libraries using NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs) and sequenced on an Illumina HiSeq 2500 or similar. Reads were mapped against the Mtb H37Rv reference genome (Genbank accession: AL123456.3) in a local Galaxy instance (Afgan et al., 2016) using the default parameters of Bowtie 2.3.4.3 (Langmead & Salzberg, 2012), featureCounts 1.6.4 (Liao et al., 2014) and limma 3.38.3 (Ritchie et al., 2015). Significantly up- or down-regulated genes (+/- 1.5 fold) with an adjusted p-value < 0.05) were analyzed using the Functional Annotation tool in the DAVID platform using an Benjamini p-value cut off of 0.1 (Huang da et al., 2009a).
Labeling of newly synthesized Mtb proteins.

To assess the levels of newly synthesized proteins in Mtb after toxin induction, triplicate VapC4 samples were induced for 12, 24 and 48 hours and azidohomoalanine (AHA, Anaspec Inc.) was added to the media to a final concentration of 50 µM and incubated for 6 h. AHA is an azide-containing methionine mimetic that is incorporated into proteins, allowing the capture or visualization of the newly synthesized proteins by a copper-catalyzed azide-alkyne cycloaddition reaction. To extract the AHA-labeled proteins, cells were pelleted, resuspended in lysis buffer (2% CHAPS, 8M Urea) and lysed in Precellys Evolution homogenizer as described in RNA isolation section. The lysate was centrifuged, and the proteins from the supernatant were linked to an alkyne-containing fluorophore (TAMRA) using the Click-IT Protein Reaction Buffer kit (ThermoFisher). Ten micrograms of protein from each sample were resolved on a 9% SDS-PAGE gel and imaged with a Typhoon FLA 9500 (GE Healthcare) image system. Fluorescence intensities from each lane were quantified using ImageJ software.

Proteomics of newly synthesized proteins

To identify newly synthesized proteins by quantitative mass spectrometry, mc2 6206 Mtb cells were harvested after VapC4 induced for 24 or 48 hrs. For Figure 18, ΔVapBC4 containing the vapC4-pMC1s plasmid were grown to an OD600 of 0.1 and induced for 48 hours along with uninduced cultures, in triplicate. Fifty ml cultures were centrifuged at 2000 g at 4 °C for 5 minutes and washed with 1X PBS two times to remove traces of the albumin-containing 7H9 media. The cell pellets were resuspended in lysis buffer (2% CHAPS, 8M Urea) and lysed using Precellys Evolution homogenizer as
described in the RNA Extraction section. The lysates were pelleted at 12000 g at 4 °C for 10 minutes and the AHA-labeled proteins contained in the supernatant were selectively captured using alkyne-coated agarose beads from the Click-iT™ Protein Enrichment Kit (ThermoFisher), following the manufacturer’s protocol.

Tryptic digests were analyzed using an Orbitrap Tribrid mass spectrometer and nanoflow LC system (Thermo Scientific), as described in Barth et al. (Barth et al., 2019). The raw LC-MS data was converted into MASCOT Generic Format (MGF) using Proteome Discover 2.1 (ThermoFisher) and searched against either the NCBI Mtb database (Accession: AL123456) together with a database of common laboratory contaminants (http://www.thegpm.org/crap/) using a local implementation of the global proteome machine (GPM Fury) (Beavis, 2006).

Spectral counts were analyzed using the QuasiSeq package (https://cran.r-project.org/web/packages/QuasiSeq/index.html), for proteins containing 15 or more spectral counts total (Lund et al., 2012). Q-values are calculated using the fdrtool package of Strimmer (Strimmer, 2008) and considered significant if below 0.05. Cys codon content of up- and down-regulated proteins was measured by estimating the codon content from the Mtb H37Rv annotated genes (GenBank accession: AL123456).

**Metabolomics - sulfur assimilation pathway**

To analyze metabolites by mass spectrometry, mc2 6206 Mtb cells containing the vapC4-pMC1s plasmid were grown to an OD600 of 0.1 and induced for 48 hours along with uninduced cultures. Equal number of cells (normalized by OD600) for each sample were centrifuged at 2000 g at 4°C for 10 min. The pellets were then transferred to
microcentrifuge tubes and centrifuged at 10,000 g at 4°C followed by removal of as much media as possible. The cell pellets were resuspended in 1.2 mL of extracting solution (Acetonitrile:Methanol:water (40:40:20), 0.5 % Formic Acid). Samples were mixed on ice. 86 µL of neutralization solvent (15% NH4HCO3) was added and mixed thoroughly. Samples were then centrifuged at 13,000 g at 4°C for 10 min. The supernatants were transferred to fresh tubes and were analyzed as previously described (Su et al., 2020).

Briefly, the sample analysis was performed on Thermo Q Exactive PLUS (Thermo Fisher Scientific, Waltham, MA) coupled with hydrophilic interaction chromatography (HILIC). HILIC separation was performed on a Vanquish Horizon UHPLC system (Thermo Fisher Scientific, Waltham, MA) with an XBridge BEH Amide column (Waters, Milford, MA). The m/z scan range was set to 72 to 1000 m/z in negative ionization mode. The MS1 data were processed using Maven (Melamud et al., 2010). Confirmation of the mycothiol retention time was aided by analysis of a panel of M. smegmatis mycothiol mutants (Rawat et al., 2002) kindly provided by the Yossef Av-Gay laboratory (University of British Columbia).

**Data availability**

The sequencing datasets generated in this study were deposited in the NCBI Sequence Read Archive under BioProject accession number PRJNA662430. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD025774.
2.4 Results

VapC4 cleaves many Mtb tRNAs within their anticodon stem loop in vitro

VapC4 expression in Mtb leads to growth arrest (Cruz et al., 2015). To understand how this toxin downregulates cell growth, we first determined how many Mtb tRNAs were cleaved by VapC4 in vitro. Eight of the complete set of 45 tRNAs (Chan & Lowe, 2009, 2016) were cleaved to completion: tRNA{Ala2-UGC}, tRNA{Ala11-CGC}, tRNA{Ala31-GGC}, tRNA{Cys21-GCA}, tRNA{Gly20-GCC}, tRNA{Phe10-GAA}, tRNA{Ser24-GGA} and tRNA{Ser26-GCU} (Figure 13). Surprisingly, five of these eight Mtb in vitro tRNA targets did not harbor an ACGC, ACUGC or ACAGC VapC4 cleavage sequence identified in earlier reports (Cruz et al., 2015; Sharp et al., 2012). In addition, 19 of the 45 tRNAs were partially cleaved (Figure 13, shaded gray). Therefore, 27 of the 45 tRNAs were fully or partially cleaved by VapC4. This data suggested that the growth inhibition associated with VapC4 toxin expression might be due to depletion of multiple tRNAs.
Figure 13: VapC4 cleaves multiple tRNAs in vitro. Cleavage assays with all 45 in vitro synthesized Mtb H37Rv tRNAs; (A) 1-24 and (B) 25-45. Individual tRNAs were cleaved by VapC4 (+ VapC4) at 37 °C for 3 hours (toxin to tRNA ratio, 5:1). No toxin controls (-VapC4) were run in parallel. Black shading denotes complete cleavage; gray, weak cleavage; white, no cleavage.
**VapC4 does not completely inhibit protein synthesis**

Since VapC4 appeared to disable so many tRNAs, we next tested if its role is to comprehensively shut down translation. In fact, others have concluded that tRNA-cleaving toxins generally act by globally inhibiting translation because their expression characteristically leads to bacterial cell growth arrest (Winther et al., 2016; Winther & Gerdes, 2011). We used a click chemistry approach for metabolic labeling to monitor *de novo* protein synthesis in Mtb cells with and without toxin expression. Incorporation of the azide-containing Met mimetic azidohomoalanine (AHA) enabled fluorescent visualization of AHA-containing proteins upon coupling to the alkyne tetramethylrhodamine (TAMRA).

Total proteins were resolved by SDS-PAGE and the newly synthesized population was visualized by fluorescent imaging (Figure 14A,B). Surprisingly, VapC4 did not completely inhibit protein synthesis even though its expression leads to growth arrest (Figure 14C).
Figure 14: VapC4 inhibits growth but does not completely inhibit translation. (A) Newly synthesized (AHA-labeled) proteins from ±VapC4 Mtb H37Rv cells at 12, 24 and 48 hrs post-induction were visualized with an alkyne-TAMRA conjugate. Each was normalized to contain 10 micrograms of total protein. (B) Whole lane fluorescence was measured by ImageJ software selecting equal areas. The average fluorescence signals from triplicates are shown. Error bars correspond to standard error of mean. (C) Growth profiles obtained from Mtb H37Rv cultures with (blue) or without (orange) VapC4 expression for 0, 24, 48 and 72 hours. Dotted lines are OD<sub>600</sub> readings for each replicate and solid lines represent the average of the triplicates. Error bars represent the standard error of the mean.
VapC4 only targets tRNA\textsuperscript{Cys} in Mtb cells

Because the effects on translation were not consistent with the large number of tRNAs cleaved by VapC4 in vitro, we next expressed this toxin in Mtb to determine if the tRNAs cleaved by VapC4 in vitro were also targeted in vivo. RNAs cleaved by VapC toxins are marked with a specific chemical tag at their 5’ end that permits facile isolation, identification and mapping of the toxin cleavage site. Therefore, we enlisted a specialized RNA-seq approach developed in our laboratory, 5’ RNA-seq, that enabled genome-wide VapC4 target identification as well as high resolution, single nt cleavage site mapping (Schifano et al., 2014). 5’ RNA-seq captures and amplifies RNAs carrying the specific chemical moieties, a 5’-phosphate (5’-P) or 5’-hydroxyl (5’-OH), that are generated upon cleavage by VapC4 and other endoribonuclease toxins.

Total RNA was isolated from VapC4 induced and uninduced cells. We created 5’-P and 5’-OH 5’ RNA-seq libraries and identified a single tRNA, tRNA\textsuperscript{CysGCA}, as the sole target of VapC4 (Figure 15A,B). This is the only Cys tRNA in Mtb, and it services both the UGC and the UGU Cys codons. The site of toxin cleavage was within the Cys anticodon (Figure 15C), thus functionally disabling this tRNA for protein synthesis. This tRNA\textsuperscript{CysGCA} target is also consistent with that identified upon ectopic expression of VapC4 in \textit{M. smegmatis} (Winther et al., 2016). Interestingly, in contrast to the MazF family of toxins, which generate a 5’-OH upon RNA cleavage, VapC4 instead appeared to generate a 5’-P upon RNA cleavage since there were no tRNA targets detected in the 5’-OH RNA-seq libraries. To prove that VapC4 directly generated a 5’-P upon RNA cleavage, instead of downstream phosphorylation of a 5’ end that was initially hydroxylated, we incubated VapC4 with total Mtb RNA and again created 5’-P and 5’-
OH 5’ RNA-seq libraries. tRNA$^{\text{CysGCA}}$ was cleaved in the 5’-P library and not the 5’-OH library (Figure 16A,B). This toxin has an absolute requirement for the sequence U G$^{34}$ C$^{35}$ A$^{36}$ A in the proper structural context to recognize and cleave only tRNA$^{\text{CysGCA}}$ (Figure 16C). Therefore, VapC4, and likely all VapC family toxins, are distinct from other ribonuclease toxins in that they generate 5’-P ends upon RNA cleavage.
Figure 15: tRNA\textsuperscript{CysGCA} is the sole VapC4 target in vivo. (A) Heatmap obtained from 5’P RNA-seq libraries (represented in fold change of internal 5’ monophosphate ends in induced vs. uninduced samples) indicating internal cleavage at a single position in only one (cysU) of the Mtb 45 tRNA genes 24 hours post VapC4 induction in Mtb H37Rv cells. (B) Bar graph showing the fold change of 5’ monophosphate ends (induced vs. uninduced) in the only Cys tRNA gene after 24 hour of VapC4 induction. (C) Representation of tRNA\textsuperscript{CysGCA} (D, D-loop; T, T\textsubscript{Ψ}C loop; ASL, anticodon stem loop). Cleavage site, yellow arrow. Anticodon positions are numbered according to standardized tRNA guidelines.
Figure 16: VapC4 cleavage generates 5’ monophosphate ends and lacks specificity in in vitro assays. Heatmap showing fold changes obtained in 5’ RNA-seq for internal 5’OH (A) or 5’P (B) ends in all 45 tRNA genes after incubating purified VapC4 with total RNA extracted from Mtb mc² 6206. (C) Probability logo (obtained using kpLogo) showing the consensus sequence observed in 100 RNA hits with highest fold change (induced vs. uninduced) in the in vitro cleavage 5’P libraries. The RNA sequence flanking the cleavage site (25 nucleotides up- and down-stream) is shown. Nucleotide positions (shown below the kpLogo) are numbered relative to the cleavage site and colored red if the nucleotide is statistically enriched at the position or black if the nucleotide is statistically enriched and its frequency is above 75%.
**tRNA^\text{Cys-GCA} cleavage leads to ribosome stalling at Cys codons**

We recently reported that MazF-mt9 toxin-mediated cleavage of a single species of tRNA, tRNA^{\text{Lys-}}^{\text{UUU}}, leads to selective ribosome stalling at the Lys codon (AAA) requiring that depleted tRNA (Barth et al., 2019). This highly specific, genome-wide ribosome stalling leads to preferential depletion of proteins containing the Lys AAA codon. Serendipitously, these ribosome stalling events were first suggested within the MazF-mt9 5’-OH RNA-seq dataset and then rigorously proven using Ribo-seq (Barth et al., 2019). In perfect agreement, we documented the same two effects for VapC4 following cleavage of tRNA^\text{Cys-GCA}.

