2020-01-01

Elucidating The Regulatory Mechanisms For The Cholesterol Catabolic Pathway Transcriptional Repressor Kstr Of Mycobacterium Tuberculosis

Chenoa D. Arico
University of Texas at El Paso

ELUCIDATING THE REGULATORY MECHANISMS FOR THE CHOLESTEROL CATABOLIC PATHWAY TRANSCRIPTIONAL REPRESSOR KSTR OF *MYCOBACTERIUM TUBERCULOSIS*

CHENOA DARA ARICO

Doctoral Program in Biological Sciences

APPROVED:

_____
Hugues Ouellet, Ph.D., Chair

_____
Charles T. Spencer Ph.D.

_____
Jianjun Sun, Ph.D.

_____
Chuan Xiao, Ph.D.

_____
Manuel Llano, Ph.D.

_____
Stephen L. Crites, Jr., Ph.D.
Dean of the Graduate School

Dedication

I would like to dedicate this dissertation and my entire academic career to my grandfather Henry Lugo Ayon, for his eternal push to make me pursue my education. He may not be able to see me complete the endeavor of graduating with my doctorates, but he was there to make sure that I began it and I know would be extremely proud of me.

ELUCIDATING THE REGULATORY MECHANISMS FOR

THE CHOLESTEROL CATABOLIC PATHWAY

TRANSCRIPTIONAL REPRESSOR KSTR

OF *MYCOBACTERIUM TUBERCULOSIS*

by

CHENOA DARA ARICO, B.S.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSPHY

Department of Biological Sciences

THE UNIVERSITY OF TEXAS AT EL PASO

December 2020

Acknowledgements

I especially would like to thank my husband Richard Arico for his enormous amount of patience and love that gave me the ability to continue with my education. I would never have been able to complete this journey without my partner and the love of my life. For my children, Christian, Madalyne, and Nathanial, I would like to thank you for the ability for you to see that bettering yourself with hard work and a tireless effort to pursue a better life is always within you, as it is your mother. To my parents, Stella Halliburton, Gary Bernhardt, and Richard Halliburton, I cannot thank you each enough for collectively making me the person that I have become. It is because of all your advice and support that I have learned the strength and determination to be better, strive to do the best I can, and never give up. My sisters Julie and Michelle, thanks for dealing with your absent sister for so long. I promise to maintain better contact with you now. For my mentor, Dr. Hugues Ouellet, you have taught me so much and continue to do so. It is because of you that I can call myself a scientist and because of your patience, kindness, and insistence that I have been able to accomplish this path. To all my friends and colleagues that have taken part in this journey with me Johnathan Abou-Fadel, Cameron Ellis, Susanna Portillo, Eva Iniguez, and all the students that I have mentored in the lab. I really do appreciate everything that you have taught me along the way.

Abstract

Mycobacterium tuberculosis (*Mtb*) is a bacterium that has been able to establish itself throughout evolution and is currently still an ever present threat resulting in approximately 2 million deaths per year worldwide. The ability for *Mtb* to utilize cholesterol as a carbon source is maintained through transcriptional regulation mechanisms not fully understood. The *Rv3574* gene codes for the self-regulated TetR-like transcriptional repressor KstR that controls the expression of large clusters of cholesterol catabolic genes in *Mtb*. This study aims to provide insights on the regulatory mechanisms of cholesterol catabolism, which could be used as therapeutic target to treat and control *Mtb* infection. To study the interaction between KstR, and its physiological ligand(s) *in vitro*, the KstR protein has been expressed and electrophoresis mobility shift assays (EMSA) conducted in the presence or absence of other potential ligands. X-ray crystallographic studies of KstR with its corresponding *kstR* DNA operator sequence as well as with its effector, 3-oxo-cholest-4-en-26-oyl-CoA ester (3OCh(25S)26-CoA), have been performed to better understand the conformational changes KstR undergoes in physiological conditions. Surface plasmon resonance (SPR) and EMSA studies revealed that the KstR regulator protein binds to the substrate 3OCh(25S)26-CoA with apparent higher affinity than that of its 3-oxo-4-cholestenoic acid precursor and SPR data has shown the ability for KstR to interact with both -S and -R stereoisomers of 3OCh26-CoA. Crystallization of KstR in its ligand bound condition with the 3OCh(25S)26-CoA, as well as bound to the DNA operator, were accomplished showing different orientations of the repressor. The possibility for KstR isoforms exists in *Mtb* strains that could provide additional signaling functionality. Therefore, an extended KstR protein and mutant derivatives were examined and tested through SPR studies and showed hypothetical phosphorylation at the N-terminus still allows for binding to KstR but at reduced affinity. The characterization of *kstR* gene

deletion studies in CDC1551 strains of *Mtb* illustrate alterations in growth that demonstrate regulatory aspects under the control of KstR are strain dependent and do not behave in similar manners. Growth curves were performed and showed *Mtb* clinical isolate CDC1551 Δ*kstR* growth is much more deficient than that of the WT. Thin layer chromatography experiments displayed how both CDC1551 WT and Δ*kstR* strains differ in how they metabolize the virulence factor PDIM after *kstR* deletion CDC1551 Δ*kstR* produces less PDIM after *kstR* deletion and complementation was unable to recover the PDIM loss to that of the WT. Proteomic studies for CDC1551 wild-type and the Δ*kstR* deletion strain showed KstR plays a role in *Mtb*, apart from cholesterol degradation, with downstream changes in enzyme quantities in various pathways, such as the glyoxylate cycle and PDIM production. In CDC1551, both deletion and over-expression conditions of *kstR* have shown to cause attenuation at the onset and in persistence with RAW264.7 macrophage infection experiments. ACDC1551 KDKstR strain have been created and has shown similar results to that of *kstR* deletion in macrophage infection studies.. Western blots, thin layer chromatography, and RAW264.7 macrophage infection studies have shown that the CDC1551 KDKstR strain behaves in the same manner as the Δ*kstR* strain and a viable model for future mouse infection experimentation. Overall, the data presented provide significant and fundamental knowledge regarding the regulatory effects of KstR deletion and over-expression in *Mtb*, and distinguishes KstR as viable a drug therapeutic target.

Table of Contents

List of Tables

x

List of Figures

Chapter 1: Background and Significance

1.1 *Mycobacterium tuberculosis* evolution and history

Mycobacterium tuberculosis (*Mtb*) is an ancient, and the more known etiological cause for the disease tuberculosis (TB). *Mtb* is an obligate aerobic, tubercle shaped bacteria, comprised of a hardy cell wall that is gram positive by nature. However, *Mtb* also presents a heavy waxy coating that prevents gram staining. Acid-fast staining is the technique used in most cases to verify presence of the bacteria. Although considered an obligate aerobe, *Mtb* is able to survive under the hypoxic environment that occurs within the granuloma during the course of infection within the lungs.

The bacterium resides among a genus characterized by the *Mycobacterium tuberculosis* bacterial complex (*MTB*C). This complex includes a group of species of *Mycobacterium* that are genetically similar in nature that consists of both animal and human infecting species *M. tuberculosis, M. africanum, M. bovis, M. microti, M. canetti, M. caprae*, and *M. pinnipedii*, **[1]** The *MTB*C is recognized to be zoonotic in nature and yet a recurring common misconception of the complexes spread and evolution is the pathogen originated from animals although increasing evidence has shown otherwise. The most common ancestor of the *MTB*C is estimated to have emerged approximately 70,000 years ago from its progenitor during the Neolithic period in East Africa, the region from where modern human populations disseminated around the same time frame.**[2]** 10-20,000 years after further migration out of the horn of Africa, later (20-30,000 years ago), the separation of two major lineages occurred, one of which spread from human to animals after the start of domestication of animals.**[3]** An example of this occurrence, is the transmission

of the complex to seals where upon their subsequent migration to the Peruvian coast allowed the spread of the *MTB*C to the indigenous human population residing in South America.**[3]**

Archeological findings have shown the presence of TB in 4000 year-old deformed human bone samples from areas such as Italy and Denmark while also finding the presence in Middle Eastern dig sites from the same periods. **[4]** This indicates the world population at that time had already seen the spread of TB alongside the migration of humans. Cases of TB have been being recorded and documented as early as the fifth century B.C by Hippocrates, dubbing the infection as phthisis which translates to consumption. **[5]** The numerous records at this time in history of those with symptoms of wasting away, chest pain, coughing, and sputum-containing blood demonstrates that TB was well established in society. The increase of TB infections during the 16th and 17th was witnessed due to the aide of the European population boom in which the formation of larger urban centers became a hotbed of a number of outbreaks. Shortly afterwards, during 19th century, TB infections peaked with the estimated mortality of one-third of the European population. **[4]** It was after this and during the latter half of the 19th century; TB infections began to decrease significantly. This can be attributed to more public health concerns and better sanitation of housing districts, although it has also been hypothesized natural selection of humans' resistant to TB may have been partially responsible for this decrease. The premise of natural selection was later debunked due to TB infections declining so rapidly. **[6]**

The discovery of *Mtb* being the causative bacterial agent for the vexing disease tuberculosis was discovered by Robert Koch in 1882. Before his discovery, TB was believed to be an inherited disease. **[7]** Unfortunately, during his studies, either the original bacteria isolated was not cultured or was lost. In 1905, Dr. Edward R. Baldwin isolated the TB strain that was serially passaged over time and has since become known as the commonly used lab strains H37Ra and H37Rv. **[8]**

Research has since then yielded a vaccination from the Bacillus Calmette–Guérin (BCG) strain derived from *M. bovis* that was first medically administered in 1921. **[9]** Nowadays, the vaccine is giving to infant children and some adults where there is a high risk and potential of acquiring the disease and is able to grant only partial protection.

1.2   Virulence and Epidemiology

The high virulence of the *Mtb* has been proven with its persistence to evolve along the evolutionary paths of humans. The pathogens presence has been a constant throughout the evolution of mankind. Unlike other infectious diseases, *Mtb* does not have distinguished virulence factors, for example toxin production. This has given the pathogen the ability to develop a number of virulence factors on many different gene levels. The genes responsible for the pathogenesis to prevail are amongst the entire genome rather than a specific set of genes that function to produce virulence. Examples of the wide range of functions are that of lipid and fatty acid metabolism (including cholesterol degradation), cell envelope proteins, secretion systems, responses to oxidative and nitrosative stresses that involve phagosomal properties and the inhibition of apoptosis, gene expression regulators, and many more. **[10]** The many ways the *Mtb* complex is able to produce a slew of virulence factors can perhaps be attributed as the reason that it has been able to establish an ongoing infection along the evolution and migration of the human species.

In 2015, a known 10.4 million people became infected with TB. Of those, 1.8 million resulted in mortality. **[11]** Infection rates have been decreasing approximately 1.5% yearly since 2000 but the cases of infection in developing countries of low income and sub-standard health practices still contain 60% of the total infected count and account for 95% of death rates. Another especially important factor at play with death rates, are the populace that are infected with HIV that become more susceptible TB. The number of HIV-positive people to die from TB infection accounted for

3

35% of all HIV deaths in 2015. **[11]** The World Health Organization has set out to eliminate TB infection rates throughout the world by the year 2030. For this goal to be accomplished, the decrease of TB infections would have to be raised from the current 1.5% a year to between 4.5-5%. Continued research to discover new methods to combat this deadly disease is key to these efforts. **[11]**

1.3   Current treatment methods and antibiotic resistance

Existing treatments for TB consist of a combination of approved front line drugs isoniazid (INH), rifampin (RIF), ethambutol (EMB), and pyrazinamide (PZA) that that are to be taken for approximately 6-9 months. Treatment consists of an intensive 2-month phase that is followed by a continuation phase lasting between 4-7 months. The length of continuation phase relies on the extent of TB infection (lung cavitation) witnessed in the patient and bacterial load in sputum cultures at the completion of the intensive phase. Side-effects such as hepatitis, dyspepsia, severe skin rash, and joint pain are associated with the administration of these drugs. Due to the high occurrences and severity of side-effects associated with TB treatment, patient adherence to the treatment regimen often becomes a large issue. **[12]** It is this lack of adherence to drug regimens, due to poor health care provision, poor drug quality, and patient ending their own treatment prematurely that has been the major underlying cause for the uprising in cases of drug resistant TB strains (More refs).

As noted previously, TB cases have been on the decline in the past 15 years. However, within this decline there has been an increase in the cases of multi-drug resistant strains. **[11]** According to WHO, Multi-drug resistant TB (MDR-TB) is a form of TB infection caused by *Mtb* that no longer responds to RIF and INH. RIF resistance occurs with mutations with *rpoB* where translation of the □□subunit of the RNA polymerase for RIF binding site is altered and affinity for the

antibiotic is rendered inactive. **[13]** The most common INH resistance occurs with mutations occurring with the *katG* gene that encodes for a catalase peroxidase enzyme used by the bacteria to activate INH. **[14]** Additionally, high INH resistance has been shown with mutations to the promoter region of the *inhA* gene which is also responsible for co-resistance with ethionamide (ETH), a structural analog of INH and a TB combating second line drug. [15] After the first line drugs have become ineffective other drugs such as fluoroquinolones and injectable antibiotics such as kanamycin and capreomycin are introduced for treatment. Once *Mtb* has become resistant to both INH and RIF, resistant to one of the fluoroquinolones and resistant to at least one of the injectables the *Mtb* strain is then considered to be as Extremely Drug Resistant (XDR-TB). Survival prognosis of those having XDR-TB is 50%. On a yearly basis, 9.5% of the MDR-TB cases have shown to have acquired some type of additional drug-resistance and 117 countries worldwide have reported to have at least one case of XDR-TB. **[11]** The implications of MDR-TB and XDR-TB being able to spread rapidly around the world with few methods available to contest the highly virulent and poorly curable TB strains, the mission of eradicating TB by 2030 could possibly be in jeopardy.

1.4   Cholesterol side-chain degradation and pathogenesis

The process of shuttling cholesterol into the *Mtb* bacterial cells is accomplished by the *mce4* gene cluster that codes for a number of membrane steroid transport proteins. **[16]** The initial steps to side-chain cholesterol degradation occur with the oxidation of the C3 carbon on ring A by extracellular ChoD or by the intracellular 3β-hydroxysteroid dehydrogenase/isomerase (3β-HSD). **[17]** The cytochrome P450s Cyp125 and/or Cyp142 then provide the oxidation reaction of the side-chain C26 forming 3-oxo-4-cholest-en-26-oic acid (3OCH26oic acid). **[18-21]** The attachment of coenzyme-A is conducted by the acyl-CoA ligase FadD19 creating 3-oxo-cholest-
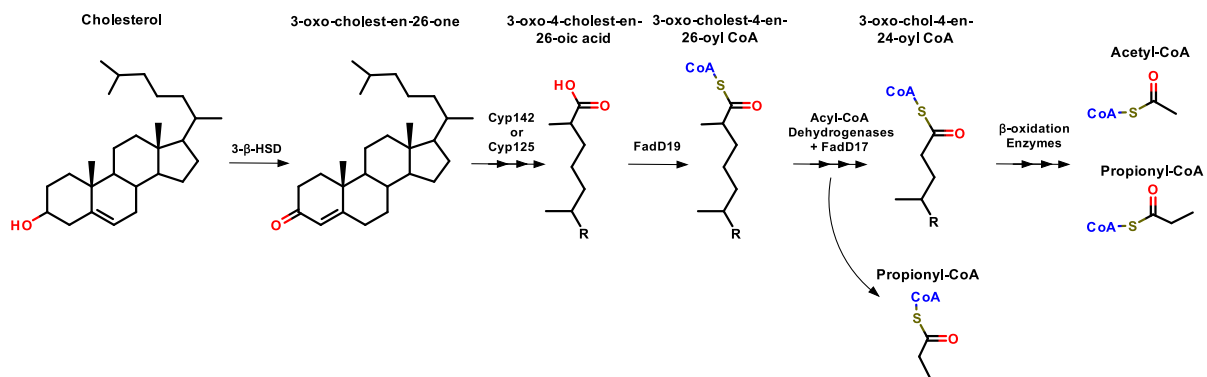
Figure 1: Simplified *Mtb* cholesterol side-chain degradation pathway. Cholesterol side-degradation in *Mtb* is carried out in three successive rounds of β-oxidation-like reactions. Degradation proceeds until the entire aliphatic side-chain is catabolized into either acetyl-CoA or propionyl-CoA, which are then routed to the TCA cycle for energy, and fatty acid and methyl-branched fatty acid biosynthesis pathways.

4-en-26-oyl CoA (3OCH26CoA). **[22]** A series of acyl-CoA dehydrogenases carry out the formation of a dioxo group on the side chain in preparation of cleavage and results in propionyl-CoA and 3-oxo-chol-4-en-24-oic acid. A second acyl-CoA ligase FadD17 catalyzes the attachment of CoA to the 3-oxo-chol-4-en-24-oic acid and the process is repeated until the entire side-chain of the cholesterol molecule has been degraded, which yields one acetyl-CoA and two propionyl-CoA molecules. The propionyl-CoA created are shuttled to the TCA cycle for acetyl-CoA and either the methylcitrate cycle or methylmalonyl pathway for propionyl-CoA. (**Fig. 1**) After the first cycle of β-oxidation of the cholesterol side-chain has been completed the onset of A/B ring degradation takes place. While the remaining sterol degradation steps remain relevant, cholesterol side-chain is the focal point of this research.

The sequential β-oxidation of the cholesterol side-chain results in shortened acetyl-CoA and propionyl-CoA carbon chains. **[23]** The acetyl-CoA formed can then be shuttled to the Citric Acid Cycle for direct energy utilization. The propionyl-CoA can be diverted to either the methyl citrate cycle, resulting in acetyl-CoA and energy or directed to the methylmalonyl pathway where methyl-branched fatty acid formation occurs. **[24]** The methyl-branched fatty acids created by this pathway are the precursors for the biosynthesis of complex virulence lipids, such as phthiocerol dimycocerosates (PDIMs) that are known to reside as free lipids on the *Mtb* cell wall. **[25]** PDIM is a known virulence factor that when deficient has shown as a protective method for *Mtb* survival *in vitro* by decreasing the host response to the bacteria. **[26]** Increased amounts of cholesterol during *Mtb* growth also increases amounts of PDIM found on the surface. **[27]** It is presumed that the cooperative link between the increased amounts of PDIM and cholesterol as a carbon source is due to the shift of side-chain degradation and the influx of propionyl-CoA to the methylmalonyl pathway.

After the introduction of *Mtb* bacilli to the lungs, alveolar macrophages ingest the bacteria through phagocytosis. This process is facilitated by the presence of plasma cholesterol-rich microdomains associated with the Tryptophan-Aspartate containing coat protein (TACO) and the reduction of cholesterol coincides with the reduction of *Mtb* intake by the cell. **[28, 29]** Additionally, the cell surface of *Mtb* presents a Receptor-C_k -like molecule that has shown to have a high affinity to cholesterol that provides stable binding to the macrophage cell membrane. **[30]** A normal occurrence for the host's response would be the formation of lysosomes where the bacteria would be degraded and used for further immunological responses. This however does not occur in the presence of *Mtb*. Phagosomes that contain *Mtb* bacilli fail or delay to perform the microbicidal functions of pH reduction through the recruitment of vacuole proton-ATPase (V-ATPase) that usually takes place in the phagosome maturation process. **[31]** This in turn prevents phagosome-lysosome fusion through preventing the acidification of the microenvironment where the bacilli are able to remain masked within the cell and begin to multiply. **[32]** During this event, cholesterol rich foamy macrophages begin to surround the macrophages containing bacteria and the formation of the granuloma is initiated. **[33]** The bacteria within the granuloma are insulated from any further immunological responses that might normally occur. It is within this sequestered environment where *Mtb* is able to persist and lay dormant for possibly decades.

The granuloma is a focal point for studies regarding *Mtb* survival during infection. It is rich with lipid nutrients where the pathogen is able to begin metabolically adapting responses to the environment to maintain persistence. **[34, 35]** Earlier transcriptional studies showed the induction of genes encoding fatty acid degradation enzymes leading to the recognition that lipid carbons are a preferred nutrient source during *in vivo* infection. **[36]** It was later discovered that *Mtb* possessed a number of genes dedicated to the uptake and degradation of host membrane cholesterol from

macrophages and that this pathway was essential for the breakdown of host membranes as a required carbon source during chronic infection. **[37-39]**

1.5   Lipid Metabolism

The cholesterol catabolic pathway functions in a multi-step process whereby once the cholesterol has been shuttled into the cell; it undergoes the initial steps of degradation, which begins, with the β-oxidation of the aliphatic side chain by key enzymes such as Cyp125. [18] Cyp125 initiates β-oxidation with the catalysis of the cholesterol and 3-oxo-4-cholesten-3-one C26 through hydroxylation. The process of side-chain degradation continues through several other enzymes until the original cholesterol molecule has been altered to androstenedione. The by-products of the β-oxidation result in the production of propionyl-CoA and acetyl-CoA, metabolic and energy utilizing molecules in *Mtb* respectively. [40]

Mtb is known to be substantially sensitive with increases in propionyl-CoA accumulation, as it is the precursor for toxic metabolites, and therefore has adapted at least three different pathways by which it can process the fatty-acid derivative. [23, 41, 42] The methylcitrate cycle consists of two isocitrate lyases (ICL) that convert propionyl-CoA to either pyruvate or succinate through the glyoxylate cycle that is shuttled to the TCA cycle. [42, 43] Propionyl-CoA carboxylase converts propionyl-CoA to methylmalonyl-CoA, which is then converted to succinate through the highly dependent vitamin B_{12} methylmalonyl pathway (MMP). [41]   Another path for methylmalonyl-CoA, rather than be converted to succinyl-CoA, is to be diverted into the metabolism of methyl-branched fatty acids, which are the foundations of the *Mtb* cell wall lipids. [44-46] *Mtb* cell wall lipids are highly complex that provide hydrophobic protection for the bacteria and act as virulence factors. [47] Methyl-branched fatty acids such as phthiocerol dimycocerosate (PDIM) a cell wall surface associated lipid produced through the MMP have

9

shown to protect *Mtb* from host innate immune responses and mutants deficient in PDIM are attenuated and killed by the host during the first two weeks of infection. **[26, 47, 48]**

1.6   The negative transcriptional regulator KstR

The catabolism of cholesterol process is controlled by two negative TetR-like transcriptional regulators, KstR and KstR2. **[49-52]**, with KstR being responsible for the breakdown of the side-chain and A/B rings while KstR2 controlling the degradation of the C/D rings. The production of the enzymes required for cholesterol catabolism resides within the auto-regulated *kstR* regulon. This regulon consists of approximately 70 genes that are involved with the uptake and degradation of the cholesterol molecule as well as play a significant role towards lipid metabolism. **[27, 53]** KstR regulated genes are known to be up-regulated in hypoxic, cholesterol-rich environments that are provided by the granuloma environment during TB infection and persistence. **[54]** TetR-like transcription regulators are proteins that often act as repressors and serve to function as prokaryotic regulatory elements for a wide range of physiological responses to environmental changes. **[55]** Transcriptional regulators in the TetR family operate by controlling genes that are involved in multidrug resistance, enzymatic reactions for catabolic pathways, pathogenicity, environmental stressors, and biosynthesis of antibiotics. The archetype for TetR-like negative regulators is homo-
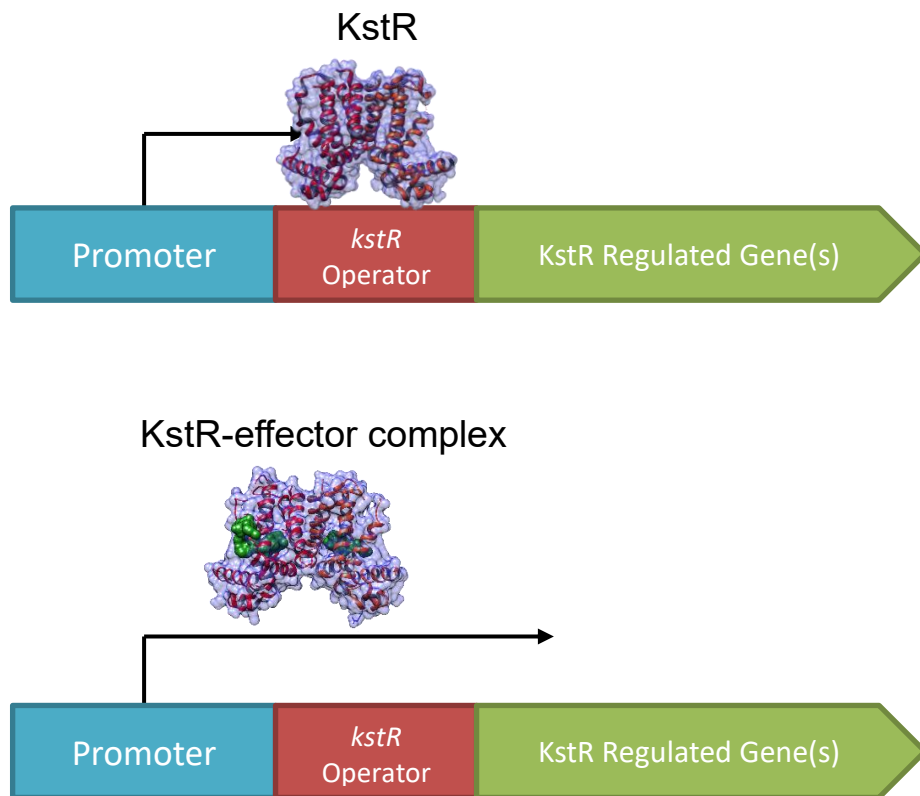
Figure 2: Depiction of KstR and effector-induced de-repression. KstR is a negative transcriptional regulator that was shown to bind to approximately 74 DNA operator sequences within the *Mtb* genome, halting production of transcripts involving cholesterol catabolism. Upon binding of an effector, the KstR-effector complex dissociates from the DNA and the expression of KstR-regulated genes is de-repressed.

dimerization where control is initiated with introduction of an effector molecule to initiate de-repression or as a two-component system. **[55]** (**Fig. 2**) Regulators that serve as two domain proteins have a DNA binding domain and a signal-receiving domain that is able to confer a signal.

Recently, another TetR-like regulator, EthR, has shown to halt the proteins ability to bind to the DNA through post-translational modification (PTM) phosphorylation of specific serine/threonine amino acids located in the N-terminal of the protein.**[56]** Like its EthR counterpart, KstR also has a number of serine's and threonine's located in the N-terminal that could serve as a mechanism for the regulation of KstR controlled genes and hence the cholesterol catabolism in *Mtb* (**Fig. 3**). When entered into the bacterial phosphorylation prediction server NetPhosBac [**57**], the N-terminal serines and threonines are chosen as possible sites of phosphorylation yet only the serine located in the 10 position is deemed a likely candidate for a PTM. To discover additional mechanisms as to how KstR is able to adapt to cellular signals would be beneficial to study further.

