

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Investigation into the biosynthesis and function of a novel sulfomenaquinone from *Mycobacterium tuberculosis*

Permalink

<https://escholarship.org/uc/item/5b15q68d>

Author

Sogi, Kimberly Makiko

Publication Date

2012

Peer reviewed|Thesis/dissertation

Investigation into the biosynthesis and function of a novel sulfomenaquinone from
Mycobacterium tuberculosis

by

Kimberly Makiko Sogi

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Chemistry
in the
Graduate Division
of the University of California, Berkeley

Committee in charge
Professor Carolyn R. Bertozzi, Chair
Professor Matthew B. Francis
Professor Daniel A. Portnoy

Spring 2012

Investigation into the biosynthesis and function of a sulfomenaquinone from
Mycobacterium tuberculosis

© 2012

By Kimberly Makiko Sogi

Abstract

Investigation into the biosynthesis and function of a sulfomenaquinone from
Mycobacterium tuberculosis

by Kimberly Makiko Sogi

Doctor of Philosophy in Chemistry
University of California, Berkeley
Professor Carolyn R. Bertozzi, Chair

Mycobacterium tuberculosis (Mtb) is the leading cause of bacterial deaths worldwide causing two million deaths annually and infects almost one third of the world population. Throughout its lifecycle, Mtb undergoes a series of changes in respiration and metabolic rates. During disease transmission from human to human, the bacteria are in a high respiratory and metabolic rate enabling rapid division. During infection, with the onset of the adaptive immune response, Mtb slows its respiratory and metabolic rates to a non-replicative state. Mtb can persist in this state for years to decades. The delicate balance of the host immune response and Mtb survival is only beginning to be understood.

In 2006, we identified a sulfomenaquinone, named S881 for its exact mass, that required the sulfotransferase 3 (*stf3*) gene for production. A Mtb *stf3* mutant exhibited a hypervirulent phenotype in the mouse model of infection. Mice infected with the *stf3* mutant, succumbed to infection more rapidly than mice infected with wild type and contained a higher bacterial burden in the lungs.

This thesis explores structural characterization of S881 and its biosynthesis and function. Chapter 1 reviews the role of menaquinone, the metabolic precursor to S881, in Mtb, while Chapter 2 discusses the characterization of the S881 structure using high resolution mass spectrometry fragmentation analysis. Chapter 3 discusses the biosynthesis of S881 and investigates the regulation of S881 production. Using the S881 mutants, Chapter 4 explores various models of Mtb infection to study the function of S881. Finally Chapter 5 describes the initial characterization of the promoter for S881 genes to understand S881 genetic regulation.

This thesis is dedicated to my family who fostered my sense of wonder at the world
and always encouraged my questioning mind

Investigation into the biosynthesis and function of a novel sulfomenaquinone from
Mycobacterium tuberculosis

Table of Contents

List of Figures	v
List of Tables	vi
List of Schemes	vii
Acknowledgements	viii

**Chapter 1: Occurrences, biosynthesis and function of menaquinone in
*Mycobacterium tuberculosis***

Introduction	1
Structure and phylogeny	2
The role of menaquinone in the electron transport chain	3
The role of menaquinone in the dormancy regulon activation	5
The role of menaquinone in disulfide bond formation in secreted proteins	5
Biosynthesis of menaquinone in mycobacteria	6
Menaquinone biosynthesis as a drug target	8
Summary	10
Thesis overview	10
References	10

**Chapter 2: Structural characterization of a novel sulfolipid produced by
*Mycobacterium tuberculosis***

Introduction	15
Results	16
S881 was enriched from Mtb total lipid extracts	16

FT-ICR enabled exact mass and elemental composition analysis	17
Tandem MS of S881 suggests that it is a polyisoprenoid	18
Confirmation of S881 structure via fragmentation of structural analogues	20
Discussion	22
Materials and methods	23
References	26
Chapter 3: Identification of the S881 biosynthetic operon	
Introduction	29
Results	31
Bioinformatic analysis of the <i>stf3</i> operon	31
Two genes, <i>cyp128</i> and <i>stf3</i> , are required for the biosynthesis of S881 from menaquinone	31
The gene <i>cyp128</i> is required for S881 biosynthesis	33
S881 is produced at similar levels during all phases of growth	33
Overexpression of S881 biosynthesis leads to an accumulation of S881 in the bacterial pellet	34
Discussion	35
Materials and methods	36
References	38
Appendix	40
Chapter 4: Using Mtb models of infection to study the function of S881	
Introduction	48
Results	49
S881 mutants showed no phenotype in the starvation model of infection	49

S881 mutants showed no phenotype in the macrophage infection model	50
S881 mutants do not show a defect in cell wall integrity	51
S881 mutants showed no phenotype when treated with inhibitors of the electron transport chain	51
S881 mutants showed no phenotype when treated with a human antimicrobial peptide	52
Treatment of Mtb with inhibitors of respiration may cause a change in S881 levels	53
Discussion	53
Materials and methods	54
References	57
Appendix	59

Chapter 5: Characterization of *cyp128* transcriptional regulation

Introduction	60
Results	61
Bioinformatic analysis of <i>rv2269c</i> and <i>cyp128</i>	61
The gene <i>rv2269c</i> acts as the promoter for <i>cyp128</i> transcription	62
Mtb transcribes <i>cyp128</i> from an alternate start site when grown in starvation conditions	64
Expression and purification of Cyp128	64
Discussion	65
Materials and methods	66
References	69
Appendix	71

List of Figures

Figure 1-1	The structure of Vitamin K and related compounds	1
Figure 1-2	The menaquinone isoforms produced by Mtb	3
Figure 1-3	The oxidation/reduction of menaquinone and menaquinol	3
Figure 1-4	The electron transport chain of mycobacteria	4
Figure 1-5	Menaquinone biosynthesis in mycobacteria	6
Figure 1-6	Inhibitors of MenA, the DHNA prenyltransferase	8
Figure 1-7	Covalent inhibitors of MenE, an acyl-CoA synthetase	9
Figure 2-1	Enrichment of S881 after anion exchange separation	16
Figure 2-2	The tandem FT-ICR mass spectrum of S881 indicates that it is a polyisoprenoid-derived molecule	17
Figure 2-3	Fragmentation of S881 on the LTQ ion-trap	18
Figure 2-4	The m/z=170-240 region of the FT-ICR MS ² spectrum of Vitamin K ₁ and S881	20
Figure 2-5	Comparison of the fragmentation patterns of geranyl sulfate and S881	21
Figure 3-1	The <i>stf3</i> operon in Mtb and the biosynthesis of S881 from menaquinone	30
Figure 3-2	Total lipid extracts from <i>M. smeg</i> engineered strains	31
Figure 3-3	Analysis of the $\Delta cyp128$ Mtb mutant strain	32
Figure 3-4	The temporal regulation of S881 in WT Mtb grown in rich media	33
Figure 3-5	Localization of S881 in WT and $\Delta rv2269c$ strains	34
Figure 4-1	The <i>stf3</i> operon and the chemical structure of S881	49
Figure 4-2	Starvation of WT Mtb and $\Delta stf3$	50
Figure 4-3	Survival of $\Delta cyp128$ mutants in macrophage infections	51

Figure 4-4	A simplified electron transport chain from mycobacteria with a selection of inhibitors tested	52
Figure 4-5	Treatment of S881 mutants with the human antimicrobial peptide LL-37	52
Figure 4-6	Levels of S881 from WT Mtb after treatment with inhibitors of respiration	53
Figure 5-1	The <i>sf3</i> operon and the proposed S881 biosynthesis from menaquinone	61
Figure 5-2	Promoter activity in <i>M. smeg</i> and Mtb using a GFP assay	62
Figure 5-3	MS analysis of total lipid extracts from the $\Delta cyp128$ complement strains	63
Figure 5-4	Starvation of WT Mtb causes changes in the <i>cyp128</i> mRNA transcript	64
Figure 5-5	Expression of a truncated Cyp128 in <i>E. coli</i>	65

List of Tables

Table 1-1	Quinone content in a selection of bacteria	2
Table 1-2	Activity of bezoxazine or quinoxaline inhibitors	9
Table 2-1	Exact mass measurements of fragment ions generated via FT-ICR tandem MS	19
Table 3-1	Bacterial strains used in this chapter	39
Table 3-2	Plasmids used in this chapter	40
Table 3-3	Primers used in this chapter	40
Table 4-1	Minimum inhibitory concentrations for WT, S881 mutants and complements treated with oxidative stress and inhibitors of cell wall biosynthesis	51
Table 4-2	Bacterial strains used in this chapter	59
Table 5-1	Bacterial strains used in this chapter	71

Table 5-2	Plasmids used in this Chapter	72
Table 5-3	Primers used in this Chapter	73

List of Schemes

Scheme 2-1	S881 is a sulfated menaquinone with observed fragmentation	20
Scheme 2-2	The structures of phylloquinone and geranyl sulfate with observed fragmentation.	21
Scheme 2-3	Synthesis of geranyl sulfate	23

Acknowledgements

There are many people who have helped and encouraged me in my years at Berkeley. First and foremost has been my advisor Carolyn Bertozzi. I deeply admire her intellect and her ability to ask the most insightful questions on any topic. She has been a great support for me throughout my difficult project, always encouraging while asking the hard questions to guide me in my experiments. I am very grateful to have had the opportunity to write an NIH grant with her. Working closely with her on the grant made me a better writer and I learned how to present scientific proposals and research projects as well as budgeting and general grant writing strategy. In addition, being in her group has taught me how to effectively communicate to a wide range of audiences; from scientists to my parents. I feel I have absorbed so much more from her than I realize, just from listening to her at group meetings and discussing science in general. Thank you for letting me explore and expand my own scientific interests and giving me the encouragement I needed throughout my graduate career.

In addition, Carolyn has created an amazing work environment. The Bertozzi lab is truly a unique group of people with an amazing array of talents. I feel fortunate to have been part of such an amazing group of scientist that are supportive but challenge you daily about your ideas to help you do the best science possible. In particular, I want to thank the TB subgroup, past and present. I thank my lucky stars every day that Jessica Seeliger joined the lab shortly after I did. It was great to have someone to learn how to do everything with, bounce ideas off of, be a great friend and cheerleader even long after she left. Stavroula Hatzios has been a great role model for me. She was always helpful and cheerful whenever I needed anything. I had the lucky chance to sit next to Sarah Gilmore for the first couple years of graduate school. I learned an incredible amount from asking her questions and bouncing ideas off of her. She was always my first go to especially in the beginning because no matter what question I asked she never made me feel insecure about what I was asking. I deeply miss all of them in the BSL3. The three of them set the standard for our BSL3 and has made it the organized, well run facility it is today. I have very much enjoyed the new ideas and projects the newer members of the TB subgroup have brought. Kimberly Beatty is always full of great ideas and a witty remark about life, Mtb or research. I wish her the best with her own group and cannot wait to see the great diagnostics they develop. Sloan is a great resource for all things microbiology. She always has the random strains of bacteria I need and a wonderful cheerful attitude. With Ben, it has been great having another chemist in the TB subgroup. His project has been really exciting and I cannot wait to see what he discovers. In the last couple of months, I have had the opportunity to collaborate with Andy. He is always willing to help in anyway and I hope he can find his passion after graduate school. I have to include both Lisa Prach and Laura Flores in my thanks to the Tb subgroup. Lisa swooped in at the perfect time and saved the BSL3. Her organization skills and patience are something I aspire to. She has become a good friend and I look forward to hearing about her grand adventures in the future. Laura has been wonderfully cheerful and helpful in taking over a difficult position. She is always ready to help with anything you ask, goes above and beyond in her tasks, and is ready and willing to adapt to any new improvements we need in the BSL3. She is a great asset to the facility and I think it will truly flourish under her direction.

Mark Breidenbach has been an incredible resource for me both for scientific and epicurean advice. He was incredibly generous with his time helping me purify a highly recalcitrant protein. We had many great nights just the two of us, wine and either oysters or desserts. I will miss our “dates” to Hog Island oyster bar. Zev Gartner has been a wonderful

mentor for me both during our overlap in lab as well as after. He always amazes me with his unconventional ideas and ways of thinking. When I am stuck on an experiment or certain results, he always has a different perspective and great ideas of where to go next. John Jewett is one of the nicest guys I know with the best sense of humor. Always helpful and kind, he had the best pranks and April Fools jokes.

In the last 6 years, I have probably spent more time in the 810 Latimer than anywhere else. It has been a wonderful room to work in. In addition to Sarah, Pam was a great lab mate and was ready with advice about science, graduate school or projects when asked. Kanna moved into 810 shortly before I did and quickly informed me of all the little quirks of the lab. He has been a great friend and lab mate throughout graduate school. We have always been able to commiserate about graduate school together and keep each other on track for all the major mile stones. He has become an invaluable resource for all things mass spectrometry as well as just about anything else I ask about. I am very thankful for his friendship and his amazing generosity with his bench, supplies, buffers, really anything I have needed. For two years I had the fortune of having an amazing undergraduate, Gabriella Fragiadakis. She is incredibly smart and would always ask great questions that really kept me on my toes. I am thankful for her patience with me when I was overwhelmed and flustered. She was a great asset to me, helping me out when I could not get out of the BSL3 in time and doing great work. The newest additions to 810, Chelsea and Paresh, have been a lot of fun. Lab is always lively and full of smiles and laughter. I love our room field trips and hope they continue in the future.

I absolutely would not have made it through graduate school without help from our support staff: Asia, Karen, Olga, and Sia. Asia has become a great friend and I truly look forward to eating lunch with her. Karen is always ready to fight for us and get things done. Olga is always so willing to help out and improve the lab. They all work so hard to make the lab run as smoothly and efficiently as it does.

Both in and out of lab, Allison Narayan and Ellen Sletten have been great friends, gym buddies and support. Allie always seemed to understand exactly what I was going through and was always a supportive shoulder to laugh or cry on or lend me her desk when I needed a quiet moment. Ellen was always so generous with her time. She was a great gym buddy because she always pushed you a little farther to make you that much stronger. In addition, I have been lucky to have started a great book club that really knows how to eat! I have truly enjoyed our monthly meetings and encourage me to not only keep reading for pleasure in graduate school but to expand my literary horizons.

I have been very lucky to have great collaborators and resources outside the lab. Cynthia Holsclaw and Julie Leary have laid the ground work for my project. Without their initial work, my project would have been very different. I had the opportunity to teach for Jamie Cate and Caroline Cane. It was a great experience to help implement a new curriculum for biochemistry lab. Through conferences, I had the opportunity to meet and converse with some amazing scientists in the field; in particular, Peter Tonge, Dean Crick, and Branch Moody for being great resources and incredibly helpful in my project, experimental design and general insight about my project and the Mtb field. I also want to thank the Chemical Biology program for the opportunity to rotate, present my research and for career development programs. I am thankful to Sarah Hubbard in the Bertozzi lab, Professor Richmond Sarpong, Professor Donald Rio, Brian Pujanauski, and Professor Christopher Chang for letting me rotate in their labs and being my guide in my first year. I would like to thank Professor Matthew Francis, Professor Christopher Chang, and Professor Daniel Portnoy. Not only were they on my qualifying exam committee but

they have been great resources to discuss my work and have had valuable insight into experiments and ideas for my project.

Finally I have to thank my family, Mom, Dad and Kenny, for being supportive, listening when I talked about science, and for being fun and adventurous. Thank you to my grandparents who enabled me to have the education I needed to get me here. My family has been essential in my graduate school success. Last but not least I have to thank Christopher Gribble who has seen me through both the best and the worst of graduate school. He has been amazingly supportive of me throughout and patient with me in every mood. I could not have done it without you.

Chapter 1: Occurrences, biosynthesis and function of menaquinone in *Mycobacterium tuberculosis*

Introduction

Menaquinone is part of a family of lipoquinones that are essential in the electron transport chains of almost all organisms (1). In *Mycobacterium tuberculosis* (Mtb), the etiological agent of the disease tuberculosis, menaquinone biosynthesis is essential for maintenance of the proton gradient and the membrane potential both of which drive ATP production (2). Until recently, it was thought that Mtb only used menaquinone in the electron transport chain. Menaquinone has been identified as being involved in a range of disparate functions in Mtb and other bacteria including disulfide formation of secreted proteins and as a signal for the rate of electron transport. A sulfated form of menaquinone, named S881, has been isolated from Mtb strain H37Rv, but the function has yet to be determined. In addition, the menaquinone biosynthetic pathway has been identified as a good target for drug development. Inhibitors of menaquinone biosynthesis show bacterial growth inhibition *in vitro*, providing new leads for antimycobacterial drugs. This chapter will focus on the functions of menaquinone in Mtb. Mtb is the leading cause of death by bacterial infections causing 2 million deaths annually and infecting approximately one third of the world population (3).

In 1943 the Nobel Prize in Medicine was awarded to Hendrik Dam for the discovery of menaquinone as an essential vitamin for baby chickens and to Edward Doisy for the elucidation of the chemical structure. Most experiments were done using chicks, where without menaquinone in their diet, the chicks grew poorly, and if injured would bleed to death (4). In 1935, the unknown factor was named vitamin K (for “koagulations” vitamin (5)) and bacteria were identified as the source. Menaquinone was isolated from putrefied fish meal in 1938 (6,7) and shown to be chemically distinct from the vitamin K form isolated from plants, synthesized two years earlier (8). To differentiate the two forms of Vitamin K, the form isolated from green plants was named phyloquinone or Vitamin K₁ and the bacterial derived form was named menaquinone or Vitamin K₂ (Figure 1-1).

