IDENTIFICATION AND CHARACTERIZATION OF *MYCOBACTERIUM TUBERCULOSIS* VapC50 TOXIN

By

JILLIAN NICOLE CORTESE

A thesis submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

and

The Graduate School of Biomedical Sciences

In partial fulfillment of the requirements

For the degree of

Master of Science

Graduate Program in Microbiology and Molecular Genetics

Written under the direction of

Dr. Nancy Ann Woychik

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

January 2017

ABSTRACT OF THE THESIS

Identification and Characterization of *Mycobacterium tuberculosis* VapC50 Toxin

by JILLIAN NICOLE CORTESE

Thesis Director:

Dr. Nancy Ann Woychik

Tuberculosis is a major health concern for a large portion of the world's population, despite its reputation as an ancient disease. One of the main concerns is the development of latent tuberculosis, where *Mycobacterium tuberculosis* remains dormant for months or years before being reactivated. Among all bacteria studied thus far, *M. tuberculosis* has the most toxin-antitoxin (TA) systems, with 88 discovered so far, 50 of which are VapBC systems. It is thought that these TA systems may aid the bacteria in establishing and/or maintaining *M. tuberculosis* cells in this nonreplicating persistent state characteristic of latent tuberculosis. TA systems have been shown to respond to various stressors that bacteria may encounter, including hypoxia, host immune responses, and nutrient starvation. VapBC, which is a type II TA system, is composed of an unstable antitoxin protein, VapB, that binds to the stable PIN domain-containing VapC toxin protein and blocks its activity. The goals of this M.S. research project were to isolate, characterize and identify the RNA target of one *M. tuberculosis* VapC toxin, VapC50. To this end, we first discovered that the annotated VapC50 gene was incorrect, and instead identified and cloned the correct VapC50 toxin. Expression of VapC50 in *Escherichia coli* resulted in growth inhibition, however, expression *Mycobacterium smegmatis* did not inhibit growth. Purification of recombinant VapC50 protein proved to

be a difficult task and was only successful after the third trial, when the protein was expressed at 30°C instead of 37°C. Since most of the analysis of VapC toxins published highlight their tRNA cleaving ability, purified VapC50 was used to conduct a tRNA cleavage assay with all 45 tRNAs in *M. tuberculosis*. None of the 45 tRNAs were cleaved by VapC50, suggesting that this toxin may instead have a novel physiological role in this pathogen.

Acknowledgements

This thesis would not have been possible without the help, guidance, and support of numerous individuals. First and foremost, I would like to thank Dr. Nancy Woychik for giving me the opportunity to be a part of your wonderful laboratory and providing guidance throughout my graduate studies, whether it be about courses, experiments, or my future career. The opportunities you gave me in lab have greatly enriched my understanding of science and helped me to grow and learn so much more than I would in just coursework alone. I would like to thank my committee members, Dr. Huizhou Fan and Dr. Jeffery Boyd for taking the time to read my thesis and provide comments and advice that will help me to become a better scientist in the future. Next I would like to thank Valdir C. Barth Jr. for all of his help over the course of my time in lab, patiently explaining protocols or concepts, and providing tRNAs and the modified pMC1s plasmid. Thank you to Melvili Cintron for all of your help over the course of my time in lab, showing me how to properly use various reagents or instruments, and for providing guidance for my experiment. I would like to thank Julia Greendyk for her help over the course of my time in graduate school, whether it be figuring out how to do new things in lab or sharing advice or information in courses and in lab. Next I would like to thank all of the members of the Woychik laboratory over the past year and a half: Melvili Cintron, Valdir C. Barth Jr., Kimberly Cruz, Julia Greendyk, Megan Chen, Tatsuki Miyamoto, and Michael DiCandia. Thank you to all of you for your help and support over the course of my time in lab and making it such a fun and rewarding experience. Lastly, I would like to thank my family, my friends, and my boyfriend for all of their love, support, and guidance during my time in graduate school.

Table of Contents

List of Tables

List of Figures

Page

Introduction

Tuberculosis and Latency

Tuberculosis, which is the result of a *Mycobacterium tuberculosis* infection, is an important topic of research for scientists and health care workers as the bacteria has developed resistance to current methods of treatment. In 2013, about 1.5 million people died from and around 9 million people developed a *M. tuberculosis* infection. Although there are drugs to treat the disease, the bacteria have evolved and adapted to treatments so that there is now multidrug resistant (MDR), extensively resistant (XDR), and more recently, completely drug resistant strains (Lee et al. 2015). Besides becoming increasingly harder to treat, tuberculosis is also able to develop a latent infection that can persist for years, and later on can be reactivated when the immune system becomes compromised. Added to the threat that *M. tuberculosis* poses to the world is the fact that about 90% of people infected end up developing latent tuberculosis, which is estimated to be one third of the world's population (World Health Organization Global Tuberculosis Report 2016). When in a latent state, the bacteria are able to survive inside a cluster of macrophages (Min et al. 2012), called a granuloma, in the lung. Granulomas are created by the immune system, hence the macrophages, to sequester the infection and limit the growth of *M. tuberculosis*, but granulomas actually become a survival niche for the bacteria to remain in a dormant state until they can be reactivated (Cruz and Woychik 2016).

Currently, a six month course of various antibiotics is the standard treatment for patients that test positive for tuberculosis. After a successful round of antibiotics, there is a 3-10% chance of disease relapse 12-24 months post treatment and evidence has been

discovered indicating that the causative bacteria may still be present in lung tissue long after treatment was deemed successful (Malherbe et al. 2016). During the infection state, there are some subpopulations of *M. tuberculosis* that slowly replicate and are naturally tolerant to all current anti-tuberculosis drugs, which can lead to development of viable, persistent, and non-culturable bacteria. In the case of non-culturable bacteria, testing for the presence of *M. tuberculosis* would come back negative, indicating that there are no bacteria present, but if the sample was tested for mRNA transcripts instead, results could come back positive. Since mRNA in *M. tuberculosis* has a very short half life, detection of mRNA transcripts would indicate that there are transcriptionally active bacteria present in these culture-negative samples. Results from this article indicate that even after six months of successful antibiotic treatment has been completed, evidence of potentially active *M. tuberculosis*, based on presence of mRNA transcripts, remains (Malherbe et al. 2016). Thus, the discovery of new and alternative treatment options remains an important global health priority.

Toxin-Antitoxin Systems

In order to survive in during latent tuberculosis, bacteria must be able to withstand stressors, which is reportedly a possible role for the toxin-antitoxin (TA) systems in *M. tuberculosis* (Cruz and Woychik 2016). TA systems are made of a stable toxin and an unstable antitoxin, which prevents toxicity of the toxin, and were first discovered in a plasmid of *Escherichia coli*, where it conferred stability to the host bacteria (Lee et al. 2015). These TA systems were thought to function as factors that could protect plasmids with a low copy number in bacteria from segregational loss during cell division. In the

event that the daughter cells were incorrectly segregated during cell division, one of these daughter cells will not contain the plasmid, but instead only contain the toxin and antitoxin. Since the antitoxin is unstable, it requires continuous transcription, but without the plasmid, no new antitoxin can be produced, so when the antitoxin is degraded, the toxin will be free to exert its effects and kill the cell (Arcus et al. 2005). It is only the daughter cell containing the plasmid that will be able to survive because the plasmid contains the genes necessary to produce the antitoxin needed to block the activity of the toxin (Lee and Lee 2016).

TA systems are present in almost all free-living bacteria. They are harbored in individual autoregulated operons with the antitoxin gene upstream of the toxin gene and a palindromic sequence in the promoter (Yamaguchi and Inouye 2011). The TA genes present on a plasmid are often transferred horizontally to other plasmids, allowing them to aid in the spread of virulence factors and antibiotic resistance in pathogenic bacteria (Lee and Lee 2016). Possible roles for TA systems include the formation of biofilm, multidrug tolerance, and arrest of growth under stress conditions (Lee et al. 2015), with reported interaction of some of these systems in the general stress response in bacteria, which is known to allow bacteria to survive stress from the environment and starvation (Wang et al. 2011).

Currently, there are five different functional categories of TA systems (designated Type I-V), which differ only in the way the antitoxin inhibits the toxicity of the toxin, with type II being the most common. Type I TA systems have the antitoxin acting as an antisense RNA that binds to the mRNA of the toxin, while the type II systems have a protein antitoxin that forms a stable complex with the protein toxin when stressors are not

present. In the case of Type III TA systems, antitoxin RNA binds directly to the toxin protein, while in type V systems, the protein antitoxin cleaves the toxin mRNA in order to prevent it from being synthesized (Hayes and Kędzierska 2014). Lastly, type IV TA systems have both antitoxin and toxin as proteins, but in contrast to type II, they have the same target, so instead of binding to each other, they compete for the same target (Sala et al. 2014).

VapBC Toxin-Antitoxin Systems

It is believed that pathogenic bacteria contain more TA systems than their non-pathogenic relatives, which is actually true for *M. tuberculosis* (Bertram et al. 2014). Several TA systems in *M. tuberculosis* have been reported to be found in genomic islands, indicating that they may have been acquired more recently in the bacteria's history (Ramage et al. 2009). Compared to other members of the mycobacteria family, which have only a couple of TA systems, *M. tuberculosis* has 88 systems (Lee et al. 2015), 50 of which are VapBC Type II systems (Sala et al. 2014), while the faster growing, nonpathogenic *M. smegmatis* only has 3 TA systems, one of which is a VapBC (Frampton et al. 2012). VapBC TA systems are characterized by the antitoxin, VapB, and the toxin, VapC, which has a PIN domain, where the toxin and antitoxin bind each other under normal conditions, but under stress the antitoxin separates from the toxin and is degraded by proteases. PIN domains are small proteins about 130 amino acids in length and have four conserved acidic residues, which possibly help make up the active site of the toxin (Lee et al. 2015). All VapC toxins have a PIN domain belonging to the pilT N-terminal domain family, which is composed of four conserved residues that form

an α/β/α sandwich structure. Although these toxins share a conserved PIN domain, the sequence identity between different toxins is actually surprisingly low (Lee and Lee 2016). Besides the four conserved acidic residues, all of the PIN domains in *M. tuberculosis* have a fifth serine or threonine residue that helps coordinate divalent cations, such as Mg^{2+}, at the active site (Cruz and Woychik 2016). All bacteria that contain TA systems with PIN domains, such as VapBC, appear to share a tendency for slow growth and dormancy, which is interesting considering the fact that *M. tuberculosis* cause latent infections most of the time, where they arrest cell growth, and are known to grow very slowly in a laboratory setting. This fact may help explain why *M. tuberculosis* has 50 VapBC TA systems, to help maintain slow growth during latency, while *M. smegmatis*, which is a fast grower in comparison, only has one VapBC TA system (Arcus et al. 2005). The antitoxin, VapB, inhibits the activity of the toxin through its C-terminal tail and is able to regulate the TA operon transcription though its DNA binding domain at the N-terminal (Winther and Gerdes 2012). Vap stands for virulence associated protein (Cruz and Woychik 2016), and these VapBC loci are tightly associated with pathogenicity and virulence factors in the genome, which may allow them to contribute to the pathogenic success of *M. tuberculosis* (Lee et al. 2015).

Mycobacterium tuberculosis VapC Toxins

Due to the relatively large number of VapBC TA systems in *Mycobacterium tuberculosis*, more than any other bacteria, it is believed that these systems play an important role in latency, based on experimental evidence and the assumption that bacterial genomes only contain what is essential to their survival. When stress conditions

normally encountered during the infection process were artificially stimulated, such as IFN-γ-stimulated macrophages and hypoxia, several of the VapBC TA systems were found to be induced, including VapBC3, VapBC7, VapBC11, VapBC15, 25, and VapBC47 (Sala et al. 2014). Three VapBC TA systems, VapBC49, VapBC3, and VapBC31, were also found to be up-regulated in persistent, or drug-tolerant bacteria (Sala et al. 2014). In a nutrient starved model, an elevated extracellular concentration of VapC4, VapC13, VapC5, VapC19, VapC27, VapC22, VapC39, VapC37, VapC41, VapC44, and VapC38 was discovered (Sala et al. 2014). Each of these discoveries provides more evidence for the potential role of VapC toxins in the establishment and/or maintenance of a *M. tuberculosis* latent infection (Sala et al. 2014).

Although the molecular targets of most of the VapC toxins in *M. tuberculosis* are unknown, the VapC toxins of *Leptospira interrogans* and *Shigella flexneri* are known to inhibit global translation by site-specifically cleaving initiator tRNAs in the anticodon loop (Winther et al. 2016). All of the *M. tuberculosis* VapC toxins studied to date have been endoribonucleases, but their cleavage targets vary and in contrast to other TA systems with this activity, VapC toxins appear to require both structure and sequence specificity to recognize their target. VapC4 of *M. tuberculosis* recognizes and cleaves discrete tRNAs in this bacteria at the anticodon stem loop (Cruz and Woychik 2016), while VapC20 was shown to cleave 23S rRNA at the Sarcin-Ricin loop (Bertram et al. 2014). Recently, the targets of 12 *M. tuberculosis* VapC toxins, previously unknown, were identified through deep sequencing and UV-crosslinking. Out of these 12 VapC toxins, 11 were discovered to cleave specific tRNAs, while the other VapC toxin was found to have Sarcin-Ricin loop cleavage activity. These experiments were conducted in

M. smegmatis, the standard laboratory substitute for *M. tuberculosis*, and showed that each VapC was very specific for its target tRNAs, with some VapCs targeting multiple tRNAs and some VapCs only targeting a single tRNA. It was also discovered that some VapCs share the same tRNA targets with each other, which indicates that there may be some redundancy in the genome. In terms of specific molecular targets, VapC4 cleaves $tRNA^{Cys44}$, VapC11 cleaves $tRNA^{Leu3}$, $tRNA^{Leu13}$, and $tRNA^{Glu10}$, VapC28 and VapC30 cleave both $tRNA^{Ser25}$ and $tRNA^{Ser28}$, VapC29 and VapC37 both cleave $tRNA^{Trp7}$, VapC15 and VapC23 both cleave $tRNA^{Leu3}$, VapC25 VapC33, and VapC39 all cleave $tRNA^{Trp7}$, and VapC26 was found to cleave 23S rRNA in the Sarcin-Ricin loop (Winther et al. 2016). Although it appears that 12 of the characterized VapC toxins cleave specific tRNAs and two seem to cleave 23S rRNA in the Sarcin-Ricin loop (Winther et al. 2016), the impact that these cleavages have on the growth and functionality of *M. tuberculosis* are not known. Nevertheless, there are more than 30 uncharacterized *M. tuberculosis* VapC toxins. Since there are an unusually high number of VapBC systems, a full understanding of their activity could provide crucial insights into latent tuberculosis.

Ni^{2+}-NTA Protein Purification

Although VapC toxins are natively produced in *Mycobacterium tuberculosis*, *E. coli* is utilized as part of the laboratory expression system to express and purify VapC toxins. *E. coli* is commonly chosen for recombinant protein expression because it is relatively easy to manipulate and grows very quickly in inexpensive media, compared to *M. tuberculosis*, which is a BSL3-level pathogen that is slow growing and requires containment. Since VapC toxins are foreign to *E. coli*, they must be inserted into a

plasmid with a controllable transcriptional promoter, allowing for induction of expression when needed (Schendel 1998). A plasmid containing a histidine tag, or six consecutive Histidine residues, can be manipulated so that this tag is present on the N-terminus or C-terminus of the VapC toxin that was inserted. With the presence of this His tag on one end of the gene, VapC toxins can be purified using resin composed of nickel ions (Ni^{2+}) immobilized by nitrilotriacetic acid (NTA) that has been covalently attached to the ions. Histidine has the ability to mediate the binding of proteins to the immobilized nickel ion, because as an amino acid, it can act as an electron donor on the protein's surface to reversibly bind to the metal ion. Nickel is commonly chosen as the metal ion for Histidine tags because nickel ions, along with copper ions, have the greatest affinity for binding to histidine. NTA is able to form a very stable complex with Ni^{2+} by binding to four of its coordination sites, leaving two sites available for the electron donor groups on protein surfaces to bind to. Once bound, the histidine tag can be dissociated from the Ni^{2+}-NTA resin by displacing the tag through increasing concentrations of imidazole or eluting the histidine tag by adding buffers of decreasing pH. Since the protein is bound to the histidine tag, when the bond between the tag and the resin is broken, the protein will be removed from the resin also, allowing it to be purified (Petty 1996).

Materials and Methods

Oligonucleotide Design

The DNA sequence for VapC50 and VapB50 was obtained from the TubercuList database (http://tuberculist.epfl.ch) by searching for the toxin's corresponding identification number, Rv3749c and Rv3750c, respectively, as designated by Sala et al. (2014). This sequence was manipulated using DNA Strider, where an opening reading frame map was created and the sequence was translated with six open reading frames. Initially, the VapC50 sequence seemed a bit uncharacteristic compared to other VapC toxins because of its unusually long length of 510 base pairs, while other toxins are normally between 300-400 base pairs. Also, the length of the antitoxin, VapB is usually longer than the toxin, but in this case, the toxin is supposedly 510 base pairs long and the antitoxin was 393 base pairs in length. Another unusual feature of this toxin sequence was its relative position to the antitoxin, because normally the VapC gene will begin shortly after the VapB gene ends, but in this case there is a gap of almost 100 base pairs between both genes. Based on these peculiarities, the sequence of the VapC50 toxin with the translated six open reading frames was closely examined to see if there maybe another potential toxin sequence present that may better fit the usual VapC characteristics. Upon examination, another potential toxin sequence was discovered closer to the end of VapB50, after a 12 base pair gap, with a shorter base pair length of 309. Taking this into account, this newly discovered potential gene was termed Short Putative VapC50 and the previously designated VapC50 was now termed Long Putative VapC50 (**Figure 1**).

The best way to determine which of these potential sequences is the best candidate for the real VapC50 toxin is by comparing their sequences to known toxins through the sequence alignment function on DNA Strider. VapC toxins are characterized by a distinct PIN domain, which is usually composed of strictly conserved Aspartic Acid (D) and Glutamic Acid (E) residues. With the exception of this PIN domain, VapC toxins do not usually share much sequence similarity, so when comparing the sequence alignment of two VapC toxins, it is most important to look for D and E residues that match up. In order to determine which of the two VapC toxin candidate genes—Putative Large VapC50 or the Short Putative VapC50—encodes the actual VapC50 toxin, both proteins were individually aligned with either VapC37 or VapC49, known toxins. We performed the alignments with either strict or loose mismatch and gap penalties to determine which VapC50 candidate was correct. The sequence that best matched both of the known toxins will be considered the true VapC50 toxin sequence and will be used to design the oligonucleotides for the rest of the experiment.