**First,** within our VapC4 5’-OH RNA-seq datasets, we saw the hallmarks of ribosome stalling at several hundred Mtb transcripts. Each stalling event is readily detected because there is a conspicuous ~15 nt distance in this 5’-OH RNA-seq dataset to an in-frame “hungry” Cys codon requiring the depleted tRNA (Figure 17A,B). In Barth et al., we used Ribo-seq to demonstrate that this 15 nt distance corresponds to the footprint of the stalled ribosome spanning its 5’ edge to the position of the hungry codon at the A-site (Figure 17C) (Barth et al., 2019). We fortuitously detect these stalled ribosomes because some transcripts that harbor them are recycled by an RNase distinct from VapC4 that happens to leave a 5’OH upon cleavage. **Second,** consistent with our earlier study, we found that the proteome in VapC4-expressing cells contained fewer Cys-containing proteins relative to control cells (Figure 18). Therefore, VapC4 expression leads to inactivation of the primary target of the toxin, tRNA^{\text{Cys-GCA}}, followed by ribosome stalling at Cys codons because the only Cys tRNA in Mtb is now in deficit. These events impact
the proteome, resulting in an overall reduction in the steady state level of Cys-containing proteins.

**Figure 17:** 5’ OH RNA-seq reveals ribosome stalling at Cys codons with single nucleotide resolution (A) top mRNA hits in 5’ OH RNA-seq libraries constructed from Mtb mc² 6206 RNA extracted after 24 hours of VapC4 induction. Cysteine codons are highlighted, approximately 15 nucleotides downstream of the 5’OH cleavage site (5’ of the green capitalized letter). The genome position and strand where the secondary
cleavage occurs is shown, as well as the Rv number of the gene containing the Cys codon. (B) Weblogo showing the consensus sequence from the top 100 mRNA hits found by 5’ OH RNA-seq. Positions are numbered relative to 5’OH cleavage site (scissor). Cysteine codons are predominantly positioned at +15 - +17 (orange underline). (C) following VapC4-mediated cleavage of tRNA\textsubscript{Cys}, ribosome stalling occurs when a transcript containing a Cys codon (UGU or UGC) reaches the A-site because the pool of tRNA\textsubscript{Cys} is depleted by VapC4 cleavage. The “5’ OH” cleavage site (scissor) represents a secondary cleavage event on the mRNA by an unknown RNase that generates a 5’ OH, enabling accurate detection Cys stalling events within our 5’ OH RNA-seq dataset.
Figure 18: Mtb expressing VapC4 preferentially downregulates translation of Cys-rich proteins. Box plot summarizing Cys codon number in newly synthesized up- or down-regulated proteins identified by quantitative mass spectrometry after 48 hr of VapC4 induction in the mc² 6206 VapBC4 deletion strain. Upregulated (n = 104, shown in red) or downregulated (n = 214, shown in blue) proteins were defined by a fold change of at least +/- 1.5 and a q-value < 0.05. Predicted outliers are represented by individual dots and asterisk represents p-value < 0.0001 in a statistical comparison using Student’s t test.
An Mtb sORF is the ortholog of a Cys-responsive attenuator in *M. smegmatis*

Our ability to definitively map the presence and precise position of stalled ribosomes on mRNAs within the VapC4 5’-OH RNA-seq dataset was exploited as a powerful tool to identify transcripts actively undergoing translation, and whose expression is thus impacted by these stalling events. We detected hundreds of stalled ribosomes on mRNAs, corresponding to 444 distinct transcripts within the VapC4 5’-OH RNA-seq dataset. The ribosomes were stalled at either of the two Cys codons, UGC and the UGU, serviced by the single Cys tRNA (tRNA\textsubscript{CysGCA}) in Mtb (Figure 17A,B).

In addition, we identified four leaderless sORFs (lacking a 5’ UTR) whose genome location and sequence revealed compelling functional clues. One of the four sORFs encoded a 29 amino acid small protein with an extremely Cys-rich carboxy-terminus: VSARIEPMLTKRAVDLCRLAGCCCCSC. We discovered that it is the ortholog of the *M. smegmatis* sORF (Ms5788A) first described as one of the many unannotated small leaderless transcripts that are unusually common in mycobacteria (Figure 19A) (Shell et al., 2015). Moreover, recent follow-up studies on the function of Ms5788A support its role in a novel mode of Cys attenuation (Canestrari et al., 2020). Ms5788A is so named because it exerts translational control of the downstream three-gene operon beginning with the Ms5788 (unknown function) followed by the *M. smegmatis* CysA2 sulfotransferase and SseC2 putative sulfotransferase both of whose upregulation is logical under Cys limiting conditions. When Cys availability is low, ribosomes are proposed to stall at the Cys codons of Ms5788A and preclude formation of a secondary structure that would otherwise block the availability of the Shine-Dalgarno (SD) ribosome binding site needed to recruit the ribosome and translate the downstream
mRNA (Figure 19B) (Canestrari et al., 2020). Distinct from the classic tryptophan attenuation in *Escherichia coli*, where the absence or presence of ribosome stalling at tryptophan codons in a small leader peptide dictates the formation of alternate secondary structures that either permit or prevent transcription termination, this *M. smegmatis* Cys-rich small protein acts as a sensor for Cys concentration and dictates whether transcripts that follow are translated.

Based on the function of the Ms5788A sORF in *M. smegmatis*, we predicted that its leaderless Mtb counterpart in front of an operon with similar architecture would also serve as a Cys-responsive attenuator. We designated this Mtb sORF as Rv0815A following the convention of newly annotated genes in Mycobrowser (also used by Canestrari et al. (Canestrari et al., 2020)) to add an A to the name of the gene it precedes, since it resides 173 nt in front of the *cysA2* (Rv0815c) and *sseC2* genes (Figure 19A). We also already knew that ribosomes stall within Rv0815A because that was the basis of its detection within our 5’ RNA-seq dataset. Therefore, stalling at its Cys-rich carboxy terminus would be expected to alter the secondary structure and thus influence expression of *cysA2* and *sseC2*. We performed conventional RNA-seq coupled with quantitative mass spectrometry to determine how expression of VapC4 influenced the abundance of the transcripts and proteins corresponding to these two genes, and additional genes downstream of other Cys-rich sORFs.

We expected upregulation of *cysA2* and *sseC2* in cells expressing VapC4, consistent with the luciferase reporter experiments for *M. smegmatis* Ms5788A (Canestrari et al., 2020). Indeed, RNA-seq performed on RNA harvested 24 and 72 hrs after VapC4 induction showed significant (all p-values < 0.0006) increases in both *cysA2*
and sseC2 transcripts relative to the uninduced controls (Figure 19B). In support of the RNA-seq data, quantitative mass spectrometry of newly synthesized AHA-labeled proteins from ±VapC4 cells also demonstrated that translation of the CysA2 and SseC2 proteins was upregulated (Figure 19C).

**Figure 19:** Downstream genes of identified Cys-containing sORFs are generally upregulated. (A,D,G,I) Genomic organization of the region surrounding the Cys-containing sORFs that occur upstream of annotated genes. Unannotated putative sORFs (red arrow) and their amino acid sequence (Cys residues highlighted in red) are shown. The Cys-containing sORFs in A, D and G have orthologs in *M. smegmatis* (Canestrari et al., 2020; Shell et al., 2015). VapC4 upregulates the transcription of genes downstream of these sORFs in Mtb mc2 6206 (B,E,H,K; all adjusted p-values ≤ 0.05; NS, not statistically significant) and translation of their corresponding proteins (C,F,I,L; solid
red, Strimmer q value ≤ 0.05; striped red, Strimmer q value ≤ 0.1; striped black Strimmer q value ≥ 0.1; ND, not detected).

**VapC4-triggered ribosome stalling at leaderless sORFs identifies additional Cys attenuation-responsive regulons**

We documented VapC4-dependent ribosome stalling on three more unannotated *Mtb* leaderless sORF mRNAs that are also implicated as Cys-responsive attenuators. Two, Rv0485A and Rv2391A, were previously recognized as orthologs of *M. smegmatis* leaderless sORFs encoding small proteins with the telltale Cys-rich tract (Shell et al., 2015) (Figure 19D,G). In agreement with a role in attenuation, the abundance of both of these *M. smegmatis* small proteins increased under low Cys conditions (Canestrari et al., 2020). We identified the third *Mtb* sORF based on its distinctive Cys-rich carboxy-terminus MQQAIQLRFILPRRLAVGCCCC and characteristic genome position (sORF starts 270 bps upstream of the *cysK1* cysteine synthase - *cysE* serine acetyltransferase operon) (Figure 19J) (Canestrari et al., 2020). In support of their role as *Mtb* Cys-responsive attenuators, our H37Rv RNA-seq data 24 hr after VapC4 induction showed significant increases in the abundance of all transcripts immediately downstream of each of the three leaderless mRNAs, consistent with active translation of these transcripts (Figure 19E,H,K). All genes were still upregulated 72 hrs post-induction but generally showed log2 fold change values lower than the 24 hr samples (Figure 19E,H,K). We observed a polar effect on the polycistronic *sirA-cysH-che1-gttB* operon (Figure 19G,H), i.e. a decrease in translation as the ribosomes traverse the polycistronic mRNA, in alignment with the findings reported by Canestrari et al. (Canestrari et al., 2020). Finally,
the transcriptional upregulation we documented for each operon downstream of these three sORFs was again supported by an increase in translation, especially for the first protein encoded by each operon (Figure 19F,I,L). For CysA2/SseC2, this was also demonstrated using direct measurement of protein quantitative mass spectrometry of newly synthesized proteins derived from AHA-labeled cells with and without VapC4 induction.

**VapC4 regulates growth through targeted manipulation of discrete pathways**

In addition to the functional insights obtained through identification of the four sORFs described above, we used the annotation tool “Database for Annotation, Visualization and Integrated Discovery” (DAVID) (Huang da et al., 2009a, 2009b) to obtain a comprehensive summary of biological themes from RNA-seq datasets derived from total RNA harvested 24 hours after VapC4 induction in Mtb H37Rv. These biological themes were consistent when DAVID was performed on RNA-seq datasets from different induction times (72 hrs) in H37Rv or 24 and 72 hrs post induction in the attenuated strain mc²6206. This DAVID analysis also reflects the global impact of the 444 mRNAs for which we identified stalled ribosomes at one or more Cys codons (96 unannotated and 348 annotated).