KstR is not shown to be an essential gene requirement for the growth of the bacteria but has been shown to be up-regulated in macrophage infection, present at the onset of immunological responses in the mouse model of infection and required for mouse infection. **[36, 58, 59]** The requirement of KstR during infection *in vivo* suggests that deregulation of cholesterol catabolism has a marked attenuation effect on the virulence of *Mtb*. It is suspected the absence of KstR promotes attenuation likely due to the over-expression of enzymes within the cholesterol catabolic pathway that alter the levels of the catalytic products. Whether this is specifically due to the over-production of a particular transcript or specific set of transcripts, it is unknown and further investigation is warranted.

Figure 3: Multiple sequence alignment and phosphorylation prediction of KstR with N-terminal 21 amino acid extension. (Top) Modeling with UCSF Chimera software of KstR with 21 amino acid extension. (Middle) Translation of the N-terminus denoting the serines and threonines predicted to be possible phosphorylation sites and the conservation through various *Mycobacterium* species. (Bottom) Probably prokaryotic phosphorylation sites of KstR when entered into the NetPhosBac 1.0 server.

13

1.6   Significance

THE CHOLESTEROL DEGRADATION PATHWAY HAS BEEN A FOCUS AS A PUTATIVE THERAPEUTIC OPPORTUNITY TO PURSUE FOR SOME TIME. [27] THE RATIONALE BEHIND THIS HYPOTHESIS STEMS FROM OTHER KNOWN *MTB* TETR-LIKE TRANSCRIPTIONAL REGULATORS SUCH AS ETHR TO HAVE ALSO BEEN SHOWN AS ADVANTAGEOUS TARGETS FOR DRUG DEVELOPMENT MODELING. [60, 61] ETHR IS ALSO A TETR-LIKE NEGATIVE transcriptional regulator that controls the transcription of the *ethA* gene encoding for the monooxygenase that is required for ETH activation by *Mtb*. **[62]** Inhibitors for EthR that render the protein incapable of forming the EthR-DNA interactions allowing for the up-regulation of *ethA* have been shown to decrease the concentration of ETH needed to decrease bacterial load in mice. **[63]**

In order to demonstrate that KstR could be a potential therapeutic candidate against *Mtb* infection, its interaction with physiological effector and the impact of disrupting this interaction on *Mtb*'s ability to adapt to intracellular growth conditions must be evaluated in detail. Indeed, the inhibition of KstR-induced adaptive mechanisms is expected to synergize with current antitubercular therapy. The introduction of an analog to disrupt KstR function alongside the therapies that are well known could provide an additional treatment method or even possibly allow for lesser quantities of medication to be used in order cure TB infection. The lowering of medication dosages would lessen the severity and occurrences of harmful side effects that afflict undertaking the grueling treatment required for TB today. Diminishing such side effects would then allow patients adhere to drug therapies to increase and hinder the nascent formation of MDR-TB and XDR-TB strains.

1.7   Hypothesis and Specific Aims

Hypothesis

The central hypothesis of this study aims to prove the negative transcriptional regulator KstR is required in *Mtb* for overall fitness, survival, and persistence during infection by maintaining a regulated state of cholesterol degradation, as the de-regulation is expected to alter the abundance of key metabolic virulence factors and energy sources, resulting in cellular stress and likely attenuation.

Specific Aims

Mycobacterium tuberculosis (*Mtb*) the bacterial agent known to cause the long-standing disease tuberculosis that is still a leading cause of deaths in the world. Due to several underlying factors, for example, treatment regimens not being followed properly by patients, strains of single and multi-drug resistant *Mtb* have begun to emerge throughout the world over the course of the past few decades. The treatment for the disease tuberculosis is well known to be a long and arduous process. In many cases, medical treatment is often lacking in developing countries and the maintenance required for curing tuberculosis infection is not followed through as needed. The emergence of drug-resistant strains and the lack of the proper treatment in many countries around the world have warranted the need for novel therapeutic approaches to staving off the disease.

Mtb has the ability to adapt to the host environment converting its nutrient intake to that of a cholesterol-rich diet. With this conversion, *Mtb* is able to survive within the lung and remains in a latent either stage or initiate a rampant infection. The pathway responsible for the degradation of the cholesterol needed for survival is maintained through a tightly knit regulatory system.

This work currently focuses on the characterization of the *Mtb* negative transcriptional regulator KstR. Although KstR has been previously described, full comprehension of how KstR is able to regulates efficiently by binding to the large number of DNA operators within the *Mtb* genome and respond to appropriate effector or inducer molecules for the act of de-repression has yet to be resolved. This proposal is aimed at identifying the physiological inducer ligand(s) of KstR and the binding characteristics of KstR to its corresponding DNA operator sequences in the presence of the proposed effector molecules. With the use of *in vitro* assays and the development of *in vivo* infection models we seek to distinguish in the importance of the different inducer ligands on the regulation and production of cholesterol catabolic enzymes required for infection persistence and maintenance.

1.7.1 Specific Aim 1: Identification of the physiological inducer ligand(s) of *Mtb* KstR.

In order to understand the transcriptional regulatory responses required for cholesterol catabolism the negative regulator KstR must interact with a ligand to initiate the release of the protein from the DNA. Establishing the ligand(s) that is responsible for this event to transpire is necessary to elucidate the characteristics of the protein and its regulatory role in the cholesterol catabolic process. We hypothesize tightly regulated utilization of host cholesterol in *Mtb*, regulated by KstR, occurs with the production of CoA esters of cholesterol side-chain intermediates.

1.7.1.1 Specific Aim 1 Sub-Aim: Distinguishing if phosphorylation is capable of providing an additional regulatory mechanism signal for de-repression of KstR through *in vitro* and *in vivo* methods.

The requirement for KstR to be able to adapt rapidly to environmental signals, further mechanisms of action would need to be exploited by the *Mtb* cell. For *Mtb* to be to begin breaking

down cholesterol and the resulting toxic catabolites, de-repression of the enzymes required must be initiated promptly after sensing the presence of cholesterol. Examination of how this is accomplished by verifying possible N-terminal phosphorylation of KstR post-translationally to inhibit binding to DNA. This discovery would enhance our comprehension as to how *Mtb* is able to respond rapidly to presence of cholesterol.

1.7.2 Specific Aim 2: Determining the impact of KstR **deletion (knockout) and depletion (knockdown) on *Mtb*'s ability to utilize cholesterol and to grow intracellularly.**

We will develop an ATc-inducible expression system that can deplete KstR and that mimics its chemical inactivation. We will apply this system to investigate the role of KstR in the regulation of cholesterol utilization activities during the course of infection. These results will provide insights into the impact of the de-regulation of cholesterol catabolic activities in both acute and chronic phases of infection.

Chapter 2: Identification of the physiological inducer ligand(s) of *Mtb* KstR

2.1 Methods

Cloning and expression of recombinant *Mtb* KstR in *E. coli*

The *Mtb* H37Rv gene *Rv3574* coding for the KstR protein along with a C-terminal hexa-histidine tag sequence was amplified with the use of polymerase chain reaction (PCR) using the primers specified in Table X and Phusion HF DNA Polymerase (NEB). The amplified coding sequences of *Rv3574* (27 ng) was inserted into the *Nco I* site of the pET-derived vector, pLIC-HK (50 ng), through the use of In-Fusion ligation independent cloning (LIC) using the Takara LIC cloning kit. The reaction was incubated at 55°C for 15 min and 25 ng of the reaction was introduced into 100 □l NEB5-□ (NEB) chemically-competent cells by incubation on ice for 30 minutes with subsequent heat-shock at 42°C for 30 seconds. The heat-shocked cells were incubated on ice for 5 minutes and 900 □l of SOC medium was added. The cells were incubated for 60 minutes while shaking at 250 rpm at 37°C for recovery. A 100-□l aliquot of the recovered cells was plated on Luria-Bertani (LB) agar plates containing 50 µg/ml kanamycin (K_{50}) then incubated overnight at 37°C. A few resulting colonies were grown and screened for the insertion of *kstR* amplicon by restriction analysis and sequencing.

1 ng of the screened recombinant plasmid DNA containing the *kstR* gene was then introduced into 50 µl *E. coli* BL21(DE3) and co-expressed with L-arabinose-inducible chaperones GroEL/ES, incubated on ice for 20 minutes, heat shocked at 42°C for 90 seconds, incubated on ice for an additional 5 minutes, and recovered with 450 µl of S.O.C. medium for 60 minutes at 37°C while shaking. 100 µl of the recovered cells were plated on LB agar plates containing K_{50} and 34 µg/ml choramphenicol (Cm_{34}). Approximately 3-4 colonies were chosen and used to

Table 1. Primers and oligonucleotide sequences used for this study

Name	Sequence
Expression	
KstRLIC-F	5'-<u>AGGAGATATACCATG</u>GCGGTACTTGCCGAGTC
KstR-X-LIC-R	5'-<u>TGAGAAGAACCCATG</u>TCAGTGATGGTGATGGTGATGGGCGCTGTCTTGATCGC
KstR-Asp-X-LIC-R	5'-<u>AGGAGATATACCATG</u>**GATGAT**GCGAAC**GAT**AAC**GACGATGA**CGCTCCCGACGCA
KstR-Ala-X-LIC-R	5'-<u>AGGAGATATACCATG</u>**G**CG**G**CAGCGAAC**G**CGAAC**G**CCGCT**GCTGCT**CCCGACGCA
KstRLIC-R	5'-<u>TGAGAAGAACCCATG</u>TCAGTGATGGTGATGGTGATGGGCGCTGTCTTGATCGC
EMSA	
EMSA*kstR*-F	IRDye700-5'-TGCCCACTAGAACGTGTTCTAATAGTGC-3'
EMSA*kstR*-R	IRDye700-5'-GCACTATTAGAACACGTTCTAGTGGGCA-3'
EMSAigr-F	IRDye700-5'-ATAGTAATGAAACGTGTTCTAGCCTGGCC-3'
EMSAigr-R	IRDye700-5'-GGCCAGGCTAGAACACGTTTCATTACTAT-3'
SPR	
SPR*kstR*-F	Biotin-5'-TGCCCACTAGAACGTGTTCTAATAGTGC-3'
SPR*kstR*-R	5'-GCACTATTAGAACACGTTCTAGTGGGCA-3'
SPR*kstR*2-F	Biotin-5'-CTGGCCGAGCAAGCGCTTGGTTGATAGT-3'
SPR*kstR*2-R	5'-ACTATCAACCAAGCGCTTGCTCGGCCAG-3'
SPRigr-F	Biotin-5'-ATAGTAATGAAACGTGTTCTAGCCTGGCC-3'
SPRigr-R	5'-GGCCAGGCTAGAACACGTTTCATTACTAT-3'
Gene deletion	
5FL*kstR*Van-F	5'-<u>TTTTTTTTCCATAAATTGGG</u>ACACCTCCAGAACGGTGA-3'
5FL*kstR*Van-R	5'-<u>TTTTTTTTCCATTTCTTGGG</u>GTACCGCCACTTTCATGACC-3'
3FL*kstR*Van-F	5'-<u>TTTTTTTTCCATAGATTGGG</u>AACCAACCGAGGACCAGTA-3'
3FL*kstR*Van-R	5'-<u>TTTTTTTTCCATCTTTTGGG</u>ACGGCTTCGTGGTGTATTC-3'
Complementation	
Promoter region-F-LIC	5'-<u>ACGCGTGCGGCCGCGG</u>CGTCGACAATTTCACATAGCAAT
Promoter region-R-LIC	5'-<u>GCCGCCAT</u>GTAGCGGGCAGCTCCT
kstR-ala-mutant-F-LIC	5'-CCGCTAC<u>ATGGCGGCA</u>GCGAAC
kstR-ala-mutant-R-LIC	5'-<u>AGATTTAAAGATCTG</u>CTAGGCGCTGTCTTGATCGC
Compl*Mtb*KstR-F	5'-CATT<u>GGTACC</u>GCGTCGACAATTTCACATA-3'
Compl*Mtb*KstR-R	5'-CATT<u>GGTACC</u>GGGCGAGCTATTGCTGAAG-3'

inoculate a 50 ml LB broth containing K_{50} and Cm_{34} for overnight incubation at 37°C while shaking 250 rpm. Four Fernbach flasks with 1 liter of LB media each was inoculated with the overnight culture and allowed to grow until reaching and OD_{600} between 0.4-0.6. Isopropyl β-D-1-thiogalactoprpmyranoside (IPTG) was added to the cultures at a 1 mM concentration for protein induction and the cultures were allowed to grow for 3 hours at 37°C while shaking at 180 rpm. After incubation, the cells were centrifuged to collect the pellets at 4000 rpm and stored at -80 °C until purification was performed.

Lysis and Purification of KstR

The frozen stored expressed *E. coli* pellets were thawed on ice and suspended in 50 ml of ice-cold lysis Buffer A (50 mM Tris-Cl, 150 mM NaCl, 10% glycerol, 2.5 mM imidazole) and 0.5 mg/ml lysozyme then incubated on ice for 30 minutes. Sonication was performed while protein homogenate was kept on ice with 0.5 second pulses for 3 minutes, 3 times, with 3 minute intervals between each round of sonication. Phenylmehtylsulfonyl fluoride (PMSF) was added in 100 µl increments at a concentration of 100 mM to the protein lysate before and between sonication breaks to prevent proteolysis. The protein lysate was then centrifuged at 10,000 rpm for 45 minutes for insoluble fraction separation and the resulting supernatant was transferred and further centrifuged at $100,000 \times g$ for 1 hour to separate the membrane-bound fraction. The supernatant was carefully removed and filtered through a 0.45 µm syringe filter (Millex) to remove larger particulates before performing purification steps.

For the purification process, two 5 ml His-trap FF columns (GE Healthcare) were attached to one another and equilibrated using 5 column volumes (50 ml) of Buffer A. The protein supernatant was loaded manually into the columns using a peristaltic pump (GE Healthcare) and the loaded columns were washed with 5 column volumes for removal of non-bound proteins and

20

debris. Throughout each step, loading and washes, the flow through was collected and used for analysis to ensure the proteins were being bound to the columns and not washing off. Using a 2110 Fraction Collector (Bio-Rad), 10 column volumes of Buffer A and 10 column volumes of Buffer B (50 mM Tris-Cl, 150 mM NaCl, 10% glycerol, 150 mM imidazole) were added to each chamber and allowed to flow through the column via peristalsis at a flow rate of 2.5 ml/min. The protein elution fractions were collected in 5 ml collection tubes and analyzed on 12.5% SDS gels for the presence of KstR in a clean manner. The fractions containing KstR were collected and dialyzed with three rounds of buffer osmosis to remove any remaining imidazole. The protein was then concentrated (Millipore 30,000 MWCO membrane) and the concentration was measured with a Bradford Protein Assay.

Electrophoretic Mobility Shift Assays with *Mtb* KstR and cholesterol catabolic pathway intermediates and other potential ligands

The end-labeled IR-Dye probes with the KstR$_{Mtb}$ operator sequences (Table 1) were annealed by heating to 100 °C for 3 minutes then allowed to slowly cool to room temperature. 100 nM of recombinant purified KstR was incubated with 20 nM of labeled probes in binding buffer (2mM HEPES pH 8.0, 7.5 mM NaCl, 1mM MgCl$_2$, 0.25% Tween-20, 2.5mM Dithiothreitol (DTT), 1 µg poly(dI-dC)) for 20-30 minutes with serial concentrations with known and other possible inducer ligands. Incubated reactions were loaded onto 5% native 75:1 polyacrylamide gels containing 0.5x TBE. Gels were electrophoresed for 1.5-2 h at 4 °C then imaged and analyzed using an Odyssey® Imaging System (LI-COR).

Crystallization, data collection and crystal structure determination of KstR **in complex with DNA operator and 3-oxo-cholesten-26-oyl-CoA**

Prior to crystallization, KstR was diluted to 10-15 mg/ml by mixing with 10 mM Tris-HCl, pH 7.5, supplemented with the CoA ester derivative (25*R*)-3-oxocholest-4-en-26-oyl-CoA, or an annealed oligonucleotide duplex 5'-C<u>TAGAAC</u>GT<u>GTTCTA</u>A-3'/5'-TTAGAACACGTTCTAG-3', representing an operon, both at 1:1 enzyme:ligand molar ratio. Crystallization conditions were determined using commercial high-throughput screening kits (Hampton Research) and a nanoliter drop-setting Mosquito robot (TTP LabTech) operating with 96-well plates, and a hanging drop crystallization protocol. Crystals were further optimized in 96-well plates for diffraction data collection. Plate-shaped KstR▪3-oxocholest-4-en-26-oyl-CoA crystals diffracting in the C2 space group grew from 11% PEG 3350, 90 mM NaSO4 and 13.3% MPD. The KstR▪DNA crystals grew from 21.7% PEG 6000, 0.2 M Na sulfate and 8% Tacsimate, pH 5.1. Prior to data collection, crystals were cryo-protected by plunging them into a drop of reservoir solution supplemented with 20-25% glycerol, then flash frozen in liquid nitrogen.

Diffraction data were collected at 100-110 K at beamline 8.3.1, Advanced Light Source, Lawrence Berkeley National Laboratory, USA. Data indexing, integration, and scaling were conducted using MOSFLM [64] and the programs implemented in the ELVES software suite [**65**]. The crystal structure was determined by molecular replacement using as a search model ligand-free KstR (PDB ID 3MNL), and using the automated molecular replacement program package implemented in CCP4 software suit. [66] The coordinates were built using the BUCCANEER **[66, 67**] and COOT [**68**] programs. Refinement was performed by using REFMAC5 software.[**66, 69**] Data collection and refinement statistics are shown in Table 2.

Surface plasmon resonance studies with KstR **and corresponding DNA operators and potential ligand substrates**

Surface plasmon resonance experiments (SPR) were carried out using the Biacore T100 (GE Healthcare) optical biosensor. The intergenic DNA fragments used for SPR experiments were purchased through Sigma (Table 1) and annealed using the technique previously mentioned for the EMSA studies. The coupling of streptavidin to a CM5 Series S Sensor Chip was performed with successive 3-min injections of 2 mg/ml streptavidin in 10mM sodium acetate (pH 4.5). Prepared biotinylated DNA fragments were injected through flow cells at 200 ng/ml until approximately 40 Resonance Unit (RU) was reached. KstR binding to DNA fragments were carried out at 25 $^{\circ}$C in HBS-EP+ (0.01 M HEPES pH 7.4, 0.15 M NaCl, 3 mM EDTA, 0.05% v/v Surfactant P20). The KstR protein-DNA interaction was determined by injection of protein in a concentration series (0.5μM-0.0039μM) at a flow rate of 50 µl/min for 7 minutes. For protein-ligand interaction runs protein was injected at a fixed 0.032μM concentration with titrating amounts (0.5μM-0.0039μM) of 3OCh(25R)26A and 3OCh(25S)26A present. Regeneration conditions were performed with 1M NaCl in 50mM NaOH at 50 µl/min for 10 s between each titration.

Data analysis for the binding studies was conducted using Biacore T100 software. The buffer injection responses from the reference cell were subtracted in order to account for any non-specific binding present. Equilibrium dissociation constants (K_D), experimental data were fitted using a minimum of three data sets for each condition for 5s before the association phase ended. Graphpad Prism 6.0 was used to curve-fit the data.

2.2   Results

Specific Aim 1: Identification of the physiological inducer ligand(s) of *Mtb* KstR.

KstR was successfully purified and able to be obtained in solution without precipitation. The ability for KstR to be purified was a trial and error as the protein was only able to be obtained with a shorter incubation time of 3 hours. Initially expression was tried overnight, but the result was a series of degraded proteins that was unusable. Further purification seemed unnecessary as the protein showed to be in a single band without the appearance of any extraneous bands of degradation products (Fig. 4).

It has been reported that 3-oxo-4-cholestenoic acid and 3-oxocholest-4-en-26-oyl-CoA are capable of binding to KstR to inhibit protein-DNA complex formation. [52, 70] The *kstR* operator is a palindromic consensus sequence that offers auto-regulation of *kstR* and would confer a higher binding affinity to the regulator rather than an operator that is only partially palindromic in nature such as the operator for *Rv3545c* which lies upstream of the *igr operon*. To determine the ability for 3OCH(25R)26A to inhibit KstR-DNA complex formation on different operators in concordance with protein-DNA binding affinities, EMSA studies were performed with both *Rv3545c* (*igr operon*) and *Rv3574* (*kstR*) operator sequences to show the interaction of the KstR-DNA in the presence of either 3OCH(25R)26A or 2OCH(25R)26-CoA. (Fig. 5, A and C) Testing the inhibition of binding between the two operators confirm 3OCH(25R)26A is capable of releasing KstR from the *igr* operator (Fig. 5, A, B) much more than that of the *kstR* operator (Fig. 5, C, D) but full release is not achieved at 100 µM for either of the operator sequences. The EMSAs displayed the formation of the multiple nucleo-protein signals present as the concentrations of 3OCH(25R)26A are increased for both the *Rv3545c* and *Rv3574* operators. (Fig.
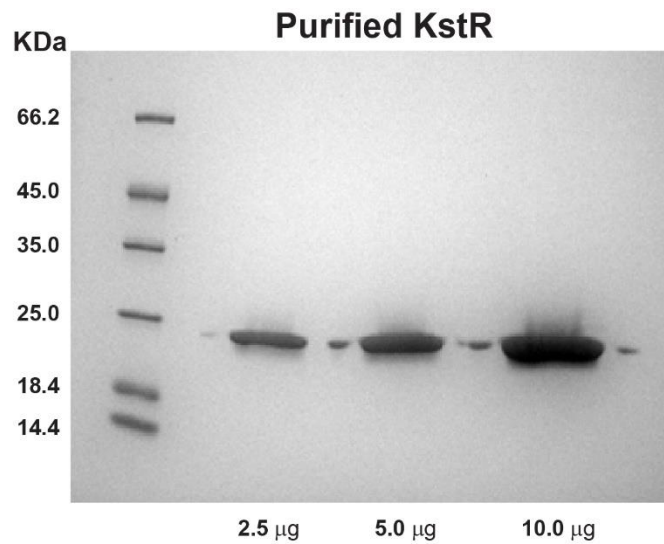
Figure 4: Purfied KstR after dialyzing and concentration. SDS PAGE gel showing increasing concentrations of purified KstR. This figure denotes the cleanliness of the purification of KstR and validated the use for the recombinant protein for the experiments in this study.
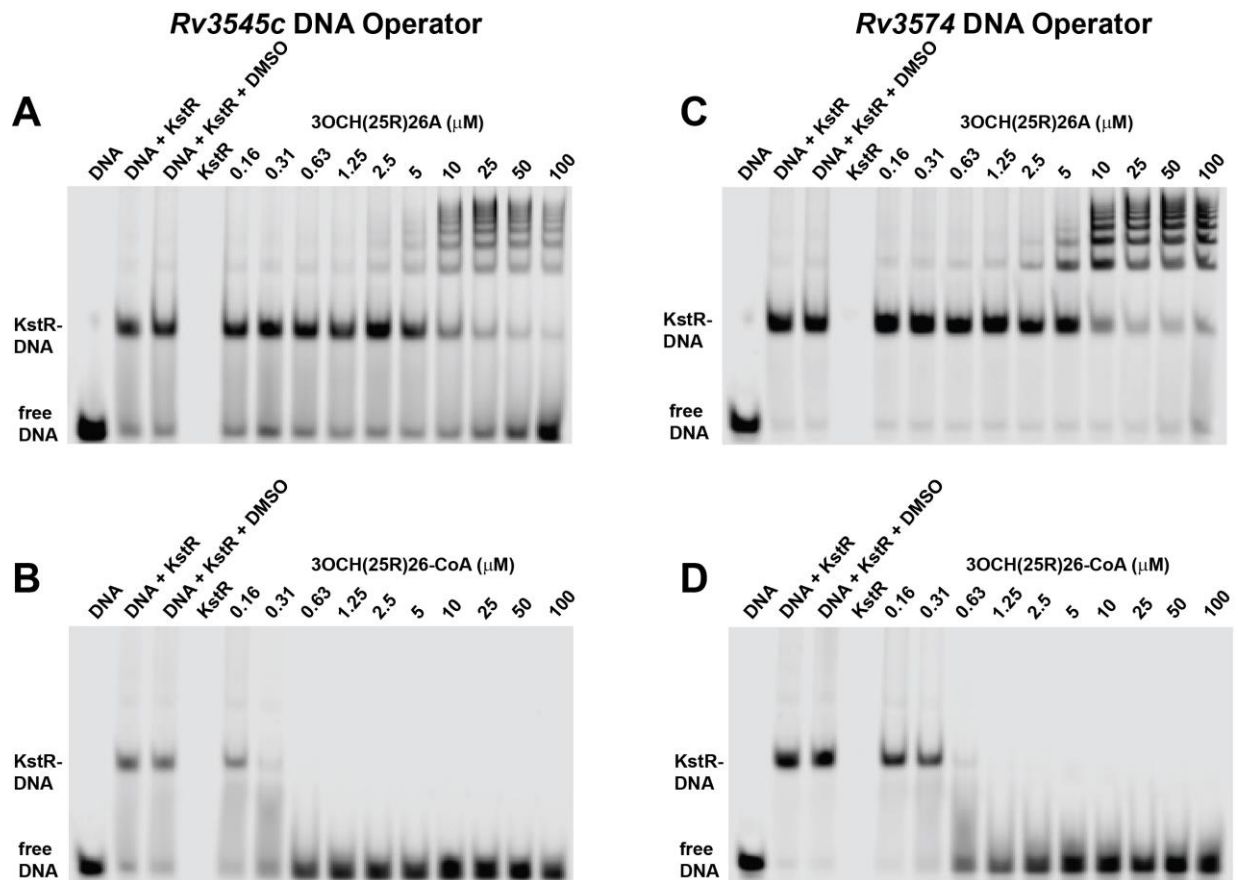
Figure 5: Electrophoretic mobility shift assay (EMSA) analysis of the inhibition of 3OCH(25R)26A and 3OCH(25R)26-CoA on known KstR DNA operators. EMSA analysis indicates the ability for 3O25R26CoA to release binding at lower concentrations than its 3O (()(25R 26oic acid counterpart. Binding reactions contained 10 nM *Rv3545c* **(A, B)** or *Rv3574* **(C, D)** DNA operator, 100 nM of purified KstR and increasing concentrations of 3OCH(25R)26A **(top panels)** or 3O25R26CoA **(bottom panels)**.

5, A, C) It also appears as a result of the increased number of KstR-DNA complexes the resulting signal for the normally occurring KstR-DNA is diminished as a consequence.

To verify the binding affinities of KstR to the different DNA operators, titrating amounts of KstR flowed over biotinylated DNA operator fragments for both *Rv3545c* and *Rv3574* for SPR experiments proved KstR to have a much higher affinity to its own operator (6.486 nM K_D) than that of the lesser palindromic sequence operator for the igr operon (43.76 nM K_D). (Fig. 6) As the *kstR* operator showed to have the higher binding capability, this operator was used for all future experimentation.