Oligonucleotides were designed to be used as primers to amplify either VapBC50+Promoter or VapC50 Nde I Hind III sequences and were ordered from Integrated DNA Technologies (IDT). The easiest way to access VapC50 is by first conducting a PCR reaction, followed by cloning, of VapBC50+Promoter from the template DNA of the whole *Mycobacterium tuberculosis* genome and then using this cloned DNA as a template to amplify VapC50 by itself. Therefore, it was necessary to design primers for amplifying VapBC50+Promoter with restriction enzyme sites in order to clone the VapC50 gene into certain plasmids. The VapC50 primers were designed with a N-terminus and C-terminus restriction enzyme site. These oligonucleotides

arrived in a dried form and since the desired storage concentration is 200 μM, based on the included data sheet, a certain amount of TE buffer was added to the tube and pipetted up and down to mix. At this concentration the oligonucleotides could be stored at -20°C, but a working stock must first be created by mixing 5 μL of the 200 μM oligonucleotides and 45 μL of TE buffer. Until they are needed for Polymerase Chain Reactions (PCR) to amplify specific DNA sequences, these working stocks were stored in the -20°C freezer. The easiest way to access VapC50 is by first conducting a PCR reaction, followed by cloning, of VapBC50+Promoter from the template DNA of the whole *Mycobacterium tuberculosis* genome and then using this cloned DNA as a template to amplify VapC50 by itself.

Short Putative VapC50

5' atgccgttccccgtagtactggacgcgtgcgtgctcgtgccacacccgctggttgatgccctgttgc
ggttcgctgacgagggcacctatcggcctctgtggtcggaggacatcctggccgagactcggcg
aaccatggtcgaccgcctcaacatcagcaccgaacgggctgactaccgcattcgctttcgggtac
aacgacctcatcgccgccatgaacaaccattacaaagatcgacatgtgctagctgcagcggtccg
agagcgcgcggaggtgatagtcacgacaaacctcaagcacttccctga 3'

Long Putative VapC50

5' atgccctgttgcggttcgctgacgagggcacctatcggcctctgtggtcggaggacatcc
tggccgagactcggcgaaccatggtcgaccgcctcaacatcagcaccgaacgggctgact
accgcattcgctttcgggtacaacgacctcatcgccgccatgaacaaccattacaaagat
cgacatgtgctagctgcagcggtccgagagcgcgcggaggtgatagtcacgacaaacctc
aagcacttccctgatgacgcgctaaagccctatcagatcaaagctttgcatcccgacgac
tttctcctcgatcagttggatttgtacgaagaggcaacgaaagcagtgatcctcgggatg
gtcgacgcctacatcgacccgccgttcacgccgcacagcctgctagatgcgctgggcgag
caggtcccacagttcgccgctaaggcacggcgtctgttcccgtccggatcgccattcggc
ctcggcgtcctgctcccattcgatcaatag 3'

Figure 1: Short and Long Putative VapC50 Sequences: DNA sequences of the Long

Putative VapC50 obtained from TubercuList for Rv3749c and the Short Putative VapC50 discovered on an open reading frame map created from the Long Putative VapC50 DNA sequence in DNA Strider.

Oligonucleotide Name	Sequence
Forward VapBC50 + Promoter	5' CCG GTT TCC AGC TCT GAT CAG CTG GTG C 3'
Reverse VapC50	5' AGC GCG TCA TCA GGG AAG TGC TTG AGG 3'
Forward VapC50 w/ Nde I site	5' CAT ATG CCG TTC CCC GTA GTA CTG GAC GC 3'
Reverse VapC50 w/ Hind III site	5' AAG CTT AGC GCG TCA TCA GGG AAG TGC TTG AGG 3'

Table 1: Oligonucleotide Sequences: The name and sequence for all of the oligonucleotides designed to be used as primers in the Polymerase Chain Reactions.

Cloning into pCR 2.1 Vector

Oligonucleotides for VapBC50 + promoter were amplified via PCR using Q5 Polymerase (New England Biolabs) and *M. tuberculosis* DNA (a generous gift of the Husson laboratory, Boston Children's Hospital) as the template. In order to verify that the oligonucleotides are not contaminated, two reactions were run, one with the oligonucleotides and *M. tuberculosis* DNA and one with only the oligonucleotides. Once the reaction was complete, the reactions were run on a 0.8% agarose gel at 150 volts for

about 20 minutes, or until the samples (i.e. dye front) had run half way down the gel. The gel was analyzed using an ultraviolet light camera to make sure that a DNA band of the desired length was amplified. Using the QIAquick Gel Extraction Kit (Qiagen), the DNA from the band of the correct length was cut out of the gel, melted, and extracted. Since Q5 Polymerase was used for the initial PCR reaction, in order to proceed with this DNA, 3'A overhangs must be added to the extracted DNA using Taq Polymerase and dATPs. VapBC50 + promoter DNA with the 3'A overhangs were then incubated with Expresslink T4 DNA Ligase and pCR 2.1 vector, from a TA Cloning Kit, for 15 minutes at room temperature. The ligated DNA was then heat shocked to transform into Mach T1 *E. coli* competent cells, which were plated on M9 Glucose Kanamycin plates and incubated at 37ºC overnight. Colonies that grew overnight on the plate were selected and grown in a M9 Glucose Kanamycin (50 mM) liquid media overnight at 37ºC in a rotating incubator. DNA was isolated from the liquid culture using QIAprep Spin Miniprep Kit (Qiagen) and tested for a DNA insert by conducting a test restriction digest with EcoRI HF in a 37ºC water bath for one hour. The reaction was then run on a 0.8% agarose gel at 150 volts and analyzed under ultraviolet light to visualize the DNA. Colonies of VapBC50 + promoter that were successfully cloned into pCR 2.1 (ThermoFisher) will resolve as two bands on the gel, one the size of the vector and one the size of the inserted DNA. Restriction enzyme digested colonies were sent for sequencing to Genewiz USA to confirm that the correct sequence was cloned into the pCR 2.1 vector. Cell stocks in 7% DMSO were made of Mach T1 colonies (ThermoFisher) with VapBC50+ promoter correctly cloned into pCR 2.1, using 1 mL of liquid culture and 75.2 µL of DMSO, stored in a -80ºC freezer.

Oligonucleotides for VapC50 Nde I Hind III were amplified via PCR using Taq Polymerase and using VapBC50 + promoter pCR 2.1 as the template. Once the reaction was complete, the reactions were resolved on a 0.8% agarose gel at 150 volts for about 20 minutes, or until the samples had run half way down the gel. The gel was analyzed using an ultraviolet light camera to make sure that a DNA band of the desired length was amplified. Using the QIAquick Gel Extraction Kit as described above. VapC50 Nde I Hind III DNA was incubated with Expresslink T4 DNA Ligase and pCR 2.1 vector, from a TA Cloning Kit, for 15 minutes at room temperature. The ligated DNA was then transformed into Mach T1 competent cells, which were plated on M9 Glucose Kanamycin plates and incubated at 37ºC overnight. Colonies that grew overnight on the plate were selected and grown in a M9 Glucose Kanamycin liquid media overnight at 37ºC in a rotating incubator. DNA was isolated from the liquid culture using the QIAprep Spin Miniprep Kit and tested for a DNA insert by conducting a test cut with Nde I and Hind III in a 37ºC water bath for one hour. The reaction was then run on a 0.8% Agarose Gel at 150 volts and analyzed under ultraviolet light to visualize the DNA. Colonies of VapC50 Nde I Hind III that were successfully cloned into pCR 2.1 will show two bands on the gel, one the size of the vector and one the size of the DNA inserted. Restriction enzyme digested colonies were sent for sequencing to Genewiz USA to confirm that the correct sequence was cloned into the pCR 2.1 vector. Cell stocks (7% DMSO) were made of Mach T1 colonies with VapBC50+ promoter correctly cloned into pCR 2.1 and stored in a -80ºC freezer.

Cloning into pMC1s Nde I Hind III Vector

A pMC1s vector, with antibiotic resistance to Kanamycin and anhydrotetracycline (ATc) inducible, that was mutated to have Nde I and Hind III sites was restriction enzyme digested with Nde I and Hind III and dephosphorylated with Calf Intestinal Alkaline Phophastase (CIP) before being run a 0.8% agarose Gel at 150 volts for 20 minutes and having the correct size band gel extracted. pMC1s Nde I Hind III was ligated with VapC50 Nde I Hind III using T4 DNA Ligase and incubated for one hour at room temperature. The ligated DNA was then heat shocked to transform into Mach T1 *E. coli* competent cells, which were plated on M9 Glucose Kanamycin plates and incubated at 37ºC overnight. Colonies that grew overnight on the plate were selected and grown in a M9 Glucose Kanamycin liquid media overnight at 37ºC in a rotating incubator. DNA was isolated from the liquid culture using QIAprep Spin Miniprep Kit and tested for a DNA insert by conducting a test cut with Nde I and Hind III in a 37ºC water bath for one hour. The reaction was then run on a 0.8% Agarose Gel at 150 volts and analyzed under ultraviolet light to visualize the DNA. Colonies of VapC50 Nde I Hind III that were successfully cloned into pMC1s Nde I Hind III will show two bands on the gel, one the size of the vector and one the size of the DNA inserted. Restriction enzyme digested colonies were sent for sequencing to Genewiz USA to confirm that the correct sequence was cloned into the pMC1s Nde I hind III vector. Cell stocks (7% DMSO) were made of Mach T1 colonies with VapC50 Nde I Hind III correctly cloned into pMC1s Nde I Hind III and stored in a -80ºC freezer.

Mycobacterium smegmatis Growth Curve

VapC50 Nde I Hind III pMC1s was transformed into MC155 *M. smegmatis* competent cells by electroporation of the DNA into competent cells at 2.5 kVolts, transferring into 1 mL of 7H9 media and incubated in a 37˚C rotating incubator for three hours. The transformation was plated on 7H9 plates and incubated in a 37˚C incubator for two to three days, or until large, irregular colonies appear on the plate. Colonies were selected and inoculated into 50 mL of 7H9 overnight in a 37˚C shaking incubator. The next day the 50 mL culture was split into two 25 mL cultures, one induced and one non-induced. One of the cultures was induced with 25mM ATC, 1 μL of 100μg/mL stock/10 mL, after it had its optical density of 0.1-0.2 at 600 nm. Once one of the cultures has been induced, the optical density of both cultures were checked on the spectrophotometer every hour for at least six hours, using 7H9 media as the blank. Between each time point the two cultures were kept in the 37˚C shaking incubator. The cultures were left in the incubator overnight after the six-hour time point and had their optical density checked the next day, just to check that the culture was still in the stationary phase. This protocol for developing a growth curve in *M. smegmatis* was completed four times for VapC50 Nde I Hind III pMC1s. Optical density data was plotted on a graph comparing induced and non-induced culture's growth along with the standard deviation amongst the four growth curves.

Cloning into pET28a Vector

The pET28a vector (Novagen/EMD Millipore), with antibiotic resistance to Kanamycin and IPTG inducible, utilized was enzyme digested with Nde I and Hind III and dephosphorylated with CIP before resolving on a 0.8% agarose gel at 150 volts for 20 minutes and excision and purification of the correct size band. VapC50 Nde I Hind III pCR 2.1 was restriction enzyme digested with Nde I and Hind III for one hour in a 37˚C water bath and then run on a 0.8% agarose gel for about 20 minutes at 150 volts. The correct size DNA band, as seen by ultraviolet light, was excised and gel extracted using the QIAquick Gel Extraction Kit. pET28a Nde I Hind III was ligated with VapC50 Nde I Hind III using T4 DNA Ligase and incubated for one hour at room temperature. The ligated DNA was then transformed into Mach T1 competent cells, which were plated on M9 Glucose Kanamycin plates and incubated at 37ºC overnight. Colonies that grew overnight on the plate were selected and grown in a M9 Glucose Kanamycin liquid media overnight at 37ºC in a rotating incubator. DNA was isolated from the liquid culture using the QIAprep Spin Miniprep Kit and tested for a DNA insert by restriction digestion with Nde I and Hind III in a 37ºC water bath for one hour. The reaction was then run on a 0.8% agarose gel at 150 volts and analyzed under ultraviolet light to visualize the DNA. Colonies of VapC50 NdeI Hind III that were successfully cloned into pET28a will show two bands on the gel, one the size of the vector and one the size of the DNA inserted. Restriction enzyme digested colonies were sent for sequencing to Genewiz USA to confirm that the correct sequence was cloned into the pET28a vector. Cell stocks (7% DMSO) were made of Mach T1 colonies with VapC50 Nde I Hind III correctly cloned into pET28a and stored in a -80ºC freezer.

Protein Expression

In order to express VapC50, VapC50 Nde I Hind III pET28a was transformed into BL21 *E. coli* competent cells using the heat shock transformation protocol for *E. coli* cells. After being plated on M9 Glucose Kanamycin plates and incubated in a 37˚C incubator overnight, one colony was selected and grown overnight in a M9 Glucose Kanamycin liquid media in a rotating 37˚C incubator overnight. The next day the optical density at 600 nm of the overnight culture was taken and since the doubling time of *E. coli* is about 30 minutes, the culture should be diluted for more accurate measurements. This overnight culture was then diluted to an optical density of 0.05-0.1 in 50 mL of M9 Glycerol Kanamycin media in a 250 mL Erlenmeyer flask. After incubating in a 37˚C shaker until the optical density is 0.4-0.5, the culture was split into two 25 mL cultures, one which was non-induced as a control and upon reaching an optical density of 0.5-0.4, one culture was induced by adding Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a concentration of 1 mM to induce pET28a expression. Once induced, both 250 mL Erlenmeyer flasks were returned to the 37˚C shaking incubator and removed every hour, for about five hours, to have their optical density taken. At each time point, 1 mL from the non-induced and induced cultures were spun at maximum speed in a micro-centrifuge for one minute before having the supernatant removed. Equal amounts of TE Buffer and 2x Laemmli Buffer, with 80 µL/1 mL of ß-(2-) mercaptoethanol mixed in, were added to the cultures to re-suspend the pellet, in an amount normalized to the cell density, x(µL) =optical density x 50, before being stored in a -20˚C freezer until needed.

The expressed samples were then analyzed on an 17.5% SDS-PAGE gel, in order to compare the protein expression in the non-induced and induced samples over time.

Once the gel was poured and solidified, the samples were heated to 95˚C for five minutes and then centrifuged at max speed for one minute. In the meantime, the gel was attached to the gel box and the gel box was filled with 1 x TGS buffer, with 2.5 µL of PageRuler Prestained ladder added to the first well after being heated to 95˚C also. After being centrifuged, 10 µl of the supernatant of each sample was added to separate wells in the gel and any wells still empty were filled with some 2x Laemmli Buffer, with 5% ß-(2-) mercaptoethanol mixed in. This gel was then run for about two hours at 220 volts and periodically checked to make sure the samples reached the bottom of the gel, but did not run off into the buffer. Next, the gel was removed from the gel box and the glass plates incasing the gel and placed in a Tupperware container, where it was covered with Coomassie blue stain (0.1% Coomassie blue in 50% methanol and 10% acetic acid) and placed on a shaker overnight. The next day the stain was poured out and the gel was covered with destain stock solution (in 40% methanol and 10% acetic acid) and put back on the shaker, where it was checked repeatedly and given new destain stock solution until the bands on the gel became clearly visible. After being rinsed with diH_2O, the gel was covered with 10% glycerol and left on the shaker for 10 minutes. Finally, the gel was placed between two wet pieces of cellophane and placed in an air dryer (Biorad) on set on max. When dry the excess cellophane was cut off and the dried gel stored for future reference.

When performing protein expression studies with toxins, a fresh transformation must be always used, so once the best time point for protein induction was determined from the test induction SDS-PAGE, VapC50 Nde I Hind III pET28a was transformed again into BL21 competent cells. After being plated on M9 Glucose Kanamycin plates

and incubated in a 37˚C incubator overnight, one colony was selected and grown overnight in 50 mL of M9 Glucose Kanamycin liquid media in a shaking 37˚C incubator overnight. The next day the optical density at 600 nm of the overnight culture was taken with a spectrophotometer and since the doubling time of *E. coli* is about 30 minutes, the culture should be diluted for more accurate measurements. This overnight culture was then diluted to an OD_{600} of 0.05-0.1 in 300 mL of M9 Glycerol Kanamycin media in two 1 L Erlenmeyer flasks. After incubating in a 37˚C shaker until the OD_{600} was 0.4-0.5, one culture was induced by the addition of 0.25 mM IPTG and placed back on the shaking 37˚C incubator until the designated time point. The culture was transferred into 200 mL centrifuge tubes and spun for 10 minutes at 4˚C and 1400 rpm, followed by removal of the supernatant. Pellets were left in the centrifuge tubes and stored in the -20˚C freezer until they were needed for protein purification.