Among downregulated gene categories, those specifically involved in respiratory ATP synthesis—“NADH dehydrogenase (ubiquinone) activity”, quinone binding”, “quinone” “ATP synthesis coupled electron transport”, “oxidative phosphorylation”—were the most affected (Figure 20). The “cell wall/cell membrane” as well as “ABC transporters” gene groups are downregulated (Figure 20). ABC transporters are substrate-
specific couriers of molecules such as amino acids, peptides, polysaccharides, proteins, ions and antibiotics into or out of the cell. “Phosphate transport” was also identified as a down-regulated category (Figure 20), based on the decreased expression of the majority of ABC phosphate transporter subunit transcripts. Although inorganic phosphate is essential for many cellular processes, it is often limiting in the environment so must be transported into cells. Consequently, phosphate limitation restricts growth and importantly, phosphate depletion triggers Mtb persistence (Rifat et al., 2009). Therefore, rather than broadly shutting down translation as previously proposed for tRNA-cleaving VapC family members (Winther et al., 2016), our data instead suggest that VapC4 controls growth and alters Mtb physiology by specifically downregulating discrete pathways that modulate growth, the mycobacterial envelope, energy generation, transport of essential factors/nutrients and synthesis of essential metabolites.
Figure 20: DAVID analysis of transcripts down- and upregulated by VapC4. (A)
VapC4 engages multiple mechanisms for growth control. DAVID Functional Analysis Tool terms\(^{39,40}\) associated with genes that were significantly (\(+/−\) 2 fold) downregulated after 24 hours of VapC4 induction by total RNA-seq analysis.

**VapC4 upregulates sulfur metabolism and oxidative stress pathways**

Overall, both RNA-seq and supporting proteomic data exposed selective upregulation of a subset of key stress survival pathways (Figure 21). The “sulfate assimilation/sulfur metabolism” categories had the highest statistical significance and are consistent with a toxin whose primary target is tRNA\(^{\text{Cys}}\). Sulfur is an essential element that possesses a variety of chemical properties—redox capacity, metal binding proficiency, nucleophility and the ability to form strong disulfide bonds—that support its critical role in Mtb metabolism. In its reduced form sulfur is used for the biosynthesis
of both Cys and methionine. Cys is particularly important as it is incorporated into proteins, coenzymes and the glutathione counterpart in Actinobacteria called mycothiol (Newton et al., 2008).

Notably, we found that all of the enzymes in the Mtb sulfur import and sulfur assimilation pathway to Cys synthesis (Hatzios & Bertozzi, 2011) were upregulated when VapC4 was expressed (Figure 22A,B). In addition, all enzymes in two alternate synthetic pathways to Cys (Agren et al., 2008; Steiner et al., 2014) were upregulated: CysE and CysK1 that catalyze synthesis of Cys from serine (Figure 22A, middle), CysM, Mec and MoeB1 or CysK2 that catalyze the synthesis of Cys from O-phosphoserine (Figure 22A, bottom right). Overall, 16 of the 19 the “cys”-assigned genes in Mycobrowser were upregulated (Dataset S4, “VapC4 DAVID_upregulated” tab).

Another upregulated DAVID category “oxidative stress/disulfide oxidoreductase” represents pathways that are functionally related to “sulfate assimilation/sulfur metabolism” (Figure 22A). The ability to sense and respond to oxidative stress is critical for Mtb survival upon infection. We found that VapC4 expression results in upregulation of other oxidative stress/disulfide oxidoreductase pathways that are consistent with an increase in sulfur metabolism shown in Figure 21A. A major end product of sulfur uptake and metabolism pathways, Cys, is required for multiple pathways in Mtb, including intra and intermolecular disulfide bond formation, coordination of metal ions, and in oxidoreductase proteins. Cys also supports the production of mycothiol (Figure 22A). Mycothiol, present at millimolar levels in the cell, is a Cys-ligated disaccharide that is essential for Mtb growth and is responsible for maintaining an intracellular reducing environment and minimizing the damaging effects of oxidizing agents such as hydrogen...
peroxide and NO (Newton et al., 2008). Increased production of mycothiol in response to oxidative stress would consume large amounts of cysteine. In support of mycothiol upregulation, transcription of the first enzyme in the pathway for its synthesis, MshA, is also strongly increased 24 and 72 hrs post VapC4 induction (Figure 19E), and newly synthesized MshA protein is also upregulated (Figure 19F). The gene encoding MshA appears to be regulated by the Rv0485A sORF because it directly follows the first gene, Rv0485, in this Cys attenuation-responsive operon (Figure 19D,E,F). More definitively, we confirmed that mycothiol and methionine production increases upon VapC4 expression (Figure 22C,D), while serine levels decrease as expected because it is a substrate for Cys (Figure 22E). Cys is not detectable by mass spectrometry because it has a very short half life. Instead, Mtb mycothiol is the storage form of Cys; it is present at levels >600 times that of Cys (Newton et al., 1996).

Independent of DAVID analysis, we also found significant upregulation of transcripts encoded by 19 previously identified genes in the alternate sigma factor SigH regulon, at 24 and/or 72 hrs post VapC4 induction (Raman et al., 2001; Sharp et al., 2016). SigH is required for virulence and regulates a network of genes that respond to oxidative and heat stress responses, both to limit cellular damage and re-establish redox equilibrium in response to these stresses (Mehra et al., 2012; Raman et al., 2001). These regulon members include two other stress-response alternate sigma factor genes (sigB, sigE) and a thioredoxin/thioredoxin reductase pair (trxB1/B2) involved in maintaining redox homeostasis and a mycothiol-dependent reductase (Rv2466c) that is required for Mtb viability in the presence of hydrogen peroxide stress (Rosado et al., 2017). The 18 SigH-regulated genes that show increased expression in response to VapC4 induction are:
Rv1221, *sigE*; Rv1471, *trxB1*; Rv1528c, *papA4*; Rv1875; Rv2466c; Rv2707; Rv2710, *sigB*; Rv3206c, *moeB1*; Rv3223c, *sigH*; Rv3913, *trxB2* Rv0140; Rv0384c, *clpB*; Rv0991c; Rv1039c, *ppe15*; Rv1259, *udgB*; Rv1801, *ppe29*; Rv3054c; Rv3463 (Sharp et al., 2016) (Dataset S3, “SigH regulon” tab). ClpB is part of a multi-chaperone protein damage repair system that is induced by cell stress and Rv0991c is an oxidation-activated molecular chaperone (Becker et al., 2020). UdgB functions in DNA excision repair by removing promutagenic uracil bases created by cytosine deamination to uracil. The high GC content of Mtb makes it especially susceptible to cytosine deamination upon exposure to reactive nitrogen and oxygen species (Kumar et al., 2011). PPE15 is required for triacylglycerol accumulation (Daniel et al., 2011), the dominant Mtb carbon source under nonreplicating conditions (Maurya et al., 2018).
Figure 21: DAVID analysis of transcripts down- and upregulated by VapC4. VapC4 mimics Cys starvation to activate the oxidative and Cu stress responses. DAVID terms associated with genes that were significantly (+/- 2 fold) upregulated after 24 hours of VapC4 induction by RNA-seq analysis. The dotted lines in both panels represent a 2-fold enrichment between the proportion of the term in the observed genes compared to the expected proportion of the term when considering all genes in the genome. The area of each circle in both panels is proportional to the number of observed genes in the corresponding category.
**VapC4 upregulates the CsoR and RicR copper resistance pathways**

Although copper (Cu) is an essential micronutrient, it is deadly at high concentrations. Macrophages increase the phagosomal concentration of Cu upon Mtb infection. Therefore, Cu resistance mechanisms are thought to help Mtb evade macrophage-mediated killing. DAVID analysis identified upregulation of “Copper” and “Copper ion” functional categories (Figure 21).

There are currently three independently regulated Mtb copper-response pathways that function in concert (Darwin, 2015), although Cu treatment itself induces many more genes beyond these defined pathways (Ward et al., 2008). In the first pathway, the copper-sensitive operator repressor CsoR controls the expression of a four-gene operon (Figure 23A). The CsoR dimer lacking bound Cu binds DNA and functions as the repressor. Cu binding to CsoR induces a conformational change, release from DNA and derepression. The second Cu resistance pathway engages the regulated in copper repressor RicR repressor which binds to a palindrome in the promoter region of several genes/operons (Festa et al., 2011) (Figure 23A). Its mechanism of de-repression is presumed to parallel that of CsoR. The third Cu resistance pathway involves the action of a single membrane protein, MctB, a mycomembrane protein thought to block entry of copper into the cell (Wolschendorf et al., 2011) (Figure 23A).

We found that all of the 14 genes in the CsoR and RicR regulons were upregulated upon VapC4 expression, but we did not observe upregulation of MctB, the sole component of the third Cu resistance pathway (Figure 23B). In addition to the dramatically upregulated, divergently transcribed RicR-regulated *mmcO* and *cysK2* transcripts, we identified marked increases in newly synthesized MmcO and CysK2
protein by quantitative mass spectrometry (Figure 23B, *white text in red box*).

Interestingly, three of the transcripts in the RicR regulon, *mmcO, lpqS* and *CysK2*, were consistently upregulated during the Wayne model of nonreplicating persistence in independent microarray experiments (Muttucumaru et al., 2004; Voskuil, 2004; Voskuil et al., 2004)

Although the mechanisms of action of many of the genes controlled by CsoR and RicR are not known, *mymT, mmcO, cysK2* and *ctpV*, encode proteins whose roles have been identified (Darwin, 2015). Mycobacterium metallothionein MymT is central to Cu homeostasis because it is able to bind and sequester up to six Cu ions (Gold et al., 2008), the Mycobacterium multicopper oxidase MmcO oxidizes toxic Cu+ to less toxic Cu2+ (Rowland & Niederweis, 2013) and the CtpV ATPase appears to export Cu (Figure 22A). CysK2 is an S-sulfocysteine synthase which catalyzes the synthesis of Cys from O-phosphoserine (Steiner et al., 2014) (Figure 22A, *bottom right*), Therefore, transcription of *cysK2* is dependent on the presence of Cu and the activity of the metabolic CysK2 enzyme leads to synthesis of mycothiol, suggesting that the redox homeostasis and Cu homeostasis pathways are interconnected in Mtb.
Figure 22: VapC4 mimics Cys starvation to activate the oxidative and Cu stress responses. (A) Transcripts involved in sulfate assimilation are upregulated in Mtb cells expressing VapC4. Red, upregulated at both 24 and 72 hrs in Mtb H37Rv cells (≥ 1.5 fold, adjusted p-value ≤ 0.05); Orange, upregulated at 72 hrs only (≥ 1.5 fold, adjusted p value ≤ 0.05). Adapted from (Agren et al., 2008; Bhave et al., 2007; Steiner et al., 2014;
Many copper response genes are metabolic enzymes involved in sulfate assimilation and oxidative stress

Among the genes identified in the proteasome mutant screen by the Darwin laboratory that fortuitously identified members of the RicR regulon were three “cys” genes, cysK2, cysNC and cysD. (Festa et al., 2011); all three were upregulated upon VapC4 expression. CysNC, CysD and CysK2 each catalyze reactions in the sulfate assimilation pathway (Figure 22A). Four “cys” genes were identified in a screen for Cu responsive genes: cysA2, cysA3, cysK2 and cysW (Ward et al., 2008). CysA2 and CysA3 are “thiosulfate sulfurtransferases”. CysW is a component of an ABC transporter for sulfate that shuttles sulfate into the cell for the sulfate assimilation pathway as illustrated in Figure 20A. Each of these “cys” annotated genes were upregulated in response to VapC4 expression as well. Curiously, cysA2 resides downstream of one of the sORFs we identified by Cys ribosome stalling (Figure 19A,B,C). As detailed above, CysA2 is the Mtb ortholog of MSMEG_5788 that resides within the well characterized M. smegmatis operon regulated by Cys attenuation through the Cys codon-rich sORF MSMEG_5788A
(Figure 19A,B,C) (Canestrari et al., 2020). Taken together, the common themes between sulfur assimilation and Cu resistance pathways are reflected by the seven “cys” annotated genes shared between the two pathways—\(cysA2\), \(cysA3\), \(cysNC\), \(cysD\) and \(cysW\)—and all are upregulated by VapC4 expression.