Once the concentrations of DNA and protein were acquired the binding capacity of the protein-DNA complex was tested with synthesized cholesterol catabolic pathway intermediates. These intermediates represent substrate products from the initial side chain degradation enzymes after the attachment of CoASH. As depicted in Fig. 7, the only intermediates to fully ablate the formation of KstR-DNA operator complex were the 3OCh(25S)26-CoA and 3OCh(25R)26-CoA. The 3β-hydroxy-5-cholen-24-oyl (5C-CoA) was able to induce release but the signal for KstR-DNA complex remained slightly. It was noted the hydroxyl group at the 3 position of the cholesterol rings for 23,24-bisnor-3β-hydroxy-5-cholen-24-oyl-CoA (5-BNC-CoA) and 5C-CoA not been altered to a ketone as the other substrates. It was decided to test whether changing the 3-hydroxy to a 3-keto was able to prevent KstR-DNA formation. To modify the hydroxyl group to that of a keto group, the addition of cholesterol oxidase and catalase was added to the binding reaction and allowed to incubate for the normal amount of time. Changing the hydroxyl to a keto group at the 3- position allowed the release of KstR from the DNA but not as well as that of the 3OCh(25)26-CoAs. After the addition of cholesterol oxidase 5-BNC-CoA proved to bind but only allowed for a slight release, yet the 5C-CoA was able to significantly change and prevent binding
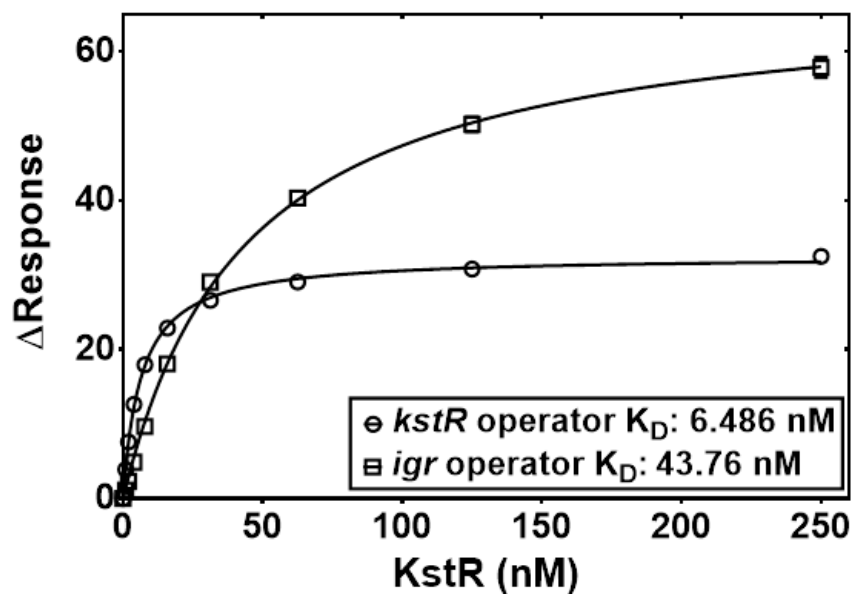
Figure 6: Surface plasmon resonance to examine the binding affinities between Rv3545c (*igr* operator) and Rv3574 (*kstR*) DNA operators. Biotin labeled DNA operators were coupled to the surface of a CM5 chip and KstR was introduced in increasing concentrations. KstR, shows a stronger interaction with the *kstR* operator than the *3545c* operator.
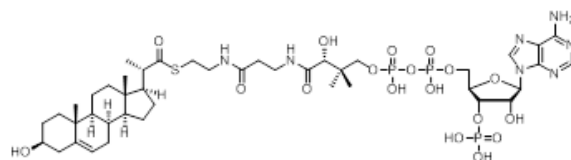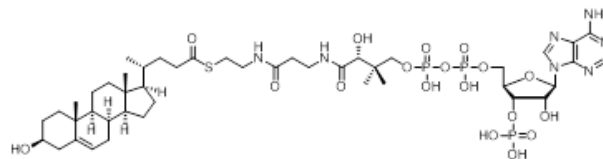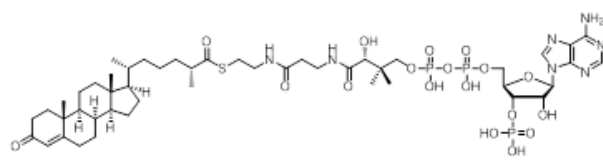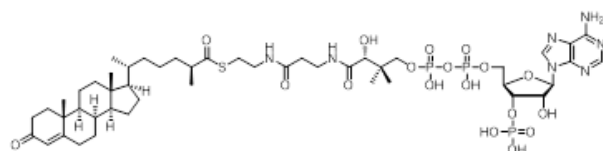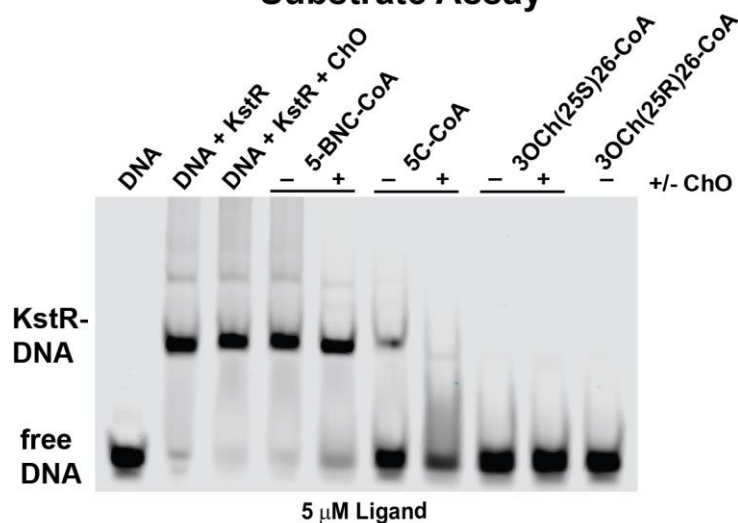
Figure 7: Electrophoretic mobility shift assay (EMSA) analysis of the inhibition of KstR - Rv3574 operator interaction upon addition of 3-OH steroyl-CoAs after conversion to their 3-oxo derivatives. Analysis indicates the failure of free CoASH to increase the poor effector activity of 3O(25R)26-oic acid. Binding reactions contained 10 nM *Rv3574* operator, 100 nM of purified KstR 5 µM 3-OH-steroyl-CoA and cholesterol oxidase (ChO).

to occur. Both outcomes proved that the 3OCh(25)26-CoA ligands remained the best suited substrates for the remainder of our studies.

Crystal structures of the (25S)-3-oxocholest-4-en-26-oyl-CoA -bound and DNA-bound forms of KstR were determined to resolutions of 1.66 Å and 2.26 Å (Fig. 8) This structure defines a range of the conformational changes between inactive, (25*S*)-3-oxocholest-4-en-26-oyl-CoA-bound KstR repressor and its active operon-bound form. Overall, KstR is a two-domain α-helical protein functioning as a homodimer. The three N-terminal helices, α1-α3, comprise the DNA binding domain (DBD) with α2 and α3 constituting classical helix-turn-helix (HTH) motif. The helices from α4 through α9 constitute the steroid binding site. The helices α8 and α9 from each monomer combine in four-helix bundle to build a dimer interface (Fig. 8).

Binding to the DNA sequence occurs through the KstR conformation more compact than that of the 3-oxocholest-4-en-26-oyl-CoA-bound KstR (Fig. 8 A, B). This conformation is achieved via increase of the buried dimerization surface from 1197 $Å^2$ in KstR-ligand complex to 1365 $Å^2$ in the KstR-DNA complex accompanied by large scale relocation of the DBD of each monomer toward the dimer interface. The distance between the arbitrary selected reference points, C□ carbon of R61, drops from 40 Å in the KstR▪3-oxocholest-4-en-26-oyl-CoA complex to 26 Å in the KstR▪DNA complex. As judged from the superimposed A-chains of both structures (Fig. 8, C), the largest conformational changes occur at the dimerization interface, including helices □8 and □9, while DBDs relocate as rigid bodies so that the HTH motifs of two monomers are spaced precisely to fit the large DNA grove, one helical turn apart from each other. The minor groove is not recognized by KstR. (Fig. 9).

Through analyzing our own crystallography data, it was noted that although the (25R) stereoisomer conformation of 3OCH26-CoA was introduced for the experiment, the ligand is

presented in the (25S) conformation in the structure itself. Despite the notably higher resolution achieved in this work, the CoA portion of the 3-oxocholest-4-en-26-oyl-CoA molecule is not visible in the electron density map (Fig. 10). This also coincides with results from the previous crystallography study performed. [70] Nevertheless, an isolated peak of the residual F_o-F_c electron density (purple mesh in Fig. 10) in vicinity of a cluster of the positively charged residues, K120, R123 and R127, in the protein surface within the reach of the CoA diphosphate group, evidence a possibility of electrostatic contributions of the CoA diphosphate moiety into binding affinity of 3-oxo-cholest-4-en-26-oyl-CoA (Fig. 10).

A previous study has shown the 3-oxo-4-cholesten-26-oyl-CoA cholesterol catabolic intermediate as being the best suitable ligand for KstR-DNA complex dissociation.[70] However, during the crystallography studies the absence of the CoA attachment to the sterol portion of the molecule was noted. To better understand the necessity of the attachment of the CoA to the 3-oxo-4-cholestenoate, competition electrophoretic mobility shift assays were performed with a homogenous solution of a fixed concentration (50 µM) of free CoASH and a titration (2.5-100 µM) of 3OCH(25R)26A as a possible solution to this circumstance. (Fig. 11, A) The EMSA showed the slight release of KstR after the addition of higher concentrations of 3OCH(25R)26A but this was to be expected as the prior EMSA studies have shown the slight release of KstR after the introduction of high concentrations of 3OCH(25R)26A. (Fig. 11, A, C) It did not show the release of DNA as has been seen with the CoA attachment of 3OCH(25R)26-CoA.

31
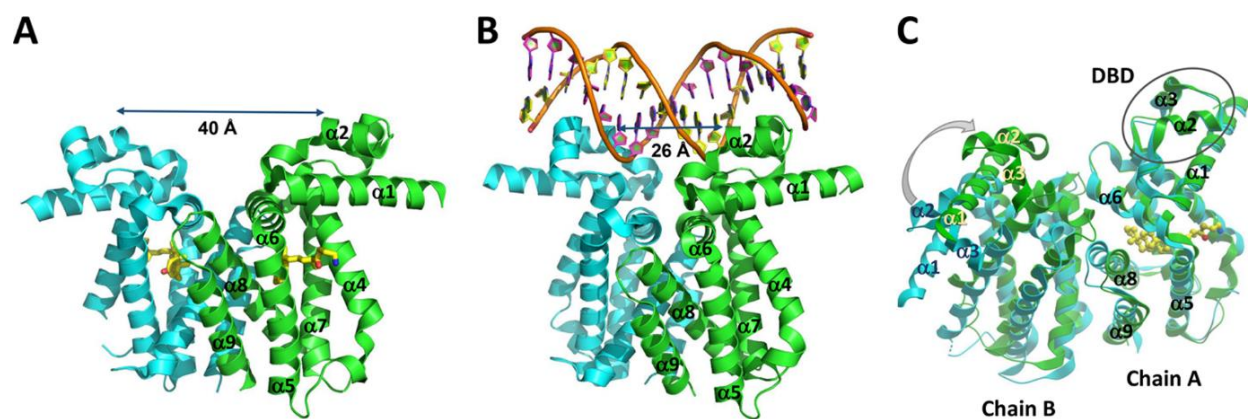
Figure 8: Overall structures of the (A) KstR▪(25S)-3-oxocholest-4-en-26-oyl-CoA (PDB ID 5AQC) and (B) KstR▪DNA (PDB ID 5FMP) complexes.
In (A) and (B) protein homodimer is shown as a ribbon colored by the polypeptide chains A (cyan) and B (green). The α-helices are numbered in the green chain. Double-stranded DNA in (B) is shown as a worm-like cartoon with the attached bases and sugar rings. The (25*S*)-3-oxocholest-4-en-26-oyl-CoA moiety in (A) is in stick mode with carbon atoms highlighted in yellow. Distances in Angstroms are between the Cα carbons of R61, chosen as a reference point. (C) The DNA-bound KstR (green) and 3-oxocholest-4-en-26-oyl-CoA-bound KstR (cyan) are shown with the A chains overlapped. The 3-oxocholest-4-en-26-oyl moiety (yellow sticks) is shown bound to the A chain. The arrow points at the direction of DBD relocation upon formation of the KstR▪DNA complex.
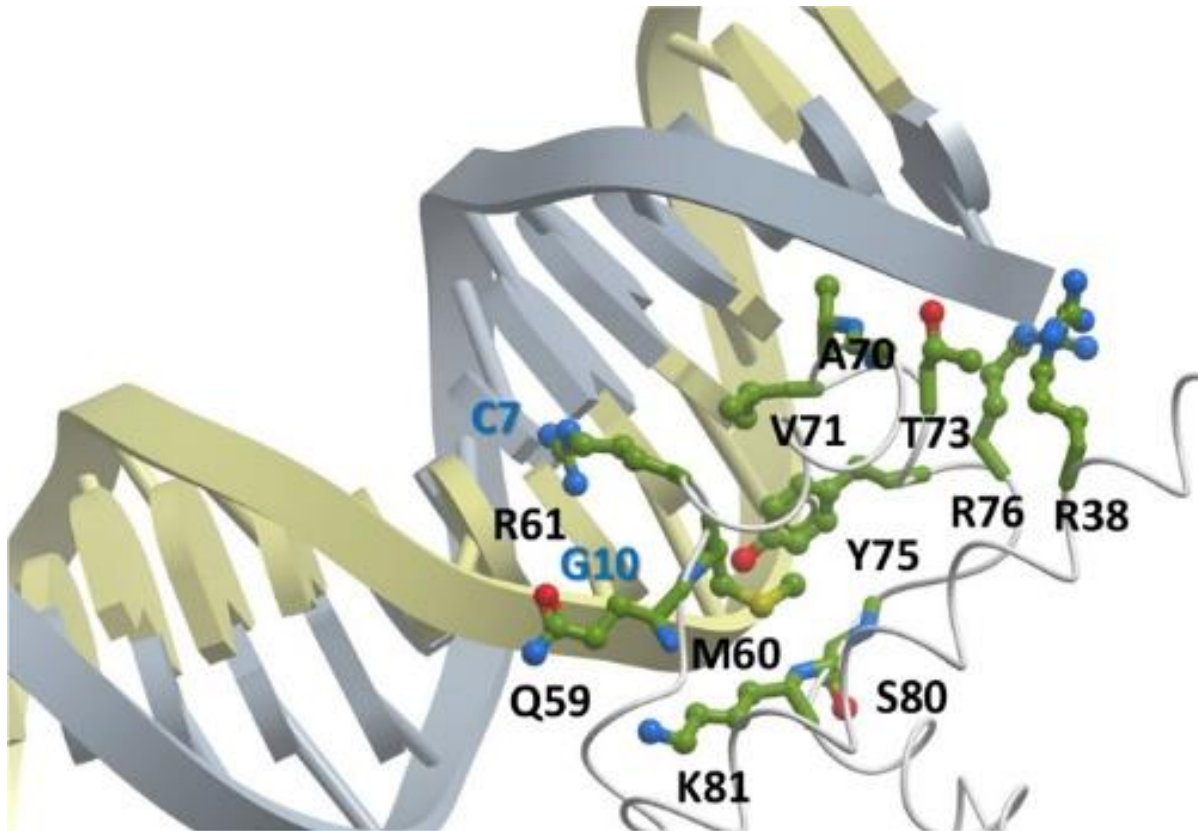
Figure 9: KstR DNA Binding Domain: The minor groove is not recognized by KstR. The only residue making H-bonding contact to a sides of the bases exposed in the major groove is R61 which interacts with G10.
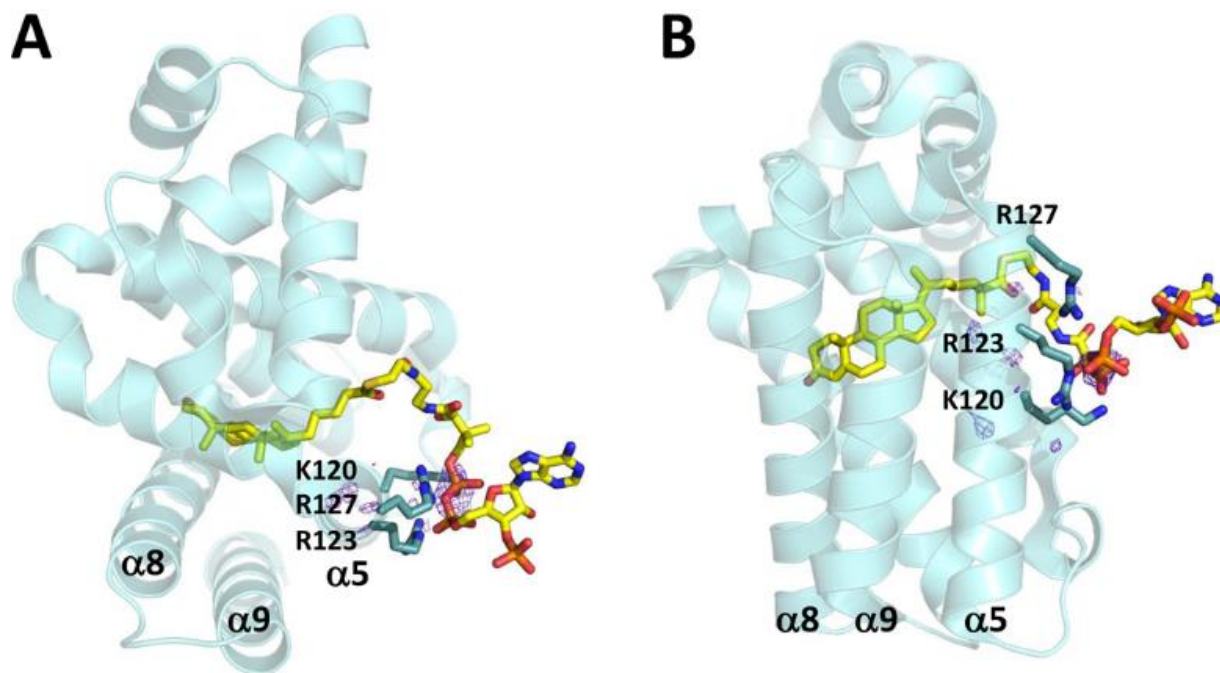
Figure 10: Modelling of the CoA moiety of 3-oxocholest-4-en-26-oyl-CoA. (stick mode: carbon = yellow, oxygen = red, nitrogen = blue, sulfur = dark yellow, phosphorus = orange) based on the crystal structure of the KstR▪(25*S*)-3-oxocholest-4-en-26-oyl-CoA complex (PDB ID 5AQC). A single protein monomer is shown as a semi-transparent cyan ribbon. A cluster of the positively charged residues, K120, R123 and R127 (stick mode with the carbon atoms in pale cyan), located on the same face of the helix α5, each separated by one helical turn, is predicted to interact with the CoA diphosphate moiety. A fragment of the residual F_o-F_c electron density map countered at 2.7 σ (purple mesh) features an isolated peak in vicinity of the clustered residues where CoA diphosphate group is fitted.

34

As the CoA attachment is required, to check if the thio-ester bond between the sterol and CoA is needed for the release of the protein from the DNA, an endogenous ethyl-ester intermediate 3OCH(25S)26-oate was extracted from a grown culture of CDC1551 and used as a ligand for EMSA experiments. (Fig. 11, B, C) The 3OCH(25S)26-oate was unable to release KstR from the operator at increasing concentrations of up to 25 µM. (Fig. 11, B) A competition assay was also done with 3OCH(25R)26-CoA at a constant concentration of 2.5 µM and increasing amounts of 3OCH(25S)26-oate ranging from 0.195-25 µM to establish if the release of the KstR-DNA complex could be prevented. (Fig. 11, C) The ethyl-ester derivative was unable to compete with the thio-ester and 3OCH26-CoA remained the preferential ligand.

As with the previously mentioned work, and our own, crystallography studies have shown the preference for the (S) stereoisomer conformation of the endogenous 3-oxo-4-cholestenoyl-CoA (3OCh26CoA). In order to synthesize both the (S) and (R) stereoisomers, (Scheme 1) 3-oxo-4-cholestenoic acid (3OCHacid) was used for the attachment of CoA using the techniques mentioned in this study. Purified recombinant KstR was used in surface plasmon experiments to determine the kinetics for inhibition of binding to the DNA operator.

Titrating concentrations of the -S (Fig 12, A) and -R (Fig 12, B) stereoisomers of 3OCH26CoA (0.976-250nM) in combination with a constant KstR concentration (6.25nM) flowed over the *kstR* DNA sequence showed that both conformations of the ligands prevented the binding of KstR to the DNA. (Fig. 12, A and B) The 3OCh(26S)CoA ligand showed the better inhibitor requiring a slightly lesser concentration preventing complex formation than the 3O25R26CoA counterpart (Fig. 12, C)
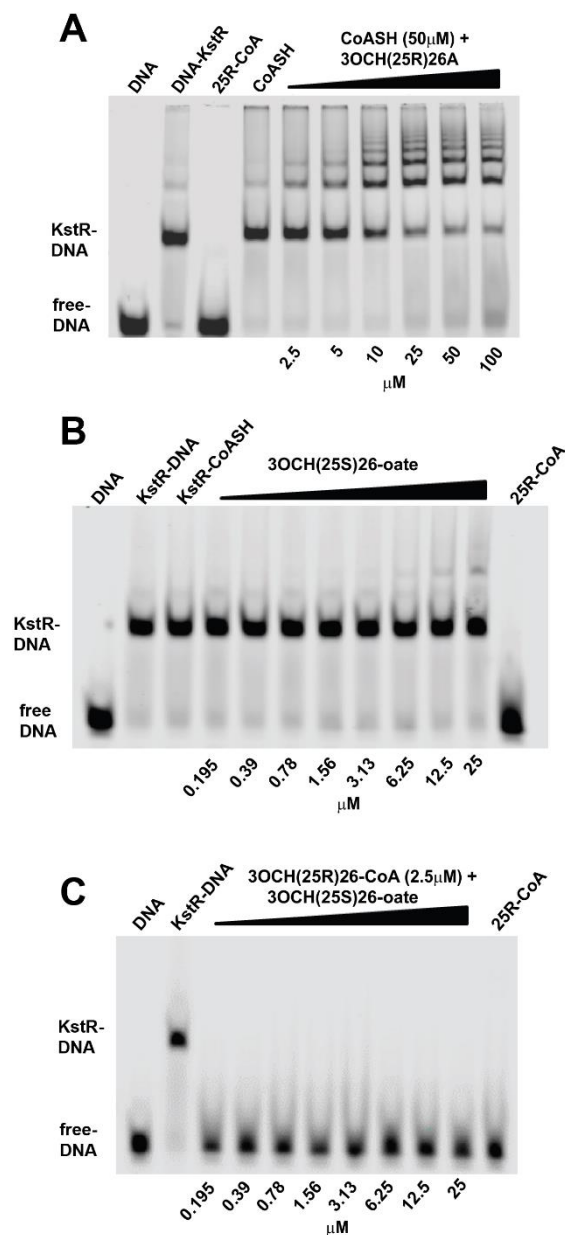
Figure 11: Inability of 3OCH(25R)26A in the presence of excess CoASH and 3OCH(25S)26-oate to prevent KstR-Rv3574 DNA operator interaction. Electrophoretic mobility shift assays were performed with 10 nM IRDye700-labeled *Rv3574* DNA operator, 100 nM purified recombinant KstR. **(A)** Increasing concentrations of 3OCH(25R)26A, and 100 µM CoASH, added when indicated**. (B)** Titrating amounts of 3OCH(25S)26-oate was added to detect inhibition of binding. **(C)** A constant amount of 3OCH(25R)26-CoA (2.5 µM) and increasing amounts of 3OCH(25S)26-oate was introduced to determine competition of binding.
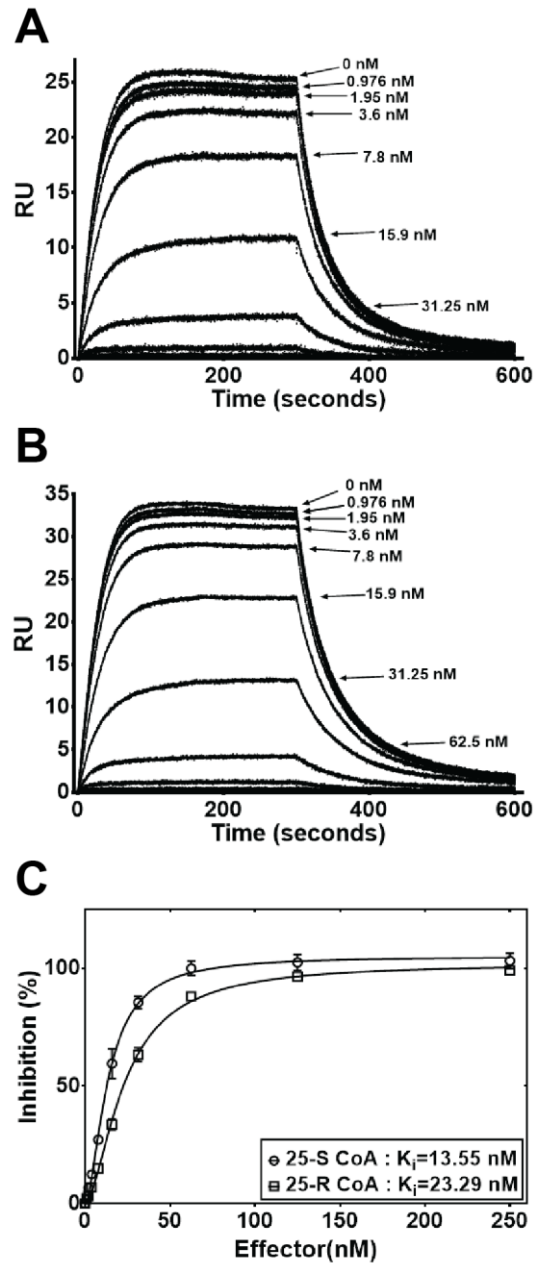
Figure 12: Analysis of the interaction of KstRMtb with Rv3574 DNA operator and effector ligands measured by SPR. Sensograms showing the serial dilutions of 25S **(A)** and 3OCH(25R)26-CoA **(B)** being flowed over a chip coated with biotinylated *kstR* operator with a constant KstR protein concentration of 6.25 nM. **(C)** Percent inhibition data showing the 3OCH(25S)26-CoA ligand is capable of preventing formation of KstR-DNA complex slightly better than the 3OCH(25R)26- CoA.