Prokaryotes utilize two quinone isoprene compounds as lipid soluble molecules found in

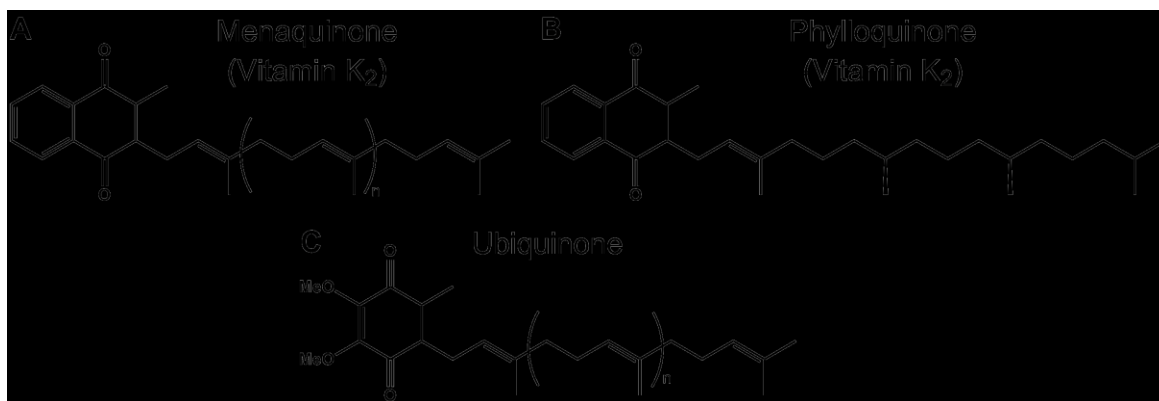


Figure 1-1: The structures of Vitamin K and related compounds. (A) Menaquinone, (B) Phylloquinone, (C) Ubiquinone.

the plasma membrane, ubiquinone and menaquinone (Figure 1-1). These compounds transfer electrons between membrane localized proteins involved in respiration. Ubiquinone is biosynthesized and utilized by many gram negative bacteria as well as mammalian mitochondria. In contrast, menaquinone is uniquely biosynthesized by bacteria, primarily gram positive bacteria, making it a possible drug target for many bacterial pathogens.

Structure and phylogeny

Menaquinone (MK) is composed of a naphthoquinone connected to an isoprene chain, the length and number of saturated double bonds varies depending on bacterial species (9) (Table 1-1). Double bonds in the isoprene chain, including the double bond in phyloquinone, are predominantly *trans* and the chiral centers were shown to be *R* with very few exceptions (4,10). A *cis* stereochemistry of the first isoprene unit is rare and may be an artifact of isolation procedures. The length of the isoprene chain and number of saturated double bonds in MK affects the redox potential of the quinone head group (11) enabling bacteria to alter the MK length and degree of saturation depending on their environment and respiratory needs. The differences in MK length and degree of saturation have been used to identify and classify bacterial species (10).

Mtb and other mycobacteria use MK-9(H₂) as the primary MK with small levels of MK-8, MK-9, MK-10 and MK-11 (10,12-14) (Figure 1-2). The levels of MK-10 and MK-11 are at the limit of detection and it is unclear of their role in respiration. Recently a study suggested that levels of MK-9(H₂) and MK-9 vary depending on the conditions of growth, which is further discussed below (15). Each of these MK isoforms is biosynthesized by the same pathway.

Table 1-1: Quinone content in a selection of bacteria

Organism	Quinone	Reference
<i>Mycobacterium tuberculosis</i>	MK-9(H ₂), MK-9,	(14)
<i>Escherichia coli</i>	Q-8 , Q-7, Q-6, Q-5, Q-4, Q-3, Q-2, Q-1, MK-8 , MK-9, MK-7, MK-6	(46)
<i>Bacillus subtilis</i>	MK-7	(47)
<i>Staphylococcus aureus</i>	MK-7 , MK-8 , MK-6, MK-5, MK-4	(48)
<i>Streptomyces</i> spp	MK-9(H₆) , MK-9(H ₈), MK-9(H ₄), MK-9(H ₂), MK-9	(15)

Quinone species in bold are the dominant quinone isolated

In 1958 Noll et al. may have found a more polar either precursor or hydroxylated form of MK in *Mycobacterium phlei* (*M. phlei*) but the structure was not determined (16). In 2006, Mougous and coworkers identified a novel sulfated MK, named S881 for its exact mass, in total lipid extracts from H37Rv Mtb (17) (Figure 1-2C). S881 is unique to the Mtb complex and is biosynthesized from menaquinone by two genes, *cyp128* and *stf3* (18); *cyp128* is annotated as a putative cytochrome P450 and *stf3* is a putative sulfotransferase (19,20). Mtb mutants lacking either *cyp128* ($\Delta cyp128$) or *stf3* ($\Delta stf3$) do not produce S881 (17,18). Interestingly, the $\Delta stf3$

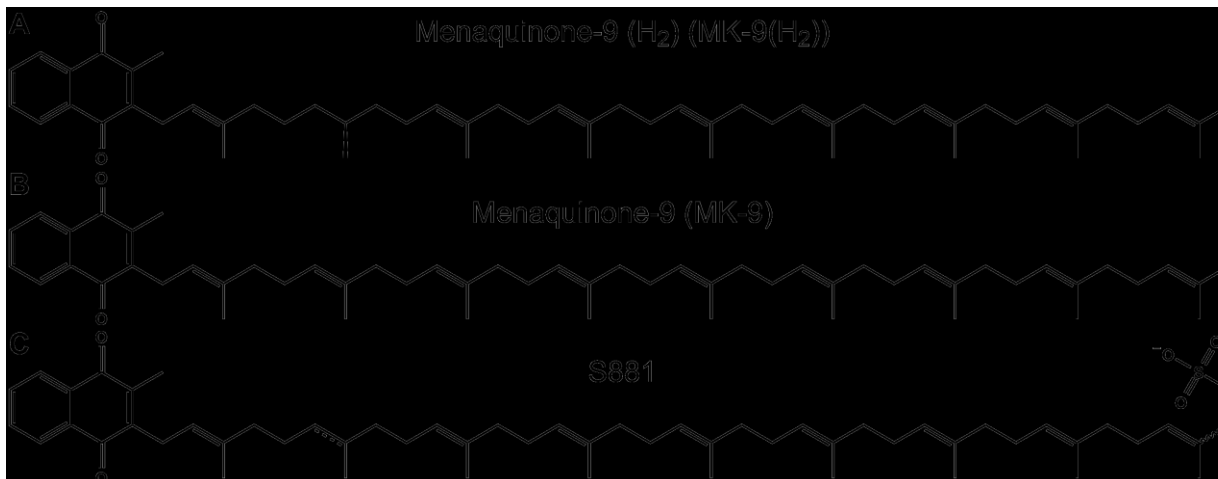


Figure 1-2: The menaquinone isoforms produced by Mtb. (A) Menaquinone-9 (H₂), (B) Menaquinone-9, (C) S881, a novel sulfated menaquinone

mutant shows a hypervirulent phenotype in the mouse model of infection. Mice infected with Δ *stf3* show increased bacterial burden in the lungs and spleen as well as increased time to death (17). The molecular basis for this highly unusual phenotype is further discussed in Chapter 4 of this thesis.

The role of menaquinone in the electron transport chain

As a component of the electron transport chain in mycobacteria, MK acts as the link between electron donating enzymes and electron accepting enzymes (Figure 1-3). The electron transport chain creates a redox loop where the electrons start at a low electron potential moving towards a higher electron potential. The electrons must flow “downhill”. For example, the mid-point potential at pH 7 (E_0') of NAD^+/NADH is -320 mV with an enzyme transferring those electrons to MK with an E_0' of -70 mV. Finally electrons are transferred to oxygen (O_2) via cytochromes and $\text{O}_2/\text{H}_2\text{O}$ has an E_0' of +820 mV. The free energy released in the process drives the formation of the proton gradient and the membrane potential. Depending on the growth conditions, Mtb utilizes a variety of electron donating enzymes and electron accepting enzymes (Figure 1-4) but MK is the only lipoquinone and small molecule that transfers electrons between membrane bound enzymes. In order to do this, MK must interact with various enzymes with different redox potentials. Based on evidence from other bacteria, one hypothesis is that the

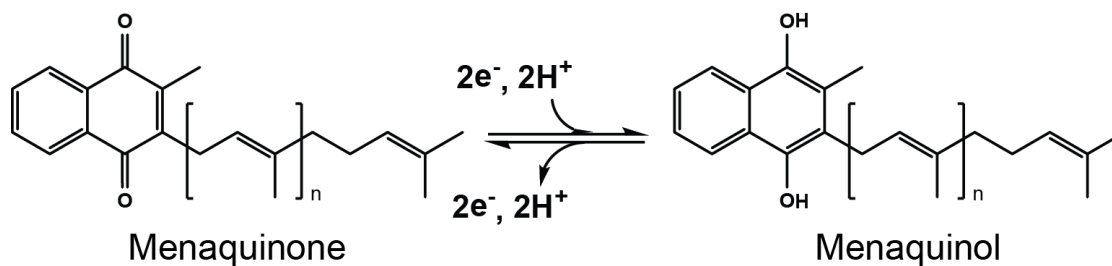


Figure 1-3: The oxidation/reduction of menaquinone and menaquinol.

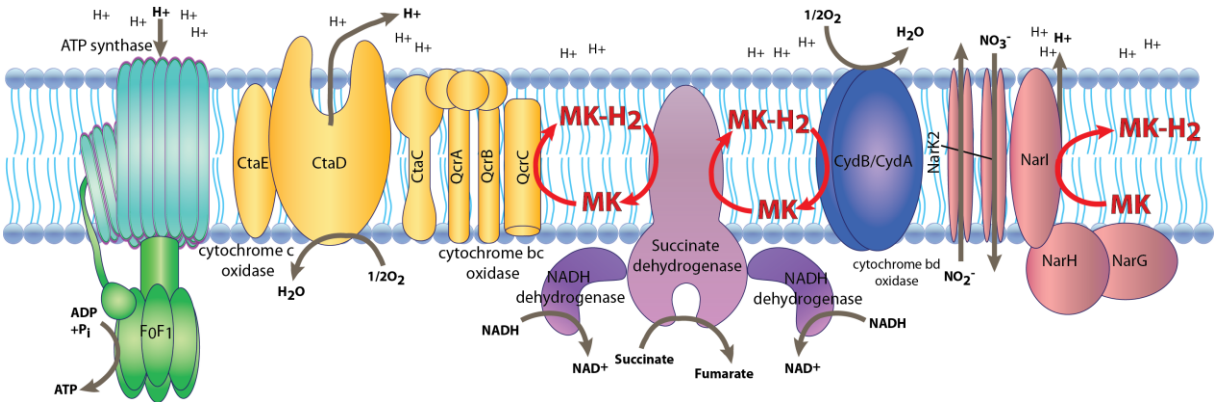


Figure 1-4: The electron transport chain of mycobacteria. Electrons are transferred from the dehydration complexes to menaquinone, shown in red. The menaquinone diffuses through the membrane to transfer electron to the oxidase or nitrate reductase complex. These enzymes create a proton gradient that the ATP synthase uses to convert ADP to ATP. Figure from ref (26).

different forms of MK, (i.e. MK9-(H₂) vs. MK9 vs. MK10) which have slightly different redox potentials (11), and interact with different enzymes depending on the growth conditions (21-24).

The mycobacterial electron transport chain can terminate in one of three possible terminal complexes: the cytochrome (cyt) *bc*₁-*aa*₃ cytochrome c oxidase complex, the cytochrome *bd* type oxidase, or the nitrate reductase, all enzymes that utilize electrons from menaquinol (25). Under aerobic conditions, electrons are transferred to the cyt *bc*₁ complexes and terminate in the *aa*₃-type cyt c oxidase for reduction of O₂ to H₂O. As the O₂ availability decreases, the bioenergetically less efficient *bd*-type menaquinol oxidase is up regulated and preferentially used to reduce O₂. The *bd*-type oxidase is essential for mycobacterial growth under microaerophilic conditions (26). As the cultures become hypoxic, the nitrate transporter (*narK*) is up regulated, increasing the reduction of nitrate to nitrite by the nitrate reductase (27). Studies of mutants with either the cyt *bc*₁ complex or the *aa*₃ type cyt c oxidase deleted show an up regulation of the cyt *bd* oxidase (28). Transcriptional profiling of the *bc*₁ mutant showed *lytB*, a gene involved in the biosynthesis of isoprenoids, a pathway necessary for MK biosynthesis, was the most highly down regulated gene (28,29). While the MK genes remain mostly unchanged, the down regulation of an upstream enzyme may play a role in how MK biosynthesis is controlled. Since it is unknown how long MK can be recycled, it is possible that down regulation of the terpenoids are simply a method for conserving resources without depleting the MK pool.

The nitrate reductase in *E. coli* has been extensively investigated, but the nitrate reductase (*narGHJ*) has not been studied biochemically in Mtb. The NarI subunit binds MK(H₂) for oxidation and transfers electrons through a series of nitrate reductase subunits via iron-sulfur clusters for the reduction of nitrate into nitrite (30). Mtb subjected to nitrous oxide (NO) stress, either *in vitro* or in mouse models, induces transcription of *narK2* the nitrate transporter (31). While the nitrate reductase genes are constitutively expressed at a low constant level, increased transcription of *narK2* results in an increase of nitrate to nitrite conversion. In a similar experiment, both terminal oxidases (cytochrome *c* complex and cytochrome *bd* complex) are down regulated after 50 days post infection in mice, indicating that an alternative electron acceptor is necessary (31), most likely nitrate. An alternative hypothesis suggests that increased

nitrate reduction is primarily for nitrogen assimilation, necessary for metabolite biosynthesis, and does not contribute to Mtb respiration (32).

The role of menaquinone in the dormancy regulon

As a result of its role in the electron transport chain, the ratio of MK-9(H2) to MK-9 may serve as a signal for inducing the dormancy regulon in Mtb (15). The dormancy regulon is a set of approximately 50 genes induced by DosR (DevR, Rv3133c), a transcription factor expressed under hypoxia, NO, and carbon monoxide (33-35). DosR is activated by two histidine kinases DosS (DevS; Rv3132c) and DosT (DevT; Rv2027c) (36-38). There are a number of stimuli known to activate DosT, but there is disagreement about DosS activation. Recently, Honaker and coworkers showed that the ratio of MK-9(H2) to MK-9 (Figure 1-2) may act as a signal to activate DosS dormancy regulon (15). There is no direct link between the ratio of MK isoforms and DosS activation suggesting that an intermediate may be required. Several bacterial systems monitor the state of the electron transport chain using quinones. One example is the ArcB system in *E. coli*, a two-component regulatory system of oxidative and fermentative catabolism. *E. coli* is somewhat unusual for a gram negative bacterium because it utilizes both ubiquinone and menaquinone (23). Under aerobic conditions ArcB is inactivated by binding ubiquinol, which is only present in aerobic environments. Ubiquinol binding causes inactivation by reduction of specific cysteines in the ArcB. As the level of oxygen is depleted, ubiquinol and ubiquinone levels decrease and MK levels increase. Since ArcB cannot utilize MK, the cysteines are oxidized, leading to activation of the histidine kinase domains and induction of the regulon (39,40). Further studies are necessary to determine the signaling pathway between MK and DosS activation.

The role of menaquinone in disulfide bond formation in secreted proteins

In addition to the function of menaquinone in the electron transport chain, mycobacteria encode a periplasmic protein DsbA that catalyzes protein disulfide bond formation in exported proteins by a MK dependent mechanism. In the *E. coli* DsbA homologue, two catalytic cysteines form disulfide bonds in secreted proteins. For regeneration, DsbA requires a second protein membrane bound protein that utilizes quinones for catalytic activity, DsbB (41). Many bacteria, mycobacteria included, contain a homologue to DsbA, but do not encode a DsbB homologue. Instead, these bacteria contain a homologue of the eukaryotic enzyme vitamin K epoxide reductase (VKOR) that may replace the role of DsbB in *E. coli*. The VKOR homologue in Mtb (Rv2968c) can complement an *E. coli dsbB* mutant ($\Delta dsbB$), but this complementation was dependent on the presence of DsbA (42). While the function of the human enzyme and the bacterial VKOR is very different, the enzymatic transformation is analogous: transfer of electrons from a thioredoxin-like protein to a quinone (43). In 2010, Li and coworkers published the crystal structure of the *Synechococcus* sp. VKOR bound to ubiquinone (44). Since mycobacteria do not produce ubiquinone, it is hypothesized that MK is used instead. A deletion of VKOR in *M. smegmatis* severely attenuates growth suggesting that this enzyme may prove to be a good drug target (43). Warfarin, an inhibitor of the human VKOR enzymes, also inhibits the mycobacterial VKOR homologue in liquid culture (43). Mycobacterial mutants resistant to warfarin show mutations in the VKOR enzymes similar to those found in humans requiring higher doses of warfarin.

Biosynthesis of menaquinone in mycobacteria

The biosynthesis of MK has primarily been studied in *E. coli*, with similar studies in *M. phlei* and *Bacillus subtilis*. Since *E. coli* produces and utilizes both ubiquinone and MK, it was possible to generate viable mutants, allowing for the assignment of the MK biosynthetic enzymes (45). Despite being essential in most Gram positive bacteria, *Staphylococcus aureus* and *B. subtilis* MK auxotrophs have been isolated in the clinic (46-48). Interestingly, the Men auxotrophs exhibit increased resistance to aminoglycosides (49). It is hypothesized that MK is involved in the active transport of aminoglycosides but the molecular details have not been elucidated. Similar studies have not been done in mycobacteria.

Only MK has been isolated from Mtb cell extracts and they lack the genes necessary for ubiquinone synthesis. This suggests that MK is the sole lipoquinone in mycobacteria (50). MK biosynthesis in mycobacteria is homologous to *E. coli*, but unlike *E. coli*, mycobacteria do not have all the *men* genes in a single operon (20). Homologues of the *E. coli men* genes (*menA-menG*) are clustered in the Mtb genome and the organization is similar to that of other mycobacteria (20,50).