Finally, Ward et al. found that 13 of the 24 genes upregulated during the Mtb Cu response are involved in the response to oxidative stress as well (Ward et al., 2008). Since many were transcription regulators for oxidative stress, they concluded that the Cu response is not specific to Cu exposure but falls within the generalized response to oxidative stress. We found that all 24 genes upregulated in response to Cu (including the 13 also involved in the oxidate stress response) were also upregulated upon VapC4 expression. Therefore, expression of VapC4 produces an expression profile that integrates Mtb sulfur assimilation, response to Cu and oxidative stress pathways. In alignment with this integrated theme, we found upregulation of VapC4 itself within the Cu response microarray data in Ward et al. (Ward et al., 2008). Therefore, Cu may be a physiological trigger for VapC4.
Figure 23: VapC4 upregulates key components of Cu homeostasis. (A) CsoR and RicR regulon proteins with known functions are illustrated; all were upregulated by VapC4 (labeled in black text). MctB was not affected by VapC4 expression (labeled with gray text). (B) Heatmap of upregulated CsoR and RicR regulon transcripts 24 and 72 hrs post VapC4 induction in Mtb H37Rv cells (≥ 1.5 fold, adjusted p-value ≤ 0.05). The proteins encoded by the *mmcO* and *cysK2* genes were also identified in quantitative mass spectrometry datasets (adjusted p values ≤ 0.05 (RNA-seq), Strimmer q-values ≤ 0.05 (QMS). Since *socA* and B were not annotated in the NCBI Mtb H37Rv database, we used their genome positions to identify read counts within the raw RNA-seq datasets (log₂FC at 24/72 hrs: *socA* 1.17/2.23, *socB* 1.17/3.59; all adjusted p-values ≤ 0.05). Illustration of Mtb Cu pathways (panel A) created with Biorender.com.

2.5 Discussion

The abundant Mtb TA systems have been suggested to be involved in the establishment of latent tuberculosis infection based on the observation that expression of *E. coli* toxin components typically leads to a reversible state of growth arrest and a
“quasidormant” metabolic state, a phenotype that has been observed with many but not all Mtb toxins (Harms et al., 2018; Masuda & Inouye, 2017). Early studies of TA toxins in *E. coli*—most of which are endoribonucleases—almost exclusively identified mRNA as the toxin target (Masuda & Inouye, 2017; Yamaguchi & Inouye, 2011; Yamaguchi et al., 2011). Therefore, for many years the prevailing view was that all endoribonuclease TA toxins in bacteria were sequence-specific “mRNA interferases” that imparted reversible growth arrest through widespread mRNA degradation that then leads to inhibition of protein synthesis. However, once our laboratory and others began to characterize the targets of TA toxins in Mtb, it was clear that this pathogen was quite different from *E. coli* and the gram-positive pathogen *Staphylococcus aureus*. Not only does Mtb employ endoribonuclease toxins for inactivation of 23S rRNA at helix loop 70 (Schifano et al., 2013; Schifano et al., 2014; Schifano & Woychik, 2014) or the highly conserved sarcin-ricin loop (Winther et al., 2013), it likely engages a large proportion of its TA systems for highly selective isoacceptor-specific tRNA cleavage (Barth et al., 2019; Cintron et al., 2019; Cruz et al., 2015; Schifano et al., 2016).

In this study we induced low level expression of VapC4 to recapitulate the physiologic effects of VapC4 produced by the not yet identified natural in vivo trigger(s) of toxin activation, thus enabling us to isolate the effects of this toxin on Mtb physiology and separate these effects from other broad stress responses. We then used a spectrum of powerful systems-level approaches to pinpoint the cellular pathways regulated by VapC4 expression (illustrated in Figure 24) in order to elucidate the molecular mechanism underlying the essential role for VapBC4 in virulence (Agarwal et al., 2018). We discovered an extensive but highly specific Mtb response based on the selection of
tRNA\textsubscript{Cys} as the primary toxin target. This toxin utilizes tRNA\textsubscript{Cys} inactivation to mimic a state of Cys starvation leading the cell to retool its metabolism toward Cys synthesis. This shift is mediated in part by ribosome stalling at tracts of Cys codons within sORFs that are projected to block the formation of secondary structures that sequester the ribosome binding site for the downstream operon, thus allowing translation of downstream genes. In conjunction with the increased production of proteins regulated by these sORFs, we identified highly specific proteome reprogramming via selective depletion of transcripts based on their Cys-codon content. The effects we observed with VapC4 are consistent with the importance of sulfur in Mtb, where inorganic sulfur is assimilated into Cys, which then serves as a precursor to a variety of key metabolites such as mycothiol and Fe-sulfur enzymes to maintain protein synthesis and a redox balance (Figure 22A) (Bhave et al., 2007; Hatzios & Bertozzi, 2011; Schelle & Bertozzi, 2006). Cys biosynthesis in Mtb is upregulated in response to oxidative stress (Pinto et al., 2004), is protective for cell survival of oxidative stress (Senaratne et al., 2006) and protects cells during the persistent phase of infection in mice (Senaratne et al., 2006). While mining the themes in upregulated genes led us to valuable insights into the metabolic changes that are expected to lead to protection against certain stresses, our data also suggested that certain features of the oxidative stress response and the response to Cu are coordinated as they both share several Cys biosynthetic genes. An equally careful analysis of downregulated genes led to many surprising modes of growth control engaged by this toxin, especially at the level of energy generation.

With over 90 Mtb TA systems, many of which are thought to be activated by stress, and a finite number of infection-associated stresses (e.g. hypoxia, oxidative stress,
nutrient limitation, nitrosative stress, low pH), it is likely that more than one toxin is activated in response to infection. We demonstrated that the VapC4 mechanism of action begins with cleavage of a single tRNA and transitions to codon-dependent proteome remodeling. These findings suggest a model in which comprehensive defense against host immune responses would enlist the coordinated activation of multiple tRNA-cleaving toxins that each target a single, discrete tRNA isoacceptor. Based on our findings with VapC4, each of these toxins would have unique effects on remodeling the proteome to activate or repress complementary pathways would allow Mtb to collectively attain comprehensive protection against host-mediated stresses. We hypothesize that VapC4 engages the integrated sulfur assimilation/oxidative stress and Cu response pathways as a survival strategy early in infection. In fact, since VapC4 expression is upregulated when Mtb cells are exposed to Cu in the growth medium (Figure 23C), Cu may be the environmental trigger for activation of VapC4 during infection. However, VapC4-mediated protection from these stresses would not fully protect this pathogen from the combined assaults of the immune response. We speculate that each toxin engages unique arms of the multifaceted Mtb stress response to infection. These toxin-induced arms of defense differ because each is defined by the codon-dependence dictated by the toxin tRNA target. A coordinated defense system engaging an armamentarium of tRNA-cleaving toxins would provide a sophisticated, multitier evolutionary adaptation for survival upon Mtb infection of the human host, one that could adapt to each stage of infection and respond to both innate and adaptive immune mechanisms.
Figure 24: VapC4 reprograms Mtb physiology to protect against oxidative and Cu stress. Cu upregulates VapC4 transcription, VapC4 toxin precisely cuts the sole tRNA$^{\text{Cys}}$ at its anticodon GC$\downarrow$A to create tRNA halves. These tRNA halves are nonfunctional in protein synthesis, resulting in ribosome stalling at Cys codons (UGU, UGC) requiring this depleted tRNA. Stalled ribosomes at Cys codons globally reduce new synthesis of Cys-containing proteins (Figure 18). A subset of transcripts with stalled ribosomes were Cys codon-containing unannotated small ORFs, where ribosome stalling at tracts of Cys codons within these sORFs is predicted to be a novel mode of Cys attenuation to regulate translation of downstream genes (as shown for sORF Rv2334A). Many of these sORFs upregulate translation of enzymes for sulfate assimilation to Cys. As macrophages attempt to eliminate Mtb with ROS/RNS and toxic Cu$^+$, the enriched pool of Cys can be incorporated into molecules needed to defend against these assaults. Key components of these two major stress response pathways are upregulated by VapC4. For oxidative stress, these include mycothiol the abundant, essential, glutathione counterpart in Mtb that regulates cellular redox status and multiple genes in the SigH regulon. For copper stress
these include multiple mechanisms controlled by copper-responsive regulators RicR and CsoR to mitigate copper toxicity.
Chapter 3: Toxin-mediated enhancement of intrinsic antibiotic resistance in *Mycobacterium abscessus*

3.1 **Abstract**

*Mycobacterium abscessus* (Mab) causes the majority of infections within the *M. abscessus* complex and is inexplicably the most intractable to clearing after aggressive and lengthy antimicrobial treatment regimens. These infections most commonly manifest as chronic lung disease or attack skin and soft tissues, and therefore take an especially high toll on those with cystic fibrosis or other chronic lung disorders. Since the sequences of hundreds of Mab clinical strains do not reveal obvious clues to explain the tenacity of these infections, Mab might instead enlist toxin-antitoxin (TA) systems to impart a phenotypic switch that enhances its ability to evade the killing action by multiple antibiotics. We found that the VapBC5 TA system is absent in the ATCC 19977 reference strain yet frequently present in Mab clinical strains. Here we uncover the detailed mechanism of action of the VapC5 toxin and track the series of downstream events it triggers that potentiate evasion of killing by first line Mab antibiotics. First, VapC5 selectively inactivates six distinct tRNAs—tRNA^{fMet}, tRNA^{Ser-CGA}, tRNA^{Met}, tRNA^{Leu-UAG}, tRNA^{Thr-UGU}, tRNA^{Thr-CGU}—through cleavage at a single site within their anticodons. This leads to transcriptional reprogramming toward activation of multiple antibiotic resistance pathways coupled with the upregulation of ribosomal protein genes. These resistance pathways encompass the dramatic upregulation of the WhiB7 master regulator of mycobacterial intrinsic resistance (typically induced by exposure to antibiotics) along with several families of transcription factors associated with antibiotic
resistance. These transcriptional changes lead to proteome remodeling toward the robust, sustained and nearly exclusive synthesis of ribosomes and other components of the translation machinery even though cells are in a state of growth arrest. Therefore, this toxin-mediated phenotypic shift is predicted to disrupt the efficacy of all first line Mab antibiotics—which attack ribosomes or require actively growing cells—as well as support robust recovery when antibiotic is removed. Consistent with this prediction, VapC5-expressing Mab cells supported an increase in persistence (aka intrinsic resistance) when exposed to amikacin and cefoxitin, representatives of the two classes of first line antibiotics. These findings demonstrate that the phenotypic changes caused by VapC5 expression—growth arrest, ribosome overproduction, activation of the WhiB7 regulon and related transcription factors—appear to override the efficacy of first line antibiotics. These findings provide a much-needed framework for the selection of shorter and more efficacious alternate treatment options for Mab infections using currently available antimicrobials whose targets are not confounded by VapC5.