37
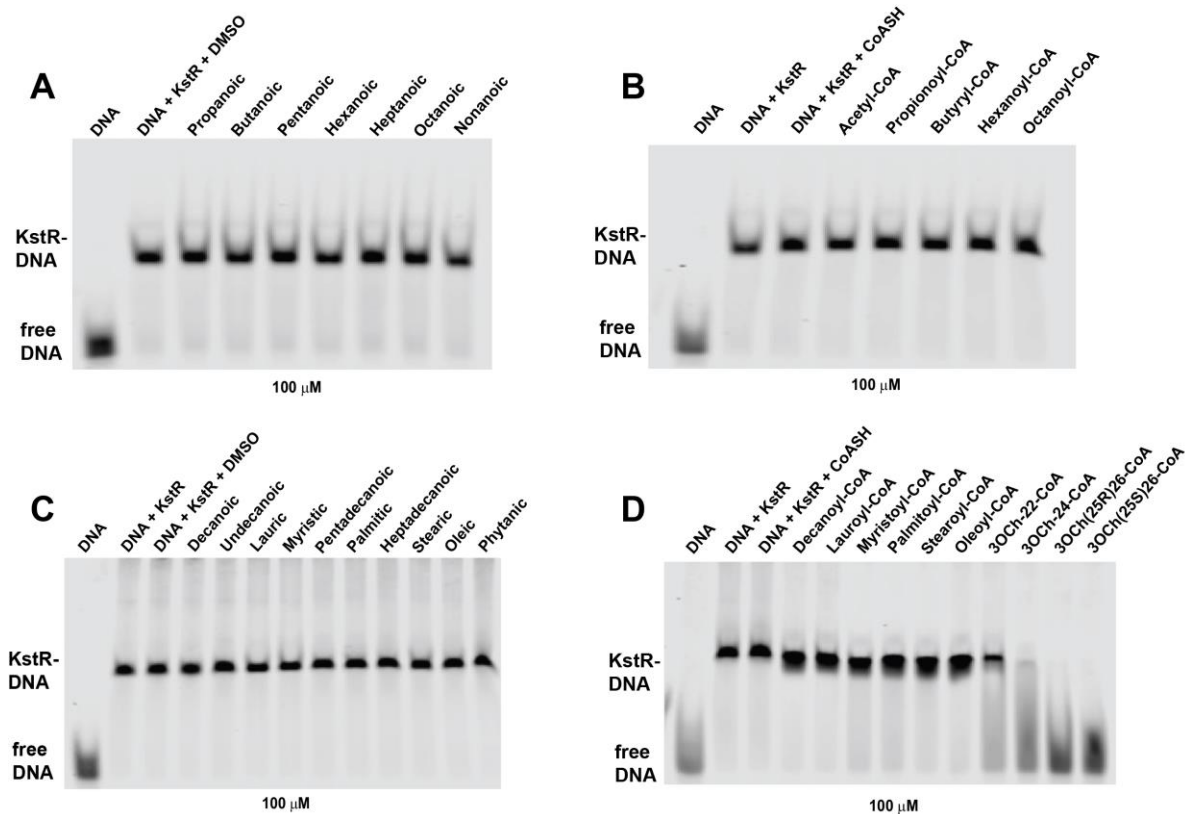
Figure 13: Inability of free fatty acids (C10-C18) and their CoA ester derivatives to prevent KstR -Rv3574 DNA operator interaction. EMSA analysis indicates that even at 100 µM the fatty acids and their CoA derivatives are very poor effector ligands. Binding reactions contained 10 nM Rv3574 operator, 100 nM of purified KstR, 100 µM of fatty acids, 100 µM fatty acyl CoAs or 5 µM steroyl CoAs. Electrophoretic mobility shift assays were performed 10 nM IRDye700-labeled Rv3574 DNA operator), 100 nM purified recombinant KstR, increasing concentrations of 3O(25R)26A, and 100 µM CoASH, when indicated.

Additional EMSA experiments with both fatty acids and fatty-CoAs were carried out to ascertain whether the cholesterol rings were required for prevention of KstR to bind to the DNA. None of the fatty acids or their respective CoA counterparts were able to fully inhibit KstR-DNA complex formation at what would be considered high physiological concentration of $100\,\mu$M. (Fig. 13) There was a slight noticeable shift in complexes where the fatty-CoAs are present (Fig. 13, D), and a slight amount of DNA can be seen being released where myristoyl-CoA palmitoyl-CoA are present in the binding reaction. This was not completely unexpected, but if they were to serve as viable inducer ligands, complex release would have occurred at these higher concentrations.

Specific Aim 1 Sub-Aim: Distinguishing if phosphorylation is capable of providing an additional regulatory mechanism signal for de-repression of KstR through *in vitro* methods.

The designing of primers, cloning, and expression were carried out for the newly designated KstR protein encoding for the additional 20 amino acids at the N-terminus. This was also completed with the replacement of alanine or aspartate amino acids mutations at each of the serine and threonine sequences within the first 10 amino acids. The expression of these proteins were carried out in the same manner the previous truncated KstR was expressed with comparable yields. These newly expressed variants of KstR as well as the previously expressed truncated KstR were used for SPR experiments testing if there would be differences with alterations in binding mechanisms with the presence of the extension as well with the mutated N-terminus. The previously used KstR protein for the experiments in the first part of this aim will now be denoted as TR-KstR.

Surface plasmon resonance studies were carried out with the same protocol as the previous experiments as to not introduce any variations in results. The binding of the four different variants were used to flow over biotinylated DNA bound to the flow cell of a CM5 chip with increasing

concentrations and the binding affinities were measured. The TR-KstR showed to have a similar binding affinity as with previous experiments with a binding affinity to the *kstR* DNA operator of 5.861 nM. Interestingly, the KstR and Ala-KstR recombinant proteins, the latter preventing the ability for any phosphorylation to occur at the N-terminus, bind to the DNA with similar affinities, 12.26 nM and 12.62 nM respectively. The Asp-KstR phosphor-mimetic recombinant mutant showed an even lesser binding affinity to the DNA, 16.25 nM, although the inhibition of binding is not very dramatic. When analyzed for statistical differences, the recombinant KstR proteins with the N-terminal extension are statistically different from that of the TR-KstR showing the TR-KstR is able to bind to the DNA at a higher rate and denoting the most difference between that of the TR-KstR and the Asp-KstR mutant. There was no discernable statistical difference between any of the recombinantly expressed proteins that have the 20 amino acid extension.

All four of the KstR protein variants were then tested to determine the inhibition of binding to the *kstR* DNA operator in the presence of 3OCh(25R)26-CoA. Interestingly, the binding of the TR-KstR to that of the extended N-terminal KstR and the Ala-KstR all show similar inhibition when presented with the ligand. On the other hand, the inhibition of the Asp-KstR mutant was significantly increased to an affinity similar to that of the 3OCh(25S)26-CoA inhibition affinity with the TR-KstR performed in previous experiments. (Figs. 14 and 14, C)
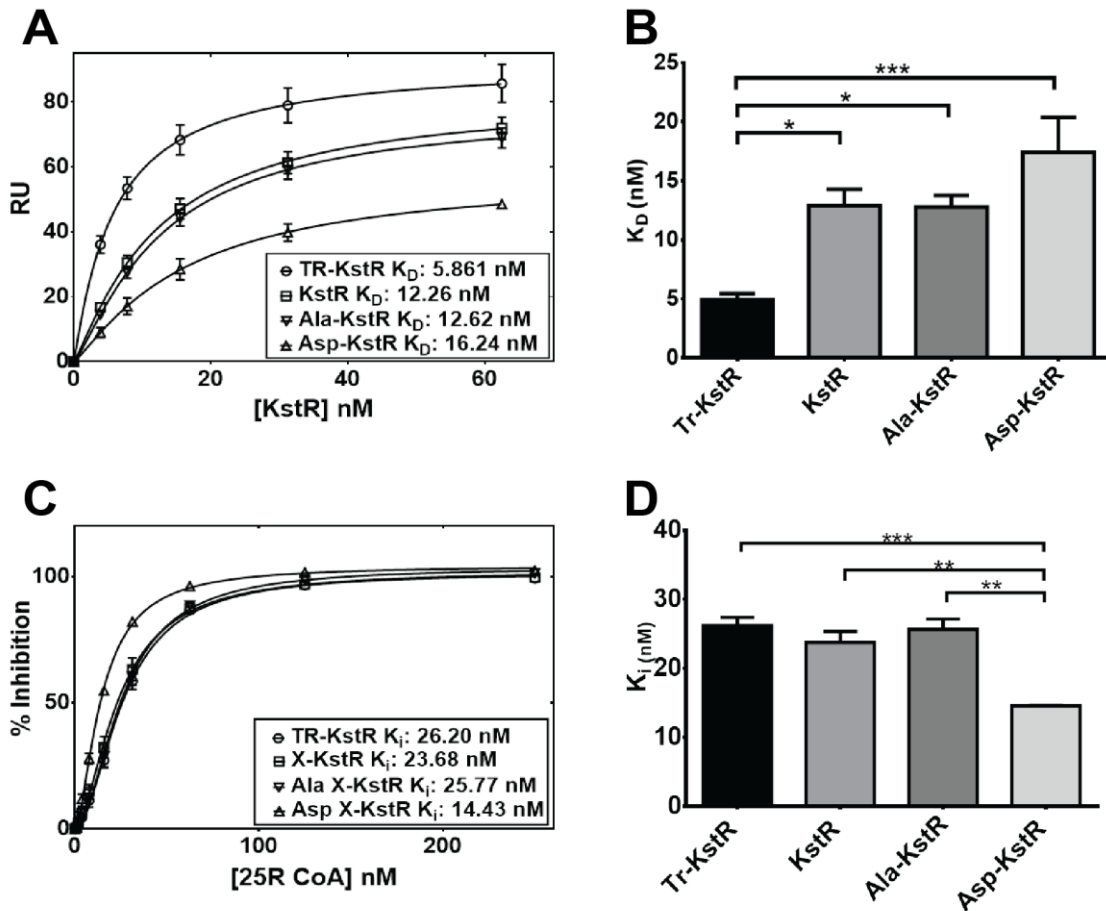
Figure 14: KstR binding and inhibition to DNA with N-terminal amino acids. (A) Surface plasmon resonance affinity studies for the truncated, full length, a phosphor-mimetic aspartate mutant, and phosphor-ablative alanine mutants of KstR. Biotinylated kstR operator DNA was bound to the chip surface and increasing concentrations of KstR variants were introduced to measure affinity. The truncated KstR has a higher affinity for the DNA (5.861 nM) and is able to bind better than the extended (12.26 nM). (B) The significance in DNA-binding differences between the full length and mutants variations of KstR. (C) A homongenous solution of KstR (6.25 nM) and increasing amounts of 3OCH(25R)26-CoA were flown over the coupled kstR operator and measured for inhibition. (D) Statistical analysis of the differences for the inhibition of KstR-DNA complex formation.

2.3 Discussion

The data obtained by our studies have confirmed that the capability of 3OCH26CoA acting as the best suitable ligand for the negative transcriptional regulator KstR at very low-level concentrations. The *in vitro* verification of this information coincides with the production of 3OCH26CoA as a known physiological substrate produced in early stages of the *Mtb* cholesterol catabolic pathway via the enzyme FadD19. The production of 3OCH26CoA at an early stage of cholesterol degradation would allow the initiation of transcription thus generating the enzymes required to break down other cholesterol and fatty acid derivatives built up in the bacterial cell that could cause potential toxicity. However, it was observed that the KstR does not bind with the same affinity strengths for different operators.

The ability for KstR to bind to a multitude of similar but not identical DNA operators suggests the cellular concentrations of potential ligands is more than likely at much lower concentrations that were observed in the our EMSA and SPR studies. The cytochrome P450 Cyp125 is the second step in the cholesterol degradation step and responsible for the oxidation of C26 of creating the 3OCH26oic acid precursor required for the turnover into 3OCH26CoA by FadD19.The upregulation of key enzymes in the degradation pathway, such as Cyp125 under the control of the *igr* operator used in our studies, showed a much weaker binding affinity to the KstR protein. (Fig. 5 and 6) This could also be translated as easier detachment of KstR from the *igr* operator with lesser concentrations of ligand. This dynamic was also seen in the EMSA with KstR and the *igr* operator. At higher concentrations of 3OCH26oic acid, was able to release KstR from the *igr* operator compared to the *kstR* operator. (Fig. 5) This coincides with previous studies performed with another operator sequence in *M. smegmatis* located upstream of the first enzyme in the cholesterol catabolic pathway 3-□-hydroxysteroid dehydrogenase/isomerase (3-□-HSD). **[51]**

42

This study showed 3OCH26oic acid was able to detach KstR from the *3-□-HSD* operator at relatively high levels and therefore was presumed to be the ligand for KstR. It is important to note that a *kstR* operator sequence upstream of *3-□-HSD* in *Mtb* is not present, and therefore the *igr* operator presented a more suitable model for our studies for the weaker binding operator.

Being that the substrate for FadD19, 3OCH26oic acid, was tested as a potential ligand and failed to show the same release of KstR from its DNA operators with and without the presence of detached CoASH, the direct attachment of CoA appears to be a requirement for the rapid conformational change in KstR. (Fig. 11) The crystal structure data failed to represent CoA structure itself bound to the surface of KstR, although there are two distinct arginine residues that may be capable of nesting the CoA adenosine moiety (Fig. 10). Interestingly enough, the 3OCH26CoA used for crystallization was in the R- stereoisomer conformation. The crystal structure denotes the 3OCH26-CoA in the S-conformation. The previous crystallographic studies done with KstR and 3OCh26-CoA also present the ligand in the S-conformation. [70] The method by which they obtained the ligand, it would have been in the R-stereoisomer form when presented to the protein for crystallization. The experiments in the previous work obtained 3OCh26-CoA through metabolic activity using recombinantly expressed FadD19 as the enzyme and commercially obtained 3OCh26-CoA and CoASH. FadD19 is known for the attachment of CoA to 3OCh26A in either stereoisomer conformation but not possess the racemization activity. This is surprising and could be caused from the positioning of the thioester bond between the cholesterol and CoA in the binding pocket once the ligand has been bound. It is suspected the thioester is reactive within the pocket with residues yet to be determined. This reactivity could initiate a shift in conformation from R- to S- which is the natural physiologically occurring cholesterol catabolic substrate conformation in *Mtb*. The introduction of an ethyl ester rather than a thioester bond

linking the acid and CoASH was also tested with the hopes in determining if the thioester bond was a requirement for KstR –DNA complex inhibition. The ethyl ester bond is non-reactive and less labile than that of the thio-ester and the inability for the substrate to act as an inducer ligand further suggests the site of the bond provides and energy to the pocket for the conformational change to occur. The requirement of the thioester bond in 3OCH26CoA also supports the EMSA results showing 3OCH26oic acid does not dissociate KstR and may be more of a necessity than that of the actual CoA itself.

The ability for all ligands binding to KstR in these studies has not been completely confirmed. It has already been established both 5BNC-CoA and 3OCh26-CoA are able to bind to KstR through crystallographic studies published by the Eltis group as well as our own unpublished crystollography. [70] The EMSA studies performed would suggest multiple ligands have the capability of binding to KstR but not have the ability for inhibiting KstR to form a protein-DNA complex. What was determined were the requirements for inhibition of complex formation. For inhibition of KstR to binding to the DNA, the ligand must possess a ketone at the 3- position of the cholesterol ring portion and a thio-ester attachment of CoA. Of the ligands tested, cholesterol acid, ethyl ester, ethyl ester-CoA, multiple fatty acids, and multiple fatty-CoAs, none of them were able to inhibit protein-DNA complex formation to an extent that proves they are inducer ligands. What was witnessed was the introduction of multiple bands of KstR-DNA complexes and shifts in the size of the complex formed. This suggests, if not confirm, that a majority of the ligands tested in this study are able to actually bind to KstR, just not inhibit the KstR-DNA complex from forming.

Looking at the binding pocket where the cholesterol moiety is nestled into KstR exhibits no plausible reason for the 3OCh(25)26A to not be capable of binding to KstR in the same manner.

(Fig. 8) The presence of the multiple bands after the introduction of the 3OCh(25)26A shows that it has the opposite effect of inhibition and increases binding in an oligomeric fashion to the DNA. This data coincides with the previous work done showing the thermos-dynamics of 3OCh(25)26A binding to KstR through isothermal calorimentric experiments. [52] Their results showed the binding temperature for KstR and 3OCh(25)26A to decrease rather than the expected rise in temperature. This would suggest that a conformational change does occur, but not in a mechanistic manner by which KstR is released or inhibited from binding to DNA. The DNA signal that is shown in the EMSAs, in what appears to be release at 100 µM of 3OCh(25)26A, may not actually be release of the DNA (Fig. 5) , but rather a sequestering of the amount of protein available to bind to remaining DNA in the reaction. This would confer a signal replicating the inhibition of binding due to the presence of ligand, but is just the inhibition caused by so many nucleo-protein aggregates forming and little protein remaining.

A growing number of studies Ser/Thr phosphorylation may serve as a mechanism for transcriptional regulators to rapidly respond and adapt to environmental changes. [56, 71, 72] PDB and uniport has listed the *kstR* gene sequence to include 63 additional nucleotides, 21 additional amino acids, upstream of the originally designated sequence due to erroneous initiation, the *kstR* gene sequence described in the genome database Tuberculist (Swiss Institute of Bioinformatics) states it is 200 amino acids in length which brings the newly denoted protein to 220 amino acids long. The additional peptide sequence at the N-terminus shows to have a serine/threonine rich region within the first 10 amino acids and this bears a similar resemblance to the Tet-R like regulator EthR. Like KstR, EthR harbors a rich serine/threonine N-terminal rich region that has been shown to be phosphorylated by at least one of the 11 kinases that are produced by *Mtb*. **[56]**

Moreover, there is increasing research documenting the ability of phosphorylation to provide an added regulatory mechanism abrogating binding for other transcriptional repressors in *Mtb*. **[73]**

To verify if this extra serine/threonine rich sequence is present with orthologs of KstR in other mycobacterium species, a multiple sequence alignment was performed. (Fig. 3) and it shows a highly conserved sequence within the first 10 amino acids rich with serines and threonines consistent with the previously mentioned EthR. When the KstR protein sequence is input into the bacterial serine threonine bacterial phosphorylation predictor site NetPhosBac 1.0, a number of serines are denoted as a possible phosphorylation site. Although the predictor only shows the 10th serine as above the calculated threshold of 0.5, the 3rd and 9th serines also conferred a higher signal and are not to be ruled out as possible phosphorylation sites. Additionally, docking studies of the additional amino acids and sequence for DNA studies were performed to assess how the interaction of the DNA and KstR may be altered, enhanced or hindered in an elongated form. Unfortunately, the results were inconclusive as to the specific binding parameters of the extension to the DNA and possible phosphorylation effects. Also noted were the other serines throughout the protein sequence that may be considered phosphorylation sites. It was realized there would be a difficulty in establishing phosphorylation of the N-terminus alone as the other potential PTM targets would give a signal for various experimentation. As growing *Mtb* in the BSL-3 environment requires the decontamination of samples, heat or other harsh methods are used to ensure the bacteria is not viable outside of containment. With this as well as the fact that PTM's are extremely labile, it proved to be a difficult task to provide experimentation with in-vivo purification of *Mtb* samples of KstR in a phosphorylated form. The daunting task of cloning point mutations of each serine and threonine at the N-terminus of KstR as well as the cloning of 10 of the 11 Ser/Thr kinases was completed, yet the fact that KstR has other possible phosphorylation sites that could impose

misleading data remained and therefore it was decided to pause the study with regards to those methods. Instead, it was decided to pursue SPR studies with different N-terminal variations of KstR.

The affinity studies completed for analyzing KstR in the truncated and elongated forms prove the extension hinders binding to the DNA operator but not by a large significant measure. The alanine mutant binding does not appear to be significantly less to KstR but nominally to the TR-KstR. By exchanging all the serines and threonines within the first 10 amino acids at the N-terminal to aspartates, the Asp-KstR mutant shows that binding is slightly less than that of the naturally occurring KstR as well as the alanine mutant, however is significantly less than TR-KstR. Reviewing the SPR, lesser binding affinity of Asp-KstR to the DNA, in a phosphor-mimetic in the presence of the 3OCh(25R)26-CoA ligand results can be deciphered as the reason for the inhibition. If KstR is able to undergo phosphorylation at the N-terminus, the ability for de-repression to occur is enhanced. Furthermore, with the evidence presenting KstR binding to DNA is relative to the palindromic sequence of the operator and weaker when the sequence is less palindromic as with the *igr* operator, the presence of the smallest amount of inducer ligand may be all that is required for de-repression of KstR to occur allowing for an additional mechanism of rapid response in transcription when cholesterol catabolism is required.

Chapter 3: Determining the impact of KstR deletion (□nockout) and depletion (knockdown) on *Mtb*'s ability to utilize cholesterol and to grow intracellularly

3.1 Methods

Mtb Strains, Media, and Culture Conditions

All experiments with *Mtb* cells were performed in a Biosafety Level-3 (BSL-3) laboratory. *Mtb* strains were routinely cultured at 37°C in 1-L roller bottles containing 100 mL 7H9 medium supplemented with 10% (*v/v*) of oleic acid-albumin-dextrose complex (OADC), 0.2% glycerol (*v/v*) and 0.05% (*v/v*) Tween-80 (7H9-OADC) or on 7H10 solid agar medium supplemented with 10% (*v/v*) OADC and 0.5 % (*v/v*) glycerol (7H10-OADC). Hygromycin (50 µg/mL), kanamycin (20 µg/mL), and zeocin (50 µg/mL) were used when necessary. For growth on defined carbon source(s), cells were first grown in liquid 7H9-OADC medium until they reached an OD_{600} of 0.8 – 1.0, washed once in minimal medium (MM: 1 g/L KH_2PO_4, 2.5 g/L Na_2HPO_4, 0.5 g/L asparagine, 50 mg/L ferric ammonium citrate, 10 mg/L $MgSO_4 \cdot 7H_2O$, 0.5 mg/L $CaCl_2$, 0.1 mg/L $ZnSO_4$, 0.1% (*v/v*) ethanol, 0.1% (*v/v*) tyloxapol), and then used to inoculate MM containing 0.2% (*v/v*) glycerol as a source of carbon at an OD_{600} of 0.1. Cells were grown until they reached an OD_{600} of 0.8 - 1.0, and then used to inoculate 100-mL cultures of MM containing either 0.2 mM cholesterol or 0.1 mM cholest-4-en-3-one at 0.02 OD_{600}. For growth in the presence of cholest-4-en-3-one, glycerol was also included at a concentration of 0.2% (*v/v*). Cholesterol and cholest-4-en-3-one were dissolved at a concentration of 100 mM in tyloxapol/ethanol (1:1), and added to

MM pre-warmed at 65°C. Growth curves were determined with 100-ml culture in 1-L roller bottles containing the MM medium and growth was monitored by measuring the OD_{600}.

Construction of the Δ*kstR* deletion mutant strain

A Δ*kstR* mutant strain was generated in the *Mtb* clinical isolate CDC1551 by homologous recombination using specialized-transducing mycobacteriophages, as described in (Jain et al, 2014). Briefly, a substrate for allelic exchange substrate (AES) at *kstR* locus was constructed by PCR-amplifying the 5' and 3' flanking regions of *kstR* from *Mtb* genomic DNA using the primers listed in Table 1. The two DNA fragments were simultaneously inserted in pYUB1471 on each side of a *hyg-sacB* cassette flanked by resolvase *loxP* sites. The resulting AES plasmid was ligated at the *PacI* site of the shuttle phasmid phAE159 and the ligation products were packaged into λ capsid heads (Gigapack III XL, Agilent Technologies, La Jolla, CA). To generate specialized-transducing mycobacteriophages, the recombinant phasmid DNA was introduced in *M. smegmatis* by electroporation and plaques were obtained at the permissive temperature of 30 ºC. After amplification, a high-titer suspension ($\geq 10^{10}$ pfu/mL) of mycobacteriophages was used to transduce *Mtb* CDC1551 at the restrictive temperature of 37 ºC. Colonies were obtained after 3-4 weeks of incubation on 7H10 plates containing hygromycin (50 µg/mL). Colonies were picked and deletion of the *kstR* gene was confirmed by Southern blotting using the 5′ flanking region of *kstR* as a probe.

Complementation of the Δ*kstR* mutants

To confirm any observable phenotype(s), the *kstR* gene with its native promoter was reintroduced in the deletion mutation strains. A 965-bp fragment encompassing the *kstR* coding region, 260 bp and 105 bp of 5′- and 3′-flanking sequences, respectively, was amplified by PCR

from *Mtb* genomic DNA using forward and reverse primers listed in Table 1. The purified amplicon was digested with *KpnI* and sub-cloned into the integrative vector pMV306:zeo. The resulting pMV306:zeo:*kstR* plasmid was electroporated into *Mtb* CDC1551 Δ*kstR* cells, and transformants were isolated after 2-3 week of incubation at 37 °C on 7H10 plates containing 50 μg/ml zeocin. Transformants were picked and grown in 15-ml cultures of 7H9-OADC medium until OD_{600} reached 1.0 -1.5 for protein extraction and western blot analysis (see below).

KstR protein knockdown and *in vitro* phenotypic characterization

To modulate the levels of cellular KstR, a protein knockdown (PKD) approach by controlled proteolysis developed by the laboratory of Dr. Dirk Schnappinger [36] was used as an alternative to *kstR* gene knockout. Briefly, the pJSC407 vector was modified to integrate a DAS+4 tag sequence upstream of the hygromycin cassette to generate pJSC407:DAS+4 plasmid. The AES was constructed by inserting the last 500 bp of coding sequences of *kstR* gene in-frame with DAS+4 tag. About 500 bp downstream of the stop codon of *kstR* was also inserted on the 3' side of *hyg* cassette. The creation of the phagemid and specialized-transducing mycobacteriophages and the generation of transductants in *Mtb* CDC1551 are as described above. Hyg[R] colonies were picked and screened by PCR and confirmed by Southern blot. To generate unmarked mutants, the *loxP-hyg-loxP* cassette was removed by electroporating the unstable pYO11:zeo plasmid, which expresses a resolvase. Zeo[R] colonies were picked and screened for hygromycin sensitivity. A Hyg[S] clone was electroporated with the pGMCKq1-10M1-sspBmyc plasmid, which allows the expression of a tagging adaptor protein, SspB, under the control of a *tetO* operator and the transcriptional repressor TetR. Clones were picked and grown in 7H9-OADC medium in the presence or absence of 150 ng/ml ATc. Cells were harvested by centrifugation for protein extraction and western blot analysis.