Protein Purification

The cell pellets prepared and frozen in from protein expression were removed from the freezer and left at room temperature for about 15 minutes to thaw. While the cell pellets are thawing, Lysis Buffer I, Wash Buffer I-III, and Elution Buffer needed for protein purification were prepared from stock solutions. Thawed cell pellets were re-suspended in 25 mL of Lysis Buffer I by vortexing and combined into a sterile 50 mL centrifuge tube, and from this point, needed to be kept on ice, in order to prevent denaturation of the protein. Next, 5 mL of 10% Triton X-100 and 3 mL of 10 mg/mL Lysozyme were added to the re-suspended pellet and incubated on ice for 30 minutes. At this point, the cells were lysed by sonication, using an amplitude of 30% and a pulse cycle of 10 seconds on, 10 seconds off for three minutes. Lysed cells were then

transferred to a 40 mL Sepcor centrifuge tube and centrifuged for 10 minutes at 15,000 rpm and 4˚C, with a second water filled centrifuge tube calibrated to balance the lysed cells. The supernatant, or 'cleared lysate', was then transferred into a sterile 50 mL centrifuge tube, with 20 µL placed in a 1.5 mL microfuge tube containing 20 µL of 2x Laemmli buffer + 80 µL/1 mL of ß-(2-) mercaptoethanol, which was stored in the -20˚C freezer. While the cell pellet was stored at -20˚C, 1 mL of 1 M Imidazole and 19 mL of Lysis Buffer I were added to the 'cleared lysate', along with 1 mL of 50% Ni-NTA slurry in a uniform suspension, which was transferred with another 1 mL of Lysis Buffer I. This solution was mixed on a rotary shaker for one hour in a 4˚C cold room and then began a cycle of centrifuging, washing, and mixing, which was performed three times. Once the one-hour mixing was completed, the mixture was spun at 1,000 rpm for three minutes in a tabletop centrifuge, had the supernatant discarded, was given 50 mL of Wash Buffer I, pH 8, and placed on rotary shaker for 10 minutes in the 4˚C cold room, after which the cycle was completed two more times, beginning with the three-minute centrifuge step. After centrifuging the mixture but before the supernatant was discarded, 20 µL of the supernatant, for each cycle, was placed in a 1.5 mL microfuge tube containing 20 µL of 2x Laemmli buffer + 80 µL/1 mL of ß-(2-) mercaptoethanol, which was stored in the -20˚C freezer.

With the completion of three cycles of centrifuging, washing, and mixing, the rest of the purification process was completed in the 4˚C cold room, unless otherwise noted. The buffer and Ni-NTA bead mixture was pipetted into a chromatography column and the flowthrough was collected in sterile 50 mL centrifuge tube, which was then pipetted into the column, but this time the flowthrough was collected in a waste tray. Twice, 25

mL of Wash Buffer I, pH 8, was pipetted down the sides of the column slowly, in order to prevent disturbing the beads at the bottom, and was drained completely from the column into the waste tray before more liquid was added. Next, 25 mL of Wash Buffer II, pH 8, was slowly pipetted down the side of the chromatography column twice and then allowed to drain completely from the column. For Wash Buffer III, pH 8, 7.5 mL was pipetted twice down the side of the column slowly and drained completely into the waste tray. Lastly, 15 mL of Elution Buffer was pipetted down the column slowly in 2.6 mL increments, but this time the flowthrough was collected in a sterile 15 mL centrifuge tube. During the addition of each type of buffer and the flowthrough in the beginning, a small drop of the liquid draining from the column was collected in a 1.5 mL tube and had an equal amount of of 2x Laemmli buffer + 80 µL/1 mL of ß-(2-) mercaptoethanol added. Each sample was stored in the -20˚C freezer until it was ready to be analyzed on an SDS-PAGE gel.

Since the protein was now eluted from the Ni-NTA beads, 10 mL of the 15 mL eluted was added to a dialysis membrane, which was pre-soaked in 1x Dialysis Buffer for two minutes. Dialysis Buffer was diluted from 10x stock, which is composed of 12g Tris-Cl, 158.7g $NaH_2PO_4 \cdot H_2O$, and ddH_2O, bringing the volume to one liter after the pH was set to 8 with 10M NaOH. The protein was added to the membrane with a 21g needle and 10 mL syringe into one of the four slots on the membrane, which was marked for reference. In order to properly add the protein into the membrane, 10 mL of protein was taken from the 15 mL centrifuge tube, released into the membrane, and the excess air was removed with the syringe so that the protein covered the whole membrane. This membrane was then soaked in two liters of 1x Dialysis Buffer overnight on a stirrer set to

speed three in the 4˚C cold room. The next day, the membrane was placed in one liter of fresh 1x Dialysis Buffer on a stirrer set to speed three in the 4˚C cold room for 30 minutes to one hour, after which it was placed in one liter of fresh 1x Dialysis Buffer with the same conditions. Once the dialysis process was complete, the membrane was removed from the 1x Dialysis Buffer and a fresh 21g needle and 10 mL syringe were used to push a small amount of air into the membrane and then extract the 10 mL of protein through the same hole it was added through. About 4 mL of the protein was added to a 10 mL test tube with a membrane built in, the concentrator, and the rest of the protein was stored at 4˚C in the meantime. The weight of the concentrator was calibrated to the weight of a water filled test tube, both of which were placed in a 4˚C Beckman centrifuge at max speed for 20 minutes. After the 20 minutes were over, the concentrator was checked to see if there was any protein still above the membrane and if so the protein was centrifuged for another 20 minutes under the same conditions. Once there is no more protein above the membrane in the concentrator, 20 µL of the flowthrough was placed in a 1.5 mL microfuge tube containing 20 µL of 2x Laemmli buffer + 80 µL/1 mL of ß-(2-) mercaptoethanol, which was stored in the -20˚C freezer, while the rest of the flowthrough was discarded. Next, the rest of the protein in the syringe was added to the concentrator and the protein was centrifuged, with same settings, in 20 minute intervals until there was only about 250-300 µL of protein left in the membrane. A 20 µL sample of the concentrated protein was placed in a 1.5 mL microfuge tube containing 20 µL of 2x Laemmli buffer + 80 µL/1 mL of ß-(2-) mercaptoethanol, which was stored in the -20˚C freezer and the flowthrough was discarded. The rest of the protein was then either kept at 4˚C in the fridge or the Beckman centrifuge until the Bradford Assay was completely set

up.

A Bradford Assay was used to provide a relative quantification of the protein purified by comparing its optical density, as part of a special program in a nanodrop, to the optical density of known concentrations of protein. The protein used to set up the concentration standards was BGG, 1.41 mg/mL, which was initially thawed after being taken out of the -20˚C freezer. In order to set up known concentrations, a 1:5 dilution of Biorad Assay Reagent was prepared, by having 3 mL of Biorad Protein Assay Reagent and 12 mL of ddH_2O mixed together gently until the liquid turned from blue to brown, which was then filtered through Whatman paper.

Standard	BGG Stock (µL)	Bradford Assay Reagent (µL)	Concentration (mg/mL)
1	2	998	0.141
2	4	996	0.282
3	8	992	0.564
4	16	984	1.128
5	20	980	1.41

Table 2: Bradford Assay Protein Concentration Standards: The amount of BGG stock and Bradford Assay Reagent used to make each of the five standards a specific protein concentration.

Once the protein concentration standards were set up based on the table above and

20 μL of the protein was mixed with 980 μL of the Bradford reagent, they were left to sit in cuvettes for a few minutes at room temperature and the Protein Bradford program was opened on the nanodrop. Optical density measurements were taken at 595 nm for this assay and directions were followed on the nanodrop to set up the protein concentration standard curve, which included blanking with the filtered and diluted Bradford Reagent, inputting the concentrations for each standard and measuring the optical density three times, and making sure the slope of the standard curve was linear. With the standards set up, the optical density of the purified protein was measured three times and based on where this optical density lined up on the standard curve with the standards, the nanodrop computed a concentration for the purified protein. To the known amount of purified protein that is leftover from the concentrating process, molecular biology grade glycerol for was added to 15% and pipetted up and down to mix with the protein. Into RNA-free tubes, the glycerol and purified protein mixture was split into 20 μL aliquots, which were stored in the -80˚C for future use. Next, the molecular mass of the protein and the measured concentration from the nanodrop Protein Bradford program were utilized to convert the protein's concentration from mg/mL into μM, which was written on the purified protein aliquots before they were placed in the freezer. Finally, the samples taken during various points in the purification process were run on an SDS-PAGE, following the same protocol described previously, in order to make sure the desired protein was actually purified.

tRNA Cleavage Assay

Since there are 45 tRNAs in *Mycobacterium tuberculosis*, there were 90 total reactions set up, 45 reactions were the negative control, which did not receive any of the purified protein, and 45 reactions which received the purified protein. The negative control reactions were composed of 1 µL tRNA (2 pm/µL), 1 µL tRNA Cleavage Buffer, 1 µL 50 % glycerol, and 7 µL DEPC H_2O. For all reactions, the glycerol component was diluted to 50% glycerol using DEPC H_2O. Cleavage reactions that received the purified protein were composed of 1 µL tRNA (2 pm/µL), 2.76 µL of 9.05 pm/µL purified protein (totaling 25pm), 1 µL tRNA Cleavage Buffer, 1 µL 50 % glycerol, and 4.24 µL DEPC H_2O. All of the reactions were placed into RNA-free tubes with filter tips and incubated for three hours in a 37˚C incubator, after which 10 µL of RNA loading dye #2 (47.5% Formamide, 0.01% SDS, 0.01% bromophenol blue, 0.005% Xylene Cyanol, and 0.5mM EDTA) was added to each reaction before they were stored in the -20˚C freezer. tRNA Cleavage Assay reactions were analyzed on a Urea PAGE gel, which was composed of 25 mL 9% Urea PAGE Stock, 170 µL 10% ammonium persulfate, and 17 µL TEMED and solidified between glass plates in the same manner as for SDS-PAGE. The tRNA cleavage assay reactions were heated to 95˚C for five minutes and centrifuged for one minute at max speed in a tabletop centrifuge, while 1x TBE running buffer was poured over into the gel box holding the gel. GeneRuler UltraLow (Thermo Scientific) was used as the ladder for these gels and was diluted 1:2 in RNA loading dye #2 before 3 µL of the dilution was added to the the first well. For each reaction, 10 µL was added to separate wells in the gel and RNA loading dye #2 was added to any empty wells before the gel was run at 220 volts for about 90 minutes, or until the dye front reached the bottom of the

gel. The gel was then removed from the gel box and glass plates and placed in a Tupperware container, where it was soaked in 100 mL 1x TBE buffer and 10 µL SYBR Gold for 10 minutes on a shaker. A picture of the gel was then taken with UV light and printed out, at which point the gel was discarded and the liquid was poured into a waste filtration apparatus.

tRNA Number	tRNA Isotype	Designation
1	Isoleucine (Ile)	$tRNA^{Ile1}$ or $tRNA^{GAT}$
2	Alanine (Ala)	$tRNA^{Ala2}$ or $tRNA^{TGC}$
3	Leucine (Leu)	$tRNA^{Leu3}$ or $tRNA^{CAG}$
4	Tyrosine (Tyr)	$tRNA^{Tyr4}$ or $tRNA^{GTA}$
5	Threonine (Thr)	$tRNA^{Thr5}$ or $tRNA^{GGT}$
6	Methionine (Met)	$tRNA^{Met6}$ or $tRNA^{CAT}$
7	Tryptophan (Trp)	$tRNA^{Trp7}$ or $tRNA^{CCA}$
8	Glutamic Acid (Glu)	$tRNA^{Glu8}$ or $tRNA^{TTC}$
9	Aspartic Acid (Asp)	$tRNA^{Asp9}$ or $tRNA^{GTC}$
10	Phenylalanine (Phe)	$tRNA^{Phe10}$ or $tRNA^{GAA}$
11	Alanine (Ala)	$tRNA^{Ala11}$ or $tRNA^{CGC}$
12	Leucine (Leu)	$tRNA^{Leu12}$ or $tRNA^{TAA}$

13	Leucine (Leu)	$tRNA^{Leu13}$ or $tRNA^{TAG}$
14	Proline (Pro)	$tRNA^{Pro14}$ or $tRNA^{GGG}$
15	Leucine (Leu)	$tRNA^{Leu15}$ or $tRNA^{GAG}$
16	Glycine (Gly)	$tRNA^{Gly16}$ or $tRNA^{TCC}$
17	Arginine (Arg)	$tRNA^{Arg17}$ or $tRNA^{TCT}$
18	Histidine (His)	$tRNA^{His18}$ or $tRNA^{GTG}$
19	Lysine (Lys)	$tRNA^{Lys19}$ or $tRNA^{CTT}$
20	Glycine (Gly)	$tRNA^{Gly20}$ or $tRNA^{GCC}$
21	Cysteine (Cys)	$tRNA^{Cys21}$ or $tRNA^{GCA}$
22	Valine (Val)	$tRNA^{Val22}$ or $tRNA^{GAC}$
23	Proline (Pro)	$tRNA^{Pro23}$ or $tRNA^{CGG}$
24	Serine (Ser)	$tRNA^{Ser24}$ or $tRNA^{GGA}$
25	Serine (Ser)	$tRNA^{Ser25}$ or $tRNA^{TGA}$
26	Serine (Ser)	$tRNA^{Ser26}$ or $tRNA^{GCT}$
27	Arginine (Arg)	$tRNA^{Arg27}$ or $tRNA^{ACG}$
28	Serine (Ser)	$tRNA^{Ser28}$ or $tRNA^{CGA}$
29	Threonine (Thr)	$tRNA^{Thr29}$ or $tRNA^{CGT}$

30	Methionine (Met)	$tRNA^{Met30}$ or $tRNA^{CAT}$
31	Alanine (Ala)	$tRNA^{Ala31}$ or $tRNA^{GGC}$
32	Glutamine (Gln)	$tRNA^{Gln32}$ or $tRNA^{CTG}$
33	Glutamic acid (Glu)	$tRNA^{Glu33}$ or $tRNA^{CTC}$
34	Valine (Val)	$tRNA^{Val34}$ or $tRNA^{CAC}$
35	Proline (Pro)	$tRNA^{Pro35}$ or $tRNA^{TGG}$
36	Asparagine (Asn)	$tRNA^{Asn36}$ or $tRNA^{GTT}$
37	Methionine (Met)	$tRNA^{Met37}$ or $tRNA^{CAT}$
38	Valine (Val)	$tRNA^{Val38}$ or $tRNA^{TAC}$
39	Leucine (Leu)	$tRNA^{Leu39}$ or $tRNA^{CAA}$
40	Arginine (Arg)	$tRNA^{Arg40}$ or $tRNA^{CCG}$
41	Glutamine (Gln)	$tRNA^{Gln41}$ or $tRNA^{TTG}$
42	Arginine (Arg)	$tRNA^{Arg42}$ or $tRNA^{CCT}$
43	Lysine (Lys)	$tRNA^{Lys43}$ or $tRNA^{TTT}$
44	Threonine (Thr)	$tRNA^{Thr44}$ or $tRNA^{TGT}$
45	Glycine (Gly)	$tRNA^{Gly45}$ or $tRNA^{CCC}$

Table 3: *Mycobacterium tuberculosis* tRNAs: All of the 45 tRNAs in *M. tuberculosis* with their number and tRNA isotype identified as both the three letter abbreviation and the full name. Also included is the isoacceptor and anticodon designation for each of the 45 tRNAs. Numbering derived from the Lowe Laboratory tRNA Database (http://lowelab.ucsc.edu/GtRNAdb/Myco_tube_H37Rv/Myco_tube_H37Rv-structs.html)

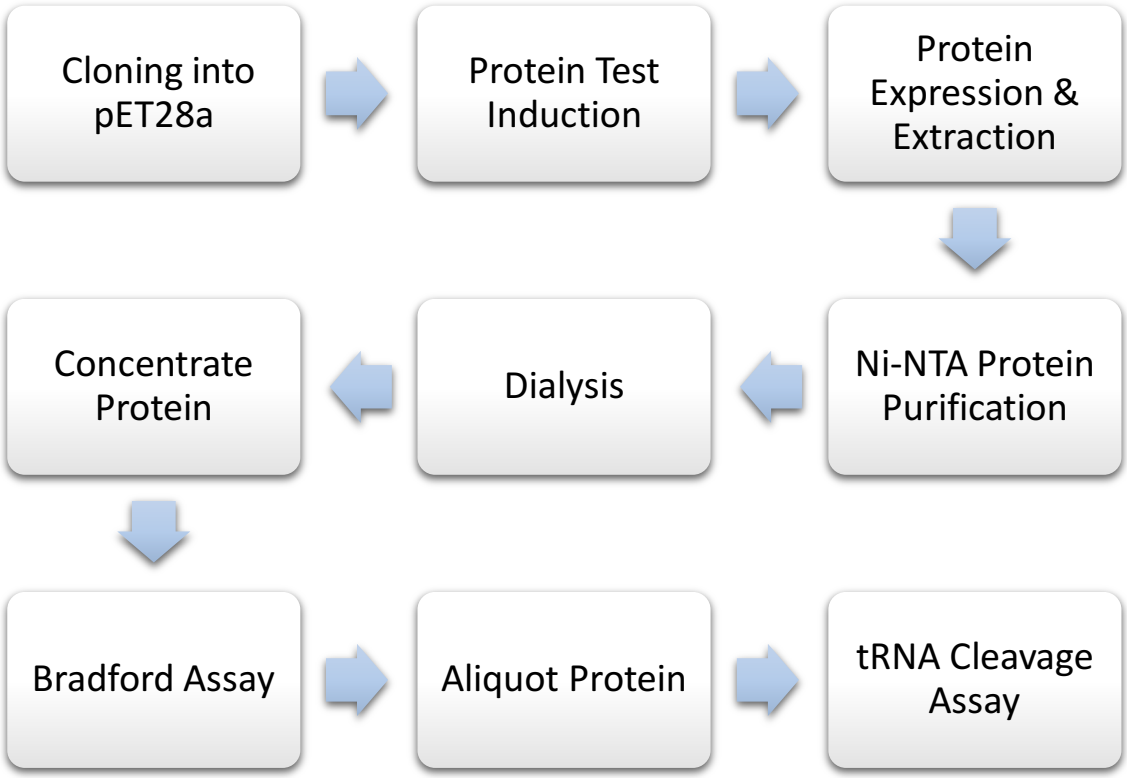


Figure 2: Protein Purification Experimental Design Schematic: A diagram depicting the stepwise progression of the experimental design in an easy to read format.

Results

Long Putative VapC50 vs. Short Putative VapC50

VapC toxins are characterized by a distinct PIN domain, which is usually composed of strictly conserved Aspartic Acid (D) and Glutamic Acid (E) residues. Besides this PIN domain, VapC toxins do not usually share much sequence similarity, so when comparing the sequence alignment of two VapC toxins, it is most important to look for D and E residues that match up. In order to determine whether the Putative Large VapC50 or the Short Putative VapC50 is the correct sequence for VapC50, both options had their sequence aligned with either VapC37 or VapC49, known toxins, and under larger or smaller mismatch and gap penalties. When the Large Putative VapC50 sequence was aligned with VapC37 under larger penalties, there only appeared to be one D and E residue that matched between both sequences (**Figure 3**). Even with a very high degree of gaps (the first 125 amino acids of Putative Large VapC50 did not align with VapC37 at all), the overall % identity was extremely low (8%). This data suggested that the Large Putative VapC50 was not related to VapC37. In order to be considered a conserved residue, these same residues should match the VapC49 sequence in the same positions with few or no gaps in the aligned sequences. Next, when the Large Putative VapC50 sequence was aligned with VapC49 under large penalties (**Figure 4**), there was a D and E residue that matched, but these residues were not in the same position as the ones that had matched to VapC37 and there was another D residue that matched for this pairing, which was not seen for VapC37. There was a very high level of gapping even with a strict gap penalty was in place and overall % identity was extremely low (6%). This data suggested that the Large Putative VapC50 was not related to VapC49 either.