3.2 Introduction

In the USA, the rapidly growing nontuberculous mycobacteria *Mycobacterium abscessus* causes the most intractable infections among the three strains comprising the *M. abscessus* complex: *M. abscessus* (referred to hereafter as “Mab”), *M. massiliense* and *M. bolletii*. Mab infections most commonly manifest as chronic lung disease or they attack skin and soft tissues and are inexplicably the most intractable to clearing after aggressive and lengthy antimicrobial treatment regimens (Benwill & Wallace, 2014;
Brown-Elliott et al., 2012; Johansen et al., 2020; Mougari et al., 2017; Nessar et al., 2012; Novosad et al., 2016; Park et al., 2017). Consequently, Mab infections take an especially high toll on individuals with cystic fibrosis (CF) or chronic lung disorders. When inhaled, Mab can accelerate inflammatory lung damage in cystic fibrosis patients as well as cause serious pulmonary disease in the immunocompromised and other individuals with underlying lung disorders (reviewed in (Benwill & Wallace, 2014; Johnson & Odell, 2014; M. R. Lee et al., 2015; Medjahed et al., 2010)). Modes of transmission are broader than once thought, as most individuals with cystic fibrosis appear to contract Mab pulmonary infections via long-lived aerosols or fomites. (Bryant et al., 2013; Bryant et al., 2016). A significant number of Mab infections are also acquired in the hospital, especially after plastic surgery, leading to skin and soft tissue infections. Mab skin infections can also occur upon exposure to contaminated syringes or needles. Nosocomial infections and suspected infections from fomites are also hard to prevent because Mab is resistant to most disinfectants and biocides (M. R. Lee et al., 2015).

Mab infections display important similarities to certain features of *M. tuberculosis* infection, namely its ability to cause chronic disease associated with granuloma formation and to prolong infection by residing within phagocytes (Johansen et al., 2020). Yet, Mab is resistant to conventional antituberculcurs, and is generally one of the most drug resistant mycobacteria. In fact, compassionate use of bacteriophage therapy is the option of last resort for gravely ill CF patients with disseminated, antibiotic resistant Mab infections (Dedrick et al., 2019). There is an urgent need for better treatment options for Mab infections because even after treatment, patient prognosis is poor, and mortality remains
high. Current treatment regimens typically last 1-2 years with an approximately 25-45% success rate (Chen et al., 2019; Diel et al., 2017; Martiniano et al., 2014; Mougari et al., 2016; Wu et al., 2018). It is not understood why Mab is so extraordinarily refractory to antimicrobial drugs. The difficulty in clearing Mab infections suggest that this bacterium enlists persistence to at least in part evade being killed when exposed to antibiotics. Persistence (often referred to as “intrinsic resistance” in Mab) refers to the ability of a bacterial pathogen to survive antibiotic exposure via a phenotypic switch, *i.e.* *without acquisition of genetic mutations.*

It is well established that activation of the toxin component of Type II toxin-antitoxin (TA) systems also confounds the efficacy of antibiotics because they lead the formation of persister cells that are not actively dividing and thus are refractory to most antibiotics (Fisher et al., 2017; Gerdes & Maisonneuve, 2012; Harms et al., 2016; Lewis, 2007, 2010; Maisonneuve et al., 2011). Yet, the presence of TA systems in any of the three Mab complex subspecies has not been documented, nor has a role for TA systems been considered as a factor undermining the efficacy of treatment. TA systems are operons comprising adjacent genes encoding two small (~10 kDa) proteins, a toxin and its cognate antitoxin that inhibits toxin activity through formation of a stable TA protein-protein complex. Stress conditions are thought to lead to lower levels of the antitoxin and thus, a preponderance of free toxin which exerts its growth-regulating and/or other functions from within the bacterial cells (Fraikin et al., 2020).

To investigate how TA systems might influence Mab physiology, virulence, and antibiotic susceptibility, we determined the mechanism of action of a representative VapC toxin, VapC5, often present in Mab clinical strains. These studies led us to uncover
a role for VapC5 in subverting the efficacy of first line Mab antibiotics without acquisition of one or more resistance mutations. They also lay the groundwork for evidence-based selection of alternate antibiotics whose potency is not abrogated by the action of this prevalent toxin.

3.3 Materials and Methods

Toxin-antitoxin Identification

VapC5 toxin (WP_074293606.1) was identified with the use of BLASTP using the known sequence for Mtb VapC5 (Rv0627). The genes were considered to be homologous if they met the criteria of a minimum percent identity of 30 across the entire gene and a minimum bit score of 40 (Pearson, 2013). Mab VapC5 compared to its M. tuberculosis counterpart held 32% identity and 48% positive (E = 2^{-5}, Bit = 100) across the genes and thus were considered homologous. In order for VapC5 to be considered a toxin within a TA system an upstream gene must be identified within the same operon to be a putative antitoxin. A putative VapB5 (WP_074293607.1) was identified upstream of the previously identified VapC5 in each clinical strain identified. This putative VapB5 also met the criteria to be considered a counterpart in M. tuberculosis (33% identity, 43% positives, E = 1^{-14}, Bit = 116).
Strains, Plasmids and Reagents

All experiments were performed using *Mycobacteroides abscessus* strain L948 (ATCC 19977). Mab L948 cells were grown in 7H9 Middlebrook medium supplemented with a final concentration of 0.05% Tween 80, 0.5% bovine albumin, 0.2% dextrose, 0.085% NaCl, 0.1% Casamino acids and 20 μg/ml kanamycin or zeocin a 25 μg/ml (for plasmid selection). All cultures were grown at 37 °C under constant agitation at 200rpm. VapC5 was synthesized with the addition of gene flanking NdeI and HindIII restriction sites by GenScript in pUC57. VapC5 was digested from pUC57 and cloned adjacent to an anhydrotetracycline (ATC) inducible promoter in the pMC1s plasmid. VapC5 expression was induced by adding a final concentration of 200 ng/ml ATC (Clontech) to the media when cells reached an OD (600nm) of ~0.1 and compared to an empty plasmid control. For co-expression experiments VapB5 (WP_074293607.1) was synthesized with the addition of gene flanking EcoRI and NdeI restriction sites by GenScript in pUC57. VapB5 was digested and cloned into an isovaleronitrile (IVN) inducible plasmid pNIT. VapB5 was then cloned into Mab competent cells already containing pMC1s-VapC5. For co-expression, cells were induced by adding ATC and IVN to a final concentration of 100ng/ml and 0.5 μM, respectively.

RNA isolation

In order to extract total RNA, ~200 ml of Mab cells were collected at 6h post-induction. Empty pMC1s vector RNA was collected at the same timepoint and utilized as controls. Cells were centrifuged at 2000 g at 4°C for 10 min. Cell pellets were resuspended in 1 ml of Tri reagent (Zymo Research) and transferred to 2 ml lysing kits.
tubes (Bertin Corp.) containing 0.1 mm glass beads. Cells were lysed with 4 cycles of 30 s at 9000 rpm with 2 min cooling periods in-between using a Precellys Evolution homogenizer (Bertin Corp.). The lysate was centrifuged at 14,000 rpm for 5 min at 4°C and RNA was isolated from the supernatant using the Direct-zol RNA Miniprep Plus extraction kit (Zymo Research). Isolated RNA was subsequently treated with 1 U of Turbo DNase for 30 min at 37°C and purified using the RNA Clean and Concentrator kit (Zymo Research). Purified RNA was eluted in 45 μl of RNase-free water and RNA concentration was measured using a BioSpectrometer (Eppendorf) with a μCuvette.

Northern analysis of tRNA levels

In order to detect Mab tRNA, specific DNA probes complementary to the 5’ or 3’ end of Mab tRNA<sup>Ser-CGA</sup>, tRNA<sup>Met</sup>, tRNA<sup>Met</sup>, tRNA<sup>Thr-UGU</sup>, tRNA<sup>Thr-CGU</sup>, tRNA<sup>Leu-UAG</sup> genes (NWO3166: 5’-TTTGAGGCGGTGCTCCTTAGGC-3’, NWO3275: 5’-TGAGTCGTGCCTCTCTCTTTAGGC-3’, NWO3111: 5’-CGAGCTGCTCCACCCCCGC-3’, NWO3279: 5’-CTTACAAGGGCAGCTCTCTACCCAC-3’, NWO3277: 5’-TTACGAGTGAATTGCTCTACCGACTG-3’, NWO2762: 5’-GCGCGCCTATTTACGCAC-3’, respectively) were radiolabeled at the 5’ end with T4 polynucleotide kinase (New England Biolabs) and [γ<sup>32</sup>P]ATP (PerkinElmer) for 1 h at 37 °C. Total RNA (1μg) from Mab was resolved on a 15% polyacrylamide, 7 M urea gel and stained with SYBR Gold (Invitrogen) to assess and ensure overall quality. The RNA was transferred to a nylon Hybond-N+ membrane (GE Healthcare) and hybridized with the <sup>32</sup>P-labeled oligonucleotides overnight at ~10°C below probe T<sub>m</sub>. To
remove non-specific signals the membranes was washed with 2X SSC and 0.1% SDS at
temperature near probe T\textsubscript{m} for 15 min, followed by a secondary was with 0.1X SSC and
0.1% SDS at the same temperature for 15 min. The membranes were then exposed to
phosphorimager screens for 2 hrs and scanned using the Typhoon FLA 9500 (GE
Healthcare) image system.

**Labeling of newly synthesized Mab proteins**

In order to assess newly synthesized protein levels following VapC5 induction,
cells were grown to an OD of ~0.1 and divided into induced and uninduced samples.
Cells were labeled with 50 μM of a methionine analog azidohomoalanine (AHA,
AnaSpec) for 2 hours at 4- or 6-hours post-induction. Prior to labeling with AHA, cells
were spun down to remove media containing casamino acids and resuspended in new
media without casamino acids. Cells were then incubated at 37°C for 30 minutes after
which AHA was incorporated into the media. Cells were pelleted by centrifugation at
4°C and washed with 1X PBST. Cell pellets were resuspended in 2% CHAPS, 8M Urea
buffer and were lysed with 3 cycles of 45 s at 9000 rpm, with 2 min cooling periods in-
between using a Precellys Evolution homogenizer (Bertin Corp.) The lysate was
centrifuged, and the supernatant was used in the Click-IT Protein Reaction Buffer kit
(ThermoFisher), which links the AHA-containing proteins with alkyne-containing
version of the fluorophore TAMRA (ThermoFisher). Two micrograms of protein for each
sample were run on an SDS-PAGE gel and scanned using the Typhoon FLA 9500 (GE
Healthcare) image system. After imaging the gel was silver stained to ensure accurate
protein loading.
Construction of 5'-dependent libraries was performed as previously described (Schifano et al., 2014). Briefly, for sequencing of RNAs containing 5’ Monophosphate ends (5’- P), total RNA from Mab containing pMC1s-Mab VapC5 (+ATC and empty vector) after 6h of induction were purified. A 5’ adapter (5’-GUUCAGAGU UCUACGUCCGACGUAUCNNNNNNN-3’) was ligated using 1 U of T4 RNA ligase 1 (New England Biolabs) at 16°C overnight. To remove the remaining free adapters, the adapter-ligated RNAs were resolved on a 6% TBE-Urea PAGE gel, excised and precipitated in isopropanol at −80°C. Purified RNAs were used in a reverse transcription reaction using Superscript IV (Thermo Fisher) and the degenerate primer (5’- GCCTTGGCACCAGAGATCCCA NNNNNNNN-3’). The resulting cDNA was loaded into a 10% TBE-Urea gel and fragments between 80 and 500 nts were excised and precipitated. The resulting cDNA libraries were amplified in 14 cycles of PCR using Phusion HF DNA polymerase (ThermoFisher). The oligonucleotides used for PCR amplification were RP1 (5’ AATGATACGGCGACCACCTATACACGCTTCAGA GTTCTACAGTCCGA-3’) and RPIX (5’-CAAGCAGAAGACGGCATACGAGA TNNNNNGTGACTGGAGTTCTTGGCACCAGAGACGCTTCCA-3’), where the N’s represent the individual Illumina barcodes for each library. The amplified libraries between 150-450 bp were gel purified from a 10% TBE PAGE gel and sequenced using a HiSeq 2500 platform or similar.