Protein extraction and western blot analysis

Cells were harvested by centrifugation at 3,500 rpm for 10 min. Cell pellets were washed twice with 5 ml of cold PBS and centrifuged. The supernatants were removed completely and the pellets were re-suspended in 0.5 ml PBS containing 1% SDS and 1 mM PMSF. The cell suspensions were transferred into 2-ml Eppendorf tubes containing 0.5 ml of 0.1-mm zirconia beads and lysed by 3 x 1-min cycles on a SoniBeast bead-beater apparatus (Biospec). The beads were settled by centrifugation at 1,000 rpm for 1 min. The cell lysates were transferred into 1.5-ml Eppendorf tubes with screw cap and boiled for 15 min to kill any residual *Mtb* cells before removal from the BSL-3 laboratory. The concentrations of the proteins in cell lysates were obtained using the BCA Protein Assay Kit (Thermo Fisher Scientific). Cell lysates (10 µg) were separated using 12% SDS-PAGE (BioRad) and then transferred to 0.2-µm nitrocellulose membranes (Pierce). The blots were washed with TBS solution containing 0.1% Tween-20 (TBST), blocked O/N with 5% skimmed milk prepared in TBST at 4 °C and incubated for 1 hour at room temperature with a primary custom polyclonal antibody raised against the purified recombinant KstR protein (Genscript) at a dilution of 1:2,000. Next, the membranes were washed with TBST and incubated for 1 hour at room temperature with horseradish peroxidase-conjugated goat anti-rabbit IgG antibody (Pierce) at a dilution of 1:10,000 dilution. The bands were visualized using Pierce ECL Western Blotting Substrate (Thermo Fisher Scientific) according to the manufacturer's protocol.

Extraction and purification and analysis of cholesterol-derived metabolites

For extraction of cholesterol-derived metabolites, cells were cultivated as described above. Aliquots of 15 mL were harvested by centrifugation at 3,500 rpm for 10 min. The supernatants were discarded and the cell pellets were washed with 5 mL of cold PBS. Cell pellets were re-suspended in 1 mL of LC-MS grade water and transferred into 15-mL glass tubes with PTFE-

lining caps containing 5 mL chloroform:methanol (2:1) containing the internal standard 7-ketocholesterol (10 µM) and then vortexed for 1 min. The extraction tubes were left in the refrigerator for at least 24 h to kill any residual *Mtb* cells. The samples were then taken out of the BSL-3 laboratory, vortexed for 1 min, and centrifuged for 15 min in a tabletop clinical centrifuge. The organic phase (bottom layer) was transferred into new 15-mL glass tubes, evaporated completely under a stream of N_2, and resuspended in 1.5 mL of LC-MS grade methanol.

Detergent (Tween-80 or tyloxapol) was removed from the lipid solutions by precipitation using a cobalt:thiocyanate solution and back-extraction with hexane, according to a published procedure.[74] The organic phase (upper phase) was dried under a stream of N_2, reconstituted in 0.5 mL of chloroform, and stored at -20 ℃ until use.

To measure the levels of cholesterol and the oxidized form, cholest-4-en-3-one, metabolites extracted from cholesterol-grown WT and Δ*kstR* cells were quantified by HPLC. Aliquots of 0.1 mL of extracted metabolites stored in chloroform were evaporated, reconstituted in 0.1 mL of DMSO: acetonitrile:5 mM ammonium acetate (2:5:3). 20.0-µL injections were analyzed on a Dionex Ultimate 3000 RS Pump HPLC system connected to a Kinetex C18, 100A column (50 × 2.1 mm i.d. stainless steel, 2.6 µm particles, Phenomenex). The metabolites were eluted at a flow rate of 0.5 mL/min (solvent A, 95% 5 mM ammonium acetate, 5% of acetonitrile + 0.1% formic acid); (solvent B, 5% 5 mM ammonium acetate, 95% acetonitrile + 0.1% formic acid) with a gradient starting at 30% B up to 1 min and then ramped up to 100% B over 6 min. The elution was maintained at 100% B up to 15 min and then ramped back to 30% B within 1 min, followed by equilibration at the same composition for 6 min before the next run. The eluted metabolites were analyzed with a charged-aerosol detector (CAD) with data collection rate at 2.0 Hz with a 5.0 sec

filter for peak shape. The retention times of cholesterol, 7-ketocholesterol, and cholest-4-en-3-one were determined by injecting pure commercial standards.

^{14}C-labeling of cellular lipids and thin-layer chromatography

Mtb cultures were grown in 7H9-OADC until OD600 reached 1.0-1.5. Cultures (10 mL) were harvested by centrifugation at 3,500 rpm for 10 min, and washed twice with 7H9 base to remove traces of carbon sources. Cells were resuspended in 1 mL of 7H9 base and used to inoculate 30-mL ink bottles containing 10-ml 7H9 medium containing 0.05% Tween-80 at an OD_{600} of 1.0. Cultures were incubated for 24 h to mimic starvation and then spiked with 4 µCi of [1-^{14}C] sodium propionate. After an additional 24 h of incubation, cells were harvested by centrifugation at 3,500 rpm for 10 min. To determine the extent of ^{14}C-propionate incorporation into cellular lipids, a 0.5-mL aliquot of the supernatant was kept and mixed with the same volume of 10% formalin for removal from the BSL-3 laboratory. The supernatants were discarded and cells were washed with 2 mL of cold PBS. After discarding the supernatants, cell pellets were reuspendend in 2 x 1 mL of methanol and transferred into 15-mL glass tubes containing 4 mL of chloroform. Tubes were vortexed for 1 min and left in the refrigerator for at least 24 h to kill any residual Mtb cells. Tubes were removed from the BSL-3 laboratory and centrifuged at maximaum speed for 5 min in a clinical centrifuge. The organic phases were collected and transferred into clean 15-mL glass tubes. The lipid extracts were evaporated under a stream of N_2 and cleaned to remove Tween-80 as described above. At the end, the radiolabeled lipid extracts were reconstituted in 0.5 mL of chloroform:methanol (2:1) and stored at -20 ℃ until use.

The incorporation of ^{14}C in count per minute (cpm) in lipids was determined by mixing 10 µL of samples with 5 mL of liquid scintillation. Labeled lipids were separated by thin-layer chromatography (TLC) and quantified by phosphorimager. For each sample, an aliquot containing

≈20,000 cpm was dried and reconstituted in 20 µL of chloroform and spotted onto a TLC plate and resolved by running in a petroleum ether/ethyl acetate solvent system (98:2). The plate was left to dry for 5 min, covered with Saran Wrap, and exposed to a phosphorImager screen for at least 3 days. The screen was scanned and analyzed using a Biomolecular Imager (UTEP Genomic Analysis Core Facility).

Proteomic profiling

A proteomic profiling was carried out the CDC1551 WT, Δ*kstR*, and two complemented Δ*kstR:kstR* strains. Cells were cultivated in 60-mL ink bottles containing 25 mL of 7H9-OADC medium supplemented with 0.2 mM cholesterol, with agitation at 125 rpm for 3 days. The cultures were harvested by centrifugation at 3,500 rpm for 10 min and cells were washed twice with cold PBS (without Ca^{2+} and Mg^{2+}). After removing the supernatants completely, cell pellets were resuspended in 1 mL of PBS containing 1% SDS and transferred into 2-mL Eppendorf tubes containing 0.5 mL of 0.1-mm zirconia beads. The cells were lysed by 3 x 1-min cycles on a SoniBeast bead-beater apparatus (Biospec). The beads were settled by centrifugation at 1,000 rpm for 1 min. The cell lysates were transferred into 1.5-ml Eppendorf tubes with screw cap and boiled for 15 min to kill any residual *Mtb* cells before removal from the BSL-3 laboratory. The concentrations of the proteins in cell lysates were obtained using the BCA Protein Assay Kit (Thermo Fisher Scientific).

For in-solution trypsin digestions, 50 µg of each protein lysate was lyophilized and reconstituted in 40 µl of 100 mM Tris-HCl (pH 8.5) containing 8 M urea. After adding 1.2 µl of 100 mM Tris(2-carboxyethyl)phosphine (TCEP), the samples were shaken for 20 min at room temperature. Next, 0.88 µl of 500 mM iodoacetamide (IAA) was added and the samples were

shaken in the dark for 15 min. The urea concentration in the samples was reduced to 2 M by adding 120 µl of 100 mM Tris-HCl (pH 8.5) and trypsin was added at a (1:30) ratio (~1.7 µg). The in-solution digestions reactions were incubated overnight at 37 ᵒC. SDS was removed from the samples using Pierce C18 spin columns (Thermo Fisher Scientific).

Peptides were analyzed with the QE Orbitrap (Thermo Fisher Scientific) along with the Dionex UltiMate 3000 RSLCnano UHPLC system (Thermo Fisher 221 Scientific). Samples were loaded in line onto a C18 PicoChip Column (75µm ID x 15 µm tip packed with 10.5 cm of Reprosil -PUR C18 3µm 120Å; 25µm x 50cm fused-silica tail, New Objective). Column was equilibrated with 95% Solvent A (95% water, 5% acetonitrile, 0.1% formic acid) and 5% Solvent B ((5% water, 95% acetonitrile, 0.1% formic acid). The column was equilibrated for 10 min with a flow rate of 0.5 µL/min with 95% solvent A, 5% solvent B. The samples were injected and loaded onto column and equilibrated for 10 min with solvent A. The elution of the peptides was completed by running a linear gradient up to 40% solvent B for 95 min, followed by a sudden increase to 95% B and let it run for 10 min. Sample was re-equilibrated for 10 minutes with 5% solvent B. Full scan spectra was collected via Xcalibur (Thermo Fisher Scientific). QE Orbitrap settings are as follows: Full MS resolution of 70,000, AGC target of 1e6, scan range from 400 to 1600 m/z; MS/MS were run with a resolution of 17,500, AGC target of 2e6, scan range from 200 to 2000, with 3-m/z isolation window. The spectra were searched using Proteome Discoverer (PD) 2.1.1.21 (Thermo Fisher Scientific) and filtered via Sequest HT with an estimated false-discovery rate (FDR) of 0.01 against sequences from *Mycobacterium tuberculosis, Escherichia coli*, human, bovine, human keratin, and porcine trypsin. A precursor mass tolerance and a fragment mass tolerance of 20 and 0.02 ppm, respectively, was used. Cysteine carbamidomethylation, methionine, oxidation and acetylation were set fixed and variable modifications, respectively.

RAW264.7 macrophage infection

The murine macrophage cell line RAW264.7 was cultivated in Dulbecco's Modified Eagle Medium (DMEM) enriched with 10% Fetal Bovine Serum (FBS). The day before the infection, three 12-well plates (days 0, 3, 5) were seeded at a density of 1×10^5 cells/well. *Mtb* CDC1551 strains (WT, $\Delta kstR$, $\Delta kstR:kstR$, and KD*kstR*) were grown in 7H9-OADC until an OD_{600} of 0.6-0.8 was reached. For *Mtb* KD*kstR*, cells cultivated in the presence or absence of 150 ng/mL ATc. The day of the infection 13-mL aliquots of *Mtb* cultures were pelleted by centrifugation at 3,500 rpm for 10 min. Cells were washed twice with 10 mL of D-PBS (without Ca^{2+} and Mg^{2+}), followed by centrifugation at 3,500 rpm for 5 minutes. To obtain single-cell suspensions, cell pellets were re-suspended in 6 mL of PBS and subjected to 3 cycles of 10-s pulse at 90% power in a Branson cup-horn sonicated filled with disinfectant, with inversion of the tubes between cycles. To remove any residual cell aggregates, the sonicated suspensions were centrifuged at 500 rpm, 5 min, and supernatants (5 mL) were transferred into new 15-mL Falcon tubes without disturbing the cells pellets. The OD_{600} was measured and adjusted to 0.0167 with DMEM supplemented with 10% FBS in order to achieve a multiplicity of infection (MOI) of 1. For *Mtb* KD*kstR* with ATc, 150 ng/mL of the inducer was also included.

The day of the infection, the culture medium in wells was removed by aspiration and the cell monolayers were washed twice by gently pipetting 1 mL of warm PBS (with Ca^{2+} and Mg^{2+}). For infection, cell monolayers in wells were gently covered with 0.6 mL of the corresponding *Mtb* suspension. The plates were centrifuged at 1,000 rpm, 10 min, and then incubated at 37 ºC for 1 h in the CO_2 incubator. Infection medium was removed by aspiration and the wells twice were washed by gently pipetting 2 mL of warm PBS (with Ca^{2+} and Mg^{2+}) to remove any extracellular

bacteria. Infected cell monolayers were finally covered with 2 mL of DMEM supplemented with 10% FBS, with 150 ng/mL ATc when needed. Plates were incubated at 37 ℃ in CO_2 incubator.

At days 0, 3, and 5, *Mtb*-infected macrophage cells were lysed to evaluate intracellular growth. The culture medium in wells was aspirated and the cell monolayers were washed twice by gently pipetting 1 mL of warm PBS (with Ca^{2+} and Mg^{2+}). The adherent cells were lysed with 0.5 mL D-PBS containing 0.1% (*v/v*) Triton X-100. A volume of 0.25 mL of the cell lysate was transferred into sterile 5-mL plastic tube with 1.75 mL of D-PBS containing 0.05% (*v/v*) Tween-80 (10^0 dilution). A volume of 0.1 mL of this solution was directly (Day 0) spread onto 7H10– OADC solid agar medium or after further dilution (Day 3 and Day 5: $10^{-1} – 10^{-3}$). Plates were put in metal canisters and incubated at 37 ℃ and colony-forming-unit (CFU) was determined after 2-3 weeks.

3.2 Results

Having conducted studies showing the best ligand for KstR de-repression to be 3OCh(25)26CoA, validation of KstR as a possible therapeutic target required in vivo studies in *Mtb*. The CDC1551 strain was chosen to conduct these in vivo experimentations as a model being it is a clinical strain that was obtained from a patient sample rather than using the H37Rv strain that has been in circulation through many laboratories for many years. Characterization of the effect for constant KstR de-repression was carried out through the gene deletion of *kstR* and subsequent reintroduction of the gene through various methods. Once kstR deletion was verified, a series of experimentation was conducted to better understand how the bacteria responds to cholesterol catabolic pathway de-regulation in its growth and infectivity.

The *ΔkstR* construct was designed and successfully transduced through general phage transduction into the CDC1551 strain of *Mtb*. A number of colonies were screened initially through

PCR and two were chosen to confirm the absence of the *kstR* gene through a Southern Blot (Fig. 15A). The clones that were chosen were grown in the presence of cholesterol and the cultures were used for protein extraction. A western blot was performed for the extracted proteins to assess if the quantity of proteins in the complements was comparable to the wild type samples (Fig. 15B). The complements did not show the same signal that was shown for the wild type, however a lesser amount of protein was present so it was decided to go forward with future experimentation. The confirmed *ΔkstR* strains were then used for the remaining experiments for these studies.

The established *ΔkstR* clones were used to a perform a growth curve to determine growth characteristics of *Mtb* after *kstR* gene deletion compared to that of the wild type strain of CDC1551. After the initial inoculation at 0.01 OD_{600} the cultures were monitored daily for growth through culture turbidity. Both strains showed to have minimal growth until approximately day 5. At that time, a marked branch between the WT and *ΔkstR* strains were seen in the CDC1551 strains regardless of what carbon source was present in the media (Fig. 15C). The cultures that had both glycerol and cholest-4-en-3-one as a carbon source showed a slight increase in growth, showing that cholest-4-en-3-one is able to be degraded without having a toxifying effect regardless of whether the *kstR* gene was present or not. Differences in deficient growth observed in the CDC1551 *ΔkstR* strain compared to that of their WT counterparts suggest that both strains differ in their ability to utilize not just cholest-4-en-3-one as a carbon source but also glycerol. Metabolite samples were extracted from the WT and Δ*kstR*, clones when grown on 0.2mM of cholesterol. These samples were examined for the depletion of cholesterol and the presence of cholest-4-en-3-one to see how the different strains would incorporate cholesterol when in the environment of constant de-repression after *kstR* deletion. (Fig. 15D) Cholesterol depletion in CDC1551 Δ*kstR* was demonstrated by 48 hours where it took the WT strain 72 hours to reach the same levels. The
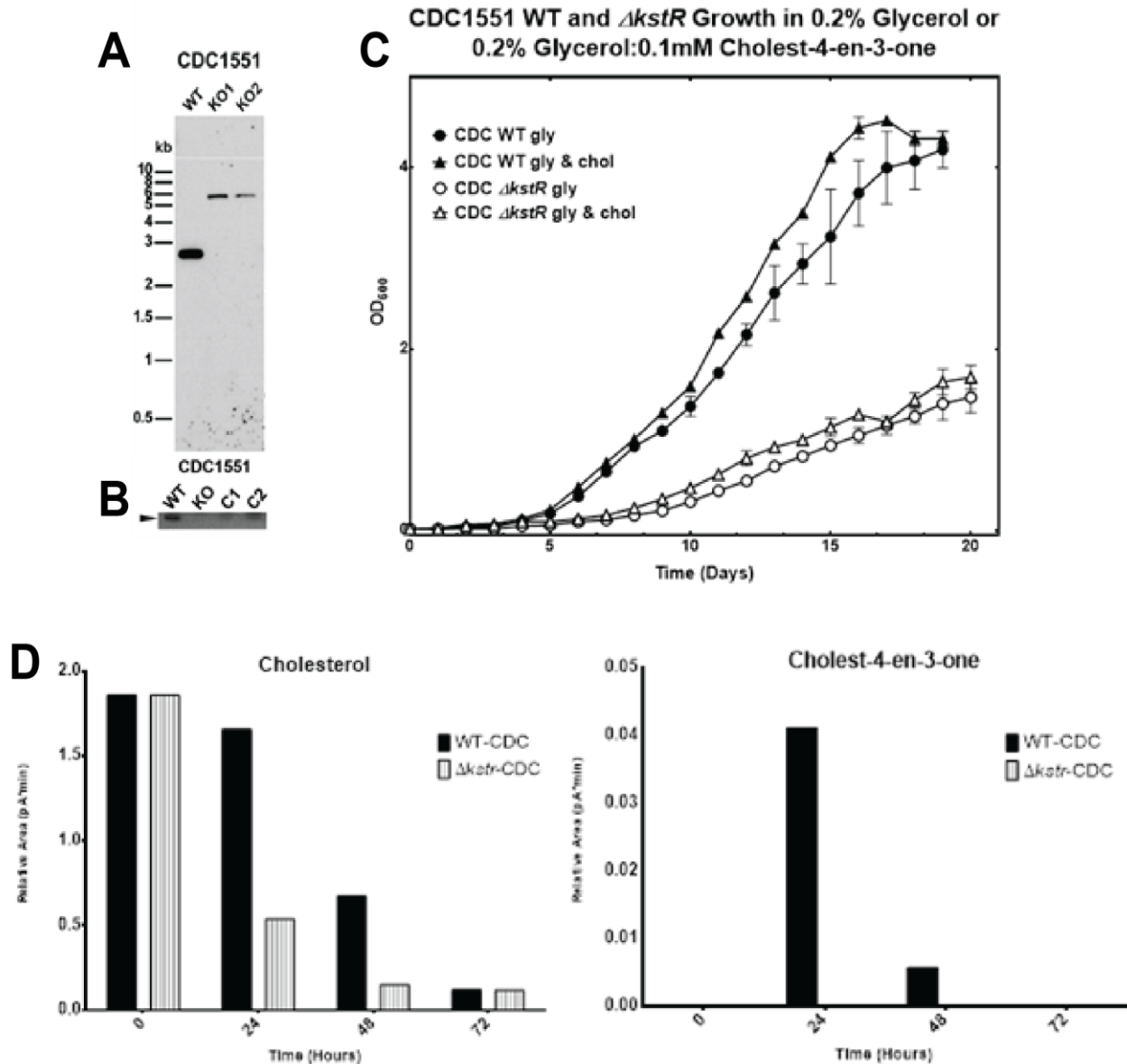
Figure 15: Comparison of Growth Curves of WT and Δ*kstR* in CDC1551: (A) Southern blot of *BamH I*-digested genomic DNA showing the hybridizing fragment of WT (2.8 kb) and of two *kstR* deletion mutants, KO1 and KO2 (5.9 kb). **(B)** Western blotting analysis of KstR expression in wild-type and Δ*kstR* strains of CDC1551. Cells were incubated in the presence of 0.2 mM cholesterol for 24 h before harvesting. **(C)** Gene deletion of *kstR* and WT in CDC551 (top) Growth curves are shown of WT (circle), Δ*kstR* (triangle) strains on minimal media containing 0.2% (v/v) glycerol (open symbols) or 0.2% glycerol (v/v) + 0.1 mM cholest-4-en-3-one (filled symbols). The growth was monitored by measuring the absorbance at 600 nm as a function of time. Both strains show a deficiency in growth when lacking kstR. **(D)** Measurement of cholesterol and cholest-4-en-3-one catabolism in CDC1551 WT and Δ*kstR*. As expected, cholesterol is degraded at a faster rate in CDC1551 Δ*kstR* than the WT and there appears to be no significant cholest-4-en-3-one build-up during this process.

CDC1551 *ΔkstR* strain also showed no signs of cholest-4-en-3-one buildup at any of the time points compared to that of the WT, although the buildup of cholest-4-en-3-one was abrogated fully by the 72 hour mark.

The confirmed CDC1551 *ΔkstR* clone used for the previous growth curve was grown and prepared for transformation of the plasmid that was constructed with the *kstR* gene under the control of its natural promoter through electroporation. The empty construct was used to ensure that any differences in growth imparted by the presence of the construct being introduced into the bacteria. Growth curves of the WT, *ΔkstR,* and *ΔkstR:kstR* in CDC1551 *Mtb* have been produced using the same media and growth conditions that were done previously for the *ΔkstR* alone (Fig. 16). Complementation of the CDC1551 *ΔkstR* yielded only partial recovery of the growth with the there is a lag in growth of approximately 5 days between the WT and ΔkstR complement regardless of the carbon source it was grown on. The endpoint of day 20 was decided on during the experiment due to the wild type CDC1551 being past the exponential phase and 4 days into stationary. Rather than allow the WT cultures to proceed to the death phase the experiment was ended with the last day for the collecting the optical density data for the WT being on day 19. It is unknown if the Δ*kstR* would be able to recover to the WT growth levels if left to grow for an additional amount of time, and assessment of the growth curves would be an approximate 8 days. The western blot performed to assess the concentrations of KstR production shows the protein concentrations are not at wild type levels and corresponds to the growth curve strains. This partial recovery could be attributed to the integration of the *kstR* gene into the *Mtb* genome at a different site than its origin and therefore lacking the genomic regulatory mechanisms, such as activation sites, that are more than likely present at the genes original placement.

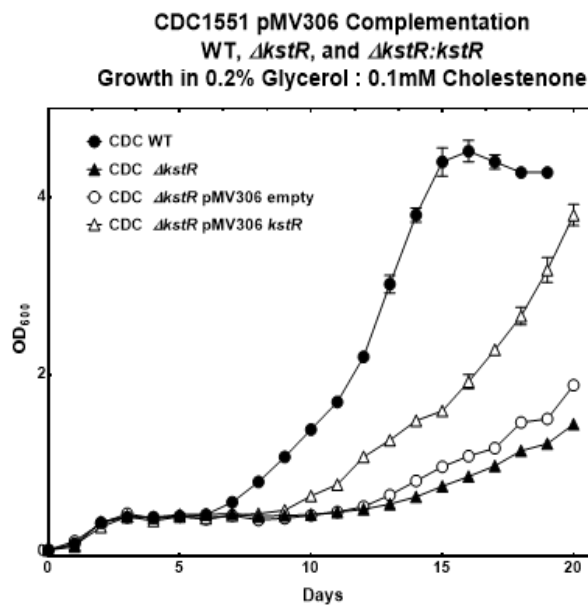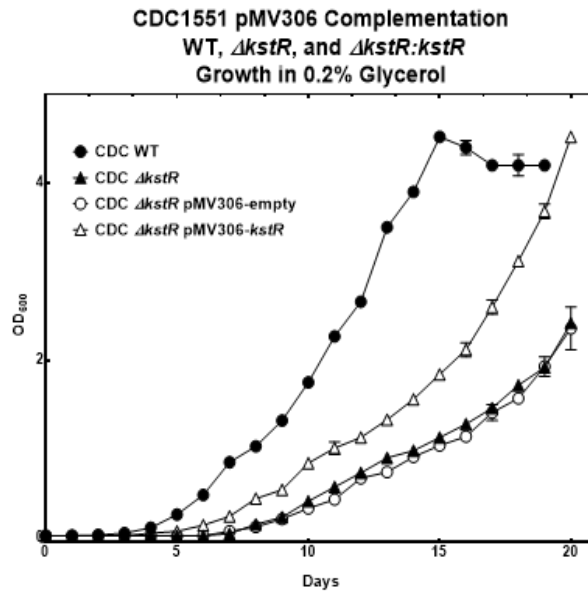**Figure 16: Growth curves of ΔkstR after complementation of the kstR gene at a different locus in CDC1551:** The growth of the WT, ΔkstR, and ΔkstR/kstR strains in glycerol alone. (top) WT, *ΔkstR*, and *ΔkstR/kstR* in both glycerol and cholest-4-en-3-one (bottom). Regardless of carbon source(s) present, only partial recovery of growth compared to the WT was achieved.

To understand how downstream effects to the deregulation of cholesterol catabolism effects PDIM production after deletion of *kstR*, lipids for CDC1551 WT, Δ*kstR*, and *kstR/kstR* were extracted after using C14 labeled propionate as the carbon source and showed the PDIM quantities. (Fig. 17) Propionate can be directly sourced to the methylmalonyl pathway and onward to methyl-branched fatty acids production by the incorporation into both the phthiocerol and mycocerosate chains. The CDC1551 WT shows a much higher PDIM signal when fed with propionate and the presence of PDIM is decreased with the deletion of *kstR*. Recovery with complementation and reintroduction of *kstR* does not occur to the same levels and is only slight.

Proteins extracted from CDC1551 WT, Δ*kstR*, and Δ*kstR/kstR* cultures grown on cholesterol were used to run mass spectrometry experiments to obtain a proteomic profile of the different mutant strains. Being KstR directly regulates the cholesterol catabolic pathway, the proteins involved in both the side-chain and A/B ring steps were analyzed. The data showed the detection of KstR regulated enzymes responsible for the side-chain degradation but none of those that are involved in A/B ring degradation. After introduction of cholesterol through the Mce4 transport system, cholesterol is catalyzed to cholestenone by the cholesterol oxidase ChoD by altering the hydroxy group at the 3 position of the cholesterol molecule to an oxo- form. The Δ*kstR* strain shows there is an decrease in ChoD as the source converting the cholesterol to cholestenone.