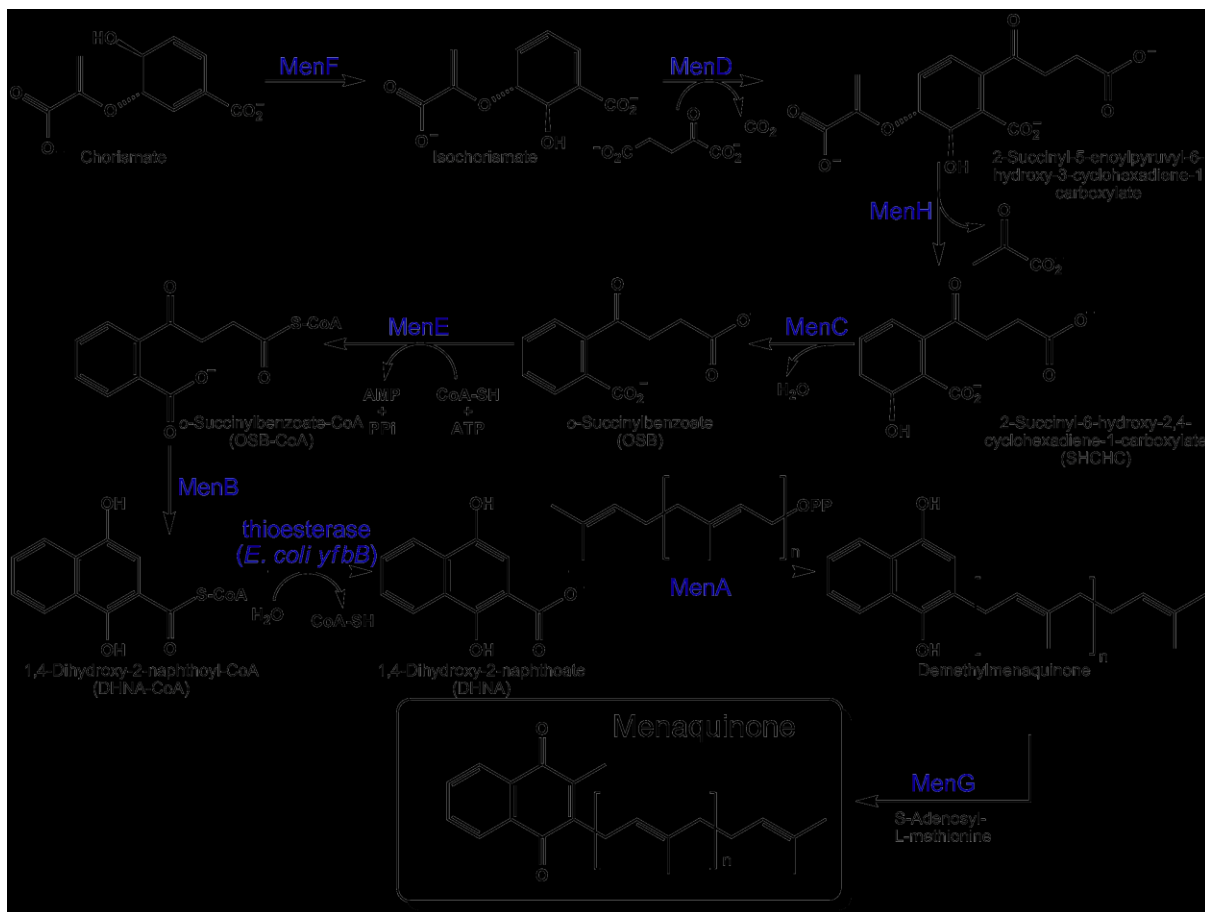


Figure 1-5: Menaquinone biosynthesis in mycobacteria (51).

Menaquinone biosynthesis begins with chorismate, derived from the shikimate pathway (4). The conversion of chorismate to isochorismate by MenF is the first committed step in the menaquinone biosynthetic pathway. MenD eliminates pyruvate and carbon dioxide and inserts α -ketoglutarate to form an intermediate, 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate (SHCHC), that has all the carbons for naphthoquinone formation (4). Elimination of water to form a benzoate by MenC produces the stable isolable intermediate O-succinylbenzoate (OSB). The conversion of isochorismate to OSB by MenC and MenD was determined using cell free extracts of *E. coli* mutants that required OSB for growth (4). The intermediate between MenD and MenC was characterized by Emmons and coworkers in 1985 as SHCHC (51). Subsequently, both MenD and MenC were purified for biochemical characterization (52,53) and an x-ray crystal structure has been determined for MenC (54,55). Next, OSB is activated by MenE forming OSB-CoA for cyclization by MenB to form 1,4-dihydroxy-2-naphthoyl-CoA (DHNA-CoA). From the crystal structure and *in vitro* biochemical experiments, a plausible mechanism has been proposed for OSB-CoA to DHNA-CoA. The mechanism is similar to a Claisen reaction, but the electrophile is the unactivated aromatic OSB carboxylate. The α -proton of the thio-ester is readily extracted allowing for attack at the OSB carboxylate forming a tetrahedral intermediate. This intermediate is converted into the DHNA-CoA by elimination of water (50).

The enzyme responsible for the conversion of DHNA-CoA to DHNA is unclear. Earlier reports indicate that MenB is responsible for hydrolyzing CoA, but an alternative enzyme, *yfbB*, located after MenD in the *E. coli* genome shows thioesterase function (1). Based on sequence alignments, Mtb does not contain a homologue to this gene and a candidate enzyme has not been identified. Following hydrolysis of CoA to form DHNA, MenA adds the prenyl group to form demethylmenaquinone. Methylation is catalyzed by MenG using S-adenosylmethionine. The final step in mycobacteria is the saturation of the second isoprene by an unidentified enzyme.

There is some conflicting data on the essentiality of the *men* genes. Sassetti and coworkers identified MenE, MenC, MenD and MenH as required for growth in liquid broth but MenA, MenB and MenG were not essential in this study (56,57). In contrast, none of the *men* genes were represented in studies that identify genes necessary for growth in the mouse infection model (56), which would suggest essentiality. As discussed later, inhibitors of menaquinone biosynthetic enzymes prevent bacterial replication and are possibly bactericidal (2).

The transcriptional regulation of the *men* genes in mycobacteria is not known, but extensive microarray data indicate that the *men* genes are regulated differentially from each other. Surprisingly, treatment of Mtb with an inhibitor of MenA does not change the transcription levels of *men* genes *menA*, *menB*, or *menH* (2). Similarly, microarrays comparing Mtb grown under hypoxic conditions to Mtb grown in oxygenated media do not show a change in transcript levels of the menaquinone biosynthetic genes (2). This is in contrast to other enzymes involved in electron transport, many of which utilize MK as a cofactor. All genes in the ATP synthase complex (*atpA-H*), subunits in the type 1 NADH:menaquinone dehydrogenase (*nuoC*, *nuoD*, *nuoF*, *nuoG*, *nuoK*, *nuoM*), *bc₁* cytochrome C reductase (*qcrC*, *qcrA*, *qcrB*) and *aa₃* cytochrome C oxidase (*ctaC*, *ctaD*, *ctaE*) are all up regulated during hypoxia compared to aerobic growth. As shown in Figure 1-4, complexes on both sides of the electrochemical gradient from MK show altered expression based on the growth conditions of Mtb, but the transcription of the *men* genes remain constant in these studies. As mentioned previously, other enzymes, such as those in the isoprene biosynthetic pathway or S881 biosynthetic enzymes, may act as regulators of MK levels. More research is necessary to determine how MK, a small molecule, is regulated in Mtb with changing environmental conditions.

Menaquinone biosynthesis as a drug target

Since MK is indirectly involved in ATP biosynthesis (5,58), it is essential in most Gram positive bacteria, mycobacteria included. As noted above, Rao and coworkers have shown that ATP production is essential for bacterial persistence during log growth and for non-replicating bacteria in both the Wayne and Loebel models of infection (59,60). There is no homologous biosynthesis of MK in humans, making the menaquinone biosynthetic enzymes possible drug targets.

To date, three of the enzymes in the MK biosynthetic pathway have been targeted for development of chemical inhibitors. Kurosu and coworkers synthesized inhibitors of Mtb MenA, the DHNA prenyltransferase. They showed antibacterial activity, validating this pathway as a target for tuberculosis drug discovery (61). A library of one hundred molecules was synthesized to afford hydrophobically substituted benzophenones. The two best hits, Allylaminomethanone-A and Phenethylaminomethanone-A (Figure 1-6), had minimum inhibitory concentrations (MIC) of 1.5 and 12.5 $\mu\text{g/mL}$ respectively against Mtb (H37Rv strain). The MIC values were in the range of 2-8 $\mu\text{g/mL}$ against other gram positive bacteria, including strains known to be drug resistant. Notably, only growth of gram positive bacteria was inhibited by these compounds. Gram negative bacteria showed no growth inhibition at 128 $\mu\text{g/mL}$ suggesting that menaquinone biosynthesis is the target of these compounds in gram positive bacteria.

MenE, an acyl-CoA synthetase (ligase), essential in Mtb (62), has been targeted with covalent inhibitors using compounds related to 5'-O-(N-acylsulfamoyl)adenosines (acyl-AMS) by Lu and coworkers (63). Previously, these acyl-AMS compounds have been used to inhibit other adenylate-forming enzymes. The inhibitors (Figure 1-7) were designed to mimic the cognate, tightly bound acyl-AMP intermediates and were inspired by sulfamoyl-adenosine natural products nucleocidin and ascamycin. The mechanism of inhibition is the creation of a covalent bond between the CoA-thiol and the acyl-AMS in the second part of the reaction catalyzed by acyl-CoA synthetase MenE. The inhibitors were tested on purified enzyme in a coupled assay using MenE and MenB monitoring the formation of DHNA-CoA at 392 nm. The initial studies of this class of inhibitors used the methyl ester analogues of the compounds shown in Figure 1-7 (63). These compounds showed modest inhibitory activity, but the carboxylate analogues had sub-micromolar inhibition of the Mtb MenE (64). The most potent inhibitor was the acyl sulfamate (Figure 1-7A), which was shown to be a competitive inhibitor with respect to

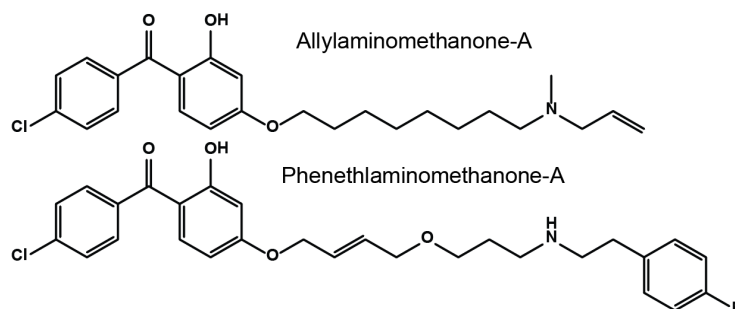


Figure 1-6: Inhibitors of MenA, the DHNA prenyltransferase. Allylaminomethanone -A and phenethylaminomethanone-A inhibited growth of Mtb (strain H37Rv) as well as other gram positive bacteria (62).

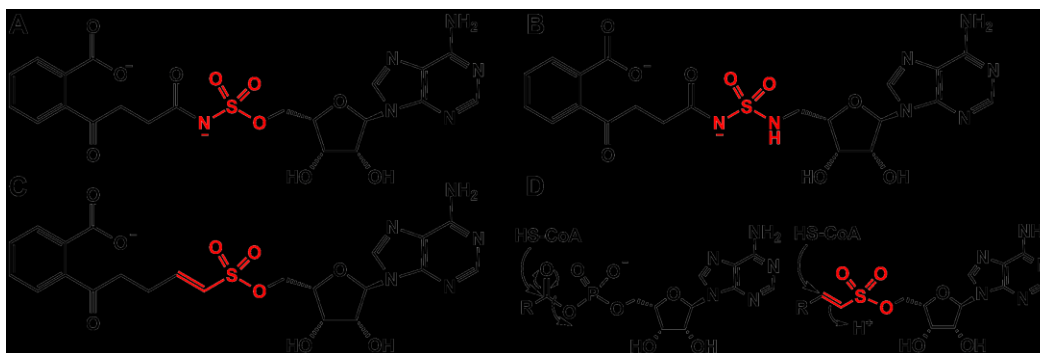


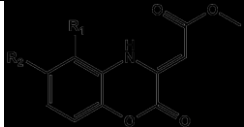
Figure 1-7: Covalent inhibitors of MenE, an acyl-CoA synthetase using a sulfamate and sulfamine functionalities (in red) (64).

ATP and a non-competitive inhibitor of OSB. The carboxylate series also showed good inhibition *in vitro* against MenE from *Staphylococcus aureus* and *E. coli*.

Li and coworkers have targeted MenB, a 1,4-dihydroxy-2-naphthoyl-CoA (DHNA-CoA) synthase (44). MenB catalyzes an intramolecular Claisen condensation from OSB-CoA to DHNA-CoA. They screened a library of over 100,000 compounds and found over 400 hits with at least 30% enzyme inhibition (44). From the hits, a series of compounds were identified that shared the 1,4-benzoxazine scaffold (Table 1-2). Using either a benzoxazine or quinoxaline backbone, thirteen compounds were synthesized and screened for enzyme inhibition and *in vitro* growth inhibition of Mtb (H37Rv). Overall, the IC_{50} values loosely correlated with the MIC values which ranged from sub-micromolar to no inhibition of growth suggesting that other factors, such as permeability or detoxification strategies, may also modulate activity. Compounds 1, 4, and 5 were found to be non-competitive inhibitors of MenB with the inhibitor binding site distinct from the active site.

Overall, inhibitors of MK biosynthesis show promising inhibition of mycobacterial growth. These inhibitors may also be useful for treating other gram-positive bacterial infections.

Table 1-2: Activity of benzoxazine or quinoxaline inhibitors. Modified from ref # (44). MenB concentration was 150 nM. (a) Compound concentration giving 50% inhibition of enzyme activity. (b) The lowest concentration of compound that inhibited visible growth of Mtb H37Rv in all replicates.



Compound	R ₁	R ₂	IC ₅₀ (μM) ^a	K _i (μM)	K _i ' (μM)	MIC (μg/ml) ^b
1	H	H	10.0 ± 1.0	9.1 ± 1.2	67.0 ± 7.9	0.64
2	Me	H	24.1 ± 1.8			25
3	H	Me	23.1 ± 1.0			>100
4	H	F	27.0 ± 3.0	11.5 ± 1.5	10.1 ± 0.9	0.63
5	H	Cl	46.3 ± 3.5	22.5 ± 1.1	18.5 ± 2.6	5
6	H	NO ₂	28.2 ± 4.4			50
7	H	EtSO ₂	17.9 ± 3.0			>100

Summary

In summary, MK plays an essential role in the electron transport chain of most gram positive bacteria including Mtb. In Mtb, MK biosynthetic enzymes are essential for bacterial viability when grown in hypoxic non-replicating conditions as well as for growth in mouse models of infection (2,61-63,65). Mycobacteria use MK for maintenance of the proton gradient and the membrane potential, which is required for ATP production and bacterial viability. Recently, MK has been identified as an indicator of Mtb respiration rate when grown under conditions that mimic infection acting as a signal for induction of the dormancy regulon (15). Additionally, MK is a precursor for S881, a novel sulfated MK that modulates the virulence of Mtb in the mouse infection model.

There are eight enzymes that biosynthesize MK from chorismate. Many of these enzymes are essential and may be good drug targets since this pathway is unique to bacteria. A variety of promising small molecules have been described that target these enzymes with MIC values in the submicromolar range. Further investigation into their efficacy in animals will aid in the development of much needed Mtb therapies. Compounds that are active against drug resistant strains of Mtb, as well as other gram positive and gram negative bacteria, are worth investigating further.

Thesis overview

Mycobacterium tuberculosis (Mtb) is a highly evolved human pathogen that produces a wide array of novel lipid-based virulence factors. We have identified a novel sulfated lipid, named S881. A mutant of Mtb that lacks the sulfotransferase *stf3* necessary for S881 biosynthesis exhibits a hypervirulent phenotype in the mouse model of infection. In Chapter 2, we elucidate the structure of S881 by high resolution mass spectrometry. Using structural analogues we confirm the structure of S881 as a terminally sulfated menaquinone-9 (H₂). Menaquinone-9 (H₂) is the most abundant lipoquinone produced by Mtb and is involved in the electron transport chain. In Chapter 3, we identify a cytochrome P450 *cyp128* that in addition to *stf3* converts menaquinone to S881 in two steps. We create a mutant with *cyp128* disrupted by a resistance cassette as well as a mutant produces a 10 fold excess of S881 over WT. Using these mutants we study the temporal regulation of S881 and its location in the Mtb cell wall. Using the S881 mutants, we explore the role S881 plays in the Mtb lifecycle in Chapter 4. We use different models of infection as well as a chemical genetic approach to study the S881 mutants. Finally, Chapter 5 explores how *cyp128* is regulated and we identify its native promoter.