The resulting FASTQ files had the adapter sequences and the first 6 nucleotides of the 5’ end trimmed using Trimmomatic (Bolger et al., 2014). Reads were trimmed to 20 nts and reads containing fewer than 20 nucleotides were excluded. The remaining
reads were mapped to Mab genome (NCBI accession: CU458896.1 with the addition of the two putative tRNAs not annotated in the Mab ATCC 19977 genome that were identified in the GtRNAdb tRNA database (Chan & Lowe, 2016)) using bowtie 1.2 applying the parameters –n 0–l20 (Langmead et al., 2009). Read counts were normalized to sequencing depth and expressed as reads per million (rpm) of mapped reads. Fold change was calculated using the ratio for each position by dividing the counts in induced sample/counts in uninduced sample. Positions that had 0 counts in the uninduced library were adjusted to a pseudo-count of 1. Reads were only considered if they had at least 5 rpm (mRNAs) and 50 rpm (tRNA) in the induced sample with a fold change >10. Weblogo (Crooks et al., 2004) was utilized to visualize nucleotide frequency of the top hits using the default parameters. Amino acid frequency was visualized using kpLogo (Wu & Bartel, 2017) of the top hits using the default parameters.

RNA-Seq

In order to deplete 16S/23S ribosomal RNA from extracted RNA, samples were subjected to the rRNA depletion method described in Culviner et al., 2020 (Culviner et al., 2020). In short, extracted RNA was incubated with biotinylated oligos designed to specifically hybridize to the 23S, 16S and 5S rRNA of a variety of bacterial species. Biotinylated rRNA was then bound to magnetic streptavidin beads, leading to the depletion of ~75% of rRNA. Approximately 100ng of rRNA depleted RNA was used to generate the libraries utilizing using NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs) and sequenced on an Illumina HiSeq 4000. The resulting sequences were mapped to the M. abscessus reference genome (NCBI
accession: CU458896.1) using the default parameters of Bowtie 1.2. Stringtie 1.3.439 and DESeq2 2.11.40.240 programs were used for transcript assembly and differential expression analysis, respectively.

**Proteomics**

To assess newly synthesized protein in Mab cultures (induced and empty vector controls, in quadruplicate) were labeled with AHA as described previously after 6 h of induction for 2 hrs. We selectively captured newly synthesized proteins using an alkyne-containing column from the Click-iT™ Protein Enrichment Kit (ThermoFisher) followed by in-column trypsin digestion. Digest were analyzed in two separate runs and combined. Data was analyzed as previously described (Barth et al., 2019). Data is presented as estimated log$_2$ ratios of induced/empty control. Q-values are calculated using the fdrtool package of Strimmer (Strimmer, 2008) with significant changes at or below a q-value of 0.05.

**Antibiotic MICs**

AMK (Sigma) and FOX were diluted into a stock solutions of 50mg/mL in sterile filtered $ddH_2O$ and stored at 4°C. ATCC19977 Mab cells were inoculated in 7H9 media (Supplemented as above) until they reached an OD$_{600nm} = 0.1$ ($1 \times 10^7$ CFU/ml). The cells were diluted so that 100 µl contained $1 \times 10^5$ cells. Antibiotics were diluted to desired concentrations (0-128 µg/ml AMK or 0-256 µg/ml FOX) and 100 µl of cells were added to applicable wells in a 96 well plate. The plates were incubated for 3 days at 37°C and the MIC was visually determined as the first clear well (Schon et al., 2020).
**Persister Cell Assay**

Mab cells containing either empty vector controls (pMC1s) or pMC1s-VapC5 were grown at 37°C until they reached an $\text{OD}_{600\text{nm}} \approx 0.1$. pMC1s-VapC5 containing cells were split into induced (+ATC 200 ng/ml) and uninduced cultures. 6hrs after toxin induction $\text{OD}_{600\text{nm}}$ measurements were taken to confirm toxicity. 1ml of VapC5 induced cultures and empty vector controls were aliquoted, serially diluted and plated for CFU/ml. 160 µg/ml of AMK or 320 µg/ml of FOX was added to each toxic VapC5 culture and the empty vector controls. After a period of antibiotic exposure 1 ml of treated cells were pelleted at 15,000g for 5 minutes. The cells were washed two times with PBST, serially diluted and plated to determine CFU/ml. The antibiotic exposure timepoints were compared to the pre-treatment timepoint to determine cell survival rate.

**3.4 Results**

**Many Mab clinical strains harbor TA systems.**

We searched the sequences of strains isolated from respiratory samples derived from Mab-infected patients that were archived in the NCBI genome database or the NCBI BioSample database (including whole-genome shotgun contigs). We identified 128 Mab strains that contained toxin genes orthologous to any of the toxin components for the ~90 annotated *M. tuberculosis* TA systems (Sala et al., 2014). We found VapBC TA systems were most often represented in the clinical strains (the *M. tuberculosis* genome harbors 50 VapC TA toxins designated VapC1-VapC50 (Sala et al., 2014)). Thirty two of
the 128 clinical strains (25%) carried an apparent VapBC5 system, while the Mab reference strain ATCC 19977 does not carry this toxin.

VapC (Virulence associated protein) toxins are endoribonucleases that are only present in bacterial pathogens. In *M. tuberculosis*, VapC toxins predominantly cleave (and inactivate) specific tRNAs at a single site in the anticodon stem loop to generate tRNA-halves (Cruz et al., 2015; Schifano et al., 2016; Schifano & Woychik, 2017; Winther et al., 2016). The Mab antitoxin VapB5 (WP_074293607.1) and toxin VapC5 (WP_074293606) were 46 and 51% similar, and 36 and 34% identical, to their counterparts in *M. tuberculosis*, respectively; Figure 24A). This degree of conservation combined with the tandem, yet out of frame, orientation of the antitoxin and toxin genes strongly suggest that the Mab VapBC5 is a functional TA system.

Therefore, in this work we investigated the detailed mechanism of action of the VapC5 toxin to determine if it contributed to Mab persistence upon antibiotic exposure.

**VapC5 expression leads Mab cell growth arrest.**

The dynamic ratios of toxin:antitoxin are thought to impart *reversible growth inhibition* in response to stress, with the free toxin acting to downregulate growth (Fraikin et al., 2020). To determine if the putative VapBC5 TA system behaves as predicted, the toxin alone or both toxin and antitoxin were expressed from plasmids in Mab ATCC 19977 cells (whose genome does not contain endogenous VapBC TA systems). Low level VapC5 expression under control of an anhydrotetracycline (ATC)-inducible promoter in the pMC1s plasmid inhibited cell growth (Figure 25B), while coexpression of VapC5 with the VapB5 antitoxin from the isovaleronitrile (IVN)-inducible promoter
of the pNIT plasmid rescued growth (Figure 25B). VapC5-mediated growth arrest was evident as early as 4h post-induction in liquid culture (Figure 25C). Therefore, Mab VapBC5 behaves as a bona fide TA system.

Figure 25: Many *M. abscessus* clinical strains contain an *M. tuberculosis* VapC5 ortholog whose expression leads to cell growth arrest. A. T-Coffee and BoxShade alignment of VapB5 and VapC5 of *M. tuberculosis* to newly identified counterparts in
Mab. B. Mab cells containing both pMC1s-VapC5 and pNIT-VapB5 were streaked on 7H9 plates containing either: no inducer, +ATC or +ATC and IVN (respectively). C. Mab cultures harboring either an empty vector (blue) or pMC1s-VapC5 (red) were grown in triplicate in 7H9 media until $\text{OD}_{600} \approx -0.1$. The pMC1s-VapC5 cultures were induced with ATC and OD$_{600}$ was measured every 2-4 hrs. Error bars represent standard deviation from the average of the three biological replicates. Asterisks represent statistical significance between control and induced in a Student’s $t$-test comparison (*, p-value < 0.02; **, p-value < 0.01).

**VapC5 targets six tRNAs, including tRNA$^{\text{fMet}}$, for cleavage within their anticodons.**

VapC toxins are structure and sequence-specific endoribonucleases (Cruz & Woychik, 2016). To identify the RNA target(s) of Mab VapC5, we used a specialized RNA seq method, 5’ RNA-seq, developed in our laboratory (Cruz et al., 2015; Schifano et al., 2014). 5’ RNA-seq identifies RNAs cleaved by VapC5 on a genome-scale and maps the toxin cleavage site to single nt resolution. VapC toxins generate a 5’ monophosphate (-P) upon cleavage of their RNA targets (Barth et al., 2021). This 5’-P moiety cleanly distinguishes toxin-generated products from the majority of intact cellular RNAs.

We performed 5’-P RNA-seq on RNA harvested from Mab ATCC 19977 cells expressing VapC5 and identified six tRNAs specifically cleaved by this toxin: tRNA$^{\text{fMet}}$, tRNA$^{\text{Ser-CGA}}$, tRNA$^{\text{Met}}$, tRNA$^{\text{Leu-UAG}}$, tRNA$^{\text{Thr-UGU}}$, tRNA$^{\text{Thr-CGU}}$ (Figure 26A). These six tRNA targets were validated by northern analysis (Figure 26B). As with other toxin
tRNases, each of the six were cleaved at a single site within their anticodon loop to generate tRNA halves (Figure 26C) (Barth et al., 2021; Cintrón et al., 2019; Cruz et al., 2015; Schifano et al., 2016). Figure 26D-I illustrates the sequence and structure of each of the six tRNA targets surrounding the VapC5 cut site. Notably, all six tRNAs were cleaved between the 2nd and 3rd nt of the anticodon sequence. Overall, the identification of six tRNA targets was unexpected since the few examples of well characterized VapCs exhibit exquisite selectivity for a single tRNA in vivo (Barth et al., 2021; Cintrón et al., 2019; Winther & Gerdes, 2011).
Figure 26: VapC5 targets multiple tRNAs within their anticodon in vivo. A. Heatmap showing the fold changes of 5’ monophosphate levels (induced/uninduced) in each position of all 47 annotated Mab tRNA genes after 6h of VapC5 induction; tRNA gene ID (from genome CU458896.1) shown in parentheses. Two additional tRNAs listed in the (Chan & Lowe, 2016) tRNA database, but not annotated in the Mab reference.
genome, were also included in our analysis. Neither of these unannotated tRNAs with low (~35) general tRNA model scores—tRNA-Ile CAT (chr1.trna18) and tRNA Lys TTT (chr1.trna47)—were targeted by VapC5. Therefore, they are not listed in the heat map. B. Northern analysis of VapC5 tRNA targets identified by 5’ RNA-seq. Each blot shown is representative of three biological replicates. Full length and positions of tRNA halves when visible are indicated. Oligonucleotides used for tRNA\textsubscript{Ser-CGA} and tRNA\textsubscript{Thr-CGU} span the cleavage site, precluding visualization of cleavage products. C. Diagram of tRNA halves produced upon VapC5 cleavage. D-I. Diagrams of Mab tRNA anticodon stem loop targeted by VapC5 from panel A (from highest to lowest fold change). Gray shaded nucleotides, tRNA anticodons; red arrows, VapC5 cleavage sites identified by 5’ RNA-seq.

**VapC5 tRNA targets share sequence and structural features.**

The selectivity of VapC toxins for their tRNA targets appears to require both sequence and structural determinants (Cintrón et al., 2019; Cruz et al., 2015). Therefore, we assessed the sequence similarities between the six full length tRNAs cleaved by Mab VapC5 and identified three identical and 26 conserved nts dispersed throughout their sequences (Figure 27). However, sequence conservation was concentrated within the regions beginning within of the D stem loop and ending in the variable stem loop moving 5’ to 3’ along the tRNA (Figure 28B). Collectively, these conserved regions comprise the vertical arm of the upside-down L-shaped tRNA tertiary structure. This arm of tRNA is predicted to interact with VapC dimers based on our published docking simulations of an
*M. tuberculosis* VapC with tRNA (Cruz et al., 2015). In addition, recognition of determinants in the sugar phosphate backbone likely contribute to cleavage specificity of VapC5 and other VapC toxins because aminoacylases recognize determinants in the tRNA sugar phosphate to ensure correct amino acid attachment to tRNA (McClain et al., 1998). Finally, tRNAs can also be post-transcriptionally modified and the presence, or lack of, a particular modification may influence recognition by VapC5. In summary, tRNA sequence, structure, modifications and sugar phosphate backbone are all predicted to enable VapC5 to distinguish six tRNAs from among the 47 Mab tRNAs and other highly structured RNA species in vivo.