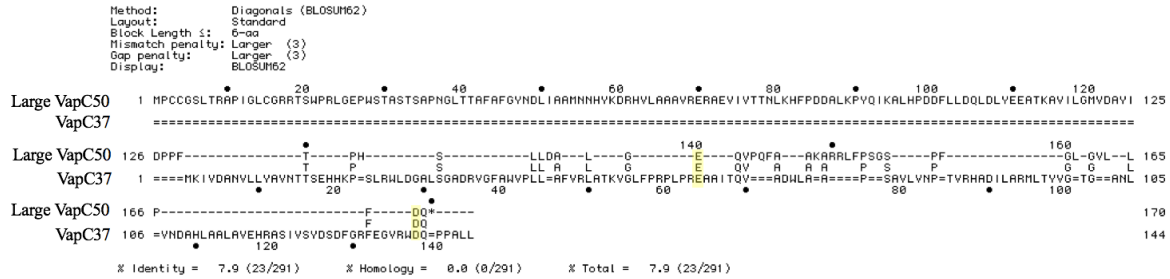


Figure 3: Sequence Alignment of Large Putative VapC50 and VapC37 with Large Penalties: Alignment of the sequences for the Large Putative VapC50 and VapC37 with large mismatch and gap penalties. Highlighted in yellow are a D or E residue that appears to match on both sequences when they were aligned.

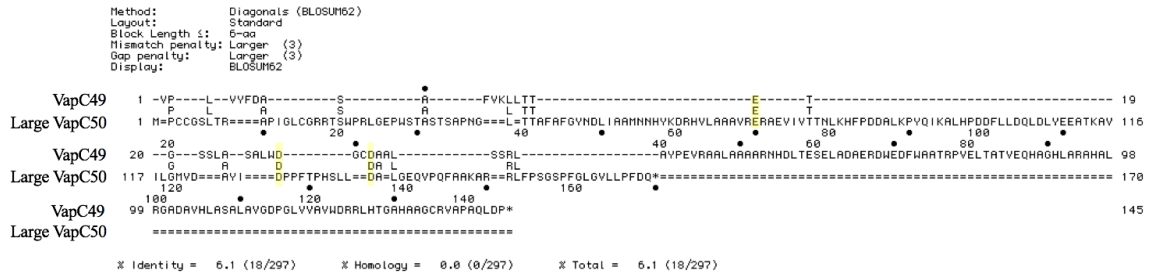


Figure 4: Sequence Alignment of Large Putative VapC50 and VapC49 with Large Penalties: Alignment of the sequences for the Large Putative VapC50 and VapC49 with large mismatch and gap penalties. Highlighted in yellow are two D residues and an E residue that appears to match on both sequences when they were aligned.

When the Large Putative VapC50 sequence was aligned with VapC37 under smaller penalties, there only appeared to be two D residues and one E residue that

matched between both sequences (**Figure 5**), but in order to be considered a conserved residue, these same residues should match the VapC49 sequence in the same positions under the same conditions. Unfortunately, when the Large Putative VapC50 sequence was aligned with VapC49 under small penalties (**Figure 6**), there was a D and E residue that matched, but these residues were not in the same position as the ones that had matched to VapC37 and there was not another D residue that matched for this pairing, which was seen for VapC37. Based on these results, the Large Putative VapC50 sequence does not appear to share a PIN domain with other VapC toxins, indicating that it may not be the correct VapC50 sequence. Ultimately, the decision for the true VapC50 sequence will depend on what the sequence alignment results for the Short Putative VapC50 sequence are, only with both sets of results can a true decision be made.

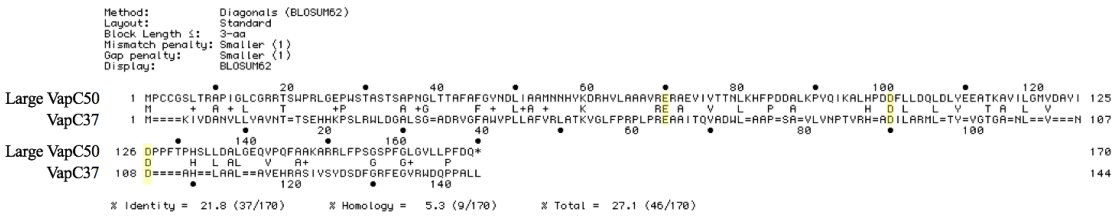


Figure 5: Sequence Alignment of Large Putative VapC50 and VapC37 with Small Penalties: Alignment of the sequences for the Large Putative VapC50 and VapC37 with small mismatch and gap penalties. Highlighted in yellow are two D residues and one E residue that appears to match on both sequences when they were aligned.

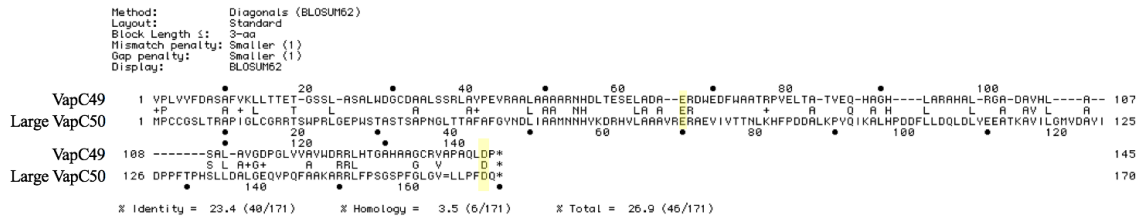


Figure 6: Sequence Alignment of Large Putative VapC50 and VapC49 with Small Penalties: Alignment of the sequences for the Large Putative VapC50 and VapC49 with small mismatch and gap penalties. Highlighted in yellow are a D and E residue that appears to match on both sequences when they were aligned.

When the Short Putative VapC50 sequence was aligned with VapC37 under larger penalties, there appeared to be two D residues and two E residues that matched between both sequences (**Figure 7**), but in order to be considered a conserved residue, these same residues should match the VapC49 sequence in the same positions under the same conditions. Unfortunately, when the Short Putative VapC50 sequence was aligned with VapC49 under large penalties (**Figure 8**), there was a D and E residue that matched, but these residues were not in the same position as the ones that had matched to VapC37 and there was not another D or E residue that matched for this pairing, which were seen for VapC37. Based on these results, the Short Putative VapC50 sequence does not appear to share a PIN domain with other VapC toxins, suggesting that it may not be the correct VapC50 sequence either.

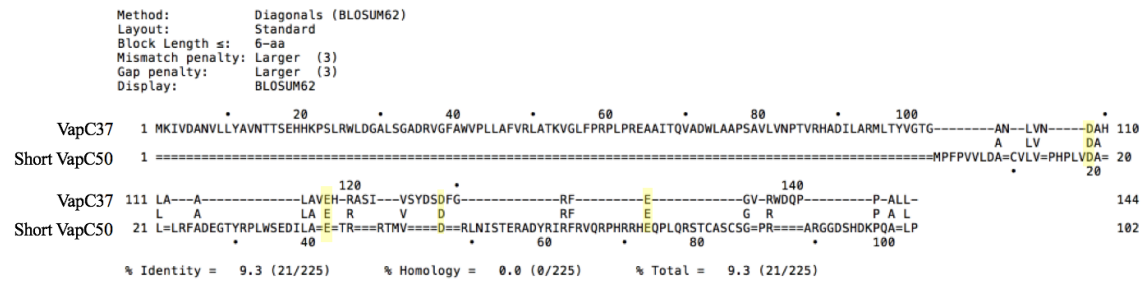


Figure 7: Sequence Alignment of Short Putative VapC50 and VapC37 with Large Penalties: Alignment of the sequences for the Short Putative VapC50 and VapC37 with large mismatch and gap penalties. All arrows point to one of the two D residues or one of the two E residues that appear to match on both sequences when they were aligned.

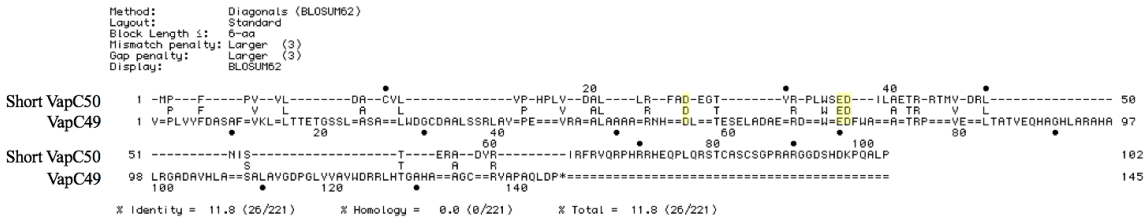


Figure 8: Sequence Alignment of Short Putative VapC50 and VapC49 with Large Penalties: Alignment of the sequences for the Short Putative VapC50 and VapC49 with large mismatch and gap penalties. Highlighted in yellow are two D residues and one E residue that appears to match on both sequences when they were aligned.

However, when the Short Putative VapC50 sequence was aligned with VapC37 under smaller penalties, there appeared to be five D residues that matched up when both sequences were aligned (**Figure 9**), but in order to be considered a conserved residue, these same residues should match the VapC49 sequence in the same positions under the

same conditions. Sequence alignment of Short Putative VapC50 and VapC49 with small penalties showed the matching of four D residues between both sequences (**Figure 10**). Surprisingly, when comparing the sequence alignment of Short Putative VapC50 with VapC37 and VapC37, there were four residues that matched both with respect to amino acid and position. These common matches, D at position 26, D at position 48, D at position 58, and S at position 83, relative to the numbering of the Short Putative VapC50 sequence, are considered to be part of the conserved PIN domain. In addition to these four residues, there are also a few other amino acids that are common in the sequence alignment of the Short Putative VapC50 with VapC37 and VapC49, but are not part of the residues that make up the PIN domain, V, A, L, and L. This is expected since VapC family members do share some sequence similarity beyond the PIN domain. In fact, the Short Putative VapC50 is clearly related to both VapC37 and VapC49 since they share an overall similarity of 33% and 29%, respectively. Therefore, based on these results, the we conclude that the Short Putative VapC50 sequence is the real VapC50 toxin sequence, and the Large Putative VapC50 was incorrectly annotated in TubercuList. From this point on, when mentioning VapC50, it is the Short Putative VapC50 that is being referred to.

Next, we PCR amplified and cloned the correct VapC50 into a TA-cloning vector. Utilizing the VapC50 sequence, oligonucleotides were designed for Forward VapBC50+Promoter, Reverse VapC50, Forward VapC50 with Nde I, and Reverse VapC50 with Hind III for future PCR reactions (**Table 1**).

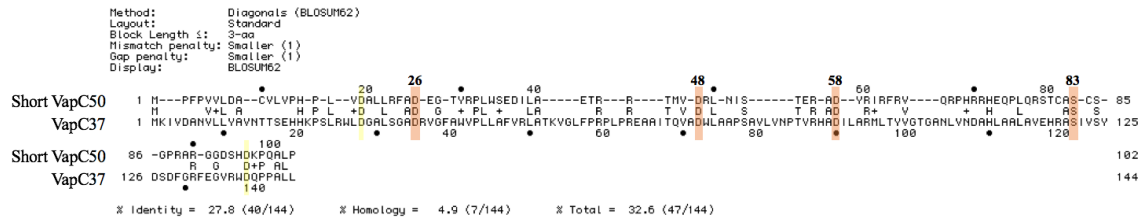


Figure 9: Sequence Alignment of Short Putative VapC50 and VapC37 with Small Penalties: Alignment of the sequences for the Short Putative VapC50 and VapC37 with small mismatch and gap penalties. Highlighted in orange are three D residues and one S residue that appears to be a part of the conserved PIN domain and is seen in the sequence alignment with VapC49 in the same position number on the Short Putative VapC50 sequence, marked in bold below the highlight. Yellow highlights indicate possible PIN domain conserved residues that are not seen in the sequence alignment with VapC49.

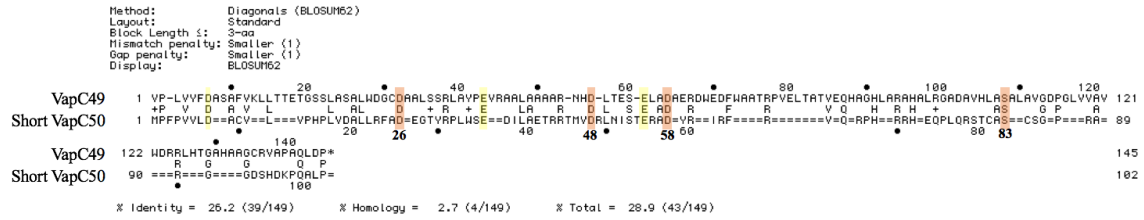


Figure 10: Sequence Alignment of Short Putative VapC50 and VapC49 with Small Penalties: Alignment of the sequences for the Short Putative VapC50 and VapC49 with small mismatch and gap penalties. Highlighted in orange are three D residues and one S residue that appears to be a part of the conserved PIN domain and is seen in the sequence alignment with VapC37 in the same position number on the Short Putative VapC50 sequence, marked in bold below the highlight. Yellow highlights indicate possible PIN domain conserved residues that are not seen in the sequence alignment with VapC37.

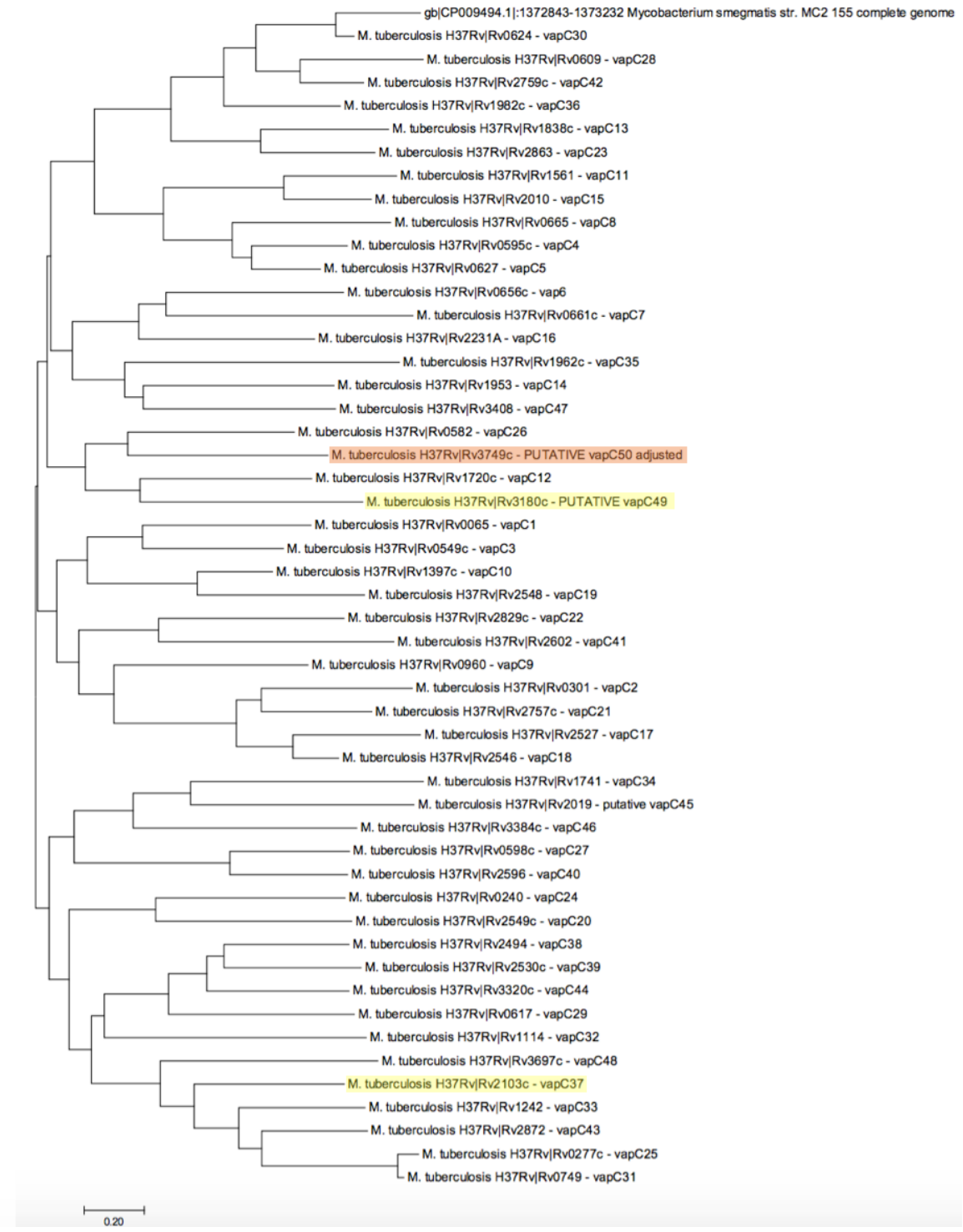
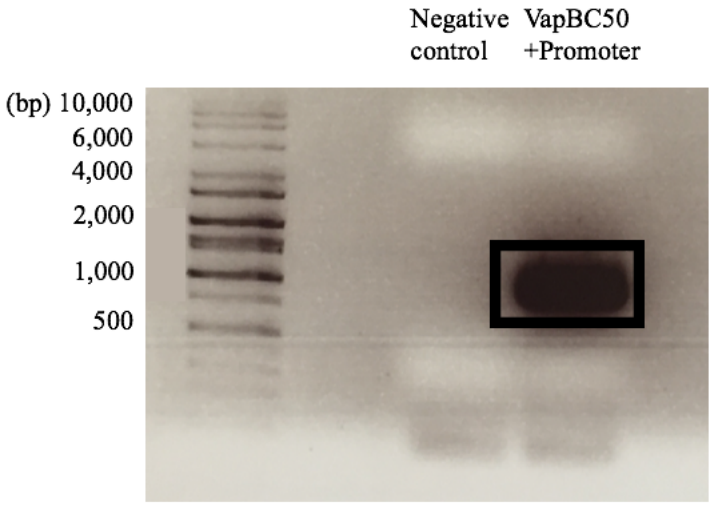


Figure 11: Evolutionary Tree of *M. tuberculosis* VapCs: This tree was made by comparing the sequence similarities of all 50 VapCs, including the correct VapC50 sequence, in *M. tuberculosis*. The orange highlight indicates the correct, or adjusted,

VapC50 sequence and the yellow highlights indicate the two VapCs utilized for the sequence alignments, VapC37 and VapC49. The phylogenic tree was created by V. Barth.