References

1. Meganathan, R. (2001) *Vitam Horm* 61, 173-218
2. Dhiman, R. K., Mahapatra, S., Slayden, R. A., Boyne, M. E., Lenaerts, A., Hinshaw, J. C., Angala, S. K., Chatterjee, D., Biswas, K., Narayanasamy, P., Kurosu, M., and Crick, D. C. (2009) *Mol Microbiol* 72, 85-97
3. World Health Organization. (www.who.int/tb/en/)
4. Bentley, R., and Meganathan, R. (1982) *Microbiol Rev* 46, 241-280

5. Kurosu, M., and Begari, E. (2010) *Molecules* 15, 1531-1553
6. McKee, R. W. (1970) From vitamin K to metabolism of cancer cells. in *Perspectives in biological chemistry* (Olson, R. E. ed.), M. Dekker Inc, New York. pp 57-81
7. Lester, R. L., and Crane, F. L. (1959) *J Biol Chem* 234, 2169-2175
8. Almquist, H. J., and Klose, A. A. (1939) *Journal of the American Chemical Society* 61, 2557-2559
9. Binkley, S. B., McKee, R. W., Thayer, S. A., and Doisy, E. A. (1940) *Journal of Biological Chemistry* 133, 721-729
10. Collins, M. D., and Jones, D. (1981) *Microbiol Rev* 45, 316-354
11. Simon, J., van Spanning, R. J., and Richardson, D. J. (2008) *Biochim Biophys Acta* 1777, 1480-1490
12. Azerad, R., and Cyrot-Pelletier, M. O. (1973) *Biochimie* 55, 591-603
13. Collins, M. D., Pirouz, T., Goodfellow, M., and Minnikin, D. E. (1977) *J Gen Microbiol* 100, 221-230
14. Campbell, I. M., and Bentley, R. (1968) *Biochemistry* 7, 3323-3327
15. Honaker, R. W., Dhiman, R. K., Narayanasamy, P., Crick, D. C., and Voskuil, M. I. (2010) *J Bacteriol* 192, 6447-6455
16. Noll, H. (1958) *J Biol Chem* 232, 919-929
17. Mougous, J. D., Senaratne, R. H., Petzold, C. J., Jain, M., Lee, D. H., Schelle, M. W., Leavell, M. D., Cox, J. S., Leary, J. A., Riley, L. W., and Bertozzi, C. R. (2006) *Proc Natl Acad Sci U S A* 103, 4258-4263
18. Sogi, K. M., Holsclaw, C. M., Fragiadakis, G. K., Leary, J. A., and Bertozzi, C. R. (2012) *Manuscript in preparation*
19. Mougous, J. D., Green, R. E., Williams, S. J., Brenner, S. E., and Bertozzi, C. R. (2002) *Chem Biol* 9, 767-776
20. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaiia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) *Nature* 393, 537-544
21. Wallace, B. J., and Young, I. G. (1977) *Biochim Biophys Acta* 461, 84-100
22. Pandya, K. P., and King, H. K. (1966) *Arch Biochem Biophys* 114, 154-157

23. Newton, N. A., Cox, G. B., and Gibson, F. (1971) *Biochim Biophys Acta* 244, 155-166
24. Maklashina, E., Hellwig, P., Rothery, R. A., Kotlyar, V., Sher, Y., Weiner, J. H., and Cecchini, G. (2006) *J Biol Chem* 281, 26655-26664
25. Boshoff, H. I., and Barry, C. E., 3rd. (2005) *Nat Rev Microbiol* 3, 70-80
26. Kana, B. D., Weinstein, E. A., Avarbock, D., Dawes, S. S., Rubin, H., and Mizrahi, V. (2001) *J Bacteriol* 183, 7076-7086
27. Wayne, L. G., and Hayes, L. G. (1998) *Tuber Lung Dis* 79, 127-132
28. Matsoso, L. G., Kana, B. D., Crellin, P. K., Lea-Smith, D. J., Pelosi, A., Powell, D., Dawes, S. S., Rubin, H., Coppel, R. L., and Mizrahi, V. (2005) *J Bacteriol* 187, 6300-6308
29. Brown, A. C., Eberl, M., Crick, D. C., Jomaa, H., and Parish, T. (2010) *J Bacteriol* 192, 2424-2433
30. Blasco, F., Guigliarelli, B., Magalon, A., Asso, M., Giordano, G., and Rothery, R. A. (2001) *Cell Mol Life Sci* 58, 179-193
31. Shi, L., Sohaskey, C. D., Kana, B. D., Dawes, S., North, R. J., Mizrahi, V., and Gennaro, M. L. (2005) *Proc Natl Acad Sci U S A* 102, 15629-15634
32. Malm, S., Tiffert, Y., Micklinghoff, J., Schultze, S., Joost, I., Weber, I., Horst, S., Ackermann, B., Schmidt, M., Wohlleben, W., Ehlers, S., Geffers, R., Reuther, J., and Bange, F. C. (2009) *Microbiology* 155, 1332-1339
33. Shiloh, M. U., Manzanillo, P., and Cox, J. S. (2008) *Cell Host Microbe* 3, 323-330
34. Voskuil, M. I., Schnappinger, D., Visconti, K. C., Harrell, M. I., Dolganov, G. M., Sherman, D. R., and Schoolnik, G. K. (2003) *J Exp Med* 198, 705-713
35. Kumar, A., Deshane, J. S., Crossman, D. K., Bolisetty, S., Yan, B. S., Kramnik, I., Agarwal, A., and Steyn, A. J. (2008) *J Biol Chem* 283, 18032-18039
36. Kumar, A., Toledo, J. C., Patel, R. P., Lancaster, J. R., Jr., and Steyn, A. J. (2007) *Proc Natl Acad Sci U S A* 104, 11568-11573
37. Honaker, R. W., Leistikow, R. L., Bartek, I. L., and Voskuil, M. I. (2009) *Infect Immun* 77, 3258-3263
38. Roberts, D. M., Liao, R. P., Wisedchaisri, G., Hol, W. G., and Sherman, D. R. (2004) *J Biol Chem* 279, 23082-23087
39. Bekker, M., Alexeeva, S., Laan, W., Sawers, G., Teixeira de Mattos, J., and Hellingwerf, K. (2010) *J Bacteriol* 192, 746-754
40. Georgellis, D., Kwon, O., and Lin, E. C. (2001) *Science* 292, 2314-2316

41. Nakamoto, H., and Bardwell, J. C. (2004) *Biochim Biophys Acta* 1694, 111-119
42. Dutton, R. J., Boyd, D., Berkmen, M., and Beckwith, J. (2008) *Proc Natl Acad Sci U S A* 105, 11933-11938
43. Dutton, R. J., Wayman, A., Wei, J. R., Rubin, E. J., Beckwith, J., and Boyd, D. (2010) *Proc Natl Acad Sci U S A* 107, 297-301
44. Li, W., Schulman, S., Dutton, R. J., Boyd, D., Beckwith, J., and Rapoport, T. A. (2010) *Nature* 463, 507-512
45. Daves, G. D., Jr., Muraca, R. F., Whittick, J. S., Friis, P., and Folkers, K. (1967) *Biochemistry* 6, 2861-2866
46. Downey, R. J. (1964) *J Bacteriol* 88, 904-911
47. Yamada, Y., Inouye, G., Tahara, Y., and Kondo, K. (1976) *Biochim Biophys Acta* 486, 195-203
48. Dunphy, P. J., Phillips, P. G., and Brodie, A. F. (1971) *J Lipid Res* 12, 442-449
49. Taber, H. W., Sugarman, B. J., and Halfenger, G. M. (1981) *J Gen Microbiol* 123, 143-149
50. Truglio, J. J., Theis, K., Feng, Y., Gajda, R., Machutta, C., Tonge, P. J., and Kisker, C. (2003) *J Biol Chem* 278, 42352-42360
51. Emmons, G. T., Campbell, I. M., and Bentley, R. (1985) *Biochem Biophys Res Commun* 131, 956-960
52. Popp, J. L. (1989) *J Bacteriol* 171, 4349-4354
53. Sharma, V., Meganathan, R., and Hudspeth, M. E. (1993) *J Bacteriol* 175, 4917-4921
54. Thompson, T. B., Garrett, J. B., Taylor, E. A., Meganathan, R., Gerlt, J. A., and Rayment, I. (2000) *Biochemistry* 39, 10662-10676
55. Palmer, D. R., Garrett, J. B., Sharma, V., Meganathan, R., Babbitt, P. C., and Gerlt, J. A. (1999) *Biochemistry* 38, 4252-4258
56. Sassetti, C. M., Boyd, D. H., and Rubin, E. J. (2003) *Mol Microbiol* 48, 77-84
57. Sassetti, C. M., Boyd, D. H., and Rubin, E. J. (2001) *Proc Natl Acad Sci U S A* 98, 12712-12717
58. Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., Boland, F., Brignell, S. C., Bron, S., Bunai, K., Chapuis, J., Christiansen, L. C., Danchin, A., Debarbouille, M., Dervyn, E., Deurling, E., Devine, K., Devine, S. K., Dreesen, O., Errington, J., Fillinger, S., Foster, S. J., Fujita, Y., Galizzi, A., Gardan, R., Eschevins, C., Fukushima, T., Haga, K., Harwood, C. R., Hecker, M., Hosoya, D., Hullo, M. F., Kakeshita, H., Karamata, D.,

- Kasahara, Y., Kawamura, F., Koga, K., Koski, P., Kuwana, R., Imamura, D., Ishimaru, M., Ishikawa, S., Ishio, I., Le Coq, D., Masson, A., Mauel, C., Meima, R., Mellado, R. P., Moir, A., Moriya, S., Nagakawa, E., Nanamiya, H., Nakai, S., Nygaard, P., Ogura, M., Ohanan, T., O'Reilly, M., O'Rourke, M., Pragai, Z., Pooley, H. M., Rapoport, G., Rawlins, J. P., Rivas, L. A., Rivolta, C., Sadaie, A., Sadaie, Y., Sarvas, M., Sato, T., Saxild, H. H., Scanlan, E., Schumann, W., Seegers, J. F., Sekiguchi, J., Sekowska, A., Seror, S. J., Simon, M., Stragier, P., Studer, R., Takamatsu, H., Tanaka, T., Takeuchi, M., Thomaidis, H. B., Vagner, V., van Dijl, J. M., Watabe, K., Wipat, A., Yamamoto, H., Yamamoto, M., Yamamoto, Y., Yamane, K., Yata, K., Yoshida, K., Yoshikawa, H., Zuber, U., and Ogasawara, N. (2003) *Proc Natl Acad Sci U S A* 100, 4678-4683
59. Rao, S. P., Alonso, S., Rand, L., Dick, T., and Pethe, K. (2008) *Proc Natl Acad Sci U S A* 105, 11945-11950
60. Gengenbacher, M., Rao, S. P., Pethe, K., and Dick, T. (2010) *Microbiology* 156, 81-87
61. Kurosu, M., Narayanasamy, P., Biswas, K., Dhiman, R., and Crick, D. C. (2007) *J Med Chem* 50, 3973-3975
62. Sasseti, C. M., and Rubin, E. J. (2003) *Proc Natl Acad Sci U S A* 100, 12989-12994
63. Lu, X., Zhang, H., Tonge, P. J., and Tan, D. S. (2008) *Bioorg Med Chem Lett* 18, 5963-5966
64. Lu, X., Zhou, R., Sharma, I., Li, X., Kumar, G., Swaminathan, S., Tonge, P. J., and Tan, D. S. (2012) *Chembiochem* 13, 129-136
65. Kurosu, M., and Crick, D. C. (2009) *Med Chem* 5, 197-207

Chapter 2: Structural characterization of a novel sulfolipid produced by *Mycobacterium tuberculosis*¹

Introduction

The disease Tuberculosis (TB) affects approximately one third of the world's population and kills approximately two million people a year (1). *Mycobacterium tuberculosis* (Mtb), the causative agent of TB, requires strict regulation of its metabolism throughout its lifecycle as well as a careful modulation of host immune response for survival. The cell wall of Mtb is rich in complex lipids many of which are involved in maintaining a delicate balance with the host immune response. Among these lipids in the Mtb cell wall are two classes of sulfolipids. In both symbiotic and pathogenic bacteria, sulfated metabolites have been shown to serve as signaling molecules between bacteria and their hosts (2-4). In addition, the sulfate modification is key to a number of mammalian extracellular signaling events including chemokine binding, cellular degradation, and modulation of signaling pathways (5-7). Mycobacteria synthesize a number of unique sulfated metabolites most of which have no defined function. The best-characterized of these molecules is the Mtb-specific sulfolipid-1 (SL-1), a tetra-acylated trehalose-based lipid localized to the outer envelope of the bacterium (8-12).

Another sulfated metabolite unique to Mtb was discovered in a genetic and mass spectrometric (MS) screen for sulfated molecules from Mtb total lipid extracts (13). The lipid was named S881 based on its measured mass of $m/z = 881.57$ in the negative ion mode. Structural characterization was hindered by the relatively low abundance of the molecule as reflected in the negative-ion mode mass spectra of Mtb total lipid extracts.

The mass spectrometry genetic screen identified sulfotransferase 3 (*stf3*) to be responsible for the production of S881. A Mtb mutant with a disruption in the *stf3* gene (Δ *stf3*) did not produce S881, but production was restored with complementation of the gene (13). The Δ *stf3* mutant was then examined for virulence in a mouse model of TB infection. Intriguingly, the Δ *stf3* mutant shows a hypervirulent phenotype in the mouse model of infection compared to wild-type Mtb. Considering this highly unusual phenotype in mice and the localization of S881 in the Mtb outer wall, it suggests that S881 may be involved in regulating Mtb pathogenicity (13).

Initial studies indicate that S881 contains a single sulfate group determined by labeling Mtb cultures with a heavy sulfate isotope (³⁴S). In this chapter, we utilize high-resolution, high-mass-accuracy, and tandem mass spectrometry (MSⁿ) to characterize the structure of S881. We further analyze structural analogs to confirm the proposed S881 structure. We show that S881 is a sulfated derivative of the most abundant quinone in mycobacteria, dihydromenaquinone-9 (MK-9(H₂)) (14).

¹ All mass spectrometry was done by Cynthia M. Holsclaw in Julie A. Leary's laboratory (UC Davis). Synthesis was done by Kimberly M. Sogi (UC Berkeley). Sarah Gilmore (UC Berkeley), Michael W. Schelle (UC Berkeley), Michael L. Leavell (UC Davis) contributed to this work.

Results

S881 was enriched from Mtb total lipid extracts

Previously, we reported the discovery of an unknown sulfated metabolite at $m/z = 881.57$, which we named S881 based on its exact mass (15). However, the low abundance of S881 in the negative-ion mode mass spectrum of a crude Mtb lipid extract (Figure 2-1A), coupled with its proximity to an isobar (Figure 2-1A, inset), precluded its structural characterization (13). Therefore, to enrich S881, we fractionated Mtb total lipid extracts on weak anion exchange resin. The resulting fractions were analyzed for the presence of S881 lacking the isobar.

The purest fraction from the anion exchange resin afforded an increase in the intensity of S881 approximately 200-fold compared to its intensity in crude lipid extracts (Figure 2-1B) and eliminated the contaminating isobar (Figure 2-1B, inset). Comparison of exact mass measurements of S881 from the enriched samples to total lipid extracts confirmed that they resulted from the same molecular species S881.

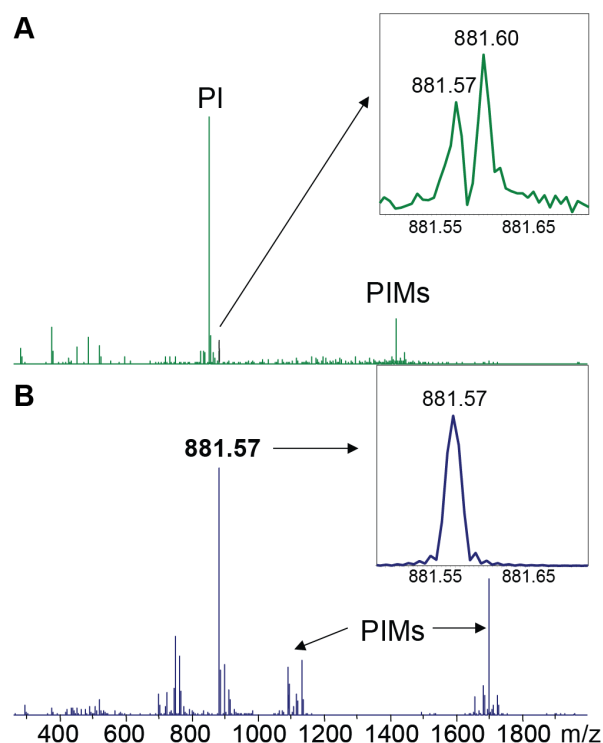


Figure 2-1: Enrichment of S881 after anion exchange separation. (A) The negative ion mode FT-ICR mass spectrum of a crude lipid extract from Mtb, containing mostly phosphatidyl inositol (PI) and phosphatidyl-inositol mannosides (PIM). The inset is a zoom indicating that S881 is a minor component of this extract, and is isobaric with another metabolite. (B) The negative ion mode FT-ICR mass spectrum of the purest fraction of S881 after separation (inset). Anion exchange separation has enriched S881 such that it is the characteristic peak and the isobar has been removed.

FT-ICR enabled exact mass and elemental composition analysis

We measured the elemental composition of S881 using the enriched fractions. The exact mass of S881 was 881.5755 Da as measured on the FT-ICR MS. The mass spectrum was calibrated internally with 12 calibration peaks, and all internal calibrants were measured to <0.7 ppm. The molecular formula generation algorithm from the DataAnalysis software (Bruker Daltonics) was used to generate putative elemental compositions for the measured mass of S881. This software estimates the number of carbon atoms in the ion based on its “M+1” isotope. Only elemental compositions with one sulfur atom were considered in our analysis, as previous studies using a heavy isotope of sulfur (^{34}S) indicate that S881 contains only one sulfate moiety (13). Elemental compositions with an odd number of nitrogen atoms were also not considered, in accordance with the “Nitrogen Rule.” Finally, compositions with sulfur and phosphate atoms without enough oxygen atoms to support both sulfate and phosphate moieties were not considered as the resulting molecule would not be organic.

Only one such elemental composition that satisfied the above requirements, $\text{C}_{56}\text{H}_{81}\text{O}_6\text{S}_1^-$, was found to be within 0.7 ppm of the measured mass of S881. This highlights the value of the high-mass-accuracy of the FT-ICR MS, as we were able to use this technique to unambiguously determine the elemental composition of S881. This elemental composition contains eighteen degrees of unsaturation, including the two on the sulfate, indicating that S881 is a highly unsaturated molecule.