**Figure 27:** VapC5 cleavage is sequence and structure specific. A. Alignment of VapC5 six tRNA targets. (Top) Alignment was generated via tRNA alignment of
anticodon position (33-35) with 32 nucleotides upstream and 37 nucleotides downstream of the anticodon; number of nucleotides upstream or downstream was selected to generate the longest length analysis on tRNAs of differing sizes. (Bottom) Consensus of similarity between the six tRNA targets. B. VapC5 target tRNA^{Ser}^{CGA} diagram demonstrating location of tRNA similarities based on the above alignment. A-C. Blue nucleotides indicate positions within the six nucleotides which have only two nucleotide possibilities. Red nucleotides indicate nucleotides which are identical in all six tRNA.

VapC5 induces WhiB7 and multiple genes that underlie intrinsic antibiotic resistance in Mab.

Genome-wide comparison of the abundance of transcripts from cells with or without VapC5 expression by RNA-seq revealed that the whiB7 transcript exhibited the highest increase in expression in the entire 4971-transcript dataset in two independent experiments each with multiple replicates (first transcript, Figure 28A). The transcription factor WhiB7 is activated by an array of antibiotics and is the central regulator of antibiotic resistance in actinomycetes (Hurst-Hess et al., 2017; Ramon-Garcia et al., 2013). In Mab, WhiB7 is known to activate the ribosomal methylase erm(41), this enzyme represents the primary mechanism of intrinsic macrolide resistance in this pathogen (Nash et al., 2009; Pryjma et al., 2017). erm(41) was significantly upregulated in our RNA-seq datasets (2nd transcript, Figure 28A). The Gcn5-related N-acetyltransferase eis2, the aminoglycoside 2′-N-acetyltransferase aac(2′) and the efflux pump tap are also regulated by WhiB7 (3rd, 4th, and 5th transcript, Figure 28A) (Hurst-
Hess et al., 2017; Miranda-CasoLuengo et al., 2016; Pryjma et al., 2017). Mab has seven other WhiB family transcription factors whose functions are not defined, only some of which are also significantly upregulated by VapC5 (“Other WhiB”, Figure 28A). The effect of VapC5 on the transcription of 13 antibiotic response genes identified by Miranda-CasoLuengo et al. through RNA-seq of erythromycin or kanamycin treated Mab ATCC 19977 is shown in the “Antibiotic Response” category in Figure 28A (Miranda-CasoLuengo et al., 2016). Finally, another RNA-seq study comparing ΔwhiB7 to whiB7-complemented strains identified 127 transcripts in the Mab WhiB7 regulon (Hurst-Hess et al., 2017). Nine of the 13 antibiotic response genes identified in Figure 28A were upregulated in the WhiB7 regulon reported by Hurst-Hess et al. (Hurst-Hess et al., 2017).

We next used the DAVID annotation tool (Database for Annotation, Visualization and Integrated Discovery) (Huang da et al., 2009a, 2009b) to help define other statistically significant up- and down-regulated functional categories of transcripts from the RNA-seq dataset that were not as readily recognizable as WhiB7. The theme within the downregulated gene categories was consistent with cells in a state of growth arrest, e.g. reduced transport (many shared with the ‘cell membrane’ category) and energy production. Genes for mammalian cell entry were also downregulated, suggesting the toxin may be active when Mab is already in the macrophage. By contrast, there were only two significantly upregulated categories: ‘transcription factor activity’ and ‘ribosomal proteins’. Among the 46 transcription factors, the most abundant subgroup contains 9 members of the MarR family (Multiple antibiotic resistance) (Grove, 2013). The next most abundant subgroup consists of 8 members of the ArsR family whose function is poorly defined in mycobacteria. However, there is one example of an ArsR family
member in that decreases susceptibility of *M. tuberculosis* cells to the front line antibiotic isoniazid (Yang et al., 2021). Seven sigma factors were among the remaining genes along with subgroups of families of transcription factors without a clear function (e.g. 7 in the GntR family, 6 in the LysR family) (Figure 28B). The second ‘ribosomal proteins’ category comprised >50% of the 57 ribosomal protein genes (Figure 28B). In fact, nearly all were upregulated to some extent but 35 met our stringent cut offs for expression level and statistical significance. Clearly, VapC5 selectively orchestrates extensive transcriptional reprogramming and upregulation of ribosomal protein genes.
Figure 28: VapC5 induces the expression of WhiB7, ribosomal proteins and multiple genes that underlie intrinsic antibiotic resistance in Mab. A. Heat map of transcript changes in 25 genes associated with Mab intrinsic antibiotic resistance in ATCC 19977 upon VapC5 expression (“VapC”, left column; all adjusted p-values < 0.0001), erythromycin treatment (“ERY”, middle column) or kanamycin treatment (“KAN” right column). “WhiB7” group, Mab transcripts in WhiB7 regulon associated
VapC5 expression reprograms the proteome to exclusively stockpile ribosomes and translation factors.

We next tracked how depletion of the six VapC5 tRNA targets influences the Mab proteome. Mab cells with or without VapC5 expression were metabolically labeled with the azide-containing Met mimetic azidohomoalanine (AHA) and newly synthesized proteins were captured on alkyne-coated agarose beads using click chemistry. Quantitative mass spectrometry enabled genome scale identification of proteins whose synthesis was up- or down-regulated after 8, 10 or 14 hrs of VapC5 expression relative to the empty vector control (Figure 29). As anticipated, upon gradual depletion of six tRNAs that included the critical tRNA^fMet, we detected thousands of proteins whose synthesis was downregulated after VapC5 expression (blue dots in Figure 28A-C). However, as protein synthesis in these growth arrested cells diminished after extended
toxin induction, fewer proteins met our statistical cutoff. Among the 1120 statistically significant proteins, 875 proteins were downregulated after 8 hrs (Figure 29A), 245 were downregulated after 10 hrs (Figure 29B). There were no proteins that met the statistically significant threshold after 14 hrs (Figure 29C). To identify and assess the physiological significance of these downregulated proteins, we used the DAVID annotation tool to sort these proteins into functional categories. DAVID analysis of the 1120 downregulated proteins in the 8 and 10 hr datasets placed them in functional categories consistent with cells in a state of growth arrest, e.g. peptidoglycan synthesis, DNA replication, energy production and metabolism.

We also identified hundreds of proteins whose synthesis increased relative to the control upon VapC5 expression (Figure 29D-F) even though cells were very tightly growth arrested (Figure 25C). DAVID analysis revealed that these proteins fall into a surprisingly small number of interrelated functional categories with ‘protein biosynthesis’ as the dominant theme (Figure 29D-F). The categories are interrelated and many of the same proteins fall into one or more of the categories. First, there was sustained new synthesis of ribosomal proteins 8, 10 and 14 hrs after VapC5 expression (Figure 29D-F). The ‘tRNA-binding’ group contains five proteins; these are ribosomal protein that interact with tRNA and are also present in the ‘ribosomal protein’ category. Of the 57 ribosomal proteins in Mab, all but four very small ribosomal proteins were detected. The four missing ribosomal proteins—L35, L34, L31 and L30 that are only 64, 47, 74 and 60 amino acids long, respectively—generate too few tryptic peptides of the length needed for reliable detection by mass spectrometry. All other essential components of the basic protein synthesis machinery, i.e. initiation factors IF-1, IF-2 and IF-3 as well as
elongation factors EF-Tu, EF-Ts, EF-G and EF-P, were detected by mass spectrometry and upregulated in the newly synthesized protein pool (represented within the ‘protein biosynthesis’ and/or ‘cytoplasm” DAVID categories, Figure 29D). Although IF-2 and EF-P were detected at levels above the control (~0.4 log₂ fold change), they were below our ± 0.58 log₂ fold change cut off (i.e. ± 1.5 fold) and excluded from the dataset used for DAVID analysis. Since GTP-bound IF-2 binds to tRNA^{fMet} and brings it to the initiation complex, VapC5-induced depletion of tRNA^{fMet} may result in coordinate signals that reduce the levels of available IF-2. EF-P is engaged only in special cases, runs of prolines residues, so may not need to be as abundant at the others. The initiation and elongation factors only transiently interact with ribosomes in the initiation or elongation complex and are continuously recycled. Therefore, they are not expected to be needed at levels as high as ribosomal proteins that stably assemble into ribosomes.

In the ‘chaperone’ category GroEL, GroES, trigger factor, DnaJ1 and DnaJ2 are significantly upregulated at 8 and 10 hr after VapC5 expression (Figure 29D,E). Although DnaK (Hsp70) is also significantly upregulated at all time points (0.81, 1.84 and 1.64 log₂ fold change, respectively), the 8 hr and 10 hr strimmer q-values did not meet our statistical cut off (≤ 0.05) for inclusion in DAVID analysis. Coordinate synthesis of chaperones with ribosomal proteins is consistent with their role in ribosome assembly and protein synthesis (Shajani et al., 2011). Trigger factor directly associates with ribosomes to promote co-translational folding of nascent polypeptides while other chaperones aid in this folding process (Deuerling et al., 2019). Many of these chaperones fall into the ‘RNA degradation’ category as well because they participate in rRNA maturation (Shajani et al., 2011) (Figure 29E). As the length of VapC5 induction
increased to 14 hrs, just two DAVID categories remained comprising all ribosomal protein (Figure 29F), four of which interact with tRNA and also sort to the ‘tRNA-binding’ category.

Finally, although not represented in the DAVID analysis because many of the corresponding genes are not annotated, we discovered that nearly all of the enzymes in the TCA cycle are upregulated 8 hrs post induction. This initial burst of energy generates ATP and NADH as well as amino acid precursors that are functionally integrated with the ‘amino acid biosynthesis’ DAVID category. Several tRNA synthetases sort to the ‘ligase’ subset along with a few unrelated metabolic enzymes with ligase activity (Figure 29D).

In summary, the predominant functional themes of newly synthesized proteins that are upregulated upon VapC5 expression center on ribosome biosynthesis, translation factors and energy production. Therefore, even though multiple tRNAs are cleaved and inactivated by VapC5, this toxin modifies Mab physiology so it retains its ability to generate energy and synthesize proteins while maintaining a state of tight growth arrest.
Figure 29: VapC5 predominantly upregulates the synthesis of ribosomal proteins.

A.-C. Cultures expressing either empty vector or VapC5 were grown to an OD of 0.1
then induced with ATC for 6h, 8h or 12h followed by AHA-labeling for 2h (n=4). Newly
synthesized AHA-labeled proteins were detected by mass spectrometry and empty vector
versus VapC5 expressing cells represented as volcano plots. Blue, down regulated
proteins; red, upregulated proteins. Horizontal dotted line indicates significance cut-off (-
log q-value =1.3) and vertical dotted line indicates fold-change cut-off (log2 ratio)
=0.58). D.-F. DAVID Functional Analysis Tool terms (Huang da et al., 2009a, 2009b)
associated with genes significantly upregulated (q strimmer < 0.05) following 8h, 10h or
14h of VapC5 induction. Circle color corresponds to its -log10 Benjamini p-value.
Diameter of each circle reflects the relative number of proteins in the corresponding
category; protein number also shown within each circle.
VapC5 expression enhances persister formation.