Cloning VapBC50+Promoter Into pCR 2.1

A



B

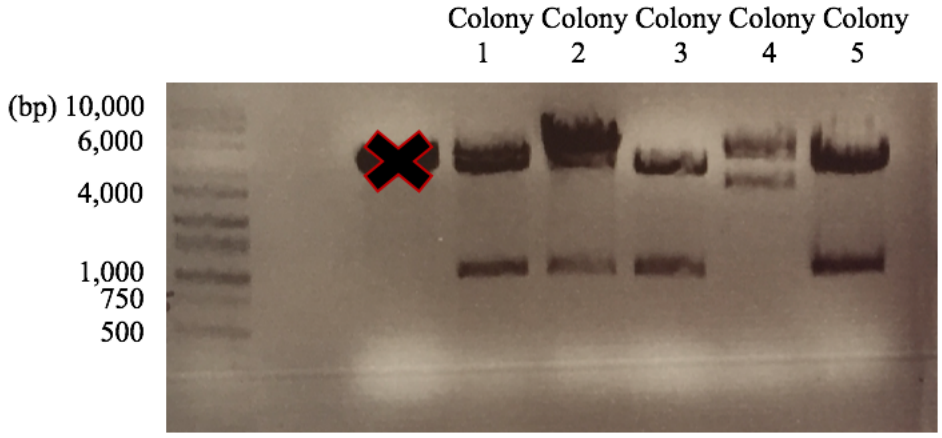


Figure 12: VapBC50+Promoter Cloning Into pCR 2.1: A. Agarose Gel analysis of VapBC50+Promoter PCR reaction and the negative control PCR reaction, which did not contain the template DNA. The DNA band with the square around it fits the correct length of VapBC50+Promoter, so it was cut out. **B.** Agarose Gel Analysis of VapBC50+Promoter pCR 2.1 Mach T1 Colonies 1-4 restriction enzyme digested with Nde I and Hind III. There was an appropriate length second band for colonies 1, 2, 3, and 5, indicating that these colonies were successfully cleaved. The band with the black and red cross placed over it was from a separate reaction that was run on this gel, but has no connection the the data being depicted.

In order to utilize the VapC50 toxin in any experiments, the appropriate sequence for this gene must be cloned into a plasmid. The easiest way to access this gene is by first conducting a PCR reaction, followed by cloning, of VapBC50+Promoter from the template DNA of the whole *M. tuberculosis* genome and then using this cloned DNA as a template to amplify VapC50. We amplified VapBC plus upstream DNA that included the promoter sequence in case we wanted to study operon regulation in the future. A band with the correct base pair length for VapBC50+Promoter was cut out of the agarose gel following PCR, suggesting that the PCR reaction amplified the correct sequence (**Figure 12 A**). The DNA from this band was extracted and cloned into pCR 2.1, followed by digestion with Nde I and Hind III, which would indicate a successful cloning if two bands showed up on the gel, one that is length of the cut plasmid minus the cloned DNA and one that is the length of the DNA being cloned. In **Figure 12 B**, all but colony 4 appear to be successfully digested by the restriction enzyme, indicated by the distinct second band of the appropriate length of the desired DNA sequence. DNA from one of

the VapBC50+Promoter pCR 2.1 Mach T1 colonies was sent for sequencing and the results indicated that this colony contained the correct sequence for VapBC50+Promoter.

Cloning VapC50 Nde I Hind III Into pCR 2.1

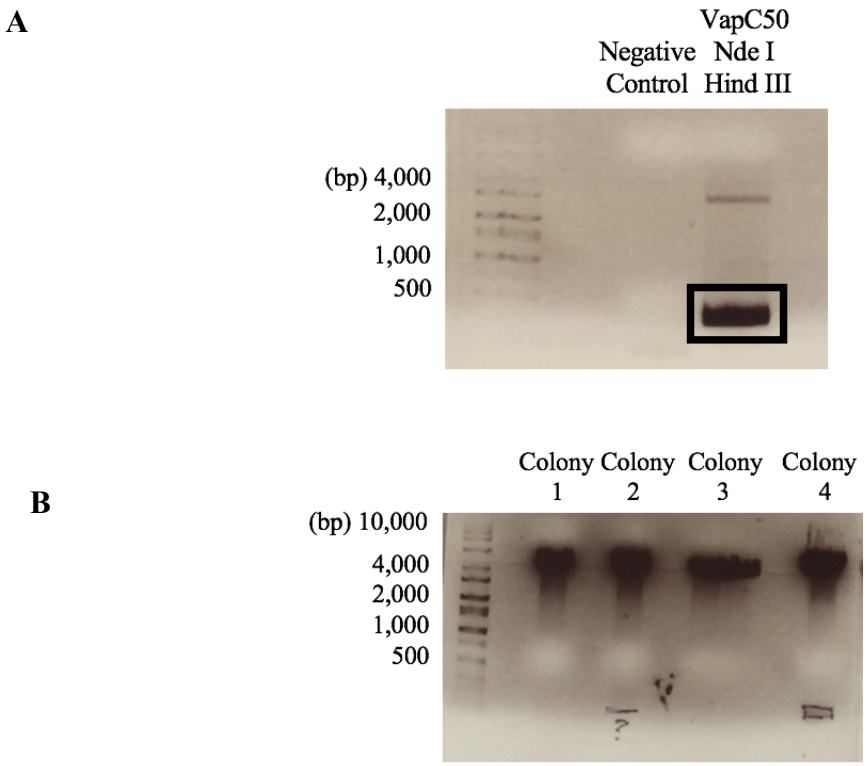


Figure 13: VapC50 Nde I Hind III Cloning Into pCR 2.1: **A.** Agarose gel analysis of VapC50 Nde I Hind III PCR reaction and the negative control PCR reaction, which did not contain the template DNA. The DNA band with the square around it fits the correct length of VapC50 Nde I Hind III, so it was cut out. **B.** Agarose Gel Analysis of VapC50 Nde I Hind III pCR 2.1 Mach T1 Colonies 1-4 restriction enzyme digested with Nde I and Hind III. Square indicates faint second band for colony 4 and question mark indicates possible second band for colony 2.

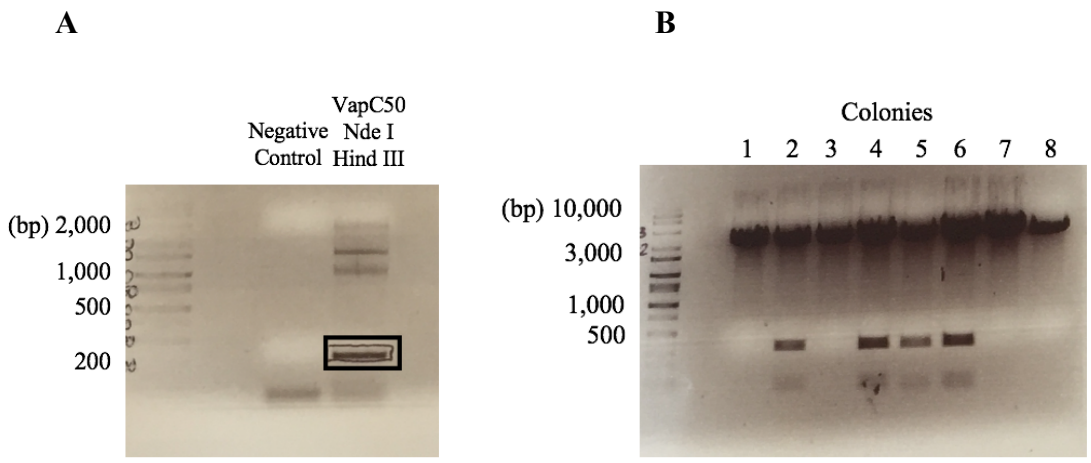


Figure 14: VapC50 Nde I Hind III Cloning Into pCR 2.1 at Lower Annealing Temperature: A. Agarose Gel analysis of VapC50 Nde I Hind III PCR reaction, with 50°C annealing temperature, and the negative control PCR reaction, which did not contain the template DNA. The DNA band with the square around it fits the correct length of VapC50 Nde I Hind III, so it was cut out. **B.** Agarose Gel Analysis of VapC50 Nde I Hind III pCR 2.1 Mach T1 Colonies 1-8 restriction enzyme digested with Nde I and Hind III. There were second bands of the correct length of VapC50 Nde I Hind III for colonies 2, 4, 5, and 5.

A band with the correct base pair length for VapC50 Nde I Hind III was excised from the agarose gel following PCR, suggesting that the PCR reaction worked (**Figure 13 A**). The DNA from this band was extracted and cloned into pCR 2.1, followed by digestion with Nde I and Hind III, which would indicate a successful cloning if two bands showed up on the gel, one that is length of the cut plasmid minus the cloned DNA and one that is the length of the DNA being cloned. In **Figure 13 B**, VapC50 Nde I Hind III pCR 2.1 Mach T1 colony 4 appears to have a second faint band, which is boxed, and

there is a possible very faint second band for colony 2, which is underlined and marked with a question mark. When VapC50 Nde I Hind III pCR 2.1 Mach T1 colony 4 was sent for sequencing, it was discovered that the VapC50 gene was correctly cloned into the plasmid, but while Hind III was present, there was no Nde I restriction site. Knowing that one of the restriction enzyme sites was missing, PCR was conducted again, but this time, instead of using the annealing temperature suggested by IDT (which did not account for the added restriction site which does not anneal to the template during early PCR steps), the annealing temperature was lowered to 50°C. This lower annealing temperature, while giving a mixture of DNA sequences as a result, would be more likely to correctly amplify the desired sequence. The agarose gel that was subsequently ran showed multiple DNA bands, as was expected, and a band the correct length of the VapC50 Nde I Hind III DNA was extracted (**Figure 14 A**). When this DNA was cloned, restriction enzyme digested with Nde I and Hind III, and run on another agarose gel, four out of the eight colonies selected showed a second band (**Figure 14 B**). VapC50 Nde I Hind III pCR 2.1 Mach T1 colonies 4 and 6 were sent for sequencing and both colonies had the correct sequence for the VapC50 gene and contained both of the restriction enzyme sites in the correct places.

VapC50 Is Not Toxic in *Mycobacterium smegmatis* MC155

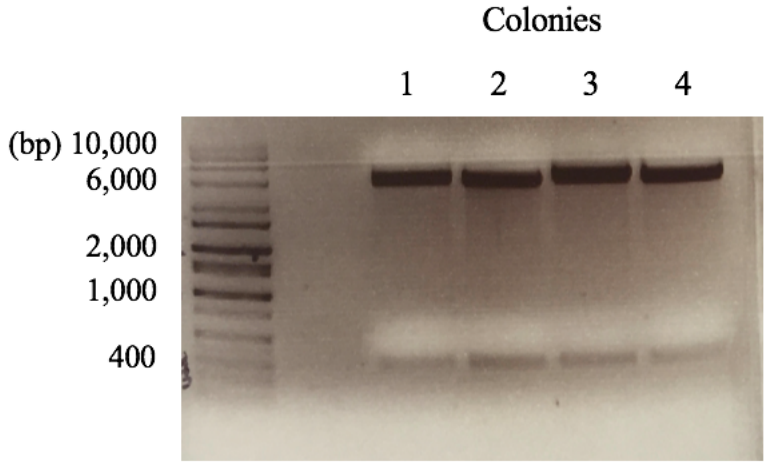


Figure 15: VapC50 Nde I Hind III Cloning Into pMC1s: Agarose Gel Analysis of VapC50 Nde I Hind III pMC1s Mach T1 Colonies 1-4 restriction enzyme digested with Nde I and Hind III. There were second bands of the correct length of VapC50 Nde I Hind III for all four colonies.

Next, we sought to clone the VapC50 toxin gene into an ATc inducible plasmid for expression in *Mycobacterium smegmatis*. The VapC50 Nde I Hind III fragment was ligated to a Nde I Hind III digested pMC1s plasmid, transformed into Mach T1, grown overnight, and four transformants were restriction enzyme digested with Nde I and Hind III. All four transformants contained the correct size DNA insert (**Figure 15**), colony 3 was sent for sequencing. The sequencing results confirmed that VapC50 Nde I Hind III was correctly cloned into the pMC1s plasmid, and DNA from this colony was then transformed into *Mycobacterium smegmatis* mc^2155, followed by the creation of four different growth curves. It is important to conduct more than one growth curve with toxins because of the possibility of suppressors masking the potential toxic effects of the

toxin. Only one growth curve needs to indicate that the toxin is toxic to the viability of *M. smegmatis* for the toxin to be considered as having a toxic effect on the growth of the bacteria. All four growth curves show a similar pattern of growth, where the Non-Induced and Induced cultures grew almost at the same rate for the whole six-hour time span. In order to get a more complete understanding of the effect of VapC50 on the growth of *M. smegmatis*, all of the time points from all four growth curves were averaged and the standard deviation for each time point of Non-Induced and Induced was calculated. All of this data was graphed in **Figure 16**, which shows that the growth pattern between Non-Induced and Induced is very similar, indicating that VapC50 does not exert a toxic effect on the growth of *M. smegmatis*. If VapC50 was toxic, there would be a noticeable difference in the growth of the Non-Induced and Induced cultures, with the Non-Induced culture following a typical growth pattern and the Induced culture growing and sharply dying off or growing very slowly from the beginning, depending on its degree of toxicity.

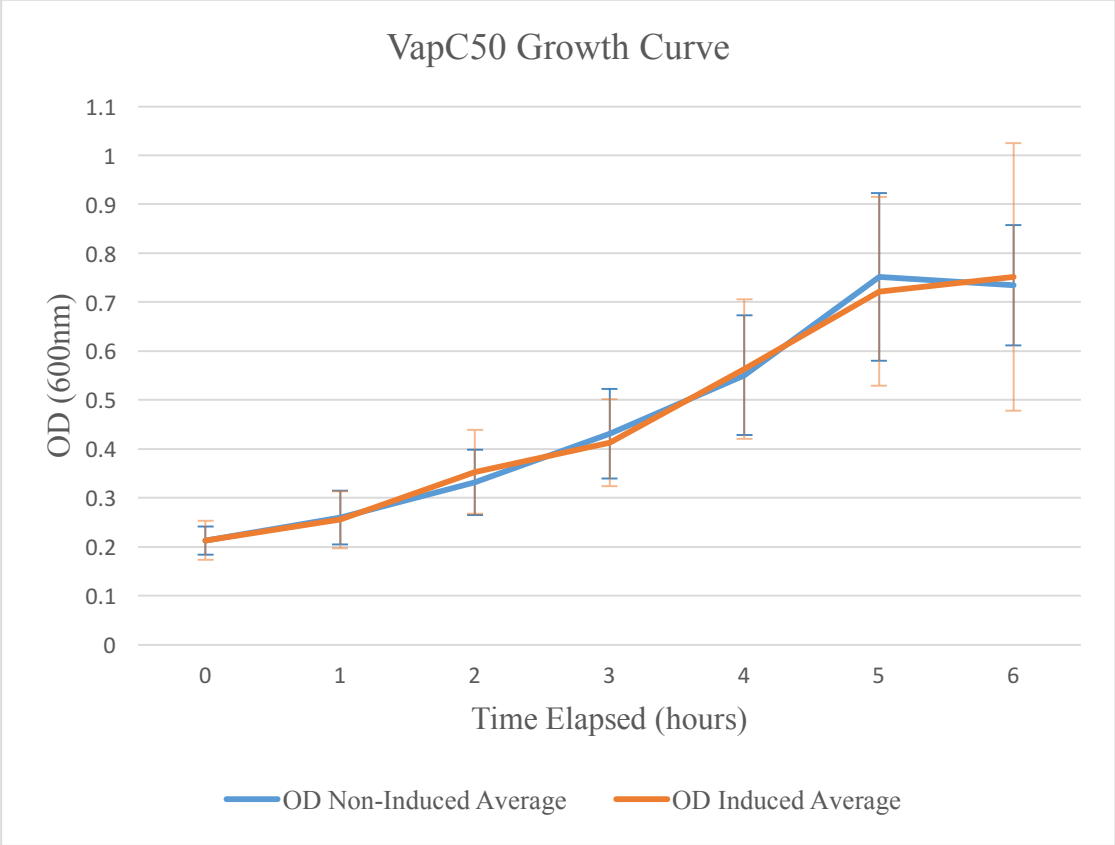


Figure 16: **VapC50 Growth Curve Average:** Growth Curve depicting the average OD values at each time point for the Non-Induced and ATC Induced VapC50 cultures and the standard deviations for each time point from each culture. Both cultures exhibit a similar growth pattern, indicating that VapC50 does not cause toxicity.