The side-chain continues to be further altered by the addition of an ester group at the end of the side-chain by either Cyp125 or Cyp142, depending on the stereo-conformation of the cholesterol molecule resulting in (25S) or (25R)-3-oxocholest-4-en-26-oate. Interestingly, the proteomic data for the Δ*kstR* shows a large fold increase for Cyp142 compared to Cyp125, indicating although the cholesterol degradation pathway is preferential to the (S)- conformation, the (R)- conformation is produced more after *kstR* gene deletion. As *Mtb* has the preference for
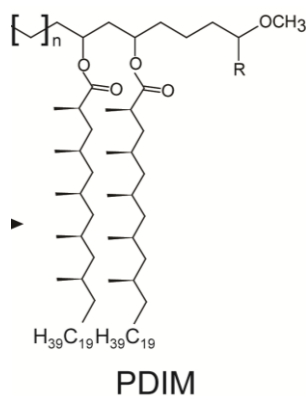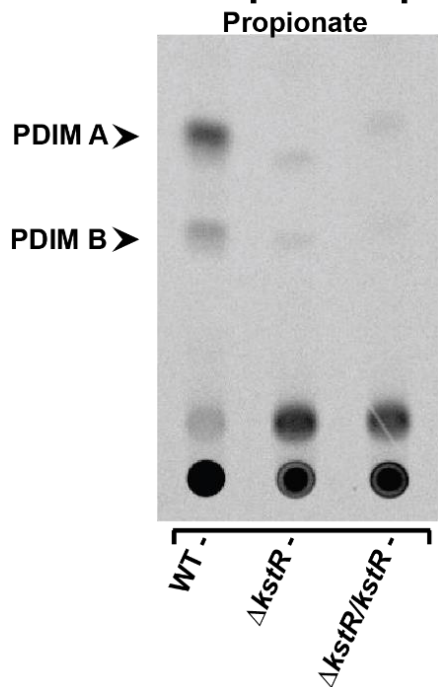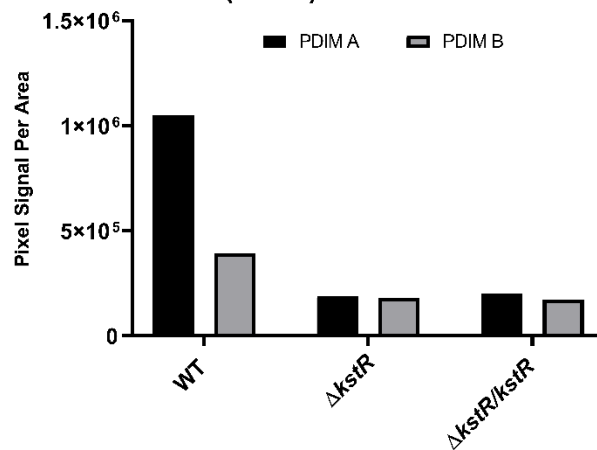
Figure 17: PDIM levels of CDC1551 WT, ΔkstR, and ΔkstR/kstR after introduction of C14-labeled carbon sources. (top) Depiction of phthiocerol dimycocerosate (PDIM). (left) The levels of PDIM are diminished in the CDC1551 *ΔkstR* and not recovered after complementation. (right) Signal quantification of the TLC plate showing the presence of PDIM.
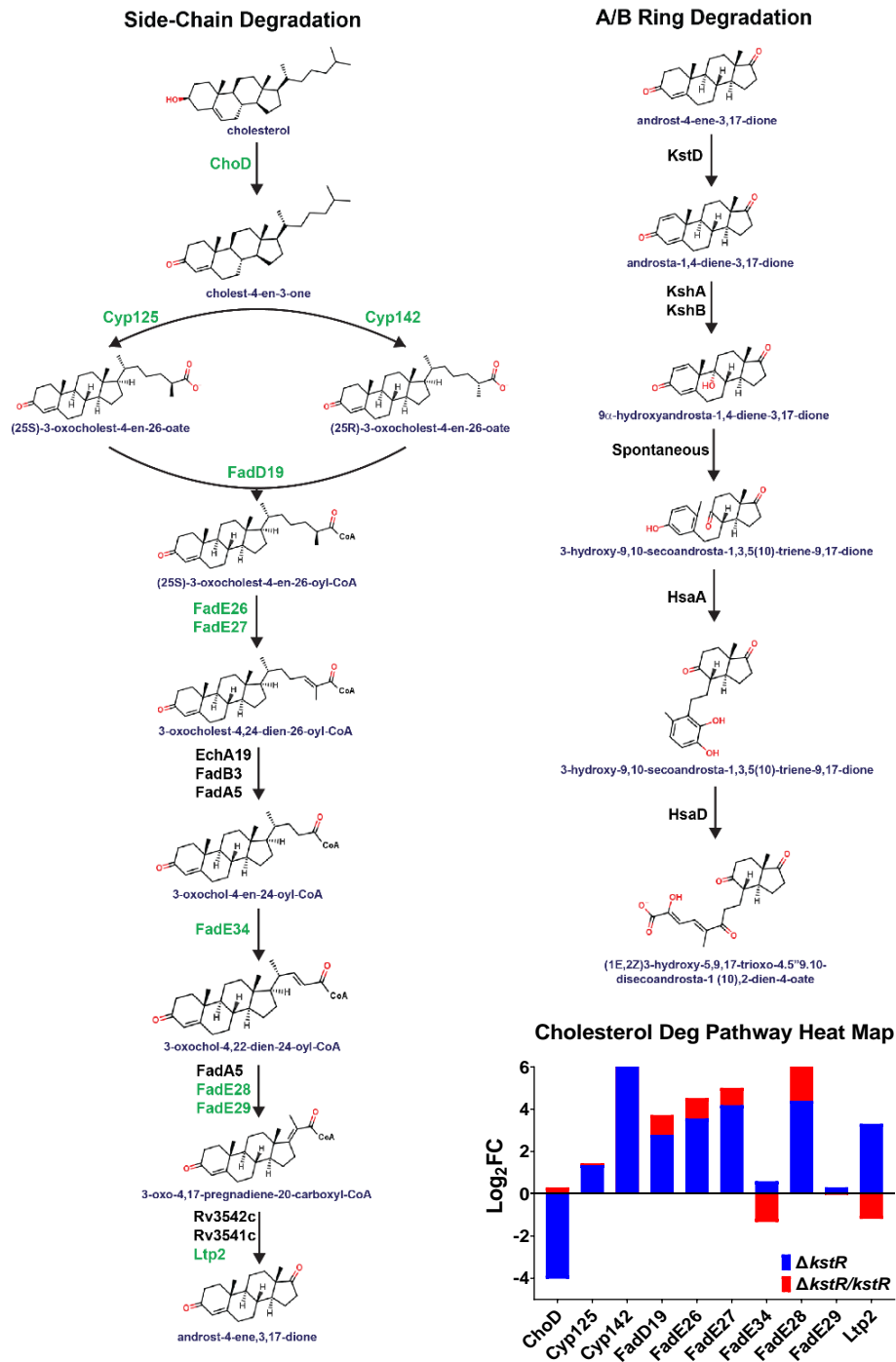
Figure 18: Representation of cholesterol catabolic pathway and detected enzymes with mass spectrometry. Depiction of the cholesterol catabolic pathway with detected enzymes within the pathway labeled in green. The initial cholesterol side-chain degradation steps are listed on the left. The A/B ring degradation steps that are known to occur after side-chain degradation are listed on the right. The fold change of detected enzymes of the pathway are shown as a heat map on the lower right.

the (S)- conformation of cholesterol, the (25R)-3-oxocholest-4-en-26-oate is converted to the (S)-state by a racemase or by the coenzyme A ligase FadD19 which also initiates the attachment of coenzyme A to the 26-carbon position of the intermediate. Once the attachment of coenzyme A has occurred, the co-expressed FadD26 and FadD27 enzymes alter the side chain of cholesterol creating a double bond between the 24 and 25 carbons. The enzymes involved in the next three enzymatic steps were not detected for any of the strains tested. Within these steps, FadA5 is the enzyme that detaches propionyl-CoA from the molecule in multiple steps throughout the side chain degradation process and surprisingly it was not present in the data. Also surprising was the lack of increase with the dehydrogenase FadE34 as it is part of the catabolic pathway that introduces a double bond in the shortened side-chain after the detachment of propionyl-CoA by FadA5. Subsequent steps in the cholesterol catabolic process that were detected to have an increase without KstR regulation were the dehydrogenase FadE28 and the lyase Ltp2. FadE28 and FadE29 are co-expressed with one another and one would expect the expression of FadE29 to also show an increase but the levels were comparable to that of the WT strain. The lyase Ltp2 reduces cholesterol to the androst-4-ene-3,17-dione form leaving only the sterol portion of the cholesterol molecule and the release of another propionyl-CoA. The sterol rings of the cholesterol then goes through ring degradation initiated by the dehydrogenase KstD. As stated previously, the remaining enzymes that perform the A/B ring degradation were not detected in any of the strains.

Focusing on the PDIM synthesis pathway, the quantities and differences in the proteins responsible for the methyl-branched fatty acids were analyzed and assessed for differences. Being that side-chain degradation final by-product is propionyl-CoA, we began looking at the production of methylmalonyl-CoA by assessing the protein presence of the subunits of propionyl-CoA carboxylase. Only two of the known subunits were detected in the proteomic data. The AccA5 and
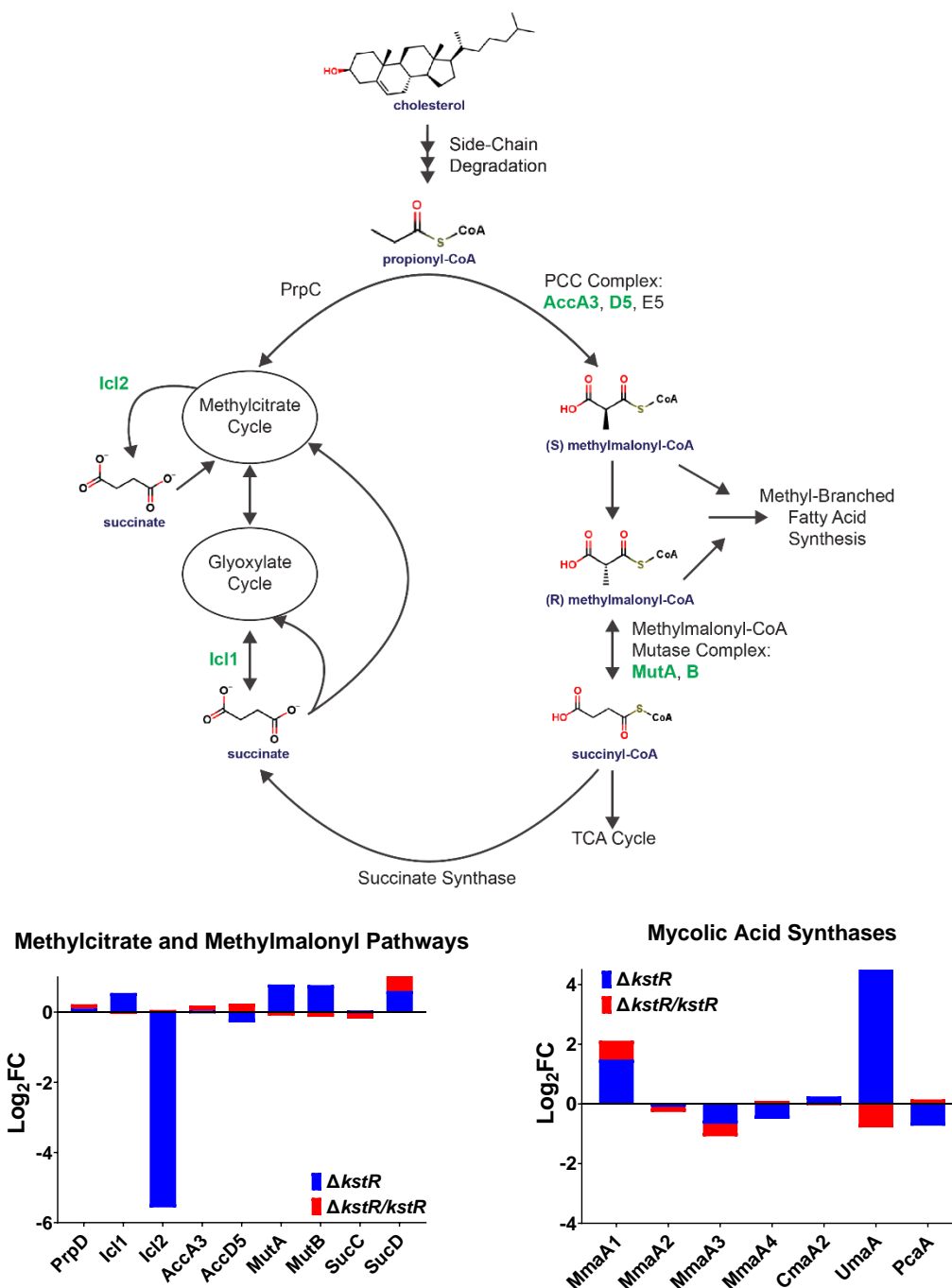
Figure 19: Presence of proteins involved in the pathways involved in propionyl-CoA utilization. (A) Pathway depiction showing how propionyl-CoA is incorporated into the metabolism of Mtb. Enzymes in green denote those detected in proteomic data. (B)Proteomic data analysis shows the increased presence of the MutA and B subunits involved in both the production of methylmalonyl-CoA and succinyl-CoA within the methylmalonyl pathway in *Mtb.* Additionally, an increase of Icl1 that catalyzes the formation of succinate was detected in the Δ*kstR* strain where Icl2 was not detected. (C) An increase of the mycolic acid synthase UmaA was detected in the Δ*kstR* strain

AccD5 subunits of propionyl-CoA carboxylase activity showed no significant differences although the AccD5 subunit did show a slight decline in the CDC1551 Δ*kstR* mutant. (Fig. 21) The product of the propionyl-CoA carboxylase is (S)-methylmalonyl-CoA. The following step in the pathway is the racemization of the methylmalonyl-CoA to the R stereoisomer. Unfortunately, the known epimerase to catalyze this reaction was not found in the studies. However, an increased presence of both the methylmalonyl-CoA mutase subunits responsible for the reversible B12 dependent conversion to succinyl-CoA was found with marked differences in quantity between the WT and Δ*kstR* for the CDC1551 strain. (Fig. 18) This would suggest succinyl-CoA is readily utilized for energy and metabolic processes in the Δ*kstR* mutant and the complement.

Propionate also can be shuttled into the methylcitrate cycle that provides a conversion source to pyruvate either through the methylcitrate cycle itself or by providing acetyl-CoA to the glyoxylate cycle. The isocitrate lyase Icl2 in *Mtb* is responsible for the production of either pyruvate or succinate while Icl1 catalyzes succinate. The production of succinate can be sourced to either the methylcitrate and glyoxylate cycles or succinate can be converted to succinyl-CoA. SucC and SucD are the subunits of succinyl-CoA synthetase that are involved with the reversible reaction of ligating and detachment of CoASH to succinate where it can then be shuttled to the mehtylmalonyl pathway feeding the production of methyl-branched fatty acids or towards the TCA cycle. [75] Looking at protein amounts of these show that a pattern occurs with quantities of Icl1 and Icl2 in relation to PDIM production. (Fig. 18). In the CDC1551 Δ*kstR* strain levels of Icl2 become undetectable while an increase of Icl1 occurs (Fig18). This result suggests the glyoxylate cycle is utilized more in an environment where KstR is not regulating cholesterol catabolism. It is logical to expect the amounts of propionyl-coA being shuttled into the methylcitrate cycle to be increased and therefore the amounts of Icl2 to increase to compensate for the higher amounts of

propionyl-Coa, however, this was not the result. It appears with the deletion of kstR, the preferred pathway for propionyl-CoA is more towards the production of methylmalonyl. This can be confirmed when looking at the quantities of both MutA, MutB, SucC and SucD. When looking at the succinyl-CoA synthetase subunits SucD, there appears to be an increase in the Δ*kstR* strain suggesting compared to the WT. The lack of change in the quantity of succinyl-CoA synthetase may indicate any excess propionyl-CoA is directed to the methylcitrate cycle initially than after passing through the methylmalonyl pathway.

Being the results showed a tendency of propionyl-CoA being incorporated into the methylmalonyl pathway and thereby shuttled to not only the synthesis of PDIM but to mycolic acid synthesis, methyltransferases (MTFs) were mined within the proteomic data (Fig. 18). Seven of the eight known MTF's were detected. While there were slight increases in MmaA1 and CmaA1, a large increase in UmaA was observed. Conversely a decrease in MmaA4 was seen. Both UmaA and MmaA4 have been shown to interact with both the fatty acid synthase complexes (FAS-I and FAS-II). [76] Besides being a functional MTF involved in the elongation of mycolic acids, UmaA has not been fully characterized with regards to regulatory response and information regarding the enzyme is not in abundance. Regardless of the lack of information regarding UmaA, it is still notable that the deletion of KstR creates an alteration in the amounts of mycolic acid synthases present and this data could be used to expand its further experimentation.

Formation of PDIM in *Mtb* consists of the synthesis of mycocerosic acid and phthiocerol independently that then merge in a final ligase reaction. For this study, reactions within the pathways of both the mycocerosic acid and phthiocerol were the focus due to the abundance of proteins present during the proteomic analysis of the CDC1551 strain and the respective mutants. Additionally, these pathways were a focus of interest due to the assimilation of methylmalonyl
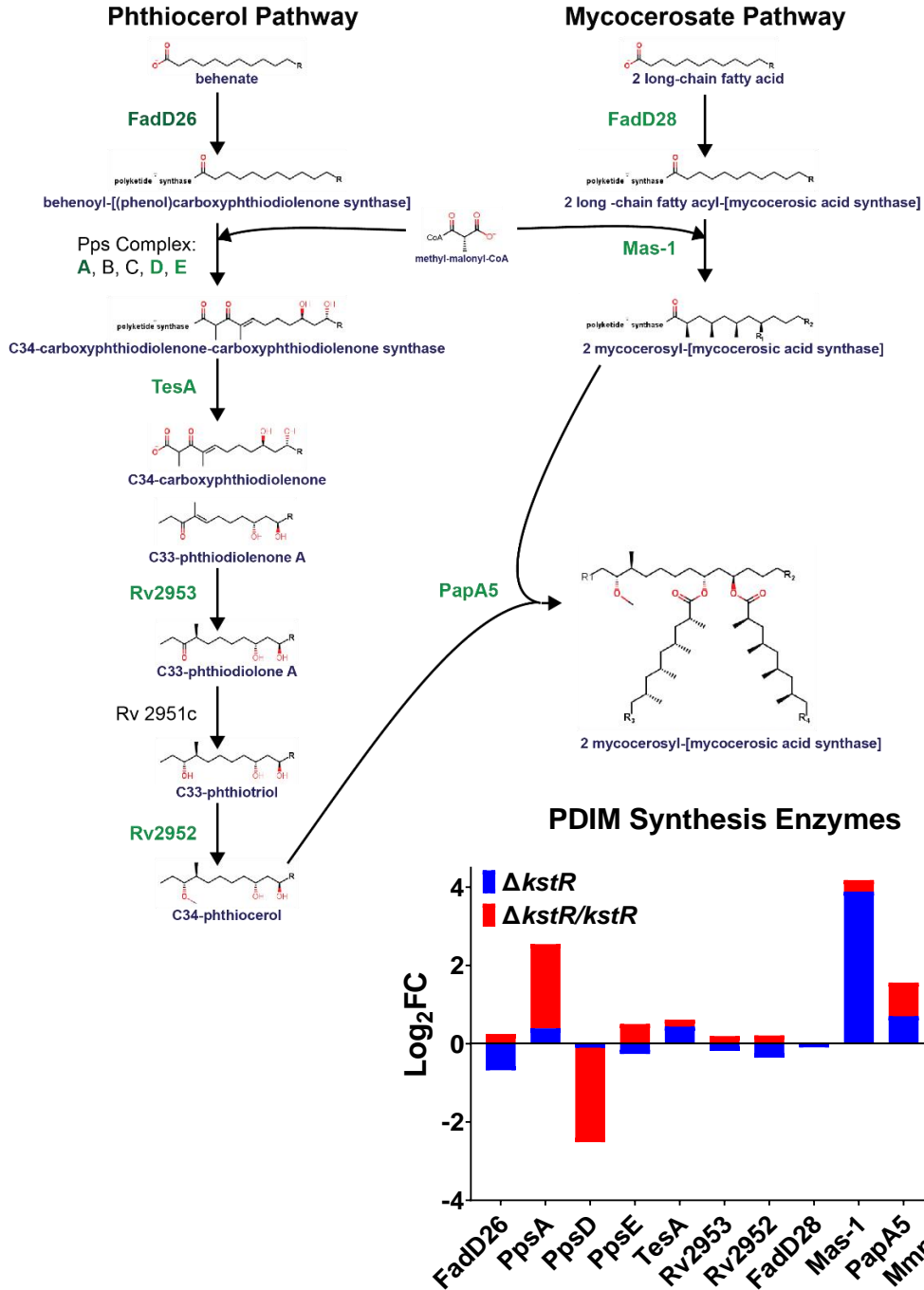
Figure 20: Mycocerosate and phthiocerol pathways result in PDIM production. (A) Representation of the mycocersosate and phthiocerol pathways that result in the formation of PDIM synthesis. All proteomic detected enzymes are labeled in green (B) Deletion of kstR resulted in an increase of TesA and decrease of enzymes in the enzymes of the pthiocerol pathway after the TesA synthesis of C33-carboxypthiodiolenone. A large increase of Mas-1 was detected for the mycocerosate pathway. Interestingly, both the steps that show an increase are where methylmalonyl-CoA is introduced.

into PDIM production. The incorporation of methylmalonyl is shuttled into both the biosynthesis of the mycocerosic acid and phthiocerol fatty chains. The elongation of mycocerosic acid and the introduction of methylmalonyl into the chain is accomplished with Mas-1. There are distinguishable differences in the presence of Mas-1 between the CDC1551 WT and Δ*kstR* strains. (Fig 23). Looking at the Δ*kstR* mutants for both strains also shows not only the opposite quantities of Mas-1 but also the opposite results of the TLC. The CDC1551 WT strain produces more PDIM and less Mas-1. The inverse amounts of Mas-1 are present for the Δ*kstR* mutants in each strain. An interesting result was seen with the *kstR* complements for each strain. There appears to be full recovery when *kstR* is reintroduced. To produce phthiocerol, a set of protein subunits PpsA-E are responsible for the integration of methylmalonyl into this pathway. Of these subunits, only the PpsD subunit was detected to have any discernable differences in quantity during this study. (Fig. 19) CDC1551 complements had a reduction in the amount of the PpsD subunit

The unification of both the mycocerosic acid and the phthiocerol fatty chains are completed by the polyketide synthase PapA5 to produce PDIM. [77] The proteomic analysis of the quantities of PapA5 correlate more with the TLC results in comparison with the presence of PDIM than with the previously mentioned proteins. The CDC1551 WT PapA5 levels were present whereas the Δ*kstR* levels were not detected (Fig. 19). The complementation of *kstR* in the CDC1551 strain showed slightly higher amounts of PapA5 to that of the WT and yet the results of the TLC showed less recovery of the end product PDIM.

Macrophage infections were carried out with the Mtb CDC1551 WT, Δ*kstR,* and Δ*kstR/kstR* with the as well as the pMV306 Δ*kstR/kstR* integrated complementation strain. (Fig. 25). For the initial infection it appears the WT strain was able to adhere and be engulfed better than the Δ*kstR* or either of the complements. This may be due to the absence of virulence factors
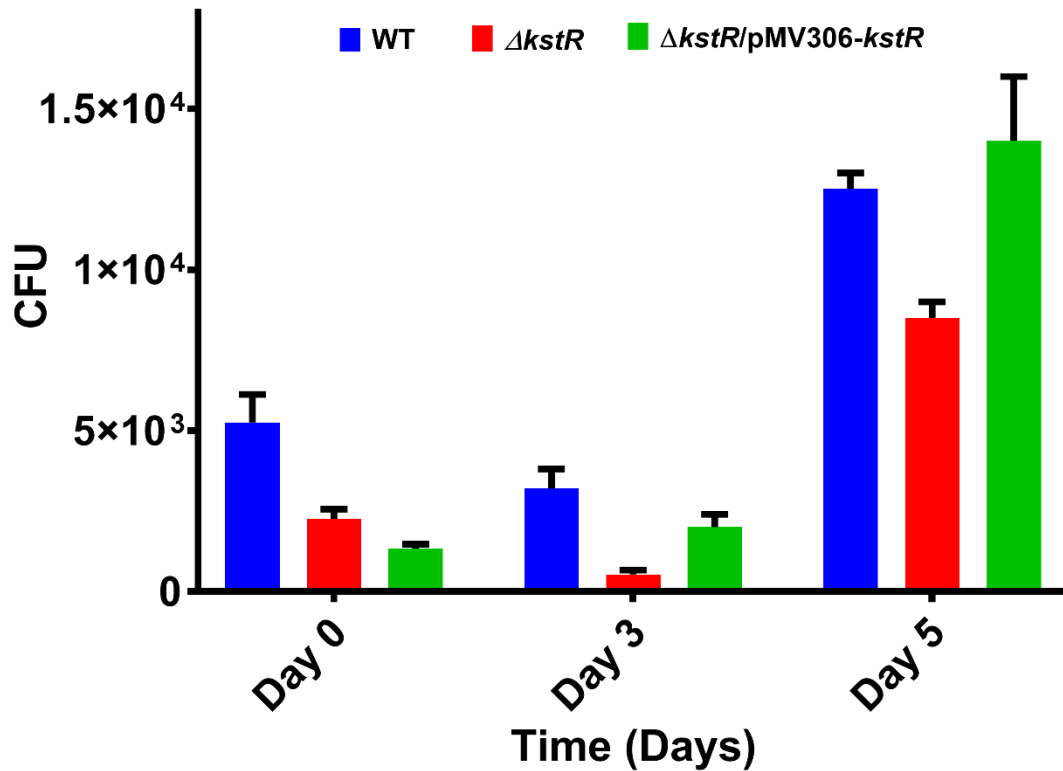
Figure 21: Infection of RAW264.7 cell line with CDC1551 WT, Δ*kstR*, and Δ*kstR*/pMV306-*kstR*: The mouse macrophage like cell line RAW 264.7 was infected with a 1:1 ratio of cells:bacteria to obtain the CFU counts. The CFU counts for the Δ*kstR* complements shows to have similar deficiency until day 5 of the infection. At day 5 the over-expression of KstR shows weaker growth in the macrophage while pMV306-kstR complement shown to express lesser levels of KstR has recovered to above WT CFU levels. This suggests that KstR is required for initial adherence but once infection has been established, lesser amounts of KstR are sufficient to promote persistence, while too much KstR with overexpression has a detrimental effect. Statistical analysis revealed no significant differences, but the trend in bacterial deficiency and recovery is visually represented.

such as PDIM which were measured and shown in Fig. 17. The PDIM levels show to be in correlation with the initial onset of infection for the macrophages with the WT being able to metabolize PDIM slightly better than the other strains tested. The Δ*kstR* and both of the Δ*kstR/kstR* complements had similar adherence and engulfment rates which also could be in relation to the lesser quantity of PDIM production that have been in witnessed in the TLC experiments. What was surprising was the recovery of infection rate of the Δ*kstR* by day 5. It was suspected the Δ*kstR* infections would be more attenuated and not recover as much as presented. This result could be an actual account of the Δ*kstR* and the WT strain actually being attenuated due to possible mutations. It appears the over-expression of KstR during infection halts infectivity compared to the other strains. With the network of genes the KstR repressor controls, having continuous repression also weakens not only the ability for Mtb to infect but also for it to replicate at similar rates.