The upregulated proteins identified by quantitative mass spectrometry revealed valuable clues to how VapC5 might manipulate Mab physiology to favor survival during antibiotic treatment. Therefore, we tested whether VapC5 influences cell survival after treatment with the aminoglycoside amikacin or the cephalosporin cefoxitin, first line antibiotics for treatment of Mab pulmonary disease (Daley et al., 2020). Mab cells containing VapC5 or empty vector were exposed to 10 X the minimal inhibitory concentration (MIC) of amikacin or cefoxitin for 12, 24 or 36 hrs, serially diluted and plated to determine CFU/ml. It is recommended to use at least several times MIC to ensure that one is measuring true persistence and not transient modes of resistance such as heteroresistance (Balaban et al., 2019). We observed an ~40-fold and ~100-fold increase in recovery of viable cells after amikacin or cefoxitin exposure, respectively, in VapC5 expressing cells compared to the control (Figure 30). This result indicates that the killing action of amikacin and cefoxitin is weakened when VapC5 is expressed, leading to a larger pool of persisters that can seed growth once the antibiotic is removed.

Mechanistically, amikacin inhibits protein synthesis by irreversibly binding to 16S rRNA and the S12 ribosomal protein components of the 30S subunit. In alignment with the activity of amikacin, our quantitative mass spectrometry data shows a VapC5-mediated burst in synthesis of the S12 ribosomal protein 8 hrs after toxin induction (1.76 log₂ fold change). This newly synthesized S12 protein is produced along with the other ribosomal proteins as shown in Figure 28, This large pool of newly synthesized
ribosomal proteins is expected to assemble into stable Mab ribosomes and build the cache of ribosomes predicted to collaborate in confounding the efficacy of this antibiotic.

Figure 30: VapC5 potentiates intrinsic resistance in Mab. VapC5 (red) or empty vector (blue) expressing cells were exposed to 10X MIC for amikacin (AMK) or cefoxitin (FOX). Survival was compared to an empty vector control. Data are averages of at least three independent experiments; error bars indicate standard error. Asterisks represent statistical significance between control and induced in a Student’s t-test comparison (*, p-value < 0.02, **, p-value < 0.0001).

3.5 Discussion

Mab is ubiquitous in the environment, therefore, clinical isolates harboring VapC5 may have acquired this TA system from neighboring bacteria since toxin genes can be in carried on plasmids, bacteriophages or within bacterial chromosomes. Acquisition of TA systems endows a competitive edge in the battle for survival under harsh conditions where competition for scarce nutrients is great and stress is frequent. In contrast to the ~90
TA systems harbored by the pathogenic H37Rv *M. tuberculosis* strain, most of the clinical 128 clinical Mab strains we identified carry only one TA system. VapBC5 was present in ~25% of these 128 strains, and in each case the only TA system we identified. However, at least in the case of VapC5, this one TA system packs a punch that is predicted to underlie the tenacity of Mab after long and aggressive treatment regimens (illustrated in Figure 30).

The Mab VapC5 tRNase inactivates six distinct tRNAs in vivo—tRNA\(^{fMet}\), tRNA\(^{\text{Ser-CGA}}\), tRNA\(^{\text{Met}}\), tRNA\(^{\text{Leu-UAG}}\), tRNA\(^{\text{Thr-UGU}}\), tRNA\(^{\text{Thr-CGU}}\)—through cleavage at a single site within their anticodons. VapC5 is far less discriminating than *M. tuberculosis* tRNase toxins that target a single, relatively rare tRNA species and trigger a cascade of events that lead to surgical reprogramming of the transcriptome and proteome contributing to pathogen virulence and/or stress survival (Barth et al., 2021; Barth et al., 2019). Yet, VapC5 results in a surprisingly precise endpoint. Initiator tRNA\(^{fMet}\) is nearly universally required for mRNA start codon recognition, the cognate UCG Ser codon for tRNA\(^{\text{Ser-CGA}}\) is the most abundant of the six Ser codons, and tRNA\(^{\text{Thr-CGU}}\) services the second most abundant Thr codon. Curiously, tRNA\(^{\text{Leu-UAG}}\) services the second rarest of the six Leu codons, and tRNA\(^{\text{Thr-UGU}}\) services one of the two rarest Thr codons. Thus, depletion of the initiator tRNA and other tRNAs for abundant codons make the strongest impact.

As with nearly all bacterial type II TA systems, and all in mycobacteria, the precise environmental triggers that activate VapC5 through selective reduction in cognate antitoxin levels remains unknown. However, here we used low level VapC5 expression to mimic natural toxin activation in a strain lacking a VapBC5 TA system. First and
foremost, VapC5 expression results in a hallmark phenotype of Type II TA systems, cell growth arrest (Figure 25C). In spite of their growth arrested state, VapC5-expressing cells still manage reprogram Mab physiology at the transcriptional and translational level.

The inactivation of multiple tRNAs, including tRNA^{fMet}, would be expected to lead to widespread ribosome stalling. Ribosome availability is known to control bacterial growth rate and bacteria cannot survive below a threshold level of ribosomes (Dai & Zhu, 2020). Thus, the extensive stalling after VapC5 cleavage of tRNAs would likely activate signals to overproduce ribosomes as cells remain growth arrested (which provides survival advantages in the presence of many antibiotics) and simply try to stay alive. But this effect on ribosome synthesis permeates into the realm antibiotic efficacy as well. The overproduction of ribosomes is a clever means—essentially via high copy suppression—to ensure that cells have enough functional ribosomes while the erm(41) and other WhiB7-activated enzymes are simultaneously inactivating them. The overproduction of ribosomes should also override the efficacy of amikacin and any other ribosome targeting antibiotic used for treatment of Mab infections. In fact, Mab and M. avium mutations in the 16S rRNA region targeted by amikacin within the ribosome active center leads to amikacin resistance (Kim et al., 2021). In addition, the stockpiling of translational capacity has been previously shown to support rapid recovery of stressed Escherichia coli cells (Mori et al., 2017). Therefore, this toxin-mediated phenotypic shift is predicted to disrupt the efficacy of all first line Mab antibiotics—which either attack ribosomes or require actively growing cells—as well as support robust recovery when the antibiotic is removed (Figure 31).
Consistent with this prediction, VapC5-expressing Mab cells supported an increase in persistence against the two classes of first line antibiotics amikacin and cefoxitin. These results indicate that the killing action of amikacin and cefoxitin is weakened when VapC5 is expressed, leading to a larger pool of persisters that can seed growth once the antibiotic is removed (Figure 30, 31). Note that this effect is over and above the already well documented intrinsic resistance exhibited by Mab ATCC 19977 cells exposed to antibiotics (Hurst-Hess et al., 2017; Miranda-CasoLuengo et al., 2016; Yam et al., 2020). However, while the mechanism enlisted by VapC5 to potentiate intrinsic resistance shares some molecular features when Mab ATCC 19977 is treated with antibiotics (i.e. WhiB7 and erm(41) activation), Figure 28A illustrates clear differences in the pattern of transcriptional reprogramming by VapC5 versus antibiotic exposure. For example, VapC5 induces expression of several other WhiB family members, suggesting that proposed therapeutic approaches that exclusively target WhiB7 inactivation may not be effective for the significant portion of clinical strains harboring VapC5. None of the seven other WhiB family members were represented in RNA-seq datasets when Mab ATCC 19977 cells were treated with erythromycin or kanamycin (Miranda-CasoLuengo et al., 2016).

Although we have not yet tested whether antibiotic treatment serve as the in vivo trigger for VapC5 activation, the phenotype resulting from VapC5 expression suggests that this is the case. VapC5 expression does not completely cripple the cell but instead is restrained enough to launch a counterattack when subjected to antibiotic stress. The acquisition of a single VapBC5 TA system in many Mab clinical strains appears to represent an adaptation that instead instills a kind of controlled chaos in the cell to trigger
two desired endpoints: stringent growth arrest coupled with the surprising shift of ribosome synthesis into overdrive in an attempt to compensate for the extreme shortage of free ribosomes and endow Mab instead with the ability to evade clearing upon antibiotic treatment, most dreaded trait of this pathogen.

**Figure 31:** Mab VapC5 toxin subverts the efficacy of first line Mab antibiotics by **activating multiple resistance pathways.** The VapC5 endonuclease TA toxin cleaves six tRNAs. The reduction in availability of these tRNAs (which includes tRNA^{fMet}) triggers a spectrum of physiological changes that aid in evasion of antibiotic killing. First, the strong cell growth arrest phenotype characteristic of VapC5 should dampen the efficacy of all antibiotics that require actively growing cells (e.g. amikacin (AMK) and cefoxitin (FOX) shown in Fig. 5). Second, the overproduction of ribosomes overrides the efficacy of all antibiotics that target ribosomes while contributing to building the strategic reserve of translational capacity that helps cells recover when antibiotics are removed. Third, dramatic upregulation of WhiB7 and its targets, along with other transcription
factors such as those in the ArsR and MarR family collaborate to resist the action of multiple antibiotics, resulting in Mab survival, and clinically, as treatment failure.
CONCLUSIONS

To this date, the majority of Mtb toxin-antitoxins remain uncharacterized. In this study, we identified targets of three VapC toxins (VapC2, VapC4, and VapC21) in Mtb through expression studies. Using specialized 5’ RNA-seq, we show that VapC toxins leave behind 5’ mono-phosphate end, allowing us to identify its target. VapC2 and VapC21 toxins are phylogenetically related and exclusively cleave and inactivate tRNA^{fMet} at a single, identical site within its anticodon loop. Since tRNA^{fMet} is essential for growth and translation, we see inhibition in both of these processes upon induction of either toxin. We also identified two conserved, contiguous amino acids (WR) at the carboxy terminal of the antitoxin that seem to play an essential role in interacting and neutralizing their cognate toxins. This site can possibly act as a drug target, enabling constitutive inactivation of tRNA^{fMet} by both VapC2 and VapC21 toxins. We also showed that VapC4 toxin exclusively cleaves tRNA^{Cys} at the anticodon sequence. Once again, expression of this toxin shows growth arrest; however, we observed a decrease in translation, but it was not inhibited completely. Upon closer look at transcripts and newly synthesized proteins, we saw upregulation of genes involved in synthesis of free cysteine, which feeds into the synthesis of Mycothiol. Mycothiol serves as a major systemic protectant by maintaining cellular redox homeostasis during oxidative stress. We also observed upregulation in genes associated in dealing with copper stresses. Oxidative and copper stresses are characteristically endured by Mtb within macrophages, and we predict that VapC4 toxin plays a vital role in mitigating these stresses and survival during early stage of infection.
Alongside with Mtb being a global threat, cases of NTM infections are rapidly growing in the USA. Amongst the uprisin of NTM infections, Mab infections are the most intractable to clearing after aggressive and lengthy antimicrobial treatment regimens. We identified 22 putative TA systems within 128 clinical Mab strains. In this study, we characterized one of those TA systems, VapBC5, which is frequently present in the clinical Mab strains. Expression of VapC5 in Mab results in growth inhibition and translation reduction. We also identified that VapC5 targets six distinct tRNAs at their anticodons, a very distinctive feature from Mtb VapC toxins that typically target single tRNA. This leads to transcriptional reprogramming toward activation of multiple antibiotic resistance pathways coupled with the upregulation of ribosomal protein genes.

We speculated that VapC5 expression causes a phenotypic shift, disrupting the efficacy of all first line Mab antibiotics. Consistent with this theory, we found expression of VapC5 led to a substantial increase in persistence (aka intrinsic resistance) upon exposure to two front line antibiotics, amikacin and cefoxitin.

Understanding mechanisms that these toxin-antitoxin systems use to facilitate survival and increase in pathogenicity of these mycobacteria will help develop new antibiotics or use currently available antibiotics more effectively in treating these infections.
References


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


https://doi.org/10.1128/JB.00763-19


https://doi.org/10.1126/science.aaf4268


system: implications for the design of novel antimicrobial peptides. *Nucleic Acids Res*, 43(15), 7624-7637. [https://doi.org/10.1093/nar/gkv689](https://doi.org/10.1093/nar/gkv689)


Mori, M., Schink, S., Erickson, D. W., Gerland, U., & Hwa, T. (2017). Quantifying the benefit of a proteome reserve in fluctuating environments. *Nat Commun*, 8(1), 1225. [https://doi.org/10.1038/s41467-017-01242-8](https://doi.org/10.1038/s41467-017-01242-8)