VapC50 Protein Test Induction at 37°C

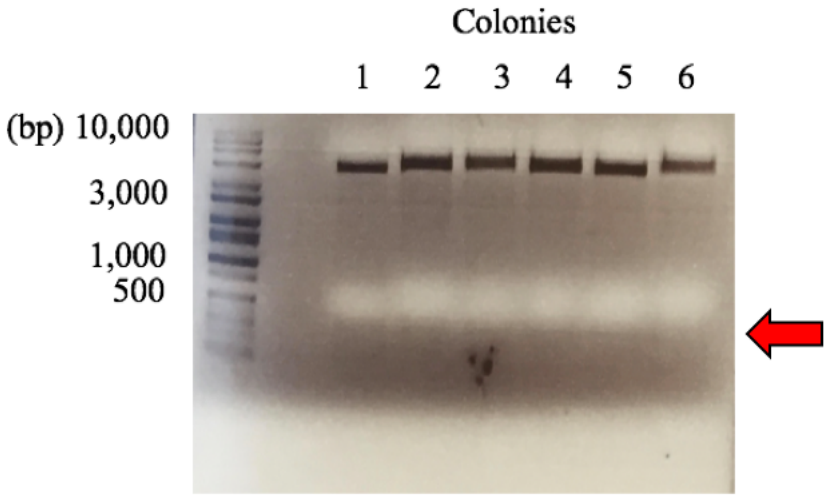


Figure 17: VapC50 Nde I Hind III Cloning Into pET28a: Agarose gel analysis of VapC50 Nde I Hind III pET28a Mach T1 Colonies 1-8 restriction enzyme digested with Nde I and Hind III. There were faint second bands of the correct length of VapC50 Nde I Hind III for all six colonies, with colony five having a slightly darker band.

VapC50 Nde I Hind III from one of the correctly sequenced colonies and cut out of pCR 2.1 was ligated with a Nde I Hind III digested pET28a plasmid, transformed into Mach T1, grown overnight, had six colonies selected, and restriction enzyme digested with Nde I and Hind III. These test cut colony DNA were run on an Agarose Gel, which showed a faint second band for all six of the colonies (red arrow), with colony 5 having the most distinct band compared to the other colonies (**Figure 17**). Colony 5 DNA was sent for sequencing and it was confirmed to have the correct VapC50 Nde I Hind III sequence in the pET28a plasmid, which is used for protein expression, due to the presence of a Histidine Tag. The OD was taken every hour for both the Non-Induced and

Induced cultures in order to make sure that the protein was being expressed, which is shown by slow and decreased growth of the Induced culture compared to the Non-Induced culture (**Figure 18**). Time points were not taken after four and a half hours due to the growth plateau for both cultures between four and four and a half hours. A growth plateau indicates that the protein has been fully expressed so there is no need to take any more samples for SDS-PAGE analysis since the time points already taken will provide enough information for the best time point of protein expression induction.

Culture samples collected at each time point for both the Non-Induced and Induced cultures were subjected to SDS-PAGE to detect the relative expression of proteins. In **Figure 19**, there was no difference in the expression profile for the Non-Induced culture across each time point, which makes sense since this culture did not receive IPTG to induce protein expression. On the other hand, there was a distinct difference in the protein expression profile for the Induced culture when compared to the Non-Induced culture, migrating at ~13.5-14 kDa, consistent with the predicted molecular mass for VapC50 of 12 kDa plus the 1 kDa Histidine tag. The abundance of this predicted VapC50 gradually increases and peaked at four and a half hours. Based on the similar protein band size and the distinct separation of growth profiles at three hours on the graph, it was determined that three hours was the best time for protein expression induction.

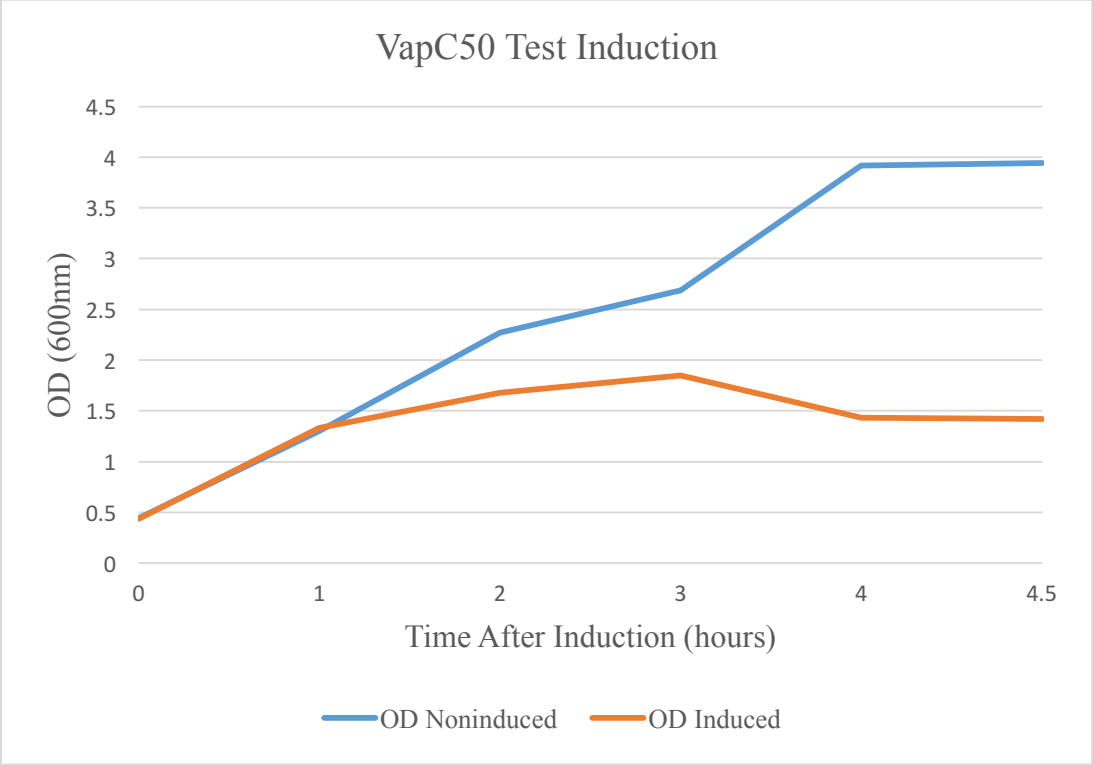


Figure 18: VapC50 Test Induction at 37°C Data: Graph depicting the test induction of VapC50 protein expression through the optical density (OD) of Non-Induced and IPTG Induced cultures over a time span of four and a half hours. The OD of the induced culture decreases as VapC50 expression increases.

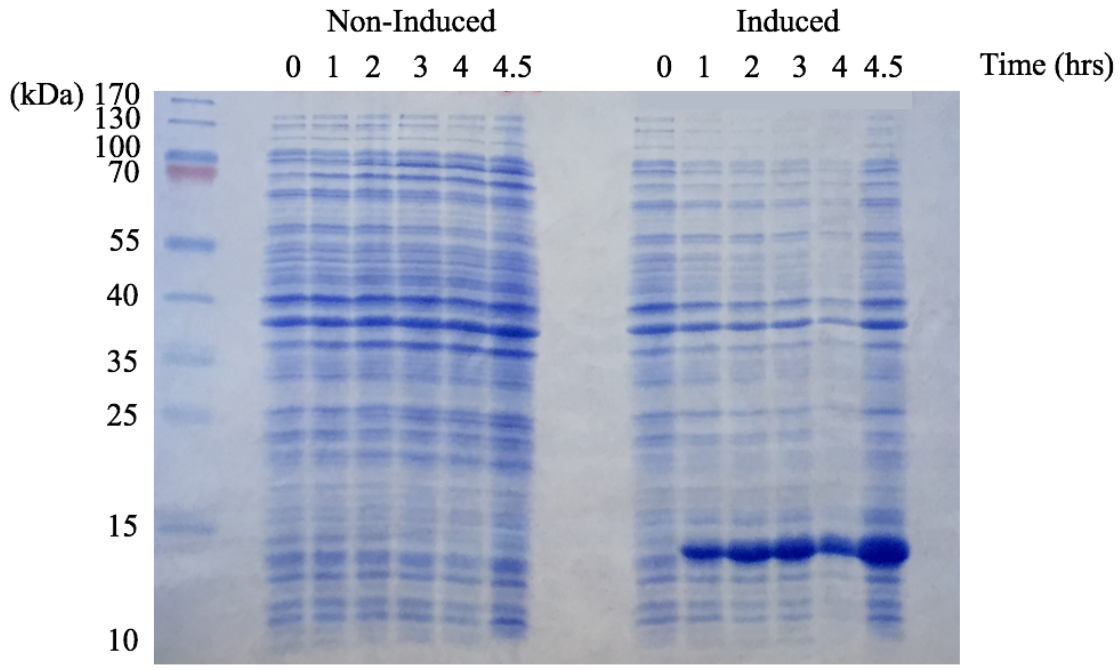


Figure 19: VapC50 Protein Test Induction at 37°C Gel: SDS-PAGE analysis of VapC50 Non-Induced and IPTG Induced protein samples over a time span of four and a half hours. Equal volumes of OD_{600}-normalized samples were loaded per lane. Details of the sample preparation and gel loading are described in the Materials and Methods section entitled "Protein Expression", pg. 18.

VapC50 Protein Purification Trial 1

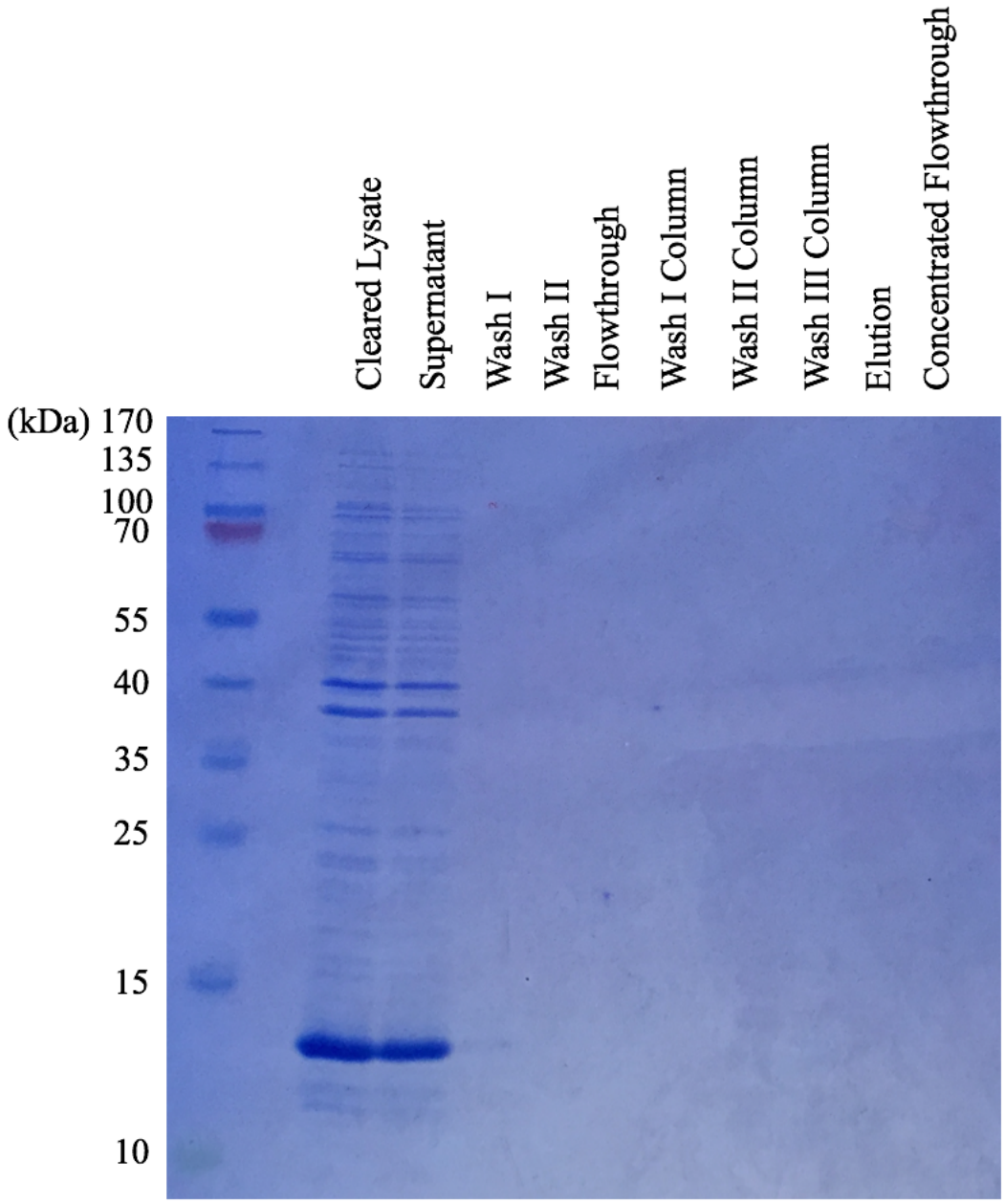


Figure 20: VapC50 Protein Purification Trial 1: SDS-PAGE analysis of samples taken during the protein purification process. If the purification was successful, we would expect the 13.5 kDa product to be visible in the Elution lane. However, VapC50 protein

bands only appear in the Cleared Lysate, Supernatant, and very faintly in the Wash I sample. Complete descriptions of the sample contents in each lane are described in the Materials and Methods section entitled "Protein Purification", pg. 20.

	Concentration (mg/mL)	Average OD
Standard 1	0.141	0.061
Standard 2	0.282	0.146
Standard 3	0.564	0.317
Standard 4	1.128	0.660
Standard 5	1.41	0.771
VapC50	0.023	0.001

Table 4: Bradford Assay for Protein Purification Trial 1: Measurements obtained for VapC50, after concentrating, based on the Bradford Assay Standards. VapC50 values are significantly below the values for the lowest standard.

VapC50 Nde I Hind III pET28a Mach T1 Colony 5 DNA was transformed into BL21 competent cells, induced with IPTG, expressed for three hours, and collected following protocol. These cell pellets were purified following the Ni-NTA Protein Purification Protocol (Materials and Methods) and samples obtained throughout the process were analyzed by SDS-PAGE to see if VapC50 was recovered. If protein was expressed and recovered, it should be present in the Cleared Lysate, Supernatant, Wash I, Wash II, Elution, and Concentrated VapC50 samples, not in the other samples due to the nature of the purification process. On the gel, VapC50 can be seen in the Cleared Lysate and Supernatant, but only very faintly in Wash I and not at all in Wash II, Elution, or

Concentrated samples (**Figure 20**). Therefore, it is apparent that VapC50 was not purified in Trial 1, especially when the Bradford Assay results are included (**Table 4**). The average OD and the concentration measured for VapC50 is very small, nowhere near the lowest standard values for the Bradford Assay. When the column from the purification process, which is supposed to catch the beads and allow the protein to elute through, was checked afterwards, there was a very small amount of beads present. With this in mind, it makes sense that the VapC50 band got very faint after the Supernatant, was not present in Wash II, and was not eluted out. A lot of beads must have been lost after centrifugation, when the supernatant is discarded, for Wash I, and more were subsequently lost for Wash II, which means there was very little or none VapC50 to be purified and measured by the Bradford Assay. Therefore, instead of pouring the supernatant out and possibly disturbing the beads enough for them to be discarded with the supernatant, for Trial 2, the supernatant will be carefully pipetted out and enough liquid above the beads will be left so that they are not disturbed.

VapC50 Protein Purification Trial 2

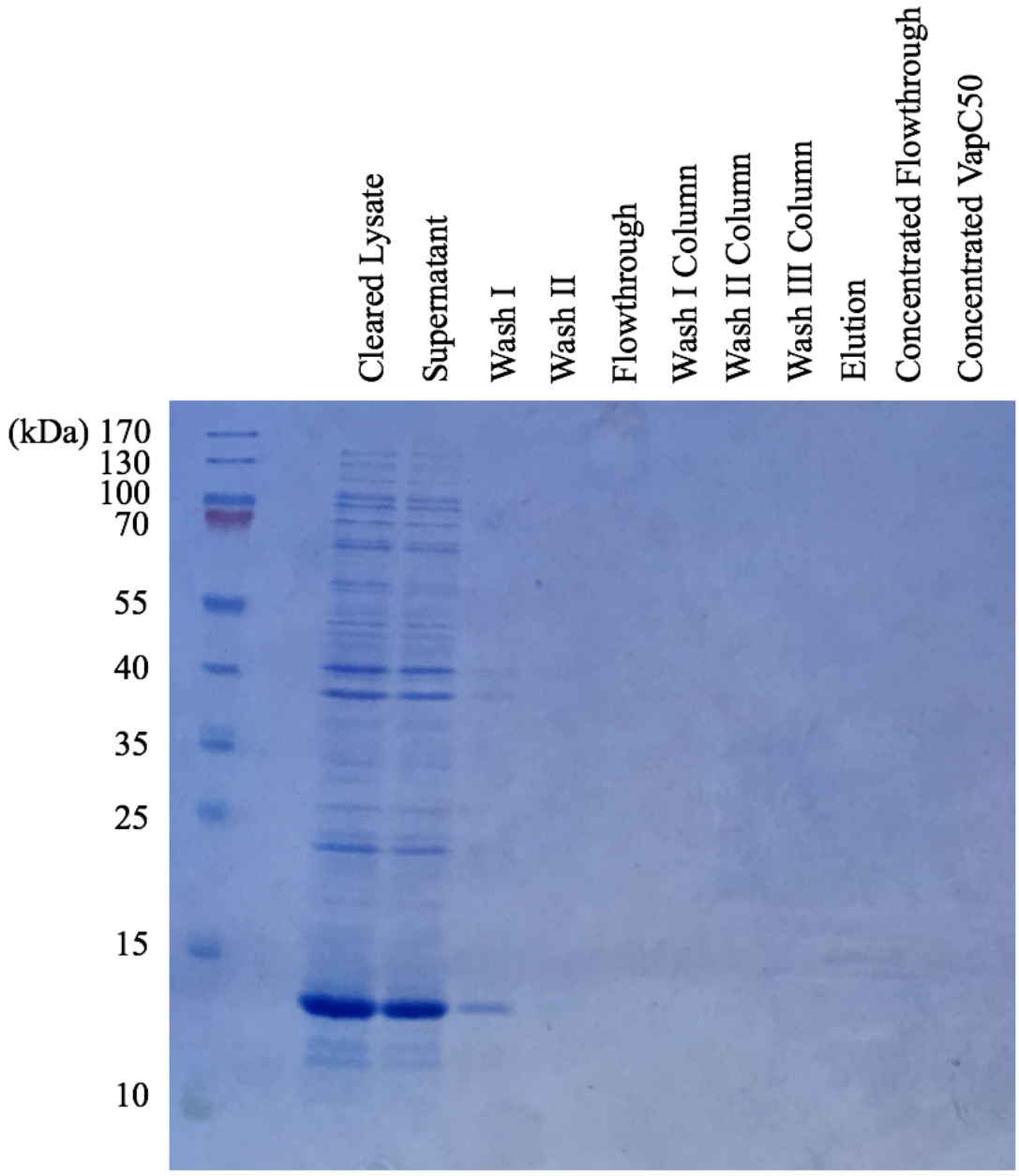


Figure 21: VapC50 Protein Purification Trial 2: SDS-PAGE analysis of samples taken during the protein purification process. VapC50 protein bands only appear in the Cleared Lysate, Supernatant, and faintly in the Wash I sample.