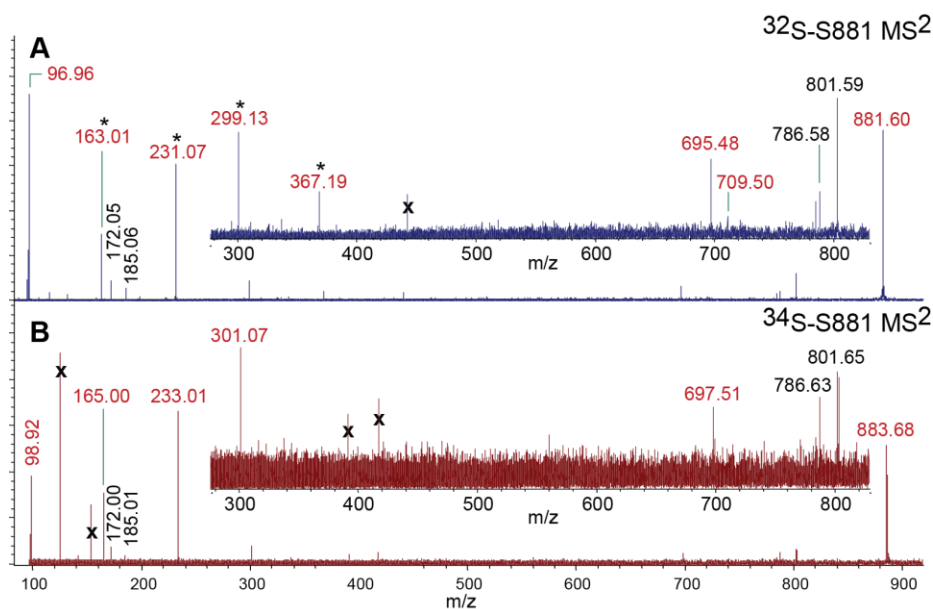


Figure 2-2: The tandem FT-ICR mass spectrum of S881 indicates that it is a polyisoprenoid-derived molecule. Ions marked with an “x” were not ejected from the ICR cell prior to activation of the parent ion, therefore these ions are not fragments of S881. (A) The MS² FT-ICR mass spectrum of S881. Fragments in red are sulfate-containing fragments as determined by exact mass measurements (Table 2-1). Fragments separated by 68 mass units, denoted by an asterisk, correspond to a difference of C_5H_8 . (B) The MS² FT-ICR mass spectrum of a ^{34}S -labeled S881. Fragments in red are sulfate-containing fragments, and are shifted by two mass units compared to the fragment ions of the ^{32}S -S881.

Tandem MS of S881 suggested that it is a polyisoprenoid

For detailed structural characterization, we analyzed S881 via FT-ICR tandem mass spectrometry (MS^n) (Figure 2-2A). The resulting MS^2 spectrum of S881 revealed ions at $m/z = 801.59$, corresponding to a loss of SO_3 , and $m/z = 96.96$, corresponding to HSO_4^- (Figure 2-2A). These fragments provide direct evidence of the sulfate modification of S881. Using the accurate mass capabilities of the FT-ICR MS, we were able to determine the elemental compositions of the fragment ions (Table 2-1). Many of the sulfate-containing fragments differed in mass by 68 mass units, corresponding to a difference of five carbon and eight hydrogen atoms (Table 2-1, Figure 2-2A).

As an additional check on our mass assignments, we also performed MS^n on a ^{34}S -labeled sample of S881 (13) to confirm the elemental compositions of the sulfate-containing fragments, as each sulfate-containing fragment ion can be distinguished by a 2-Da shift in the MS^2 spectrum of the ^{34}S -labeled S881 (Figure 2-2B). This analysis indicated that the fragment ions at $m/z = 97, 163, 231, 299,$ and 695 are indeed sulfated, while the fragment ions at $m/z = 172, 185, 786,$ and 801 do not contain sulfate (Figure 2-2B). Furthermore, this analysis confirms that the fragment ion at $m/z = 96.95$ is not $H_2PO_4^-$, a fragment commonly seen in MS^n spectra of phosphorylated molecules. In combination with our exact mass measurements, these data unambiguously confirm that S881 contains a sulfate and not a phosphate.

In order to obtain more fragment coverage, we performed MS^n of S881 on an LTQ ion trap MS. The LTQ allows more stages of MS^n , and therefore more structural information can be gleaned using this mass analyzer (Figure 2-3). We performed MS^3 on the fragment at $m/z = 801$, yielding fragment ions without the sulfate residue, and compared this MS^3 spectrum to the MS^2

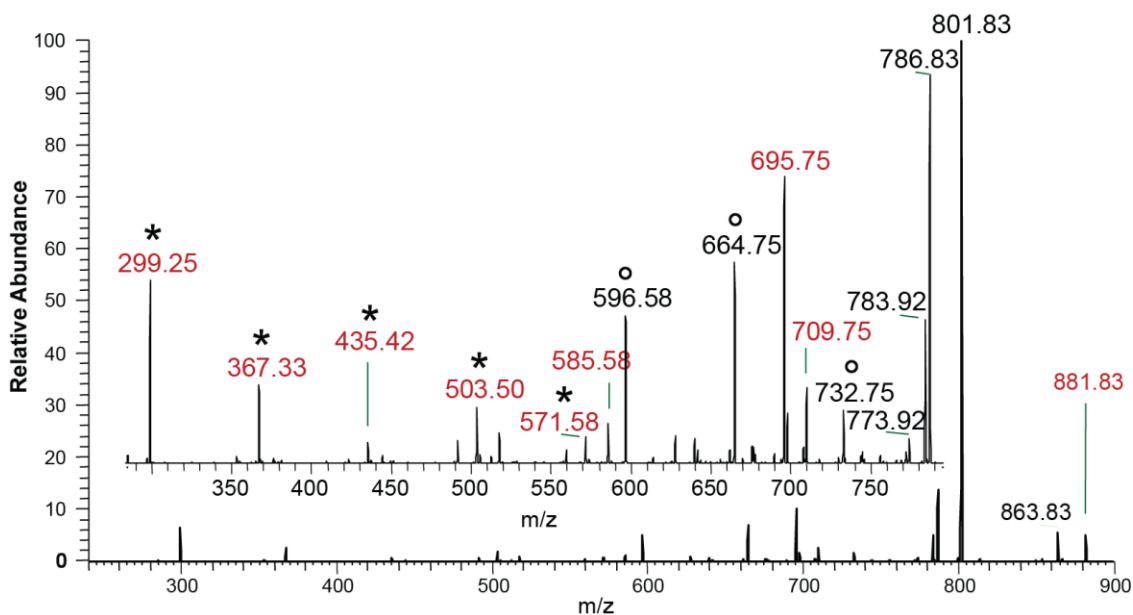


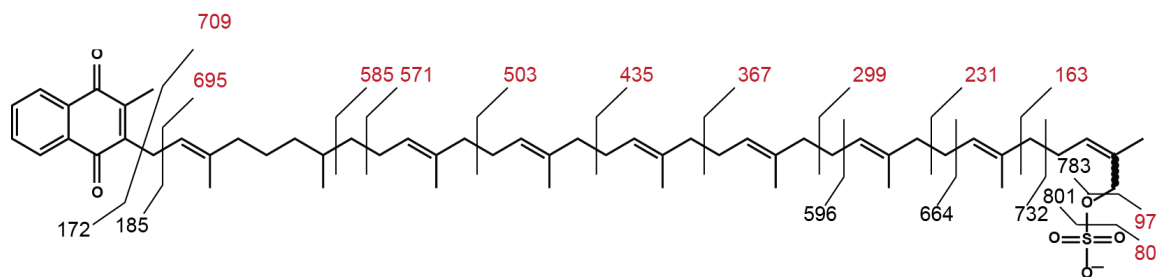
Figure 2-3: Fragmentation of S881 on the LTQ ion-trap. The MS^2 spectrum of S881 indicates that there are two series of fragments separated by 68 mass units, one containing sulfate (denoted by *) and the other does not contain sulfate (denoted by °). Sulfate-containing fragment ions are indicated in red.

spectrum of S881 to identify ions that contained the sulfate modification. This comparison yielded a series of sulfate-containing ions differing by 68 mass units at $m/z = 299, 367, 435, 503,$ and 571 (Figure 2-3). Another series of fragment ions also differing by 68 mass units, but lacking the sulfate residue, were measured at $m/z = 596, 664,$ and 732 (Figure 2-3). The spectrum of the fragment ion at $m/z = 299$ yielded MS^3 fragment ions at $m/z = 231, 163,$ and 97 , (data not shown) consistent with the FT-ICR MS^2 spectrum. MS^3 of $m/z = 367, 435, 503,$ and 571 yielded MS^3 fragment ions at successive losses of 68 mass units. In combination with the assigned elemental compositions of the fragment ions obtained on the FT-ICR MS (Table 2-1), these data indicate that the sulfate-containing fragment ions of S881 derive from a sulfated hydrocarbons with a repeating unit of C_5H_8 .

A successive loss of 68 mass units is consistent with that of a polyisoprenoid chain, which typically fragments via random allylic cleavage between the individual isoprene units in the chain (16). Mycobacteria contain many polyisoprenoid-derived molecules including lipid II, the cell wall precursor (17), polyprenyl phosphate, the saccharide carrier (18), and dihyromenaquinone-9 (MK-9(H_2)), the sole lipoquinone in the mycobacterial electron transport chain (14). A sulfated MK-9(H_2) is the only polyisoprenoid that matched the exact mass and predicted elemental composition. MK-9(H_2) consists of a 2-methyl-1,4-naphthoquinone moiety alpha-linked to a chain of nine polyisoprene units with the second isoprene unit from the

Measured S881 Fragment Mass	Assigned Elemental Composition	Theoretical Exact Mass	ppm
79.9571	SO ₃	79.9568	-3.75
96.9597	HSO ₄ ⁻	96.9601	4.13
163.0060	C ₅ H ₇ O ₄ S ₁ ⁻	163.0071	6.75
172.0521	C ₁₁ H ₈ O ₂ ⁻	172.0530	5.23
185.0608	C ₁₂ H ₉ O ₂ ⁻	185.0597	-5.94
231.0689	C ₁₀ H ₁₅ O ₄ S ₁ ⁻	231.0697	3.46
299.1327	C ₁₅ H ₂₃ O ₄ S ₁ ⁻	299.1321	-2.01
367.1952	C ₂₀ H ₃₁ O ₄ S ₁ ⁻	367.1937	-4.09
695.4969	C ₄₄ H ₇₁ O ₄ S ₁ ⁻	695.5067	14.09
783.5649	C ₅₆ H ₇₉ O ₂ ⁻	783.6074	54.24
786.5704	C ₅₅ H ₇₈ O ₃ ⁻	786.5945	30.64
801.5888	C ₅₆ H ₈₁ O ₃ ⁻	801.6180	36.43
881.6075	C ₅₆ H ₈₁ O ₆ S ₁ ⁻	881.5759	-35.84

Table 2-1: Exact mass measurements of fragment ions generated via FT-ICR tandem MS.



Scheme 2-1: S881 is a sulfated menaquinone with observed fragmentation. Red values denote sulfate-containing fragments.

naphthoquinone being saturated. MK-9(H₂) is the most abundant quinone in mycobacteria (14), however mycobacteria also produce smaller amounts of MK-9 and MK-8 (19). There is no evidence that mycobacteria produce ubiquinone (14,20) indicating that menaquinone is the sole quinol electron carrier in the mycobacterial respiratory chain.

Confirmation of S881 structure via fragmentation of structural analogs

Given S881's high degree of unsaturation and its MSⁿ fragmentation pattern suggesting a polyisoprenoid structure, we sought to determine if S881 is a sulfated derivative of MK-9(H₂) (Scheme 2-2). Mass spectrometry has been a valuable tool in the structural determination of isoprenoid quinones, however the vast majority of these studies have been performed in the positive ion mode (14). Since the electrospray ionization of sulfated molecules is inefficient in the positive ion mode, FT-ICR MSⁿ was performed on the 2-methyl-1,4-naphthoquinone (more

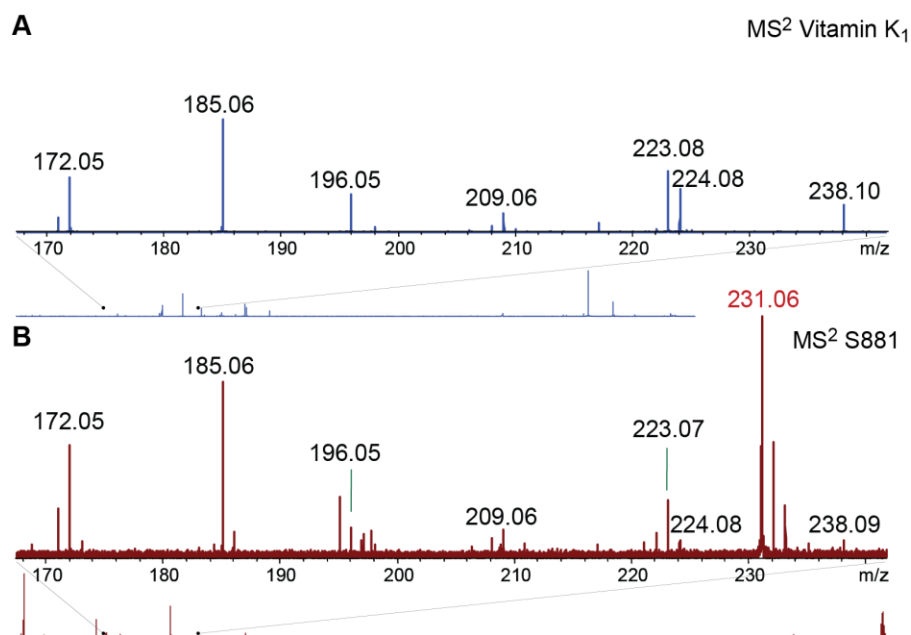
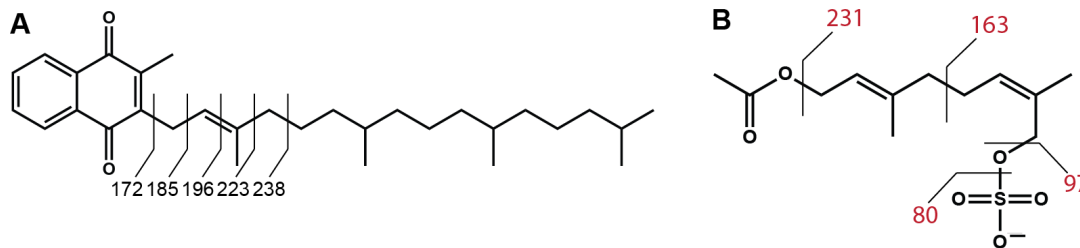


Figure 2-4: The m/z=170-240 region of the FT-ICR MS² spectrum of (A) vitamin K₁ and (B) S881.



Scheme 2-2: (A) The structure of Vitamin K₁ (phylloquinone) with observed negative-ion-mode MS² fragmentation. (B) The structure of geranyl sulfate with observed fragmentation.

commonly known as phylloquinone) using negative ion mode detection for comparison with the MS² spectrum of S881 (Figure 2-4). First we compared the spectrum of phylloquinone to S881 to study the non-sulfated fragments from S881. Phylloquinone revealed ions at $m/z = 185$ and 223 (Figure 2-4A), which are consistent with the characteristic fragment ions at $m/z = 187$ and 225 observed in the positive mode spectra of polyisoprenoid menaquinones (14). The phylloquinone fragment ions at $m/z = 172$, 185 , 196 , 223 , and 238 in the MS² spectrum obtained in the negative ion mode correspond to fragmentation of the isoprene unit alpha-linked to the naphthoquinone moiety (Figure 2-4A, Scheme 2-3A). These ions are also present in the MS² spectrum of S881 (Figure 2-4B), clearly supporting a structure containing a naphthoquinone moiety.

The MS² spectrum of S881 suggests that the sulfate residue is on the opposite end of the polyisoprenoid chain from the naphthoquinone moiety (Scheme 2-2). To confirm if this is the

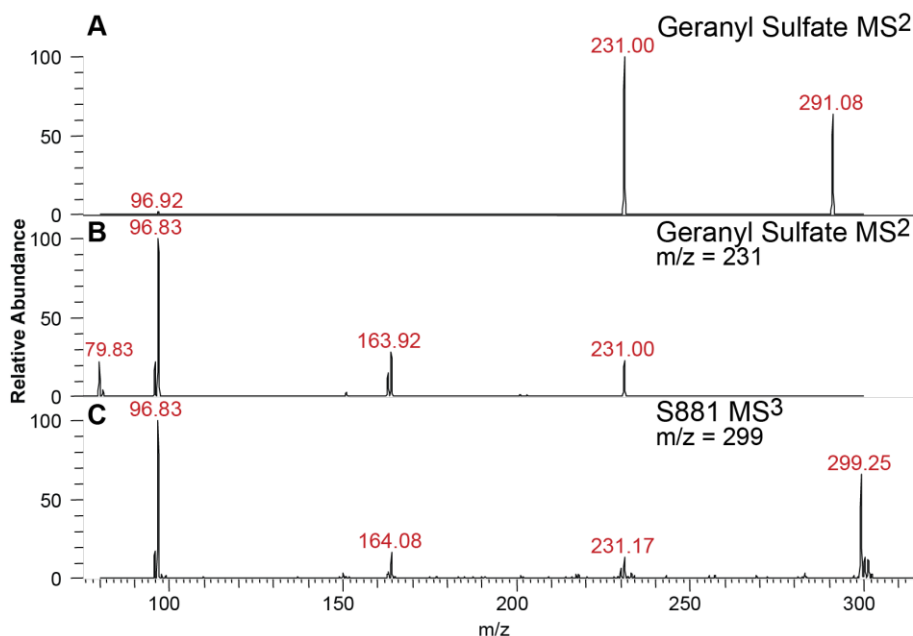


Figure 2-5: Comparison of the fragmentation patterns of geranyl sulfate and S881. (A) The MS² spectrum of geranyl sulfate. (B) The MS³ spectrum of $m/z = 231$ from geranyl sulfate. (C) The MS³ spectrum of the $m/z = 299$ from S881.

case, we synthesized a geranyl sulfate derivative to determine if its fragmentation pattern was consistent with that of S881 (Scheme 2-3B). Confirmation of the sulfate position on the geranyl sulfate was determined via 2D $^1\text{H-NMR}$. The MS^2 spectrum of geranyl sulfate yields the expected fragment at $m/z = 97$ corresponding to HSO_4^- , as well as a fragment at $m/z = 231$ also present in the MS^2 spectrum of S881 (Figure 2-5A). Conspicuously absent from this spectrum was the expected fragment at $m/z = 163$ observed in the MS^2 spectrum of S881 (Figure 2-2A). Geranyl sulfate contains an ester bond (Scheme 2-3B), and fragmentation in the MS^2 experiment is preferred at this bond rather than at the carbon-carbon bond between the two isoprene units. However, MS^3 of the $m/z = 231$ fragment ion from geranyl sulfate yielded a fragment at $m/z = 164$ (Figure 2-5B). This fragment is also present in the MS^3 spectrum of the $m/z = 299$ fragment ion of S881 (Figure 2-5C). This data shows that S881 is a terminally sulfated derivative of MK-9(H_2) (Scheme 2-2).