As deletion of genes and selection thereafter can result in altered phenotypes and frame shift mutations, an alternative approach to deplete the KstR using a tet-operator was attempted. The *kstR* gene in CDC1551 was successfully replaced with the *kstR* with an additional DAS+4 tag at the C-terminus of the protein. After the successful recombination was confirmed, transformation of the plasmid pGMCKq1-10M1-sspBmyc in both the WT and knockdown KstR (KDKstR) strain was completed and verification of the presence of KstR translation and depletion through western blot was performed. (Fig. 26) The presence of KstR was not changed with or without the addition of ATc to the growth culture in the CDC1551 WT. The KDKstR clone shows depletion of the protein despite not having any ATc added to the culture and although the band is very light, a faint band remains after introduction of ATc. The western blot shows the tet-operator in control of the translation of sspB to be "leaky" thereby allowing the SspB protein to be expressed and reducing the amount of KstR before induction of the knockdown system. It was decided to pursue further

73
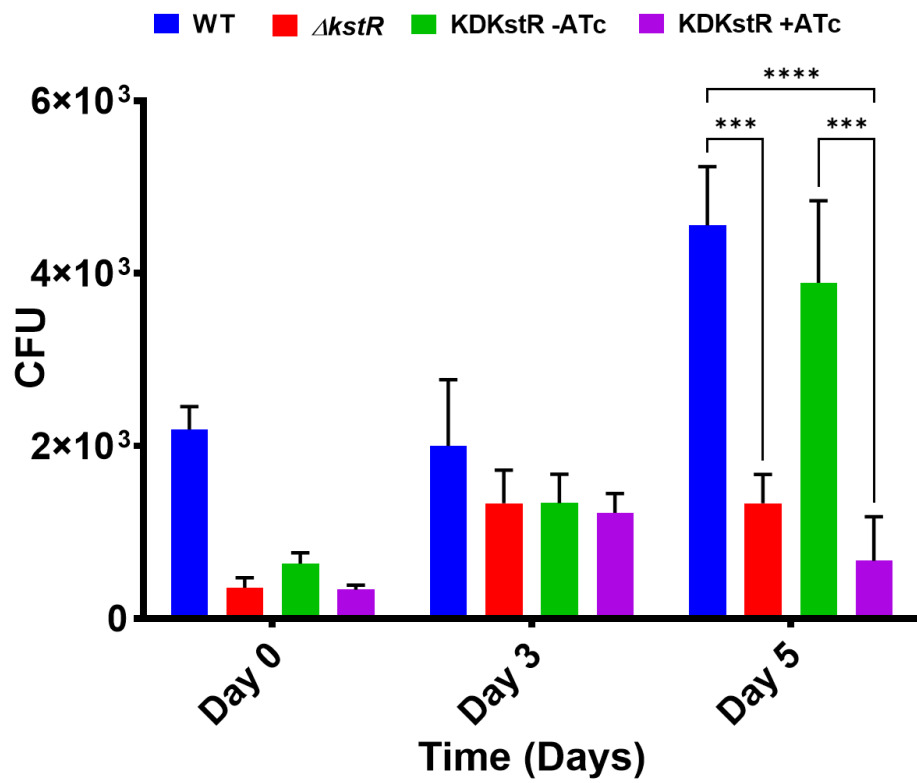
experimentation to better understand the effects of the reduction of KstR and determine if the knockdown model would be a functional alternative to the Δ*kstR*.

Both the CDC1551 WT and KDKstR were also fed with C^{14} labeled propionate to perform TLC studies. The WT PDIM levels are much higher than that of the KDKstR after the introduction of ATc. Additional studies showing the PDIM amounts of each the WT without the addition of ATc displayed the quantity of PDIM was also decreased with the KDKstR strain. (Not shown) It did not increase or decrease after the addition of ATc. Therefore, regardless of the presence of ATc, the PDIM quantities for the KDKstR were reduced.

Infection studies with RAW264.7 murine macrophages were conducted to assess the virulence after deletion and/or depletion of KstR in CDC1551. (Fig. 28) Colony forming units (CFUs) were measured at days 0, 3, and 5. The Δ*kstR* and KDKstR with ATc constructs CFU counts were much less than the WT and were not able to recover by day 5. The KDKstR infectivity without ATc present in the media was stunted at both days 0 and 3 but the CFU count began to recover by day 5. As it was previously confirmed KDKstR had a diminished but not complete abolishment in the presence of KstR when ATc was not presented to the bacteria. This suggests that although KstR is not required for infectivity, it does play a significant role during the onset of infection. It appears by day 5 the mechanisms of adaptation aided by the tight regulation imparted by KstR is not able to be overcome with the complete absence or depletion of KstR, yet the KDKstR that was confirmed to have a lesser amount of KstR was able to modify the responses and persist. Being that both the Δ*kstR* and KDKstR with ATc constructs behaved in similar manners during the infection studies, both of them seem to be viable models for future experimentation and research.

3.3 Discussion

The studies conducted thus far have shown the deletion of *kstR* attenuates the growth of *Mtb* but the extent of attenuation is strain dependent. The clinical isolate CDC1551, collected in the mid 1990's, is well documented to be less virulent compared to that of H37Rv. [78] However, CDC1551 demonstrates a marked and significant decrease in cell density after the removal of *kstR* from the genome in both glycerol and glycerol/cholest-4-en-3-one containing medias. (Fig. 16). [58] Further investigation as to how both strains are able to metabolize glycerol and cholesterol is required.

After complementation of *kstR* was carried out, the same protocol for growth culture was completed with the CDC1551 WT and *ΔkstR*, *ΔkstR/kstR* mutant strains. Upon completion of the 20 day cultures, it was observed full complementation of *kstR* was not achieved. (Fig. 18) This deficiency could be due to the implementation of the restored *kstR* within the *Mtb* genome. As previously stated, the regulatory aspects of KstR are well defined. Implementation of the gene with the integrative plasmid pMV306 is at a site where the likelihood of additional regulatory mechanisms (i.e. activation sites) required for *kstR* regulation to occur properly was unable to be attained. Being that *kstR* is auto-regulated and responsible for the transcription of many other *Mtb* genes implies there are a multitude of components that would be required for the entire regulatory system to function properly. The misplacing of the gene from its position of origin would indeed result in a diminished transcriptional effect.

How the different strains utilize the different carbon sources a key factor on how they adapt to the environment changes and stressors to survive. All previous studies have been completed only with the H37Rv Mtb strain. [54] The final by-product of cholesterol catabolism is propionyl-CoA which can lead to the buildup of toxic intermediates detrimental to *Mtb* survival. [40]

75

Propionyl-CoA can be used as either an energy source with further catabolism to acetyl-CoA in the methyl citrate cycle or for the incorporation into the methylmalonyl pathway which can then be shuttled to either the TCA cycle through manipulation to succinyl-CoA or the methylmalonyl can be shuttled into the production of methyl-branched fatty acids such as PDIM. [40] The initial AccA3 Previous studies have stated both methylmalonyl mutase subunits MutA and MutB responsible for the conversion of either (S)-methylmalonyl or succinyl-CoA display control by KstR and the proteomic data appear to coincide with the studies. [54] Without *kstR*, there is an increase in the amounts of both MutA and MutB subunits. (Fig. 21) As virulence factors such as PDIM have been shown to be important for infectious stages such as phagocytosis. [79] Both CDC1551 and H37Rv strains of Mtb show to react to excess pools of propionate in different manners. The increased amounts of MutA and MutB Δ*kstR* strains would indicate the high likelihood of an increased presence of PDIM in the. The CDC1551 Δ*kstR* is unable to assimilate propionate into PDIM as efficiently as the H37Rv Δ*kstR* and this contrasts with their respective wild types. Neither strain was capable of regaining the WT production of PDIM, whether it was increasing or decreasing the amounts, after deletion and reintegration of the *kstR* gene. Again, this more than likely is related to the placement of *kstR* after complementation and the lack of interaction with additional regulatory factors.

The proteomic data correlates with the thin layer chromatography results. It is important to note that the KstR regulator itself was not detected in the proteomic analysis. There were a number of other transcriptional regulators present, yet the key regulators known to control cholesterol catabolism such as the central hub genes Rv0081 and Rv0324 were absent. These genes also play a major role in hypoxic responses in *Mtb*. KstR is also known to be a part of the enduring hypoxic response (EHR) in *Mtb* H37Rv that is independent of the DosR response for hypoxia. [80] The

EHR consists of 230 genes that are found to be induced at days 4 and 7, and among this group cholesterol catabolic pathway genes *fadD19*, *cyp125,* and *kstR* are present, among others. In our studies, these enzymes were found to be over-expressed without *kstR* regulation. These results would be expected, but there doesn't appear to be any additional regulatory mechanisms that rescue an over-expression response for these enzymes. With the absence of *kstR*, the induction of a constant hypoxic response state may be how *Mtb* copes with a lacking key regulatory factor. In order to establish a more complete transcriptional profile after the removal of *kstR*, RNA-seq should be performed to determine how *Mtb* is able to respond without a key regulator in both cholesterol catabolism and hypoxic response.

The production of the two fatty acid chains phthiocerol and mycocerosate conjoined by PapA5 to produce PDIM, is decreased to non-detectable levels in the CDC1551 Δ*kstR*. PapA5 is known to be controlled by the hypoxic response regulator Rv0081. [54] The levels of PapA5 showed to be recovered with the Δ*kstR* complements in the CDC1551 strain. Although KstR has not been proven to be directly involved with PDIM synthesis, the downstream effects of producing high amounts of propionyl-CoA from the constant side-chain degradation of cholesterol certainly plays a role in other regulatory mechanisms in *Mtb*. [54, 81] Although KstR has been previously over-expressed for Chip-seq experimentation, the phenotypical changes that would occur have not yet been assessed. [54]

Previous studies have been completed using an over-expression system to map the Mtb transcriptional network. [54, 82, 83] The data was obtained without the presence of cholesterol or any cholesterol alternative as an additional carbon nutrient that may give a better insight as to how KstR responds to different environments but provides a valid baseline to determine the transcriptional capabilities of KstR. With over-expression of KstR, it has been previously shown

that KstR binds to 138 sites within the *Mtb* H37Rv genome with 35 of the binding sites having a repression effect. There are 10 transcription factors that are able to bind in the *kstR* region and only 2 of them have shown to cause an induced differential expression outcome on media lacking cholesterol. [83] The previous studies with over-expressing KstR were never longer than a 24 hour time frame and served the purpose of getting the initial responses of RNA signals for *Mtb*. [54] As KstR has shown to take part in the regulation of MutA, MutB, and other PDIM production enzymes such as PapA5, it can be assumed that there would be a decreased amount of PDIM. Methylmalonyl production would be inhibited and therefore the downstream effect of PDIM would be decreased. qRT-PCR experimentation determining the transcription data for which enzymes are being produced during over-expression conditions would need to be performed to how the data translates with regards to PDIM production.

Macrophage infections with *Mtb* CDC1551 WT, Δ*kstR*, and both complementation construct strains and Δ*kstR*/pMV306-kstR were carried out and displayed the differences in over-expression and slight expression of KstR was able to effect both initial infection rates as well as the later stages of infection. Both showed a lessened CFU count at the initial stage of adherence and engulfment. This could be the result of the diminished amounts of PDIM production that has been seen throughout these studies. PDIM is a well-known methyl-branched fatty acid that plays a part in preventing the immune response during infection and contributes to the phagosomal escape of *Mtb*. [84-86] Although phagosomal escape in Mtb has been proven to be depended on the presence of the ESAT6-1 pore forming substrate EsxA, there was no significant changes in the quantification of this protein in the Mtb CDC1551 strains for the proteomic studies we performed. [86-88] The inability for Δ*kstR* to replicate at WT levels shows there is a failure to evade the host response. This could be due to an inability for these strains to produce PDIM at WT levels and

therefore phagosomal escape within the macrophages. A constant increased amount of KstR would suppress the enzymes that are needed for cholesterol catabolism and the enduring hypoxic response necessary for long term survival. Surprisingly, the Δ*kstR*/pMV306-*kstR* complement strain was able to regain WT CFU levels even with reduced amounts of KstR shown in other experiments. It appears KstR is required for a robust response at the onset of infection, but after the infection has become established, even reduced amounts of KstR are sufficient to persist within the macrophage.

We know from these experiments a lack of KstR diminishes the amounts of PDIM, and it more than likely creates a de-regulated environment within the bacterial cell that leads to attenuation at the onset of infection. Additionally to PDIM, it is highly likely other virulence factors required for the onset of a normal infection of macrophage are diminished. However, higher KstR levels would produce a continuous hyper-regulated response for not only cholesterol catabolism but for the hypoxic response. This situation may prohibit *Mtb*'s ability to maintain a persistent infection.   As phenotypic studies have not yet been completed for the deletion or the over-expression of KstR, this data provides novel information that can be used for further determining if KstR is a valid therapeutic drug target.

The complete removal of *kstR* in fact does alter the growth and pathogenicity of *Mtb* during infection as seen with this study. This phenomenon has already been established with *kstR* being required for in vivo mouse model infection in prior studies.**[58]** In order to bypass the deleterious effect that occurs with the absence of *kstR* in *Mtb*, a knockdown approach is a much more appreciable method to study the de-regulation of the *kstR* regulon. With the addition of ATc in cell culture and doxycycline in the animal model, the KstR protein will be degraded at various time frames during infection. The degradation of KstR post-translationally will in essence mimic the *kstR* deletion and commence with the de-regulation of all the genes involved with *kstR* regulatory

mechanisms. This approach was used for initial experimentation to see if it could be used to bypass the required and essential regulatory role *kstR* plays during infection. [58]

Initial results from our experimentation showed the addition of the DAS+4 tag to the C-terminus of the gene resulted in a significantly reduced amount of KstR in western blots without the presence of ATc added to the media used for growth. However, KstR was indeed absent after the addition of ATc. Thin layer chromatography also verified similar PDIM deficiency characteristics between the CDC1551 Δ*kstR* strain and the KDKstR strain after ATc induction the ClpP proteolytic system. TLC experiments without the addition of ATc also presented a diminished quantity of PDIM for the KDKstR strain. This shows that there is a regulatory dysfunction with the KDKstR strain after manipulating kstR to also transcribe with the addition of a degradation tag. The ClpP protease system was confirmed to be either "leaky" or there are other proteases produced by CDC1551 that is able to recognize the DAS+4 tag and therefore diminish the amount of KstR present in normal growth conditions. The effect of KstR degradation in the strain without induction of ClpP protease was evident as well with the PDIM profile experiments. The PDIM profiles of KDKstR before and after adding ATc showed PDIM to be disrupted and unformed compared to the WT profile.

Macrophage infections carried out with the WT, Δ*kstR*, and KDKstR were carried out with the addition of ATc to clarify if KDKstR would be able to recover infection rates with the disruption of KstR post-translationally. The macrophage infections results offered great insight as to how KstR ablation through post-translational degradation is able to behave similarly to the Δ*kstR* deletion strain. It also offered confirmation as to how even a diminished amount of KstR is able to recover infection after initial onset infection deficiency. This infection data provides us confirmation the ClpP protease degradation system is a valid model to use for the KDKstR strain

in the mouse model for future experimentation. The ability to shut KstR off once the infection has been established would offer a novel approach in determining its regulatory functions during persistence and its role with the enduring hypoxic response in the mouse model. This methodology will provide essential data regarding at which phases during infection the presence of KstR is able to inhibit transcription, and importantly, for establishing and maintaining infection.

Chapter 4: Overview and Final Discussion

Mycobacterium tuberculosis has been an ongoing major health threat at a global capacity for hundreds of years. This threat does not appear to be subsiding with the emergence of Multidrug-resistant (MDR-TB) and Extensively Drug Resistant (XDR-TB) in recent years. According to the Texas State Health Services, there was an increase 1,161 new cases of TB diagnosed in 2019 compared to the previous year with 1,129 reported in 2018. Of the new cases in 2019, 11 of them were assessed to be MDR-TB. There is a significantly higher concentration of diagnoses for the Hispanic community (53%), and in the border cities such the city of El Paso, TX. For these reasons, coupled with the global threat already described, it is important for the continuation and further research for novel anti-tubercular drugs and therapeutic targets. Our research focuses on determining which key factors within cholesterol catabolism are elements that would provide additional knowledge and data to further potential therapeutic approaches.

Previous work has provided information regarding KstR being the transcriptional regulatory factor that is responsible for Mtb's ability to respond to the requirement for cholesterol catabolism during infection. [27, 49, 52, 53, 58] At the beginning of this study, the ligand for KstR had not been determined. We suspected the ligand for KstR would be a cholesterol catabolite that was produced during the early phases of the cholesterol degradation pathway and that it would more than likely have the attachment of Coenzyme A. The attachment of CoA was shown to be essential for ligand binding and release of DNA to other TetR like transcriptional regulators with the same structure. [51, 56] At the onset of experiments to determine ligand binding and release, we saw that KstR was able to bind to a number of cholesterol derivatives, but actual release from the DNA was not able to be achieved unless the presence of the attachment of CoA was located at the side chain of the cholesterol catabolite was achieved. It was proven the (25S)-3-oxocholest-4-

en-26-oyl-CoA (3OCh(25S)26-CoA) ligand to be the best target to bind to KstR and inhibit the KstR-DNA binding complex. (Fig. 5, 6, &7) When 3-oxo-4-cholestenoic acid was introduced to the EMSA the binding exhibited multiple nucleo-protein complexes rather than release, showing that it does in fact bind to KstR but it does not inhibit binding to the DNA, rather it promotes binding capabilities of the protein. (Fig. 5) Although binding of fatty acids and fatty acid-CoAs have been hypothesized to be able to assist in regulatory mechanisms and be possible ligands for KstR, the EMSAs conducted with them did not show release of the DNA-KstR complex, nor did they show the ability to produce nucleo-protein complexes which would suggest the mechanism of repression. (Fig. 13)

Other intricacies of binding were also observed during the EMSA studies. It was seen the -oxo group at the 3 position of the cholesterol derivative A-ring was preferred over an -OH group regardless of the size of the cholesterol-side chain length. (Fig. 7) Additionally, between the presence of 3-oxo-4-cholestenoic acid (3OCHacid) and 3OCh(25S)26-CoA, KstR has preferential binding to the latter. Inhibition of binding to DNA was not achieved with 3OCH(25S)26-oate showing the requirement for a thioester bond and the attachment of CoA to prevent DNA-KstR complex formation. (Fig. 11)

Crystallography studies were achieved with KstR and 3OCh(25S)26-CoA showed how the ligand and protein interact with one another as well as to the DNA operator. (Fig. 8-10) The 3OCh(25S)26-CoA ligand is able to bind to each homodimer of the KstR complex in a pocket that consists of the α4-α9 helices. (Fig. 8) Interaction and binding of 3OCh(25S)26-CoA to KstR increases the distance between the two DNA binding domains to 40 Å preventing the KstR-3OCh(25S)26-CoA complex from binding to the DNA. (Fig. 8) The CoA attachment portion of 3OCh(25S)26-CoA was not visualized in the crystallography studies implying that it is "wobbly"

and does not form a stable bond with any part of KstR. (Fig. 10) Interestingly, when looking at key amino acids at the top of KstR where the -CoA portion could be attached to a similar pattern of amino acids appeared to present themselves at more towards the DNA binding domain portion of the protein as well. This may offer a stability to the –CoA portion for attachment, but further studies for this would have to be completed through mutant variants of KstR.

Remarkably, the ligand stereoisomer conformation sent for crystallography was 3OCh(25R)26-CoA, yet when the data was analyzed the 3OCh(25S)26-CoA was presented. As cholesterol catabolism reactions occur in the S-stereoisomer conformation, it would be understandable for the transcriptional regulator to prefer the ligand in this form. After deliberation and questioning with chemists, it was agreed upon KstR has racemization capability where the thio-ester bond is located in the binding pocket. The attachment of –CoA provides the reactive thio-ester bond that would allow the deprotonation and switching of conformation via formation of the carbanion intermediate. To determine which amino acids in the pocket are responsible for the aide in shifting the equilibrium to the preferable enantiomer, point mutation expression experiments altering amino acids that are located where the thio-ester bond sits would have to be conducted along with additional crystallography studies.

During the course of the studies, it was discovered there was the possibility for KstR, similar to the EthR model, may have an extended N-terminus that has phosphorylation signaling properties. [56] The mycobacterial network research that has been conducted and presented online shows multiple binding sites for KstR in the 5' UTR of the *kstR* gene. [83] The binding of KstR to multiple binding sites may allow KstR to be transcribed in multiple isoforms within the bacteria in response to environments or stressors. For this reasoning it was decided to create and express an extended isoform of KstR (X-KstR) with an addition 20 amino acids that possesses a signal

sequence that has a the high likelihood of phosphorylation activity and those that mimic (Asp-KstR) or prevent (Ala-KstR) phosphorylation from occurring. Surface plasmon resonance studies with these new protein constructs showed the extension with the signaling sequence intact and the phosphor-ablative version was able to bind to the *kstR* operator at rates that were less than with the shortened version of KstR and the phosphor-mimic isoform of the protein bound at even less rates more than likely to the repulsion of charge. (Fig. 14) When the 3OCh(25R)26-CoA ligand was introduced to the experimental design, the Asp-KstR isoforms binding to DNA was slightly inhibited or had weaker binding compared to the other three protein constructs. (Fig. 14)

In vivo studies were carried out to determine how Mtb CDC1551 and H37Rv strains would behave lacking the regulatory functions of KstR. The gene was removed by homologous recombination and replaced with a hygromycin cassette for selection purposes. The deletion of *kstR* was confirmed through southern blot experiments. (Fig. 15) The complementation of KstR was initially carried out with an integrative vector with *kstR* under the control of its native promoter at the *attB* site in the *Mtb* genome and western blots were carried out to verify the presence of KstR. (Fig. 15) The recovery of KstR was not achieved to WT levels but it was decided to go forward with experimentation. (Fig. 15) The use of other promoters, such as the hsp60 promoter, at this time were not thought to be a good decision based on not wanting to over produce KstR and create an over-expression model.

Growth curves were performed for Mtb CDC1551 with the WT, Δ*kstR*, and Δ*kstR/kstR* to determine how each strain would respond to the lack of KstR regulation as well as with the complementation although it was previously determined the amounts of KstR were less than the WT. (Fig. 18) The Δ*kstR* strains for both the CDC1551 grown on both glycerol and glycerol:cholest-4-en-3-one conditions were reduced compared to that of the WT. The

complementation for CDC1551 did show a significant growth increase even with decreased amount of KstR expression yet the growth was still not up to wild type levels. It is important to note that for growth studies, glycerol was used as the main carbon source yet this is not normally the nutrient that is most abundant during infection. *Mtb* clearly prefers cholesterol as a carbon source and seems to utilize it first before switching its metabolism to glycolysis and this environment does not mimic that of a living host. The lack of recovery and production of KstR in the complement strains can more than likely be attributed to the placement of the gene. The gene was placed at a different locus where the interactions with other regulatory factors such as activators and other transcriptional regulator normally occurring for signaling events to take place would be hindered. This would prevent normally occurring transcriptional events to take place and therefore recovery of the production of KstR would be impossible with this method of complementation.

Studies to determine how each of these strains would cope with the toxic effects of cholest-4-en-3-one, cultures were grown and metabolites extracted at multiple time points to measure the quantities of cholesterol cholest-4-en-3-one. (Fig. 17) The CDC1551 Δ*kstR* strain was able to catabolize cholesterol in the predicted manner, faster than the WT with no additional cholest-4-en-3-one build up. If there would be an excess store of propionyl-CoA then it would have to be shuttled into various pathways to be able to overcome the toxic effects of cholest-4-en-3-one.

We decided to assess the amounts of phthiocerol dimycocerosate (PDIM) through thin layer chromatography studies with the CDC1551 WT, Δ*kstR,* and Δ*kstR/kstR* strains after introducing C^{14}labeled acetate, propionate, and stearate. (Fig. 20) PDIM is a known virulence factor that is attributed to macrophage phagosomal escape and survival of Mtb. [26, 86] It was found there were significant changes of PDIM production was after the introduction of C^{14}-labeled

propionate. CDC1551 WT had higher levels of PDIM and the Δ*kstR* levels were reduced. The for CDC1551 Δ*kstR/kstR* to recover PDIM levels to the WT levels is more than likely is a result of an altered level of regulation capabilities of Δ*kstR* being in a different locus of the gene.

Proteomic analysis was performed for the CDC1551 *Mtb* strain with the respective WT, Δ*kstR,* and Δ*kstR/kstR* strains after growth on cholesterol. (Fig. 21-23) The proteomic data confirmed there were major differences on how each copes with the lack of KstR regulation. The results for key enzymes in the cholesterol catabolic pathway were noted to be upregulated with both Δ*kstR Mtb* strains. We analyzed at enzymes that dictate what happens with what would be the larger pool of propionyl-CoA with the constant de-repression of cholesterol degradation pathway enzymes. It was previously discovered that both MutA and MutB displayed regulation by KstR and it appears this study has helped to validate those findings. [54] There was over-expression of both the subunits that are responsible for the precursor metabolite that is shuttled to the pathways for methyl-branched fatty acids. Another enzyme not previously shown to be KstR regulated was PapA5, yet the deletion of *kstR* in CDC1551 abated expression of the enzyme. PapA5 is the acyltransferase known for acylating phthiocerol with mycocerosic acid connecting the two branches to create PDIM. [77]

It was found there were only particular enzymes that were show to have differential expression and any one pathway was not directly affected by the removal of KstR. The most striking result was the absence in detection of Icl2 in the CDC1551 Δ*kstR* strain. (Fig. 22) H37Rv is believed to have the Icl2 gene disrupted by a frame shift mutation that does not allow the entire gene to be encoded as one gene therefore it is denoted at two separate genes. [89] There is a downstream ORF which possesses a ribosome binding site (RBS) which would allow the transcription of the second half of the known enzyme. In CDC1511 the two segments are

transcribed as one gene. It is plausible that the deletion of *kstR* provoked a frame shift mutation allowing for CDC1551 to transcribe the two genes which may account for the excess quantity seen in the knockout. The mechanisms into this phenomenon are worth studying in the future at a later date.