	Concentration (mg/mL)		Average OD	
	1	2	1	2
Standard 1	0.141	0.141	0.190	0.144
Standard 2	0.282	0.282	0.286	0.160
Standard 3	0.564	0.564	0.654	0.474
Standard 4	1.128	1.128	0.989	1.008
Standard 5	1.41	1.41	0.945	0.971
VapC50	*0.254	0.077	0.002	0.036

Table 5: Bradford Assay for Protein Purification Trial 2: Measurements obtained for VapC50, before (1) and after (2) concentrating, based on the Bradford Assay Standards. VapC50 values are way below the values for the lowest standard and the the asterisk indicates a likely machine error in measurement.

In VapC50 Protein Purification Trial 2, the sample was carefully handled during supernatant discarding and resulted in almost all of the beads that were started with being deposited into the column, which means that the chance of purifying the protein was dramatically increased. Despite this correction of an error made during Trial 1, the SDS PAGE gel looked almost the same as the previous one, the only difference was that the Wash I band was darker (**Figure 21**). However, the protein concentration was very low (**Table 5**). Logic would indicate that the concentration and OD should both go up after being concentrated, but since the results for the concentration contradict this, it was assumed that the first concentration value was a machine measuring error. Since the Bradford Assay results showed such low values that did not even reach the lowest standard value and since the SDS Page Gel did not show the presence of a purified protein band, it was decided that VapC50 was not purified in Trial 2 either.

VapC50 is Present in The Pellet

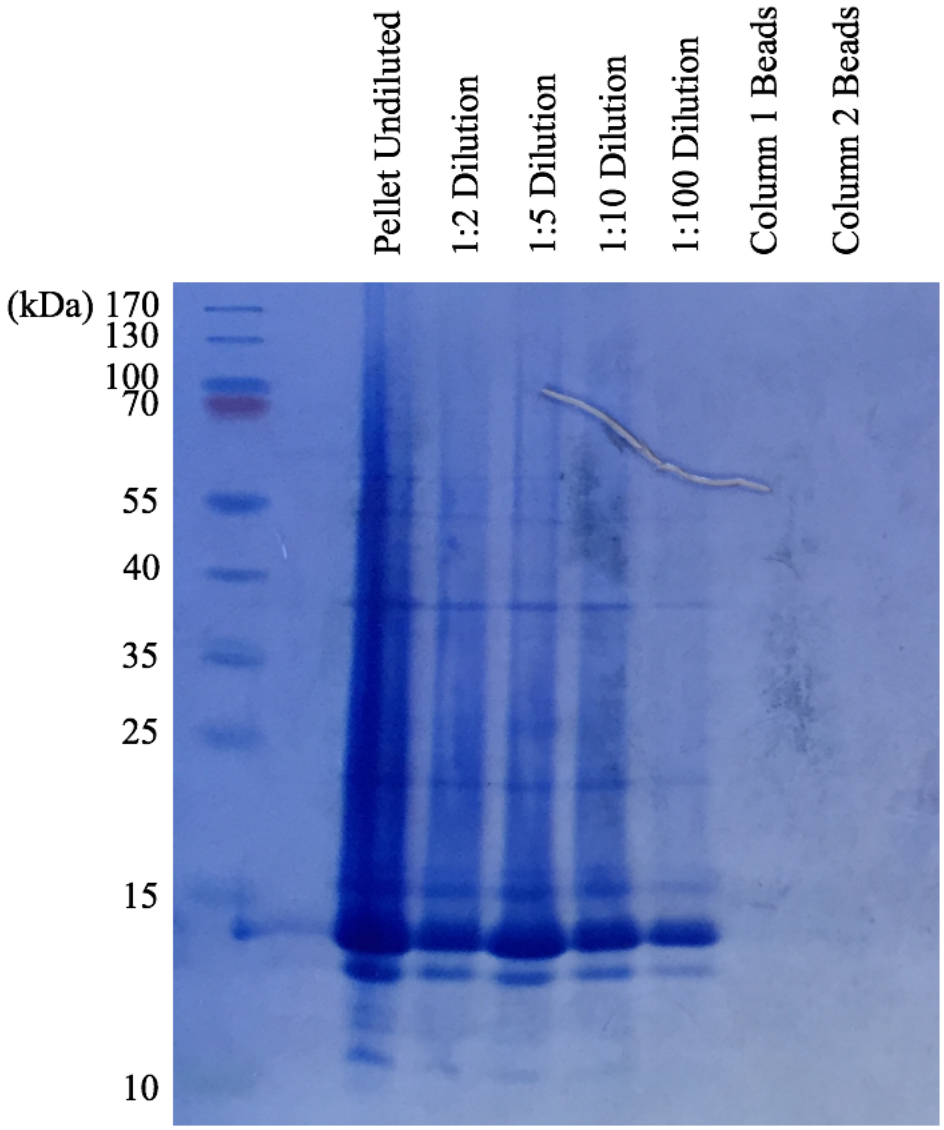


Figure 22: VapC50 Protein Purification Sample Analysis: SDS-PAGE analysis followed by Coomassie Blue staining of VapC50 Pellet Undiluted, Diluted 1:2, Diluted 1:5, Diluted 1:10, Diluted 1:100, and Column 1 and 2 Beads. VapC50 bands only showed up for all of the pellet samples, undiluted and diluted, not for the column bead samples.

After two unsuccessful attempts to purify VapC50, several other options were considered to attempt to purify this seemingly difficult protein. The first and easiest thing to check was if the protein is even in the extracted pellet and if it was present in the beads trapped in the column. A VapC50 pellet from the centrifuging step after sonication in Trial 2 was removed from the -20°C freezer and thawed at room temperature. With a sterile applicator stick, a small chunk of this thawed pellet was placed in a microfuge tube and re-suspended with 50 μL of 2x Laemmli buffer. Calculated to a volume of 50 μL a portion of undiluted re-suspended pellet was added to a new microfuge tube and diluted 1:2 with 2x Laemmli buffer. Following this dilution, the pellet was then diluted 1:5, 1:10, and 1:100 in a similar manner and then prepped like previous samples for SDS-PAGE analysis. Using a clean metal scrapper, a small clump a beads were removed from the column from Trial 1 and Trial 2, placed in a clean microfuge tube, and given an equal volume of 2x Laemmli buffer and visualized by SDS-PAGE and Coomassie staining. All of these samples contained large quantities of the ~14 kDa band corresponding to VapC50, even at a dilution of 1:100, indicating that the expressed VapC50 protein is insoluble and predominantly in the pellet, likely because it accumulated in inclusion bodies in *E. coli* cells upon overexpression (**Figure 22**). Consistent with our two unsuccessful purification attempts, samples taken from the beads in the Ni-NTA columns contained no detectable protein, which means that there was no soluble VapC50 available to bind to the beads. There was precedent for the tendency of VapC proteins to become insoluble when expressed at high levels in *E. coli*. Earlier expression and purification of other *M. tuberculosis* VapC toxins by other laboratory members demonstrated that it is common for VapC proteins to be heavily present in the pellet, with only a small portion

of that protein soluble and available for purification. However, this small amount was sufficient for the small scale experiments we use. This inability of VapC50 to bind could be attributable to not enough of the protein is being solubilized or some innate folding of the protein that is blocking the Histidine tag from being accessible to the Ni-NTA beads. In order to purify VapC50, some alterations to the usual Protein Purification protocol were going to need to be made.

VapC50 Protein Expression at 30°C

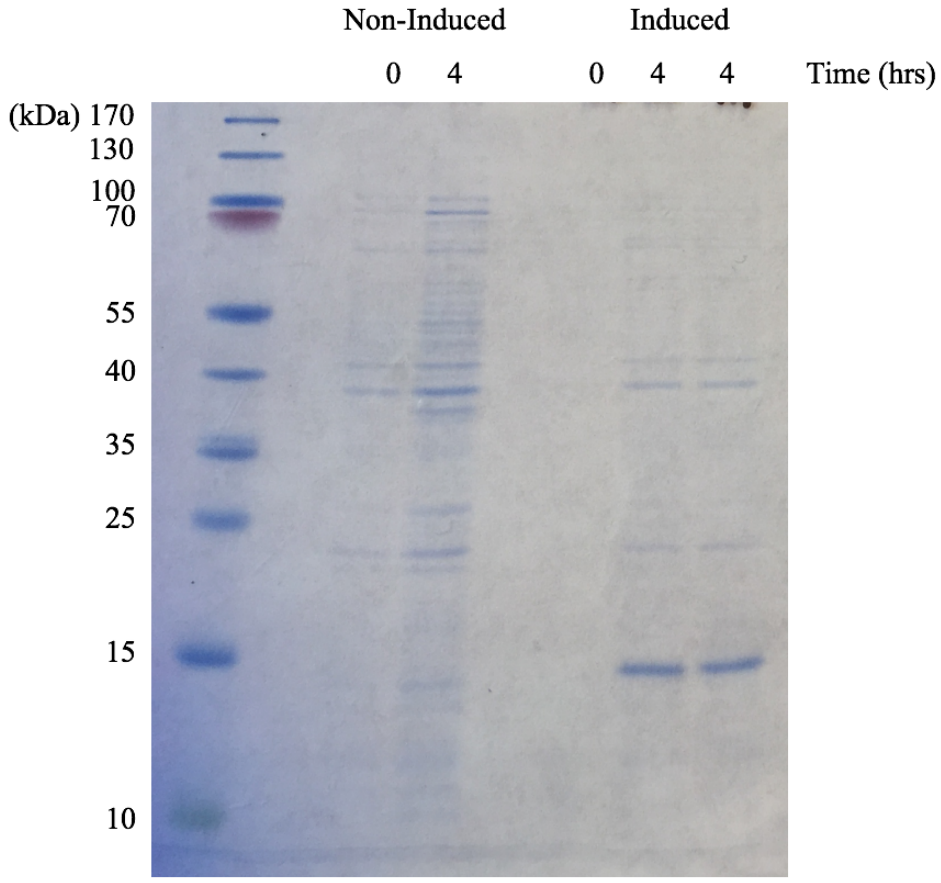


Figure 23: VapC50 Protein Expression at 30°C: SDS-PAGE analysis of VapC50

Non-Induced and IPTG Induced samples expressed at 30°C and taken at zero and four hours. There were two flasks of IPTG Induced VapC50, accounting for the two four-hour lanes on the gel.

While many options were being considered, such as designing oligonucleotides to place the Histidine tag on the C-terminus instead of N-terminus and possible having to purify under denaturing conditions, it was suggested that expressing the protein under a lower temperature might help improve the solubility. Another lab member had had better purification results with a tricky protein when they expressed the protein in culture at 30°C instead of 37°C. Since there is warm room set to 30°C and easily accessible, it was decided to express the protein under these conditions and then attempt to purify before any further measures were taken. VapC50 was expressed following the same protocol as previously described, but this time the only difference was in the temperature the experiment was conducted at. Even though it is unlikely that the protein will not be expressed at 30°C, just to make sure, a Non-Induced culture was set up too and samples were taken at time of induction and at four hours, when the protein was collected. These samples were analyzed by SDS-PAGE followed by Coomassie Blue staining and while more faint then at 37°C, the VapC50 band was present in both four hour Induced samples and not in any of the Non-Induced samples (**Figure 23**). High levels of VapC50 could result in solubility problems, so having less protein expressed might prove beneficial in increasing the solubility of the protein, allowing more protein to be purified.

VapC50 Protein Purification Trial 3

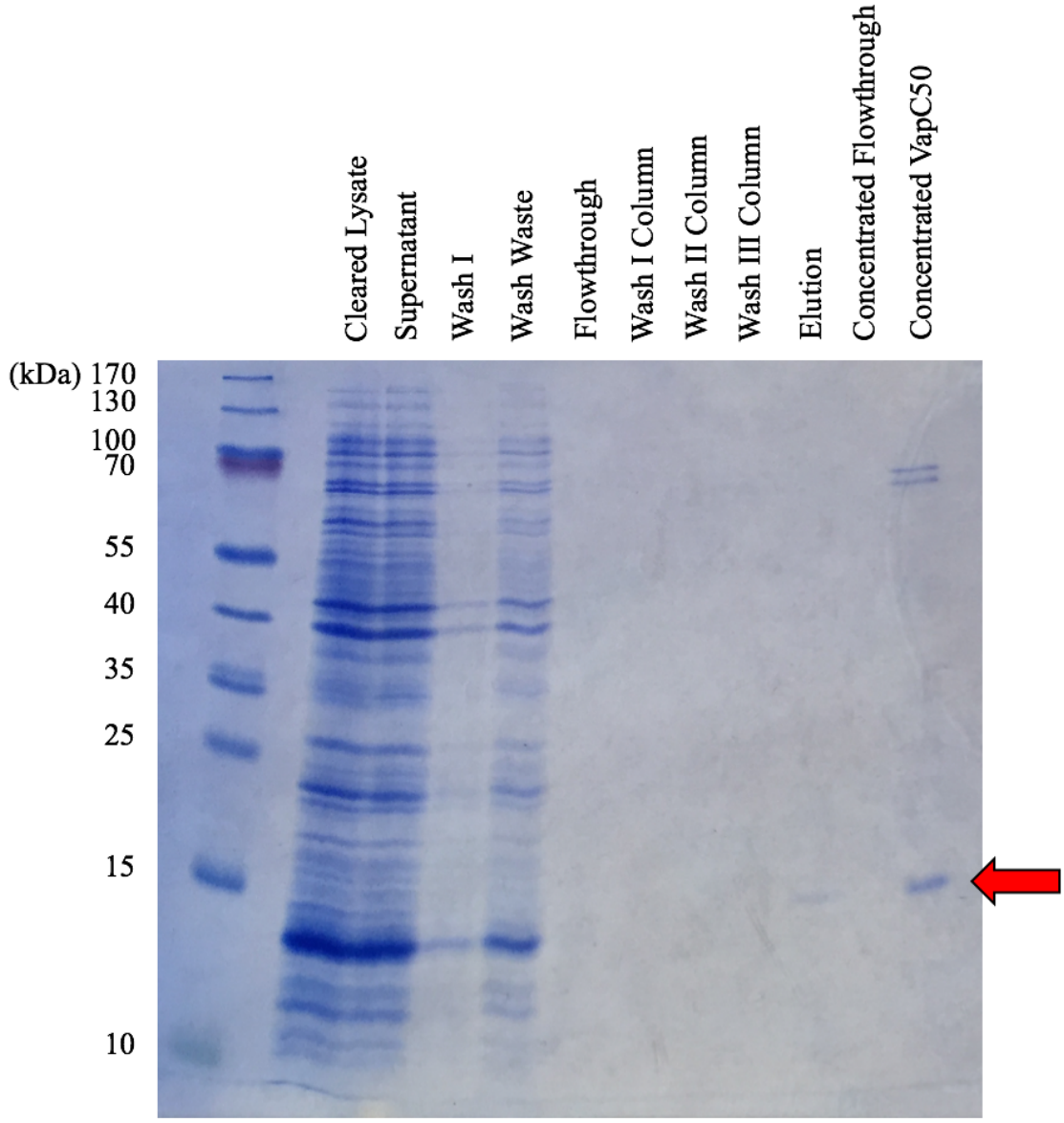


Figure 24: VapC50 Protein Purification Trial 3: SDS-PAGE analysis of samples taken during the protein purification process. VapC50 protein bands appear in the Cleared Lysate, Supernatant, Wash I, Wash I waste, faintly in Elution, and the Concentrated VapC50 samples. The two arrows point to the two bands indicating that VapC50 was

actually purified this time, due to its presence in the Elution and Concentrated VapC50 samples.

	Concentration (mg/mL)		Average OD	
	1	2	1	2
Standard 1	0.141	0.141	0.063	0.075
Standard 2	0.282	0.282	0.139	0.139
Standard 3	0.564	0.564	0.262	0.266
Standard 4	1.128	1.128	0.549	0.513
Standard 5	1.41	1.41	0.747	0.775
VapC50	0.066	0.106	0.016	0.040

Table 6: Bradford Assay for Protein Purification Trial 3: Measurements obtained for VapC50, before (1) and after (2) concentrating, based on the Bradford Assay Standards. VapC50 values are still below the values for the lowest standard, but the highest values obtained from any of the trials and much closer to Standard 1.

The pellet collected from protein expression at 30°C was purified following the same protocol detailed above, except for a few alterations. Our laboratory purchased a new and very efficient mechanical cell lysing machine, the Precellys Evolution. Unlike the other two trials, this time the protein could be seen in the elution and concentrated VapC50 samples, marked by the two arrows on the gel, indicating that the protein was actually purified (**Figure 24**). Also supportive of the actual purification of VapC50 is the Bradford Assay results, which increased after the concentration step and are the highest values obtained from all of the trials (**Table 6**). While still below the lowest standard for the Bradford Assay, these values come the closest to that standard and this is the only trial in which the VapC50 band actually showed up in the purified samples. Based on the

measured concentration of 0.106 mg/mL from the Bradford Assay and the sequence determined mass of 12 kDa + 1 kDa tag, the actual concentration of purified VapC50 was calculated to be 10 pM/μL. While still smaller than the desired yield, there is still enough protein to work with, so Protein Purification Trial 3 was deemed successful.