Discussion

In this study, we provide evidence that the structure of the Mtb metabolite S881 is a sulfated MK-9(H_2). To our knowledge, this metabolite is unique to Mtb and represents a novel modification to the common quinone scaffold. In addition, this structure is distinct from previously characterized mycobacterial sulfated metabolites, including the carbohydrate-based SL-1 and its biosynthetic precursors (8,15,21,22), and the sulfated glycopeptidolipids from *M. avium* and *M. fortuitum* (23,24). Sulfur-containing menaquinones have been reported in the thermophilic bacteria *Hydrogenobacter thermophilus* (25) and *Caldariella acidophila* (14), however these quinones contain reduced sulfur in the naphthoquinone moiety, as opposed to a sulfate modification on the polyisoprenoid chain.

Elucidation of the chemical structure of S881 will allow us to further probe its biosynthesis. The menaquinone biosynthesis in Mtb is homologous to the well characterized pathways in *M. phlei* and *E. coli* (20,26). We hypothesize that only two steps would be required for sulfation of MK: oxidation of the terminal isoprene and sulfation of the resulting alcohol. Previously, we identified a sulfotransferase *stf3* as necessary for the production of S881 in addition to the universal sulfate donor PAPS (13). Intriguingly, a putative cytochrome P450 monooxygenase, *cyp128*, is directly upstream of *stf3* in the Mtb genome (27). Given the genomic clustering of these genes, we hypothesize that Cyp128 is responsible for oxidation of the terminal position of the polyisoprenoid chain of MK enabling sulfation by Stf3. Characterization of Cyp128 *in vitro* and analysis of the metabolomic profile of a *cyp128* mutant will be key to understanding its role in S881 biosynthesis.

Previous studies have shown that bacterial sulfated metabolites are known to be key regulators of bacteria-host communication (3,28). The structure of S881, coupled with its localization to the outer cell wall of the bacterium and its possible role as a negative regulator of virulence, poses many questions regarding a molecular function for the metabolite. The location of S881 in the outer leaflet of Mtb makes it poised to interact with host immune cells and modulate the immune response.

An alternative hypothesis is that S881 is a precursor to another virulence factor. Allylic sulfate esters, under acidic conditions, are excellent leaving groups which would enable S881 to be cyclized into a terpenoid. The terpenoid would not have been identified in our mass spectrometry screens which were focused on molecules containing a sulfate moiety. The Mtb genome encodes a pair of class I and class II diterpene synthases, Rv3377 and Rv3378 which

together produce a diterpene edaxadiene from geranylgeranyl diphosphate (29-31). There is no equivalent enzyme(s) in the region around the S881 biosynthetic genes, but further MS studies are underway to identify any MK-9(H₂) metabolites that are absent from S881-deficient Δ *stf3* mutants.

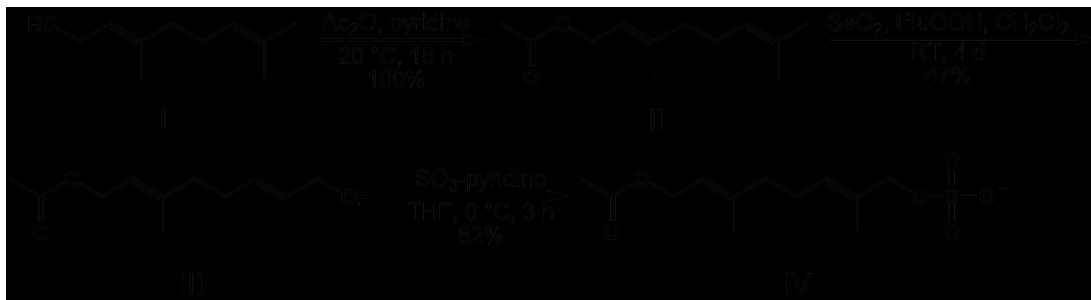
Another putative function of S881 is to alter the levels of MK in the bacteria. Since menaquinone is the sole quinol electron carrier in the mycobacterial respiratory chain (14), it is an essential compound to the survival of the bacterium and therefore a key drug target (20,32). Many recent studies have shed light on the branched nature of the mycobacterial respiratory chain and its remarkable adaptation to low-oxygen conditions via alternative terminal electron acceptors (33-37). However, little is known about the regulation of quinone availability in Mtb. S881 could be an intermediate in a mechanism of non-transcriptional regulation of quinone availability, and sulfation could be the signal for transport to the outer cell envelope. The conversion of MK-9(H₂) to S881 and the subsequent secretion of S881 could disrupt the redox balance of the respiratory chain, possibly slowing the respiratory capacity of the cell.

In conclusion, we have utilized high-resolution, high-mass-accuracy, and tandem mass spectrometry to successfully characterize the structure of the Mtb metabolite S881 as a sulfated derivative of MK-9(H₂). This study highlights the utility and value of MS to the study of previously uncharacterized metabolites. The structures of S881 represent a novel class of metabolites, and broaden our general understanding of prokaryotic sulfated molecules. The structural elucidation of S881 represents a key breakthrough in the study of its molecular function, and further insight into its mechanism of action will aid our understanding of its contribution to the virulence and life cycle of Mtb.

Materials and Methods

S881 extraction and enrichment

Approximately five grams of Mtb H37Rv cells (provided by Colorado State University) were extracted in 100 mL 1:1 CHCl₃/MeOH at room temperature for 2 h and filtered 3 times by vacuum filtration. The filtrate was concentrated to dryness and resuspended in 50 mL CHCl₃. The CHCl₃ was extracted with water (2 x 25 mL), and the water layer was extracted with 50 mL CHCl₃. The organic layers were combined and concentrated. A total of 40 mg of lipid extract was resuspended in 4:1 CHCl₃/MeOH and passed over anion exchange resin (AG4-X4, 100-200 mesh, biotechnology grade, free-base form, Bio-RAD) that had been pre-charged with 400:100:0.6 CHCl₃/MeOH/AcOH. The void volume was collected and the column was washed



Scheme 2-3: Synthesis of geranyl sulfate

with 4:1 CHCl₃/MeOH and the wash was collected. The column was eluted by a gradient of 2-5 mM triethylamine in 4:1 CHCl₃/MeOH. Fractions were concentrated, and kept at -20 °C until MS analysis.

Chemicals and reagents

All chemical reagents were of analytical grade, obtained from commercial suppliers, and used without further purification unless otherwise noted. With exception of reactions performed in aqueous media, all reaction vessels were flame-dried before use and performed under an N₂ atmosphere. Tetrahydrofuran (THF) was distilled under N₂ from Na/benzophenone immediately before use and CH₂Cl₂ was dried over CaH₂ and distilled under a N₂ atmosphere. Flash chromatography was carried out with Merck 60 230-400 mesh silica gel. Analytical thin layer chromatography (TLC) was performed on glass-backed Uniplate GHLF silica gel plates and visualized by staining with ceric ammonium molybdate. Organic extracts were dried over Na₂SO₄, and solvents were removed with a rotary evaporator at reduced pressure (20 torr) unless otherwise noted. ¹H and ¹³C NMR spectra were obtained with 400 MHz or 500 MHz Bruker spectrometers. Chemical shifts are reported in ppm referenced to the solvent peak. Coupling constants (*J*) are reported in Hz.

Synthesis of II (Scheme 2-1)

Acetic anhydride (1.2 mL, 13.0 mmol) was added to a solution of geraniol (I) (1.0 g, 6.5 mmols) in pyridine (13.0 mL) at room temperature and left to stir overnight. The reaction was quenched with water (5.0 mL) and was extracted with diethyl ether (3 x 50 mL). The combined organic extracts were washed with water (3 x 50 mL) and brine (1 x 50 mL), and then dried and concentrated. The product was purified by silica gel chromatography (20% ether/hexanes), resulting in a yellow oil (1.3 g, 100%). ¹H NMR (CDCl₃, 400 MHz): δ 1.55 (s, 3H), 1.63 (s, 3H), 1.66 (s, 3H), 2.01 (m, 5H), 2.05 (m, 2H), 4.53 (d, 2H, *J* = 7.2), 5.02 (t, 1H, *J* = 6.4), 5.22 (t, 1H, *J* = 6.8). ¹³C-NMR (CDCl₃, 400 MHz): δ 16.3, 17.6, 20.9, 25.5, 26.2, 39.4, 61.2, 118.2, 123.6, 131.7, 142.1, 171.0.

Synthesis of III (Scheme 2-1)

Selenium dioxide (44.4 mg, 0.4 mmol) was dissolved in a mixture of CH₂Cl₂ (6.0 mL) and 70% *tert*-butyl hydrogen peroxide in water (0.2 mL, 1.6 mmol) and stirred for 30 min. Compound II (0.160 g, 0.8 mmols), dissolved in CH₂Cl₂ (1.2 mL), was added to the reaction and stirred for 24 h at room temperature. When TLC analysis showed depletion of starting material, the reaction was quenched with water (5 mL) and the organic layer was extracted with water (3 x 25 mL) and dried. The desired product (191.8 mg, 12%) was purified from a mixture of at least six compounds by column chromatography (gradient of 20 to 100% Et₂O/hexanes). ¹H NMR (CDCl₃, 400 MHz): δ 1.55 (s, 3H), 1.63 (s, 3H), 1.66 (s, 3H), 2.01 (m, 5H), 2.05 (m, 2H), 4.53 (d, 2H, *J* = 7.2), 5.02 (t, 1H, *J* = 6.4), 5.22 (t, 1H, *J* = 6.8). ¹³C-NMR (CDCl₃, 400 MHz): δ 16.3, 17.6, 20.9, 25.5, 26.2, 39.4, 61.2, 118.2, 123.6, 131.7, 142.1, 171.0.

Synthesis of IV (Scheme 2-1)

Sulfur trioxide-pyridine complex (215.7 mg, 1.4 mmol) was added to a solution of compound III (191.8 mg, 0.9 mmol) in THF (3.6 mL) at 0 °C. The reaction was stirred until TLC analysis showed depletion of starting material. The reaction was quenched with MeOH (2 mL), concentrated on a rotary evaporator, and the product was purified directly by column chromatography (gradient of 20 to 30% MeOH/CHCl₃), yielding the pyridinium salt IV (233.4 mg, 82% yield) as a white solid. ¹H NMR (MeOD, 500 MHz): δ 2.09 (s, 3H), 2.12 (s, 3H), 2.43 (s, 3H), 2.50 (m, 2H), 2.58 (m, 2H), 4.11 (s, 2H), 4.93 (d, 2H, *J* = 22.5), 5.74 (t, 1H, *J* = 7.0), 5.88 (t, 1H, *J* = 6.5). ¹³C-NMR (MeOD, 500 MHz): δ 14.0, 16.3, 27.2, 40.1, 50.0, 55.4, 59.5, 75.2, 125.3, 130.2, 132.1, 139.0.

Sample preparation for MS analysis

Immediately prior to analysis, the S881 enriched fractions were dried and resuspended in 1 mL of 2:1 CHCl₃/MeOH for MS analysis. Vitamin K1 was purchased from Fisher Scientific and was diluted with 2:1 CHCl₃/MeOH to a final concentration of 5 μM for MS analysis. Geranyl sulfate solid (Scheme 2-1, IV) was diluted to a final concentration of 20 μM in 2:1 CHCl₃/MeOH.

ESI-FTICR MS

Mass spectra were obtained on an Apex II FT-ICR MS (Bruker Daltonics, Billerica, MA) equipped with a 7 T actively-shielded superconducting magnet. Ions were introduced into the ion source via direct injection at a rate of 2 μL/min. Ions were generated with an Apollo pneumatically-assisted electrospray ionization source (Bruker Daltonics) operating in the negative ion mode, and were accumulated in an rf-only external hexapole for 0.5-2 s before being transferred to the ion cyclotron resonance cell for mass analysis.

For MSⁿ, ions were isolated by a correlated harmonic excitation field isolation sweep, and cleanup shots were used to eject ions not ejected from the initial sweep. Isolated ions were collisionally activated via sustained off-resonance irradiation collision-induced dissociation (SORI-CID) at 1.2 to 1.5 kHz above the cyclotron frequency of the ion of interest, using argon as the collision gas. Fragment ions were excited for detection after a delay of 2-3 s to allow the residual argon to pump away.

Mass spectra consist of 256k (MSⁿ spectra) to 1M (exact mass spectra) data points and are an average of 24-32 scans. The spectra were acquired using XMASS version 6.0.0 or 7.0.8 (Bruker Daltonics). For accurate mass measurements, spectra were internally calibrated with at least four known compounds. MSⁿ spectra were calibrated externally. DataAnalysis 3.4 (Bruker Daltonics) was used to determine elemental compositions.

ESI Linear-Ion Trap MS

Additional mass spectra were obtained on an LTQ ion trap MS (ThermoFinnigan, San Jose, CA) operating in the negative ion mode. Ions were introduced into the ion source via direct injection at a rate of 5 μL/min. For MSⁿ experiments, the precursor ions were isolated with an isolation width of 1-3 Da, the ions were activated with a 25% normalized collision energy for 100 ms, and the *q_z* value was maintained at 0.250. Spectra are an average of 40-100 scans, acquired using Xcalibur, version 1.4 (ThermoFinnigan).

References

1. World Health Organization. (<http://www.who.int/tb/en/>)
2. Ehrhardt, D. W., Atkinson, E. M., Faull, K. F., Freedberg, D. I., Sutherland, D. P., Armstrong, R., and Long, S. R. (1995) *J Bacteriol* **177**, 6237-6245
3. Roche, P., Debelle, F., Maillet, F., Lerouge, P., Faucher, C., Truchet, G., Denarie, J., and Prome, J. C. (1991) *Cell* **67**, 1131-1143
4. Shen, Y., Sharma, P., da Silva, F. G., and Ronald, P. (2002) *Mol Microbiol* **44**, 37-48
5. Chapman, E., Best, M. D., Hanson, S. R., and Wong, C. H. (2004) *Angew Chem Int Ed Engl* **43**, 3526-3548
6. Armstrong, J. I., and Bertozzi, C. R. (2000) *Curr Opin Drug Discov Devel* **3**, 502-515
7. Hemmerich, S. (2001) *Drug Discov Today* **6**, 27-35
8. Goren, M. B., Brokl, O., and Das, B. C. (1976) *Biochemistry* **15**, 2728-2735
9. Goren, M. B. (1970) *Biochim Biophys Acta* **210**, 127-138
10. Mougous, J. D., Petzold, C. J., Senaratne, R. H., Lee, D. H., Akey, D. L., Lin, F. L., Munchel, S. E., Pratt, M. R., Riley, L. W., Leary, J. A., Berger, J. M., and Bertozzi, C. R. (2004) *Nat Struct Mol Biol* **11**, 721-729
11. Seeliger, J. C., Holsclaw, C. M., Schelle, M. W., Botyanszki, Z., Gilmore, S. A., Tully, S. E., Niederweis, M., Cravatt, B. F., Leary, J. A., and Bertozzi, C. R. (2012) *J Biol Chem* **287**, 7990-8000
12. Gilmore, S. A., Schelle, M. W., Holsclaw, C. M., Leigh, C. D., Jain, M., Cox, J. S., Leary, J. A., and Bertozzi, C. R. (2012) *ACS Chem Biol*
13. Mougous, J. D., Senaratne, R. H., Petzold, C. J., Jain, M., Lee, D. H., Schelle, M. W., Leavell, M. D., Cox, J. S., Leary, J. A., Riley, L. W., and Bertozzi, C. R. (2006) *Proc Natl Acad Sci U S A* **103**, 4258-4263
14. Collins, M. D., and Jones, D. (1981) *Microbiol Rev* **45**, 316-354
15. Mougous, J. D., Green, R. E., Williams, S. J., Brenner, S. E., and Bertozzi, C. R. (2002) *Chem Biol* **9**, 767-776
16. Hermansson, K., Jansson, P., Low, P., Gallner, G., Swiezewska, E., and Chojnacki, T. (1992) *Biol Mass Spectrom.* **21**, 548-553
17. Rick, P. D., Hubbard, G. L., Kitaoka, M., Nagaki, H., Kinoshita, T., Dowd, S., Simplaceanu, V., and Ho, C. (1998) *Glycobiology* **8**, 557-567