Over-expression of KstR would hinder the ability for *Mtb* to catabolize fatty acids and cholesterol as well as de-regulate the enduring hypoxic response (EHR). KstR is known to play a role in EHR. [54] As KstR is not considered to be a part of the central hub for *Mtb* hypoxic response, but cholesterol degradation is believed to play a larger role in hypoxia and dormancy states. [90] Macrophage infection studies showed the inability for the over-expression of KstR to infect as well as the WT, Δ*kstR*, or even the pMV306-*kstR* complement that has less KstR expression than the WT strain. This is more than likely due to the prevention of the transcription of enzymes to utilize fatty acids as a carbon source and produce virulence factors that are known to be protective to *Mtb*, such as PDIM.

As we have seen through our studies, choosing the correct model for KstR studies has been challenging. Another approach we took was that of a knockdown system where KstR would be able to be degraded post-translationally. It was believed in the beginning the KDKstR strain would be able to replace the WT strain in studies, as without addition of ATc, KstR would be able to be transcribed naturally under the control of its native promoter and be able to respond naturally to environmental stimuli. This was not the case as the KDKstR strain showed attenuation of KstR even in the absence of ATc. (Fig. 26-28) This issue was able to be overcome though because KstR was still present even in lesser amounts. KstR is required for the mouse model of infection. [36, 58, 59] With even lesser amounts of KstR present, hypothetically, the KDKstR would allow infection to transpire and KstR would be able to be degraded and "shut off" with the addition of

ATc during the course of infection. This would allow various time frames of infection to be tested for possible treatment effects by de-regulating KstR through derepression.

During the course of these studies, novel and significant information regarding the characterization of how the negative transcriptional regulator KstR effects the fitness and survival in different strains of *Mtb*. These studies have established ligands that both promote the attachment or inhibition of KstR binding to its prospective operators is key to determining which types of structures to focus on to initiate a particular desired effect. As both deletion of *kstR* and depletion of KstR both produce an attenuated effect both conditions are compelling models for de-repression states of not only cholesterol catabolism but other regulatory mechanisms through interactions with other pathways and networks. This can be achieved successfully with chemicals that are able to mimic the ligand substrates that are similar to 3OCh(25S)26-CoA. On the other hand, repression of KstR control of cholesterol catabolism and regulatory functions can be induced with ligands that are similar or mimic the 3O(25R)26-oic acid function of the formation of nucleoprotein complexes as the over-expression of KstR model showed attenuation of infection as well. With high throughput screening of testing chemicals or molecules similar to these catabolites, therapeutic drugs may be designed to weaken *Mtb* and perhaps used as a drug treatment or therapy.

References

1. Cole, S.T., et al., *Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.* Nature, 1998. **393**(6685): p. 537-544.

2. Coscolla, M., et al., *Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans.* Nature …, 2013.

3. Wirth, T., et al., *Origin, spread and demography of the Mycobacterium tuberculosis complex.* PLoS Pathog, 2008. **4**(9): p. e1000160.

4. Smith, I., *Mycobacterium tuberculosis Pathogenesis and Molecular Determinants of Virulence.* Clinical Microbiology Reviews, 2003. **16**(3): p. 463-496.

5. Haas, F., and S. S. Haas, *The origins of Mycobacterium tuberculosis and the notion of its contagiousness.* Tuberculosis, 1996. **W. N. Rom and S. Garay (ed.)**(Tuberculosis): p. p. 3-19.

6. Lipsitch, M. and A.O. Sousa, *Historical intensity of natural selection for resistance to tuberculosis.* Genetics, 2002. **161**(4): p. 1599-1607.

7. Lerner, B.W.L.a.K.L., *Robert Koch.* World of Microbiology and Immunology, 14 Apr. 2013. **Detroit: Gale, 2006.**(Web).

8. Steenken, W., W.H. Oatway, and S.A. Petroff, *BIOLOGICAL STUDIES OF THE TUBERCLE BACILLUS III. DISSOCIATION AND PATHOGENICITY OF THE R AND S VARIANTS OF THE HUMAN TUBERCLE BACILLUS (H37).* The Journal of Experimental Medicine, 1934. **60**(4): p. 515-540.

9. World Health, O., *BCG vaccine.* Wekly epidemiological record, 2004. **79th**(No. 4).

10. Forrellad, M.A., et al., *Virulence factors of the Mycobacterium tuberculosis complex.* Virulence, 2013. **4**(1): p. 3-66.

11.   World Health Organization, W., *Global Tuberculosis Report 2016.* Global Tuberculosis Report 2016, 2016.

12.   Awofeso, N., *Anti-tuberculosis medication side-effects constitute major factor for poor adherence to tuberculosis treatment.* Bulletin of the World Health Organization, 2008: p. 240-240.

13.   Goldstein, B.P., *Resistance to rifampicin: a review.* J Antibiot, 2014. **67**(9): p. 625-630.

14.   Cade, C.E., et al., *Isoniazid-resistance conferring mutations in Mycobacterium tuberculosis KatG: Catalase, peroxidase, and INH-NADH adduct formation activities.* Protein Science, 2010.

15.   Abate, D., et al., *Isoniazid and rifampicin resistance mutations and their effect on second-line anti-tuberculosis treatment.* The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease, 2014. **18**(8): p. 946-951.

16.   Mohn, W.W., et al., *The Actinobacterial mce4 Locus Encodes a Steroid Transporter.* Journal of Biological Chemistry, 2008.

17.   Uhia, I., et al., *Initial step in the catabolism of cholesterol by Mycobacterium smegmatis mc2 155.* Environ Microbiol, 2011. **13**(4): p. 943-59.

18.   Capyk, J.K., et al., *Mycobacterial cytochrome p450 125 (cyp125) catalyzes the terminal hydroxylation of c27 steroids.* The Journal of biological chemistry, 2009. **284**(51): p. 35534-35542.

19.   Rosłoniec, K.Z., et al., *Cytochrome P450 125 (CYP125) catalyses C26-hydroxylation to initiate sterol side-chain degradation in Rhodococcus jostii RHA1.* Molecular microbiology, 2009. **74**(5): p. 1031-1043.

20.   Johnston, J.B., H. Ouellet, and P.R. Ortiz de Montellano, *Functional redundancy of steroid C26-monooxygenase activity in Mycobacterium tuberculosis revealed by biochemical and genetic analyses.* The Journal of biological chemistry, 2010. **285**(47): p. 36352-36360.

21.   Ouellet, H., et al., *Mycobacterium tuberculosis CYP125A1, a steroid C27 monooxygenase that detoxifies intracellularly generated cholest-4-en-3-one.* Molecular microbiology, 2010. **77**(3): p. 730-742.

22.   Casabon, I., et al., *Actinobacterial acyl coenzyme A synthetases involved in steroid side-chain catabolism.* Journal of bacteriology, 2013. **196**(3): p. 579-587.

23.   Muñoz-Elías, E.J., et al., *Role of the methylcitrate cycle in Mycobacterium tuberculosis metabolism, intracellular growth, and virulence.* Molecular Microbiology, 2006. **60**(5): p. 1109-1122.

24.   Savvi, S., et al., *Functional Characterization of a Vitamin B12-Dependent Methylmalonyl Pathway in Mycobacterium tuberculosis: Implications for Propionate Metabolism during Growth on Fatty Acids.* Journal of Bacteriology, 2008. **190**(11): p. 3886-3895.

25.   Minnikin, D.E., G. Dobson, and I.G. Hutchinson, *Characterization of phthiocerol dimycocerosates from Mycobacterium tuberculosis.* Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism, 1983. **753**(3): p. 445-449.

26.   Day, T.A., et al., *Mycobacterium tuberculosis strains lacking the surface lipid phthiocerol dimycocerosate are susceptible to killing by an early innate host response.* Infection and Immunity, 2014. **82**(12): p. 13.

27.   Ouellet, H., J.B. Johnston, and P.R. de Montellano, *Cholesterol catabolism as a therapeutic target in Mycobacterium tuberculosis.* Trends in microbiology, 2011. **19**(11): p. 530-539.

28.   Gatfield, J. and J. Pieters, *Essential role for cholesterol in entry of mycobacteria into macrophages.* Science, 2000.

29.   Munoz, S. and B. Rivas-Santiago, *Mycobacterium tuberculosis Entry into Mast Cells Through Cholesterol-rich Membrane Microdomains.* Scandinavian journal of …, 2009.

30.   Kaul, D., et al., *Cholesterol-sensor initiates M. tuberculosis entry into human macrophages.* 2004.

31.   Russell, D.G., et al., *The macrophage marches on its phagosome: dynamic assays of phagosome function.* Nature reviews. Immunology, 2009. **9**(8): p. 594-600.

32.   Rohde, K., et al., *Mycobacterium tuberculosis and the environment within the phagosome.* Immunological reviews, 2007. **219**: p. 37-54.

33.   Russell, D.G. and D.G. Russell, *Who puts the tubercle in tuberculosis?* Nature Reviews Microbiology, 2007.

34.   Griffin, J.E., et al., *Cholesterol catabolism by Mycobacterium tuberculosis requires transcriptional and metabolic adaptations.* Chemistry & biology, 2012. **19**(2): p. 218-227.

35.   Peyron, P., et al., *Foamy Macrophages from Tuberculous Patients' Granulomas Constitute a Nutrient-Rich Reservoir for M. tuberculosis Persistence.* PLoS Pathogens, 2008.

36.   Schnappinger, D., et al., *Transcriptional Adaptation of Mycobacterium tuberculosis within Macrophages: Insights into the Phagosomal Environment.* The Journal of experimental medicine, 2003. **198**(5): p. 693-704.

37.   der Geize, V.R., K. Yam, and T. Heuser, *A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into Mycobacterium tuberculosis survival in macrophages.* Proceedings of the …, 2007.

38. Nesbitt, N.M., et al., *A thiolase of Mycobacterium tuberculosis is required for virulence and production of androstenedione and androstadienedione from cholesterol.* A thiolase of Mycobacterium tuberculosis is required for virulence and production of androstenedione and androstadienedione from cholesterol, 2010.

39. Pandey, A.K. and C.M. Sassetti, *Mycobacterial persistence requires the utilization of host cholesterol.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(11): p. 4376-4380.

40. Lee, W., et al., *Intracellular Mycobacterium tuberculosis exploits host-derived fatty acids to limit metabolic stress.* The Journal of biological chemistry, 2013. **288**(10): p. 6788-6800.

41. Savvi, S., et al., *Functional characterization of a vitamin B12-dependent methylmalonyl pathway in Mycobacterium tuberculosis: implications for propionate metabolism during growth ....* Journal of …, 2008.

42. Upton, A.M. and J.D. McKinney, *Role of the methylcitrate cycle in propionate metabolism and detoxification in Mycobacterium smegmatis.* Microbiology, 2007.

43. McKinney, J.D., et al., *Persistence of Mycobacterium tuberculosis in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase.* Nature, 2000. **406**(6797): p. 735-738.

44. Jain, M., C.J. Petzold, and M.W. Schelle, *Lipidomics reveals control of Mycobacterium tuberculosis virulence lipids via metabolic coupling.* Proceedings of the …, 2007.

45. Rainwater, D.L. and P.E. Kolattukudy, *Isolation and characterization of acyl coenzyme A carboxylases from Mycobacterium tuberculosis and Mycobacterium bovis, which produce multiple methyl-branched ....* Journal of bacteriology, 1982.

46. Russell, D.G., et al., *Mycobacterium tuberculosis wears what it eats.* Cell host & microbe, 2010. **8**(1): p. 68-76.

47. Cox, J.S., et al., *Complex lipid determines tissue-specific replication of Mycobacterium tuberculosis in mice.* Nature, 1999. **402**(6757): p. 79-83.

48. Goren, M.B., O. Brokl, and W.B. Schaefer, *Lipids of putative relevance to virulence in Mycobacterium tuberculosis: correlation of virulence with elaboration of sulfatides and strongly acidic lipids.* Infection and immunity, 1974: p. 142-149.

49. Kendall, S.L., et al., *Cholesterol utilization in mycobacteria is controlled by two TetR-type transcriptional regulators: kstR and kstR2.* Microbiology, 2010. **156**(5): p. 1362-1371.

50. Casabon, I., et al., *Regulation of the KstR2 regulon of Mycobacterium tuberculosis by a cholesterol catabolite.* Molecular microbiology, 2013. **89**(6): p. 1201-1212.

51. García-Fernández, J., et al., *Characterization of the KstR2 regulator responsible of the lower cholesterol degradative pathway in Mycobacterium smegmatis.* Environmental microbiology reports, 2014.

52. García-Fernández, E., F.J. Medrano, and B. Galán, *Deciphering the Transcriptional Regulation of Cholesterol Catabolic Pathway in Mycobacteria IDENTIFICATION OF THE INDUCER OF KstR REPRESSOR.* Deciphering the Transcriptional Regulation of Cholesterol Catabolic Pathway in Mycobacteria IDENTIFICATION OF THE INDUCER OF KstR REPRESSOR, 2014.

53. Kendall, S.L., et al., *A highly conserved transcriptional repressor controls a large regulon involved in lipid degradation in Mycobacterium smegmatis and Mycobacterium tuberculosis.* Molecular microbiology, 2007. **65**(3): p. 684-699.

54.   Galagan, J.E., et al., *The Mycobacterium tuberculosis regulatory network and hypoxia.* Nature, 2013. **499**(7457): p. 178-183.

55.   Ramos, J.L. and M. Martínez-Bueno, *The TetR family of transcriptional repressors.* The TetR family of transcriptional repressors, 2005.

56.   Leiba, J., et al., *The Mycobacterium tuberculosis transcriptional repressor EthR is negatively regulated by Serine/Threonine phosphorylation.* Biochemical and Biophysical Research Communications, 2014. **446**(4): p. 1132-1138.

57.   Miller, M.L., et al., *NetPhosBac - a predictor for Ser/Thr phosphorylation sites in bacterial proteins.* Proteomics, 2009. **9**(1): p. 116-25.

58.   Sassetti, C.M., D.H. Boyd, and E.J. Rubin, *Genes required for mycobacterial growth defined by high density mutagenesis.* Molecular Microbiology, 2003. **48**(1): p. 77-84.

59.   Rengarajan, J., B.R. Bloom, and E.J. Rubin, *Genome-wide requirements for Mycobacterium tuberculosis adaptation and survival in macrophages.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(23): p. 8327-8332.

60.   Frénois, F., et al., *Structure of EthR in a ligand bound conformation reveals therapeutic perspectives against tuberculosis.* Structure of EthR in a ligand bound conformation reveals therapeutic perspectives against tuberculosis, 2004.

61.   Routh, M.D., et al., *Structures of AcrR and CmeR: insight into the mechanisms of transcriptional repression and multi-drug recognition in the TetR family of regulators.* Biochimica et Biophysica Acta (…, 2009.

62.   Baulard, A.R., et al., *Activation of the pro-drug ethionamide is regulated in mycobacteria.* Journal of Biological Chemistry, 2000. **275**(36): p. 28326-28331.

63.   Willand, N., et al., *Synthetic EthR inhibitors boost antituberculous activity of ethionamide.* Nature medicine, 2009. **15**(5): p. 537-544.

64.   Leslie, A.G.W., *Recent changes to the MOSFLM package for processing film and image plate data.* Recent changes to the MOSFLM package for processing film and image plate data, 1992.

65.   Holton, J. and A.-T. of the of, *Automated protein crystal structure determination using ELVES.* Proceedings of the National Academy of …, 2004.

66.   D, C.-C.P. and Biological, *The CCP4 suite: programs for protein crystallography.* Acta crystallographica. Section D, 1994.

67.   Biological, C.-K., *The Buccaneer software for automated model building. 1. Tracing protein chains.* Acta Crystallographica Section D: Biological …, 2006.

68.   Emsley, P. and C.-K. Biological, *Coot: model-building tools for molecular graphics.* Acta Crystallographica Section D: Biological …, 2004.

69.   Murshudov, G.N. and V.-A.A. Crystallographica …, *Refinement of macromolecular structures by the maximum-likelihood method.* Acta Crystallographica …, 1997.

70.   Ho, N.A., et al., *The Structure of the Transcriptional Repressor KstR in Complex with CoA Thioester Cholesterol Metabolites Sheds Light on the Regulation of Cholesterol Catabolism in Mycobacterium tuberculosis.* The Journal of biological chemistry, 2016. **291**(14): p. 7256-7266.

71.   Leiba, J., et al., *A novel mode of regulation of the Staphylococcus aureus catabolite control protein A (CcpA) mediated by Stk1 protein phosphorylation.* Journal of Biological Chemistry, 2012. **287**(52): p. 43607-43619.

72.   Canova, M.J., R. Veyron-Churlet, and Z.C.-I. …, *The Mycobacterium tuberculosis serine/threonine kinase PknL phosphorylates Rv2175c: Mass spectrometric profiling of the activation loop phosphorylation sites and …. …*, 2008.

73.   Cohen-Gonsaud, M., et al., *The Mycobacterium tuberculosis Ser/Thr Kinase Substrate Rv2175c Is a DNA-binding Protein Regulated by Phosphorylation.* Journal of Biological Chemistry, 2009. **284**(29): p. 19290-19300.

74.   Ouellet, H., et al., *Mycobacterium tuberculosis CYP125A1, a steroid C27 monooxygenase that detoxifies intracellularly generated cholest-4-en-3-one.* 2010. **77**(3): p. 730-742.

75.   Bonds, A.C. and N.S. Sampson, *More than cholesterol catabolism: regulatory vulnerabilities in Mycobacterium tuberculosis.* Current opinion in chemical biology, 2018. **44**: p. 39-46.

76.   Cantaloube, S., et al., *The Mycobacterium Tuberculosis FAS-II Dehydratases and Methyltransferases Define the Specificity of the Mycolic Acid Elongation Complexes.* PLoS ONE, 2011. **6**(12): p. e29564.

77.   Yu, J., et al., *Both phthiocerol dimycocerosates and phenolic glycolipids are required for virulence of Mycobacterium marinum.* Infection and …, 2012.

78.   Manca, C., et al., *Mycobacterium tuberculosis CDC1551 induces a more vigorous host response in vivo and in vitro, but is not more virulent than other clinical isolates.* Journal of immunology (Baltimore, Md. : 1950), 1999. **162**(11): p. 6740-6746.

79.   Quigley, J., et al., *The Cell Wall Lipid PDIM Contributes to Phagosomal Escape and Host Cell Exit of Mycobacterium tuberculosis.* mBio, 2017. **8**(2): p. 17.

80.   Rustad, T.R., et al., *The enduring hypoxic response of Mycobacterium tuberculosis.* PloS one, 2008. **3**(1).

81. Upton, A.M. and J.D. McKinney, *Role of the methylcitrate cycle in propionate metabolism and detoxification in Mycobacterium smegmatis.* Microbiology, 2007. **153**(12): p. 3973-3982.

82. Rustad, T.R., et al., *Mapping and manipulating the Mycobacterium tuberculosis transcriptome using a transcription factor overexpression-derived regulatory network.* Genome biology, 2014. **15**(11): p. 502.

83. Turkarslan, S., et al., *A comprehensive map of genome-wide gene regulation in Mycobacterium tuberculosis.* Scientific data, 2015. **2**(1): p. 150010.

84. Giovannini, D., et al., *A new Mycobacterium tuberculosis smooth colony reduces growth inside human macrophages and represses PDIM Operon gene expression. Does an heterogeneous population exist in intracellular mycobacteria?* Microbial pathogenesis, 2012. **53**(3-4): p. 135-146.

85. Cambier, C.J., et al., *Mycobacteria manipulate macrophage recruitment through coordinated use of membrane lipids.* Nature, 2014. **505**(7482): p. 218-222.

86. Quigley, J., et al., *The cell wall lipid PDIM contributes to phagosomal escape and host cell exit of Mycobacterium tuberculosis.* The cell wall lipid PDIM contributes to phagosomal escape and host cell exit of Mycobacterium tuberculosis, 2017.

87. Simeone, R., et al., *Cytosolic access of Mycobacterium tuberculosis: critical impact of phagosomal acidification control and demonstration of occurrence in vivo.* PLoS …, 2015.

88. microbiology, R.-D.G., *The ins and outs of the Mycobacterium tuberculosis-containing vacuole.* Cellular microbiology, 2016.

89. Nandakumar, M., C. Nathan, and K.Y. Rhee, *Isocitrate lyase mediates broad antibiotic tolerance in Mycobacterium tuberculosis.* Nature communications, 2014. **5**: p. 4306.

90. Soto-Ramirez, M.D., et al., *Cholesterol plays a larger role during Mycobacterium tuberculosis in vitro dormancy and reactivation than previously suspected.* Tuberculosis (Edinburgh, Scotland), 2017. **103**: p. 1-9.

Appendix

Figure 22: KstR **Ser/Thr phosphorylation prediction calculations**

```
>Sequence        220 amino acids
#
# netphosbac-1.0a prediction results
#
# Sequence                 # x    Context     Score   Kinase    Answer
# ----------------------------------------------------------------------
# Sequence                 2 S    ---MSSANT   0.163   main        .
# Sequence                 3 S    --MSSANTN   0.478   main        .
# Sequence                 6 T    SSANTNTSS   0.112   main        .
# Sequence                 8 T    ANTNTSSAP   0.215   main        .
# Sequence                 9 S    NTNTSSAPD   0.303   main        .
# Sequence                10 S    TNTSSAPDA   0.552   main        Y
# Sequence                28 S    VLAESELGS   0.185   main        .
# Sequence                32 S    SELGSEAQR   0.391   main        .
# Sequence                46 T    ILDATMAIA   0.128   main        .
# Sequence                51 S    MAIASKGGY   0.173   main        .
# Sequence                73 T    VAVGTLYRY   0.305   main        .
# Sequence                80 S    RYFPSKVHL   0.598   main        Y
# Sequence                87 S    HLLVSALGR   0.386   main        .
# Sequence                94 S    GREFSRIDA   0.264   main        .
# Sequence               100 T    IDAKTDRSA   0.275   main        .
# Sequence               103 S    KTDRSAVAG   0.511   main        Y
# Sequence               109 T    VAGATPFQR   0.361   main        .
# Sequence               132 T    NPLLTEAMT   0.121   main        .
# Sequence               136 T    TEAMTRAYV   0.113   main        .
# Sequence               145 S    FADASAASE   0.468   main        .
# Sequence               148 S    ASAASEVDQ   0.253   main        .
# Sequence               159 S    KLIDSMFAR   0.530   main        Y
# Sequence               171 T    NGEPTEDQY   0.177   main        .
# Sequence               182 S    ARVISDVWL   0.464   main        .
# Sequence               187 S    DVWLSNLLA   0.441   main        .
# Sequence               194 T    LAWLTRRAS   0.346   main        .
# Sequence               198 S    TRRASATDV   0.553   main        Y
# Sequence               200 T    RASATDVSK   0.397   main        .
# Sequence               203 S    ATDVSKRLD   0.332   main        .
# Sequence               219 S    GDQDSA---   0.381   main        .
#
   MSSANTNTSSAPDAPPRAVMKVAVLAESELGSEAQRERRKRILDATMAIA    #     50
   SKGGYEAVQMRAVADRADVAVGTLYRYFPSKVHLLVSALGREFSRIDAKT    #    100
   DRSAVAGATPFQRLNFMVGKLNRAMQRNPLLTEAMTRAYVFADASAASEV    #    150
   DQVEKLIDSMFARAMANGEPTEDQYHIARVISDVWLSNLLAWLTRRASAT    #    200
   DVSKRLDLAVRLLIGDQDSA                                 #    250
%1 .........S........................................    #     50
%1 ...............................S..................    #    100
%1 ..S...............................................    #    150
%1 ........S...................................S..       #    200
%1 ....................
```

101

Table 2. Crystallography data collection and refinement statistics

Protein	KstR	KstR
Ligand	(25S)-3-oxocholest-4-en-26-oyl-CoA	DNA
PDB ID	**5AQC**	**5FMP**
Data collection		
Space group	C2	C2
Cell dimensions		
a, b, c (Å)	127.3, 67.4, 49.9	80.5, 62.8, 96.6
α, β, γ (°)	90.0, 105.7, 90.0	90.0, 89.9, 90.0
Molecules in AU	2	2
Wavelength	1.11587	1.11587
Resolution (Å)	1.66	2.26
R_{sym} or R_{merge} (%)	6.5 (64.9)[1]	5.8 (25.1)
$I / \sigma I$	8.6 (1.2)	8.9 (1.7)
Completeness (%)	92.4 (62.2)	87.3 (50.1)
Redundancy	3.6 (2.6)	2.4 (1.7)
Refinement		
No. reflections	42163	19214
R_{work} / R_{free} (%)	18.4/22.6	20.2/27.0
No. atoms		
Protein	2791	2817
Ligand	66	650
Solvent	227	96
Mean B value	26.7	33.1
B-factors		
Protein	27.4	35.4
Ligand	24.2	26.5
Solvent	38.2	34.0
R.m.s deviations		
Bond lengths (Å)	0.020	0.012
Bond angles (°)	1.940	1.484

Vita

Chenoa D. Arico, was born on June 3, 1976 in El Paso, TX as the only child of Stella Halliburton and Gary Bernhardt. The oldest of 3 siblings, she became the first grandchild to pursue her bachelor's degree at The University of Texas at El Paso (UTEP). She acquired her Bachelor of Science degree in Cellular and Molecular Biochemistry in Spring 2013 while being awarded the NSF-Student Mentoring to Achieve Retention Scholarship (NSF-SMARTS). After graduation, Chenoa was accepted into a prestigious Students Accelerating in Biomedical Research (SABR) program at Texas Tech University Health Sciences Center (TTUHSC), before beginning her doctoral program in Biological Sciences with a Pathobiology focus at UTEP.

For the first two years of her time in the graduate program, she was awarded the honorable Louis Stokes Alliances for Minority Participation (LSAMP) Bridges to Doctorates Fellowship to pursue her studies. Her application towards the National Science Foundation Graduate Research Fellowship Program awarded her an honorable mention. Additionally, she was awarded the Dodson Research Grant to provide funding for her laboratory experimentation. During her graduate studies she served as the Co-Chair for the Biological Sciences Graduate Student Association (BSGSA).

Her doctorate research has focused on elucidating the regulatory mechanisms of the cholesterol catabolic pathway controlled by the transcriptional regulator KstR in Mycobacterium tuberculosis which her technique with surface plasmon resonance led to collaborations with multiple principal investigators that have yielded published manuscripts. She has shown her proficiency in presenting her research in both oral and poster exhibitions and has had the great privilege towards mentoring other students and peers in both the laboratory and classroom setting.

Permanent address: 6920 Imperial Ridge Dr.

                          El Paso, TX, 79912

This thesis/dissertation was typed by Chenoa D. Arico.