VapC50 Does Not Cleave Any of The 45 *Mycobacterium tuberculosis* tRNAs

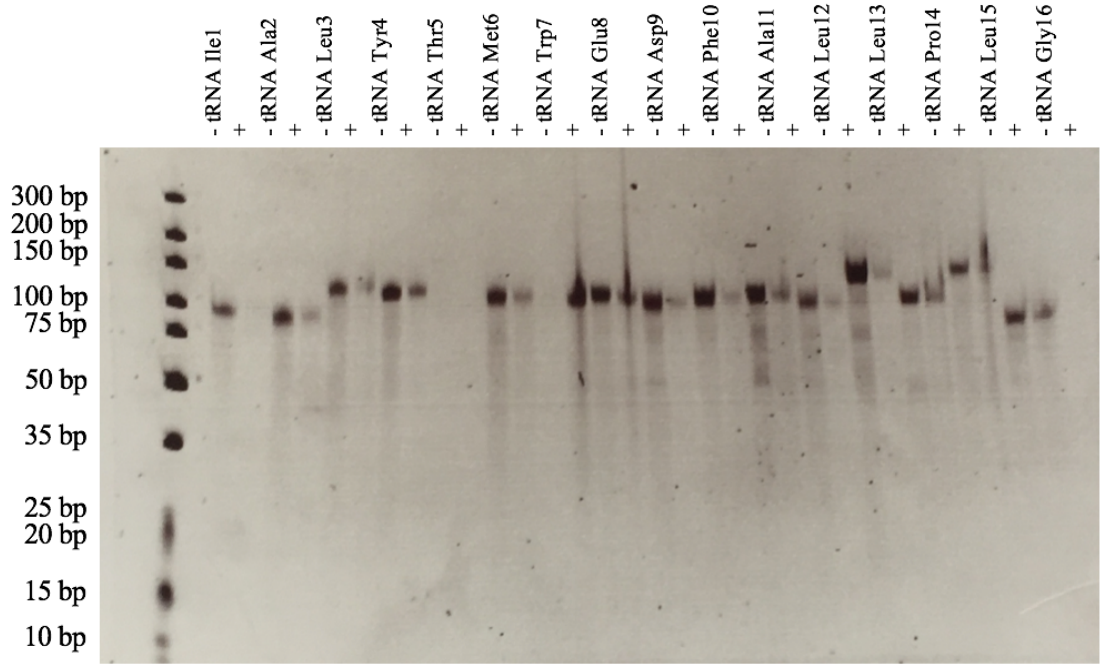


Figure 25: tRNA Cleavage Assay of tRNAs 1-16: Urea PAGE Gel analysis of the tRNA Cleavage Assay of *M. tuberculosis* tRNAs 1-16. There does not appear to be any toxin-specific cleavage bands for the reactions with VapC50, indicating that this toxin is unlikely to be cleaving tRNAs 1-16.

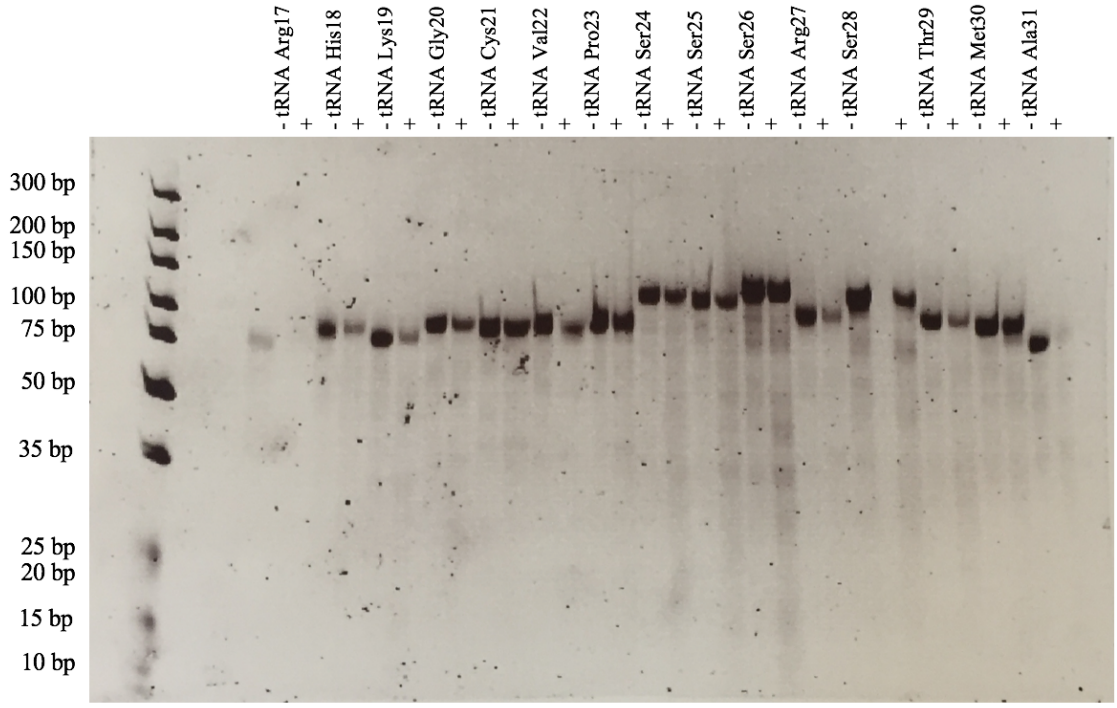


Figure 26: tRNA Cleavage Assay of tRNAs 17-31: Urea Page Gel analysis of the tRNA Cleavage Assay of *Mycobacterium tuberculosis* tRNAs 17-31. There does not appear to be any toxin-specific cleavage bands for the reactions with VapC50, indicating that this toxin is unlikely to be cleaving tRNAs 17-31.

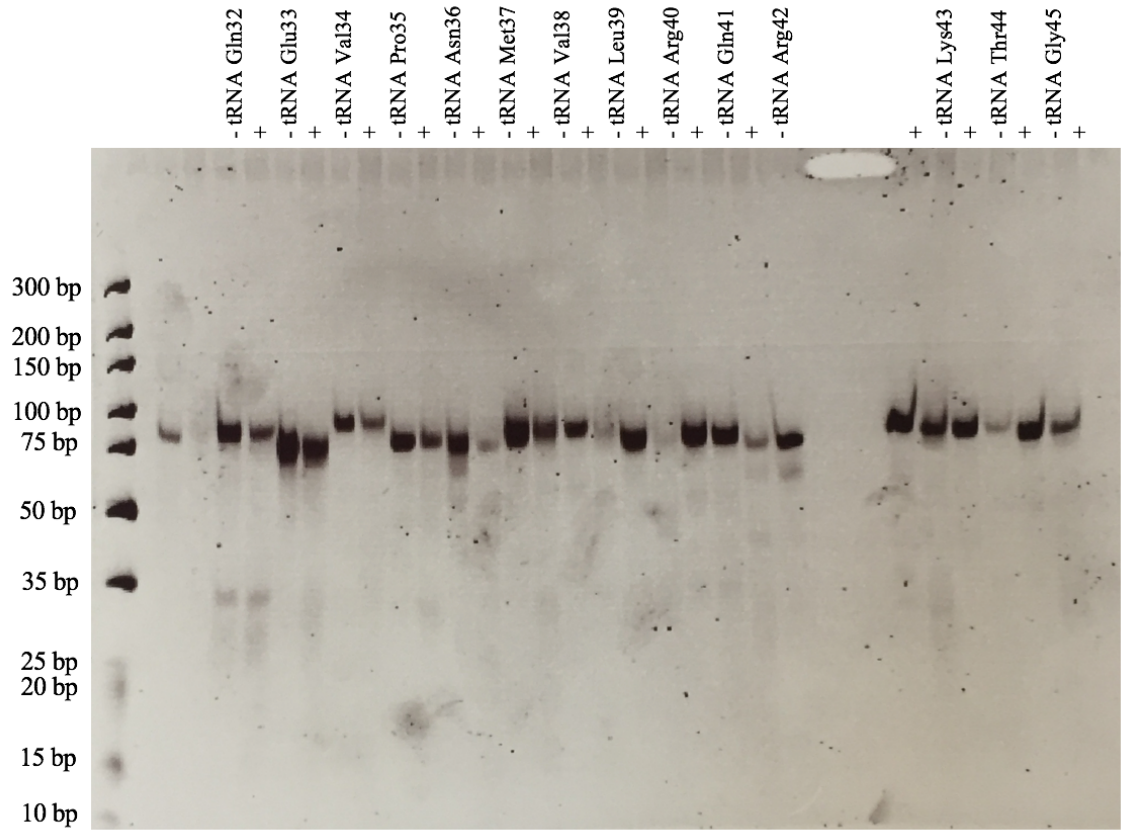


Figure 27: tRNA Cleavage Assay of tRNAs 32-45: Urea Page Gel analysis of the tRNA Cleavage Assay of *Mycobacterium tuberculosis* tRNAs 32-45. There does not appear to be any toxin-specific cleavage bands for the reactions with VapC50, indicating that this toxin is unlikely to be cleaving tRNAs 32-45. The band closest to the ladder is due to a loading error and the two small bands seen for tRNA 32 Gln are present for both the negative control and the VapC50 reaction, so these bands are not from tRNA cleavage.

Since a good portion of the published data on VapC toxins has implicated their role in preferentially cleaving specific tRNAs in of *M. tuberculosis*, conducting a tRNA cleavage assay with the newly purified VapC50 was considered a logical next step. Negative Control reactions, which did not contain the toxin, and reactions with VapC50

for each of the 45 tRNAs in of *M. tuberculosis* were run on a Urea PAGE gel to check for possible VapC50 induced tRNA cleavage. If there were cleavage of a tRNA by VapC50, there would be a distinct small base pair length second band in that lane under the band for the tRNA. This second band would only be present in the reactions conducted with VapC50 present and not in the negative control reactions, because there would be nothing present to cleave the tRNA. Out of all of the Urea PAGE gels, it is only in **Figure 27** that a possible distinct small base pair length second band is seen for the $tRNA^{Gln32}$. Unfortunately, this band is not a cleavage product since it is present in both the negative control reaction and the VapC50 inclusive reaction for this tRNA. For tRNAs 1-16 (**Figure 25**), some of the tRNA bands are very light or do not even show up at all, such as for tRNA 5 Threonine, which is probably due to human error, loading error, or not enough tRNA added to the reaction. For the most part the tRNAs in **Figure 26** all appear on the gel, but again there are no secondary bands present, which together with the other gels indicates that VapC50 does not cleave any tRNAs in this experiment. It is possible that there is no cleavage appearing on the gel because the yield of purified VapC50 was low, but this would not make sense since enough VapC50 was added to the appropriate reactions so that there was 25pM present, which is more than enough protein to induce cleavage based on previous experiments. Therefore, until further experiments are conducted, VapC50 does not appear to cleave any of the 45 tRNAs in *M. tuberculosis*.

Discussion

The discovery that the sequence believed to be VapC50 was actually not VapC50 was an unexpected finding. This annotation error was confirmed because the incorrectly annotated VapC50 gene/protein showed several irregularities: it was in the wrong genome position (too distant from the antitoxin gene), was much larger than other *M. tuberculosis* VapC toxins and lacked the PIN domain essential for catalysis of the RNA cleavage reaction. Unless one where to closely exam the sequence and analyze, one might not realize that the supposedly correct sequence is in fact wrong, a mistake that could lead to experiments with the wrong. It is important to learn from this discovery that information presented in journals and by other scientists should not automatically be assumed to be correct, since mistakes happen and sometimes details are missed or not closely examined. Keeping this in mind, it was important to present all the pieces of the experiment, no matter how simple they may seem, and carefully examine them to make sure that every part was conducted as correctly as possible and that the results are presented accurately.

Despite the fact that all four of the VapC50 growth curves in *Mycobacterium smegmatis* indicated that VapC50 is not toxic, meaning it does not cause growth arrest, it is important to remember that *M. smegmatis* is not the natural bacterium for this toxin, but is as close as one can get without utilizing a BSL 3 laboratory. The growth curves were conducted in *M. smegmatis* and not in *E. coli* because as a member of the same *Mycobacterium* family, *M. smegmatis* is evolutionarily closer to *Mycobacterium tuberculosis* than *E. coli* is. A recent article by Winther et al. (2016) mentioned that out of 12 different VapCs tested, the majority of them showed weak or no inhibition of

growth when grown in *M. smegmatis*. Although a few of these non-toxic VapCs showed no protein expression 60 minutes after induction, the majority of the VapCs tested did show a detectable level of protein expression after 60 minutes. It is possible that the functions of these expressed VapCs are not involved in inhibiting cell growth or that the VapC proteins are not functional in *M. smegmatis*, due to the absence of their cellular target (Winther et al. 2016). This notion is very interesting considering that in *M. smegmatis* VapC50 showed no growth arrest, but when the protein was expressed, there was a detectable level of protein 60 minutes after induction, coinciding with the growth inhibition. Thus, VapC50 appears to cause toxicity in *E. coli*. Therefore, it is possible that although VapC50 does not exhibit a toxic effect on the growth of *M. smegmatis*, as with *E. coli*, it may also inhibit growth of *M. tuberculosis*.

The major road block toward characterization of the RNA target(s) of VapC50 was the purification of recombinant toxin, which proved to be tricky and arduous. Thankfully it only took three trials to purify the protein and obtain enough for tRNA cleavage analysis. We believe the success of the third expression/purification experiment was due to the expression of the protein at a lower temperature, 30°C instead of 37°C.

Since the majority of functional analyses on VapC toxins published thus far have indicated the role of VapC in selectively cleaving specific tRNAs in *Mycobacterium tuberculosis*, assaying for tRNA cleavage activity was the next logical step to analyze the purified VapC50. Based on the results of the tRNA cleavage assays, it appears that VapC50 does not cleave any of the 45 tRNAs in *M. tuberculosis*. Although this is true it is not known if a second trial of this protocol would yield the same or different results, but it is very possible that VapC50 might have another role to play. The only other

published function of VapCs in *M. tuberculosis* is cleaving the 23S rRNA in the Sarcin-Ricin loop. However, perturbation of this highly evolutionarily conserved, functionally critical rRNA domain is always lethal. Since we did not observe toxicity when VapC50 was expressed in *M. smegmatis*, we can safely conclude that this toxin does not target the Sarcin-Ricin loop. It is also possible that VapC50 might have a yet to be discovered function—perhaps cleaving mRNAs at specific secondary structures—since most VapC toxins are know to require both a specific consensus sequence within a precise structural context. Alternate functions for VapC toxins beyond tRNA and Sarcin-Ricin loop cleavage is plausible since there are over 30 VapCs in *M. tuberculosis* that have not yet had their function identified. The precise role of VapC50 will be determined in follow-on experiments in the Woychik laboratory. Knowledge of the functions of all 50 VapCs in *M. tuberculosis* could prove crucial in understanding the molecular aspects of latent tuberculosis and provide a means for dealing with such an important global health concern.

References

Arcus, Vickery L., Paul B. Rainey, and Susan J. Turner (2005). "The PIN-domain toxin-antitoxin array in mycobacteria." TRENDS in Microbiology 13(8): 360-365.

Bertram, Ralph, and Christopher F. Schuster (2014). "Post-Transcriptional Regulation of Gene Expression in Bacterial Pathogens by Toxin-Antitoxin Systems." Frontiers in Cellular and Infection Microbiology 4: 6.

Cruz, Jonathan W., and Nancy A. Woychik (2016). "tRNAs Taking Charge." Pathogens and Disease 74.2: ftv117.

Frampton, Rebekah et al. (2012). "Toxin-Antitoxin Systems of *Mycobacterium Smegmatis* Are Essential for Cell Survival." The Journal of Biological Chemistry 287(8): 5340–5356.

Hayes, Finbarr, and Barbara Kędzierska (2014). "Regulating Toxin-Antitoxin Expression: Controlled Detonation of Intracellular Molecular Timebombs." Toxins 6(1): 337–358.

Lee, In-Gyun et al (2015). "Structural and Functional Studies of the *Mycobacterium Tuberculosis* VapBC30 Toxin-Antitoxin System: Implications for the Design of Novel Antimicrobial Peptides." Nucleic Acids Research 43(15): 7624–7637.

Lee, Ki-Young, and Bong-Jin Lee (2016). "Structure, Biology, and Therapeutic Application of Toxin-Antitoxin Systems in Pathogenic Bacteria." Toxins 8(10): 1-33.

Malherbe, Stephanus T., et al. (2016). "Persisting positron emission tomography lesion activity and *Mycobacterium tuberculosis* mRNA after tuberculosis cure." Nature Medicine 22(10): 1094-1100.

Min, Andrew B et al (2012). "The Crystal Structure of the Rv0301-Rv0300 VapBC-3 Toxin—antitoxin Complex from *M. Tuberculosis* Reveals a Mg^{2+} Ion in the Active Site and a Putative RNA-Binding Site." Protein Science: A Publication of the Protein Society 21(11): 1754–1767.

Petty, Kevin J. (1996). Metal-Chelate Affinity Chromatography. Current Protocols in Molecular Biology. John Wiley & Sons, Inc. 10.11.10-10.11.24.

Ramage, Holly R., Lynn E. Connolly, and Jeffery S. Cox (2009). "Comprehensive Functional Analysis of *Mycobacterium Tuberculosis* Toxin-Antitoxin Systems: Implications for Pathogenesis, Stress Responses, and Evolution." Ed. Susan M. Rosenberg. PLoS Genetics 5(12): e1000767.

Sala, Ambre, Patricia Bordes, and Pierre Genevaux (2014). "Multiple Toxin-Antitoxin Systems in Mycobacterium Tuberculosis." Toxins 6(3): 1002–1020.

Schendel, Paul F. (1998). Expression of Proteins in *Escherichia coli*. Current Protocols in Molecular Biology. John Wiley & Sons Inc. 16.1.1-16.1.3

Wang, Xiaoxue, and Thomas K. Wood (2011). "Toxin-Antitoxin Systems Influence Biofilm and Persister Cell Formation and the General Stress Response." Applied and Environmental Microbiology 77(16): 5577–5583.

World Health Organization (2016). "Global Tuberculosis Report 2016." WHO Press.

Winther, Kristoffer S., and Kenn Gerdes (2012). "Regulation of Enteric *VapBC* Transcription: Induction by VapC Toxin Dimer-Breaking." Nucleic Acids Research 40(10): 4347–4357.

Winther, Kristoffer et al. (2016). “VapCs of *Mycobacterium tuberculosis* cleave RNAs essential for translation.” Nucleic Acids Research 44(20): 9860-9871.

Yamaguchi, Yoshihiro, and Masayori Inouye (2011). "Regulation of Growth and Death in Escherichia Coli by Toxin-antitoxin Systems." Nature Reviews Microbiology 9(11): 779-90.