18. Kaur, D., Brennan, P. J., and Crick, D. C. (2004) *J Bacteriol* **186**, 7564-7570
19. Dunphy, P. J., Phillips, P. G., and Brodie, A. F. (1971) *J Lipid Res* **12**, 442-449
20. Truglio, J. J., Theis, K., Feng, Y., Gajda, R., Machutta, C., Tonge, P. J., and Kisker, C. (2003) *J Biol Chem* **278**, 42352-42360
21. Converse, S. E., Mougous, J. D., Leavell, M. D., Leary, J. A., Bertozzi, C. R., and Cox, J. S. (2003) *Proc Natl Acad Sci U S A* **100**, 6121-6126
22. Domenech, P., Reed, M. B., Dowd, C. S., Manca, C., Kaplan, G., and Barry, C. E., 3rd. (2004) *J Biol Chem* **279**, 21257-21265
23. Khoo, K. H., Jarboe, E., Barker, A., Torrelles, J., Kuo, C. W., and Chatterjee, D. (1999) *J Biol Chem* **274**, 9778-9785
24. Lopez Marin, L. M., Laneelle, M. A., Prome, D., Laneelle, G., Prome, J. C., and Daffe, M. (1992) *Biochemistry* **31**, 11106-11111
25. Ishii, M., Kawasumi, T., Igarashi, Y., Kodama, T., and Minoda, Y. (1987) *J Bacteriol* **169**, 2380-2384
26. Bentley, R., and Meganathan, R. (1982) *Microbiol Rev* **46**, 241-280
27. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaiia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) *Nature* **393**, 537-544
28. Lee, S. W., Han, S. W., Sririyanyum, M., Park, C. J., Seo, Y. S., and Ronald, P. C. (2009) *Science* **326**, 850-853
29. Prach, L., Kirby, J., Keasling, J. D., and Alber, T. (2010) *FEBS J* **277**, 3588-3595
30. Mann, F. M., Xu, M., Chen, X., Fulton, D. B., Russell, D. G., and Peters, R. J. (2009) *J Am Chem Soc* **131**, 17526-17527
31. Mann, F. M., Pristic, S., Hu, H., Xu, M., Coates, R. M., and Peters, R. J. (2009) *J Biol Chem* **284**, 23574-23579
32. Rao, S. P., Alonso, S., Rand, L., Dick, T., and Pethe, K. (2008) *Proc Natl Acad Sci U S A* **105**, 11945-11950

33. Weinstein, E. A., Yano, T., Li, L. S., Avarbock, D., Avarbock, A., Helm, D., McColm, A. A., Duncan, K., Lonsdale, J. T., and Rubin, H. (2005) *Proc Natl Acad Sci U S A* **102**, 4548-4553
34. Boshoff, H. I., Myers, T. G., Copp, B. R., McNeil, M. R., Wilson, M. A., and Barry, C. E., 3rd. (2004) *J Biol Chem* **279**, 40174-40184
35. Matsoso, L. G., Kana, B. D., Crellin, P. K., Lea-Smith, D. J., Pelosi, A., Powell, D., Dawes, S. S., Rubin, H., Coppel, R. L., and Mizrahi, V. (2005) *J Bacteriol* **187**, 6300-6308
36. Voskuil, M. I., Visconti, K. C., and Schoolnik, G. K. (2004) *Tuberculosis (Edinb)* **84**, 218-227
37. Shi, L., Sohaskey, C. D., Kana, B. D., Dawes, S., North, R. J., Mizrahi, V., and Gennaro, M. L. (2005) *Proc Natl Acad Sci U S A* **102**, 15629-15634

Chapter 3: Identification of the S881 biosynthetic operon

Introduction

Mycobacterium tuberculosis (Mtb), the etiological agent of tuberculosis disease, is a highly specialized human pathogen that has evolved specific mechanisms to survive in its natural environment: the human macrophage (1,2). Mtb infection begins with the inhalation of the bacteria into the lungs where they are phagocytosed by alveolar macrophages. Once inside the phagosome, Mtb has evolved mechanisms to halt phagosome maturation avoiding the host's natural bactericidal processes. Failure to eliminate the bacteria induces the innate immune response causing leukocytes to traffic to and surround the infected cells. The leukocytes provide a barrier to prevent the spread of infection forming a large cell mass within the lung called a granuloma. Within the granuloma, Mtb adapts to the changing environment by altering its respiration and metabolism (3,4). Upon initial infection, Mtb is in a high oxygen environment and has a high metabolic rate enabling rapid replication. The granuloma, on the other hand, has been shown to have a range of oxygen tensions from microaerobic to hypoxic (5). As an obligate aerobe, Mtb tailors its rate of respiration and metabolism to survive in the granuloma where it can persist intracellularly for long periods of time.

As a result of its environment, Mtb has evolved to coordinate its pathogenicity to its metabolism and respiration pathways. An increasing number of studies are finding that the carbon source utilized by Mtb *in vivo* correlates with the production of lipid based virulence factors (6,7). For example, Mtb uses host lipids such as cholesterol as its primary carbon source during infection (6). When Mtb changes from metabolizing carbohydrates, commonly used during *in vitro* culture, to cholesterol, Mtb increases the production of lipid-based virulence factors found in its outer layer as determined by metabolomics and mass spectrometry (8).

Our knowledge of the Mtb metabolism and respiration pathways are primarily based on our knowledge of analogous pathways in *E. coli*. Because *E. coli* and Mtb occupy very different environmental niches, each has evolved to optimize different biological pathways. On a genomic level, these homology-based approaches have failed to identify functions for almost 40% of the Mtb genome (9) and a little less than 30% of known biochemical transformations have no gene ascribed to them (10). The use of genetics and transcriptomics is greatly enhanced and complemented when used in conjunction with analysis of the unique small molecules and metabolites of Mtb.

The use of DNA microarrays has enabled the analysis of how Mtb responds and adapts to a specific stimuli. But transcriptomics only gives a snap shot of the changes at the protein level and does not give any information on the state of metabolism or respiration. With the development of high resolution mass spectrometry and liquid chromatography, we can begin studying the vast array of unique small molecules produced by Mtb. With the use of metabolomics, we can get an accurate view of the active processes of Mtb at a given time point. Using this technology, we can understand more about how Mtb survives in the hostile environment of a phagosome.

As previously mentioned, the carbon source for Mtb metabolism is linked to the levels of lipid based virulence factors on the bacterial surface (8). Many of the lipid based virulence factors are unique to Mtb and the structural characterization of the lipids led to the identification and biochemical characterization of the enzymes involved in their biosynthesis (11). In 2006, using a genetic and mass spectrometry approach, we sought to identify novel sulfated

metabolites produced by Mtb (12). With bioinformatics, we identified four sulfotransferases encoded in the Mtb genome. The genetic mass spectrometry approach entailed making single gene deletion mutants in each of the identified sulfotransferases and analyzed the total lipid extracts from each strain by high resolution mass spectrometry. As a result of this screen, we discovered a novel sulfated metabolite, which we named S881 based on its exact mass, in the total lipid extracts of Mtb (13). Subsequent experiments using a Mtb *stf3* mutant ($\Delta stf3$) showed that *stf3* is required for S881 production. Further investigation of this strain in a mouse infection model displayed a hypervirulent phenotype, one of the first examples of this highly unusual phenotype. In addition to an increased time to death, the $\Delta stf3$ strain showed increased bacterial burden in the lungs and spleen of infected mice but had no growth phenotype *in vitro*. Using high resolution mass spectrometry we determined the structure of S881 as an unprecedented terminally sulfated form of menaquinone, the sole lipiquinone utilized in the electron transport chain of Mtb (14).

Initial studies of S881 focused on identifying the S881 biosynthetic genes. We identified a cytochrome P450 *cyp128* in addition to *stf3* necessary for S881 production by engineering a model organism *Mycobacterium smegmatis* (*M. smeg*) to produce S881. Next, we generated a *cyp128* mutant in Mtb and showed that it did not produce S881. To better understand S881 regulation, we explored how S881 production changes during different phases of growth *in vitro*. Finally, using a mutant that produces increased S881 over wild type (WT), we studied S881 transport to the outer bacterial envelope.

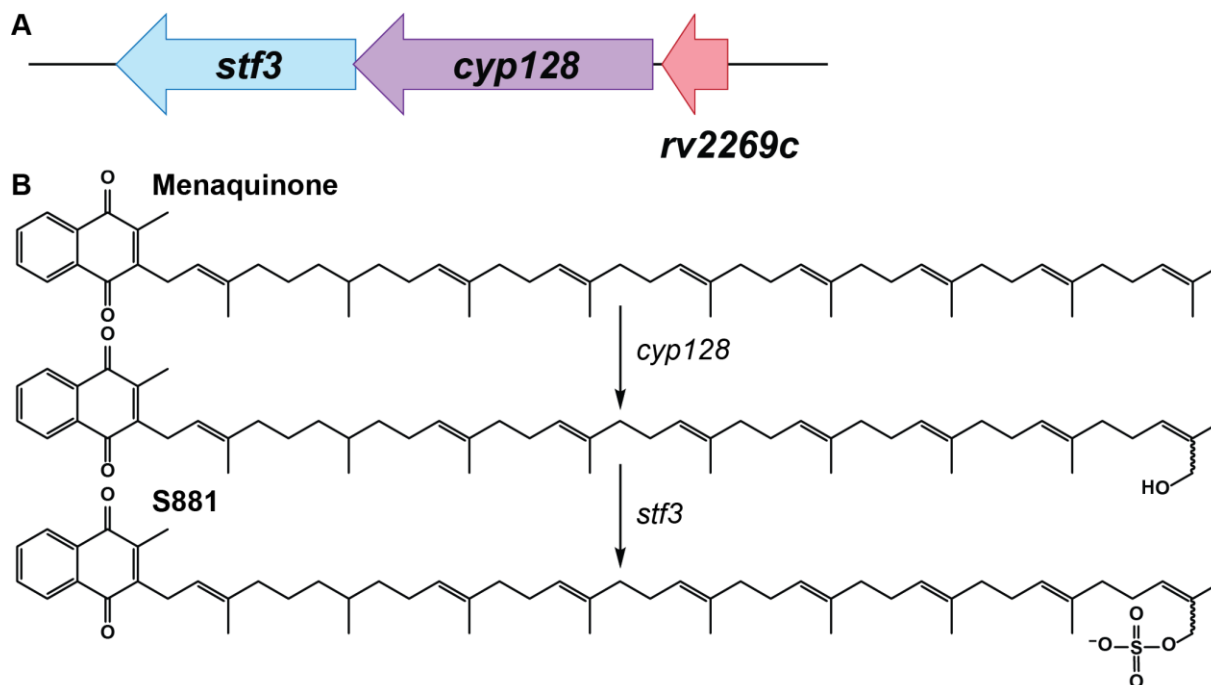


Figure 3-1: (A) The *stf3* operon in Mtb. (B) The putative biosynthesis of S881 from menaquinone.

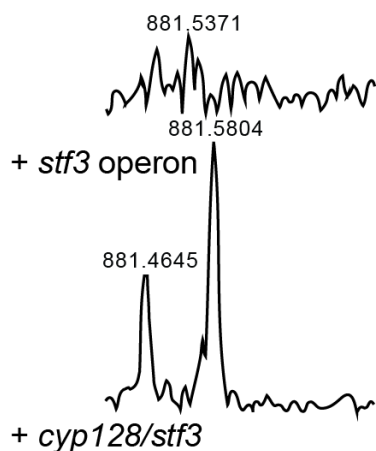
Results

Bioinformatic analysis of the *stf3* operon

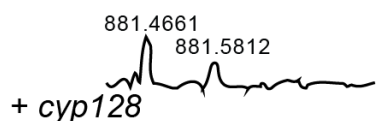
The putative *stf3* operon contains three genes: *stf3* (*rv2267c*) the putative sulfotransferase required for S881 biosynthesis (13), *cyp128* (*rv2268c*) a putative cytochrome P450, and *rv2269c* a hypothetical protein of unknown function (9). *Rv2269c* is a very short putative gene (333 bp) that is not predicted to encode any known conserved protein domains and has low homology to other proteins. A prediction of protein structure suggests that it may have a short alpha helix fold but is mostly unstructured.

A BLAST search of *cyp128* indicates that there are two possible start sites. The most commonly start site, as annotated in H37Rv genome, is 294 bp (98 aa) longer than the other. An orthologue of *cyp128* is found in *M. bovis* and *M. canettii* and *M. kansasii* encode a homologue to Mtb *cyp128*. The *M. kansasii* homologue is 78% identical and 86% similar with the largest differences in the N and C-termini. The Erdman strain of Mtb produces a reduced level of S881 (data not shown) due to a single nucleotide deletion (C at 1140 bp of 1470 bp) causing an early truncation in the Erdman *cyp128* (18).

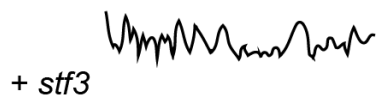
WT *M. smeg*



+ *cyp128/stf3*



+ *cyp128*



+ *stf3*

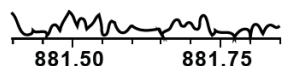


Figure 3-2: Total lipid extracts from *M. smeg* engineered strains.

The sulfotransferase *stf3* was identified in a bioinformatics study to identify Mtb sulfotransferases by homology to the human carbohydrate sulfotransferase, GlcNAc-6-sulfotransferase GST3 (19). It is unique to the Mtb complex, but there are homologues in *M. canettii* and *M. kansasii*.

Two genes, cyp128 and stf3, are required for the biosynthesis of S881 from menaquinone

The putative *stf3* operon includes three genes: *rv2269c*, a putative protein of unknown function, *cyp128*, a putative cytochrome P450, and *stf3*, a putative sulfotransferase that is required for the production of S881 (9,13,19). In addition, there is a cytochrome P450, *cyp124*, adjacent and anti-parallel to *stf3* that has biochemically been characterized as methyl-branched lipid ω -hydroxylase with no activity on menaquinone *in vitro* (20). To determine which genes are required for S881 biosynthesis, we engineered *M. smeg*, a fast growing relative of Mtb, to produce S881. *M. smeg* produces the same menaquinone analogues as Mtb but does not encode homologues for any of the genes hypothesized to be involved in S881 biosynthesis. *M. smeg* was transformed with episomal plasmids containing the genes *Rv2269c* (pKMS101), *cyp128* (pKMS102), or *stf3* (pKMS103) individually, as well as plasmids

containing all three genes (pKMS104) or with *cyp128* and *stf3* (pKMS105). Total lipid extracts from all strains were analyzed by high resolution mass spectrometry for the presence of S881. As expected, wild type (WT) *M. smeg* and the strains expressing each gene individually (pKMS101, pKMS102, or pKMS103) did not have a peak corresponding to S881 (Figure 3-2). Alternatively, total lipid extracts from the *M. smeg* strains expressing plasmids pKMS104 (*stf3* operon) and pKMS105 (*cyp128/stf3*) were analyzed by high-resolution mass spectrometry and both exhibited peaks corresponding to the exact mass of S881 (Figure 3-2). Fragmentation of the S881 peak confirmed its identity. Only two genes, *cyp128* and *stf3*, are required for S881 biosynthesis from menaquinone. The addition of the putative protein *Rv2269c* showed no effect on S881 production as analyzed by mass spectrometry.

Strains of *M. smeg* expressing S881 genes also showed an increase in a sulfated MK-9(H₄), which has two isoprene units saturated (results not shown). It is unclear if MK-9(H₄) is naturally produced in *M. smeg* or it is an artifact of S881 engineering. The sulfated MK-9(H₄) does not appear to be present in Mtb total lipid extracts.

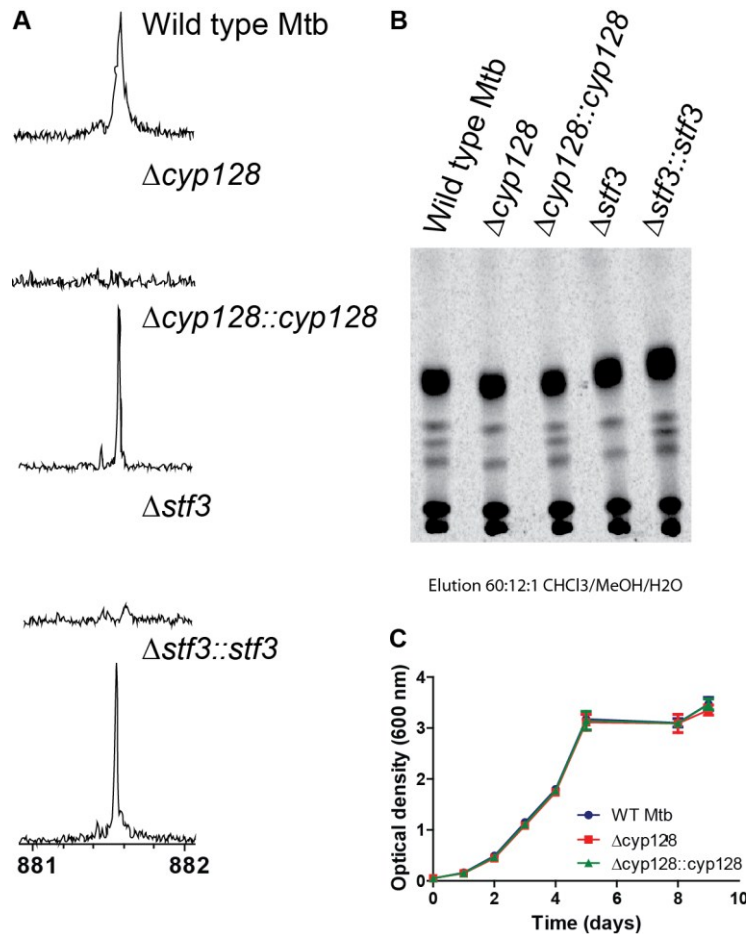


Figure 3-3: Analysis of the $\Delta cyp128$ Mtb mutant. (A) MS analysis of total lipid extracts from Mtb strains zoom on the region of $m/z = 881$. (B) Radio-TLC analysis of Mtb strains grown with ^{35}S -sulfate. (C) Growth curve of $\Delta cyp128$ and complement strains compared to WT.

The gene cyp128 is required for S881 biosynthesis

A *cyp128* gene deletion mutant in Mtb H37Rv ($\Delta cyp128$) was made using phage transduction to introduce a hygromycin cassette interrupting *cyp128*. Despite being annotated as essential by Sassetti et al. (21,22), the $\Delta cyp128$ mutant grew readily in rich broth. The mutant was confirmed by PCR amplification of genomic DNA. To determine if *cyp128* was necessary for S881 production, total lipid extracts from $\Delta cyp128$ were analyzed by high resolution mass spectrometry. As compared to WT H37Rv, total lipid extracts from the $\Delta cyp128$ mutant showed a loss of S881 production (Figure 3-3A). The total lipid extracts of the previously characterized $\Delta stf3$ mutant were done in parallel for comparison. S881 production was restored in the $\Delta cyp128$ mutant when full length WT *cyp128* was introduced on an integrating plasmid. The loss of S881 production in $\Delta cyp128$ was further confirmed by labeling sulfated metabolites with radioactive sulfate ($^{35}\text{SO}_4^{2-}$). The wild type bacteria, $\Delta cyp128$, and complement were incubated overnight in the presence of ^{35}S sulfate in PBS and the cell pellets lipid extracts were analyzed by TLC. Cultures of the $\Delta stf3$ mutant and complement were included as controls. The $\Delta cyp128$ and $\Delta stf3$ mutants showed no spot corresponding to S881 but complementation restored S881

production (Figure 3-3B). Additionally, the *in vitro* growth of the $\Delta cyp128$ mutant in rich media showed no difference in growth rate to wild type or the complemented strain (Figure 3-3C).

S881 is produced at similar levels during all phases of growth

Using WT Mtb, we wanted to determine if S881 was temporally regulated during Mtb growth *in vitro*. Previously, our lab has shown that Sulfolipid-1 (SL-1), a highly abundant sulfoglycolipid found in the outer cell envelope of Mtb, is more abundant during late-log phase and decreases with transition to stationary phase (unpublished observation). To determine if S881 is similarly regulated, we grew WT Mtb to mid-log phase, washed in tween free media and diluted to OD₆₀₀ of 0.2 in 25 mL with one flask per time point (T=0). Total lipid extracts were taken on day 1 (~OD₆₀₀ 0.3; early log phase), day 3 (~OD₆₀₀ 0.7, late log phase) and day 7 (~OD₆₀₀ 2.0, stationary phase). Extracts were analyzed by FT-ICR MS (Figure 3-4). The levels of S881 normalized to OD and did not change much through the different phases of growth, indicating that the S881 biosynthesis is not highly regulated during *in vitro* growth.

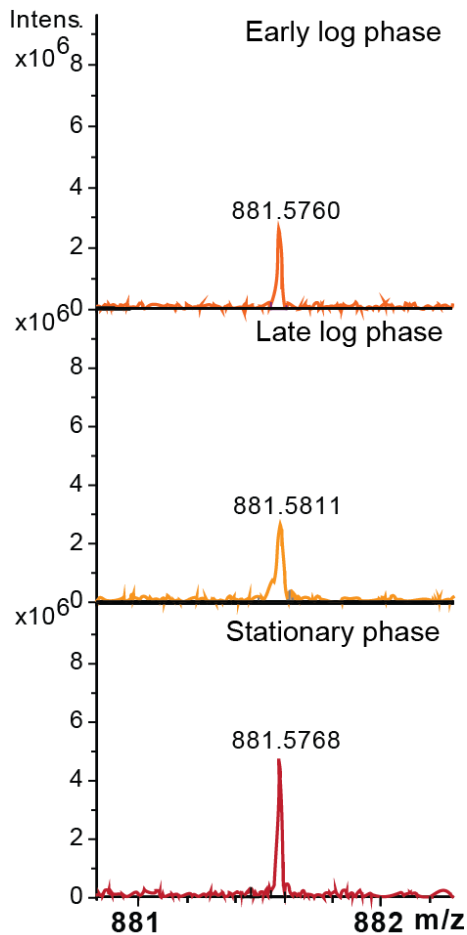


Figure 3-4: The temporal regulation of S881 in WT Mtb grown in rich media.

Overexpression of S881 biosynthetic genes leads to an accumulation of S881 in the cell pellet

Characterization of the *rv2269c* mutant ($\Delta rv2269c$) showed a surprising increase in S881 production. All attempts to complement the phenotype were unsuccessful. The increase of S881 production is most likely due to a polar effect as a result of the hygromycin cassette interrupting *rv2269c*. The promoter for the hygromycin cassette is most likely driving transcription of either or both *cyp128* and *stf3*. Previously we had shown that S881 is found primarily in the outer cell envelope of Mtb (13). Using the overexpression mutant $\Delta rv2269c$, we wanted to study how S881 is transported from the plasma membrane to the outer cell envelope. To do this, we used a crude fractionation of the Mtb cell envelope using Hexanes to remove the outer non-covalently attached lipids from the cell surface (Hexanes layer). The remaining Mtb cell pellet was extracted with 1:1 chloroform/methanol to extract the remaining lipids (pellet extract). The Hexanes extract of WT Mtb contains most of the S881, with a very small amount remaining in the pellet extract (Figure 3-5A). The $\Delta rv2269c$ mutant, shows WT levels of S881 in the Hexanes extract, but shows an accumulation of S881 in the pellet extract (Figure 3-5B). This suggests that S881 is actively transported across the cell envelope by an as yet to be identified lipid transporter.

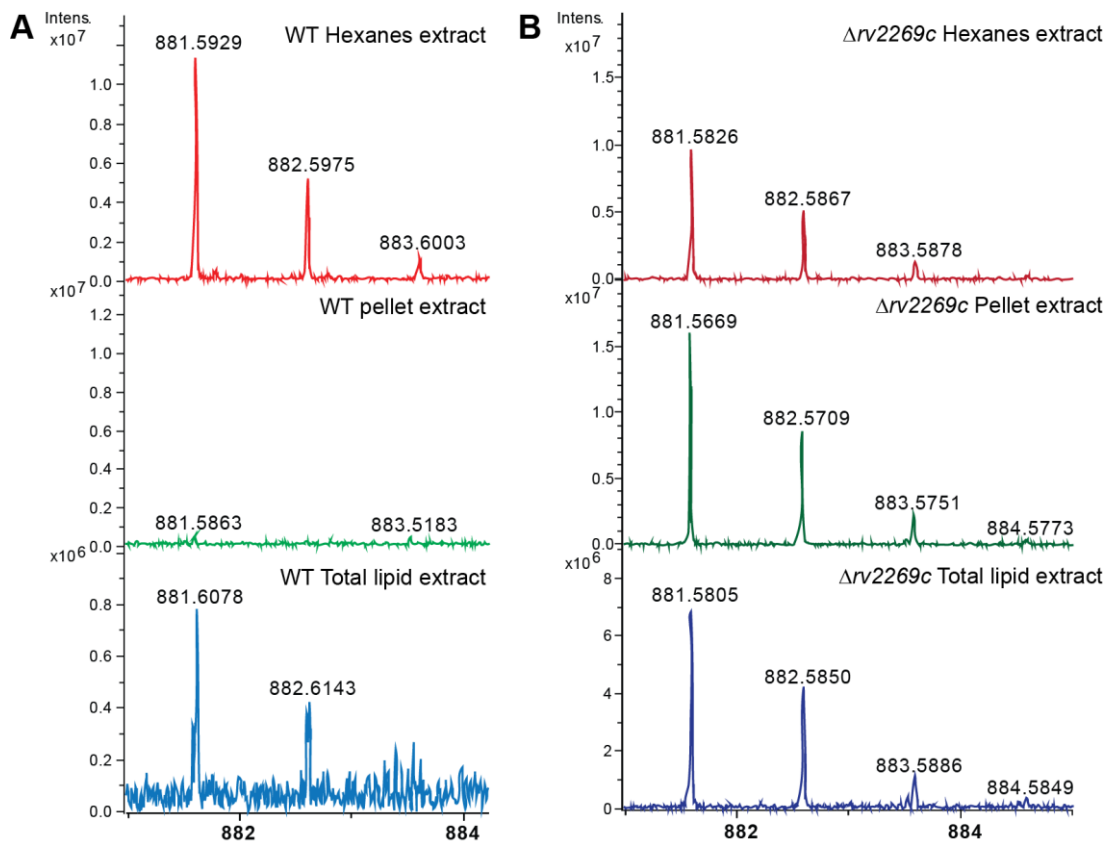


Figure 3-5: Localization of S881 in WT and $\Delta rv2269c$ strains. The Hexanes extract contains non-covalently attached outer cell envelope lipids. The pellet extract contains the remaining cellular lipids associated with the cell pellet. Total lipid extracts were used to compare total S881 levels. (A) WT Mtb. (B) $\Delta rv2269c$ mutant.

Discussion

We have demonstrated that the biosynthesis of S881, a novel sulfated menaquinone, is encoded by two genes: *cyp128*, a cytochrome P450 and *stf3*, a sulfotransferase. Recombinant expression of *cyp128* and *stf3* in *M. smeg*, a model organism that produces menaquinone, is sufficient to engineer S881 biosynthesis. Previously we had shown that *stf3* was required for S881 biosynthesis. Similarly, a Mtb mutant lacking *cyp128* ($\Delta cyp128$) does not produce S881 validating it involved in the pathway from MK to S881. We confirmed this result by analysis of total lipid extracts by high resolution mass spectrometry as well as metabolically labeling S881 with ^{35}S -sulfate and analyzing by TLC.

Initial experiments, studying the temporal production of S881 indicate that during *in vitro* growth S881 biosynthesis is constant. Unpublished RNA-seq data from David Sherman's lab (personal communication with Tige Rustad) shows very low levels of RNA transcript for the *stf3* operon. In fact, there appears to be higher transcript levels in the antiparallel direction. Further studies using conditions that mimic intracellular infection are needed to determine if S881 production is increased.

We hypothesized that the $\Delta stf3$ mutant would show an accumulation of the intermediate hydroxylated menaquinone. All attempts to identify a hydroxylated form of menaquinone were unsuccessful. No ion was seen in positive or negative total lipid extracts from Mtb or in strains of *M. smeg* over expressing the *cyp128*. Further attempts to identify it by radiolabeling or by analysis of Mtb lipids by HPLC also proved unsuccessful. We hope to study the biochemistry of Cyp128 and in the process make a standard for hydroxylated menaquinone to aid in identification by mass spectrometry.

Finally, using a mutant that produces an excess of S881, we hypothesize that S881 is actively transported to the outer cell envelope. The overexpression mutant $\Delta rv2269c$ exhibited an accumulation of S881 in the pellet extracts but similar levels of S881 in the hexanes extracts, suggesting that the levels of S881 on the cell surface are regulated by a lipid transporter. Mtb has dedicated almost 6% of its genome to lipid synthesis, degradation and transport (9). No lipid transporter has been identified for S881, but Mtb encodes a couple of lipid transport families. There are 14 MmpL proteins, three of which are known to be involved in the transport and biosynthesis of other Mtb specific lipids. Another family, including over 100 proteins, are the MmpS transporters, often associated with MmpL proteins. Only one has been characterized, MmpS4, in *M. smeg* as a glycopetidolipid transporter with no known native substrate (23). A family of lipid transporters encoded by the mammalian cell entry (Mce) operons, are hypothesized to transport host lipids, like cholesterol, that are used as a carbon source for Mtb growth (6). Further research is needed to identify the S881 transporter as it may be important in the regulation of S881 and its presentation to the host immune system.

The active transport of S881 across the cell wall may be another method of regulating how S881 is presented. If the function of S881 is to modulate the electron transport chain by removing menaquinone, the lipid transporter may either expedite S881 removal or provide a scaffold for S881 biosynthesis. Conversely, if S881 is involved in host-pathogen interactions, the lipid transporter may regulate how S881 is presented to the host cell and how much is presented. Further studies will be needed to identify the lipid transporter and its role in S881 regulation.

In conclusion, we have elucidated the S881 biosynthetic genes. *M. smeg* mutants engineered to produce S881 will be a valuable tool in studying the role S881 has on the electron

transport chain of the bacteria, as well as provide a facile system to isolate S881 for further studies in the host immune response to S881 presentation. The Mtb mutants in S881 will be used to further study the function of S881 and its role in hypervirulent phenotype in the mouse model of infection.

Materials and Methods

Bacterial strains and growth media

Mycobacterial strains used in these studies were Mtb H37Rv (gift from the Riley lab, UC Berkeley) and *M. smeg* mc²155 (ATCC 700084). Mtb was grown in 7H9 (liquid culture) or 7H11 (solid media) with 10% OADC supplement, 0.5% glycerol, and 0.05% tween-80 unless stated otherwise. *M. smeg* was grown in 7H9 (liquid) or on 7H10 (solid) media with 10% ADC, 0.5% glycerol, and 0.05% tween unless stated otherwise. Media and supplements were from BD Biosciences. Antibiotics were used for selection at 20 µg/mL kanamycin or 50 µg/mL hygromycin for mycobacteria. Cloning and plasmid propagation were performed in *E. coli* TOP10 and XL-1 Blue strains. Mtb genes were amplified from H37Rv genomic DNA. All *E. coli* cultures were grown in LB growth medium with 50 µg/mL kanamycin or 100 µg/mL hygromycin for selection.

Construction of Mtb knockout mutants

The $\Delta cyp128$ mutant strain was created by homologous recombination as previously described (15). Briefly, specialized transduction phage phKMS109 was incubated with concentrated Mtb strain H37Rv cells for 4 h at 39 °C. Cells were then plated on 7H11 plates containing hygromycin. Colonies were picked and screened for the disruption by PCR, which confirmed the replacement of 1279 bp of *cyp128* (amino acids 32 through 457) with a hygromycin resistance cassette. The $\Delta cyp128::cyp128$ complementation strain was created by cloning the *cyp128* gene plus 333 bp upstream of the gene from Mtb strain H37Rv into the mycobacterial expression vector pMV306, a derivative of the pMV361 vector (16) with a multiple cloning site in place of the expression cassette. The resulting plasmid was electroporated into $\Delta cyp128$ and transformants were selected on 7H11 kanamycin-containing plates. The $\Delta rv2269c$ mutant was created as described above using phKMS108. Complementation of *rv2269c* was done with both a strong constitutive promoter (pGS) and with the putative native promoter (1 kb upstream of *rv2269c*).

Extraction and MS analysis of lipids

Mtb strains or *M. smeg* strains were grown to mid-log phase. Cultures were washed and diluted in Tween-free media to OD₆₀₀ of 0.2 and grown for two doublings. Cells were harvested, and resuspended in 4 mL 1:1 chloroform/methanol and extracted overnight at RT. Alternatively the pellet was resuspended in 2 mL hexanes per 50 mL culture, and spun for 2 min at 3500 rpm. The upper organic phase was removed and added to an equal volume of 1:1 chloroform/methanol. The remaining cell pellet and aqueous phase were extracted in 4 mL per 50 mL culture of 1:1 chloroform/methanol and incubated at room temperature overnight. Cell

debris was pelleted by centrifugation and the supernatant decanted. All extracts were stored at -20° C until analysis. All extractions were repeated in at least three independent experiments.

High-resolution FT-ICR mass spectra were obtained on an Apex II FT-ICR mass spectrometer (Bruker Daltonics, Billerica, MA) equipped with a 7-tesla actively-shielded superconducting magnet. Ions were introduced into the ion source via direct injection at a rate of 1 μ L/min. Ions were generated with an Apollo pneumatically-assisted electrospray ionization source (Bruker Daltonics, Billerica, MA) operating in the negative ion mode. The ESI source tuning parameters were for the mass range of m/z 300-1000: the capillary voltage was set to 4.5 kV, the capillary exit voltage was -300 V, the skimmer 1 voltage was -20 V, and the skimmer 2 voltage was set to -7 V. The ions were accumulated in an rf-only external hexapole for 1 s before being transferred to the ion cyclotron resonance cell for mass analysis.

High-resolution mass spectra consist of 512k data points and are an average of 24 scans. The spectra were acquired using XMASS version 7.0.8 (Bruker Daltonics, Billerica, MA). For accurate mass measurements, spectra were internally calibrated with at least four known compounds bracketing the region of the ions of interest.

Additional MSⁿ spectra were obtained on an LTQ ion trap mass spectrometer equipped with an electrospray ionization source (ThermoFinnigan), operating in the negative ion mode. Ions were introduced into the ion source via direct injection at a rate of 5-10 μ L/min. CID was used for MSⁿ experiments. The precursor ions were isolated with an isolation width of 1–3 Da, the ions were activated with a 26% normalized collision energy for 100 ms, and the q_z value was maintained at 0.250. Spectra are an average of 100 scans, acquired using Xcalibur, version 1.4 (ThermoFinnigan).

Mtb and M. smeg ³⁵S metabolic labeling and lipid analysis by thin-layer chromatography (TLC)

Mtb or *M. smeg* strains were grown in 7H9 to late-log phase. Two generations prior to labeling, strains were diluted to an OD₆₀₀ of 0.3. For ³⁵S labeling, cells were resuspended at OD₆₀₀ ~1 in 10 mL PBS with 1% acetate and 100 μ Ci ³⁵S-sulfate (PerkinElmer). After shaking incubation at 37 °C overnight, cell pellets were extracted in 1:1 chloroform/methanol as described above. Solvent was removed by evaporation and the extracts resuspended at 1/10 or 1/20 the original volume of 1:1 chloroform/methanol. An equal volume of each fraction was spotted on silica plates (HPTLC Silica Gel 60, EMD Chemicals) and developed in 60:12:1 chloroform/methanol/water. Plates were analyzed by phosphorimaging (GE Biosciences Typhoon).

Measurement of Mtb strains growth kinetics

Wild type *Mtb*, S881 mutants and compliments were grown to mid-log phase in 50 mL in roller bottles. Clumps were removed by gentrification and the cultures were diluted to an OD₆₀₀ of 0.05 with a final volume of 50 mL in 250 mL shaker flasks. A 1 mL aliquot was taken every 24 hours and measured by OD₆₀₀.

References

1. Russell, D. G. (2001) *Nat Rev Mol Cell Biol* **2**, 569-577
2. Nathan, C. (2009) *Cell Host Microbe* **5**, 220-224
3. Russell, D. G., VanderVen, B. C., Lee, W., Abramovitch, R. B., Kim, M. J., Homolka, S., Niemann, S., and Rohde, K. H. (2010) *Cell Host Microbe* **8**, 68-76
4. Boshoff, H. I., and Barry, C. E., 3rd. (2005) *Nat Rev Microbiol* **3**, 70-80
5. Russell, D. G. (2007) *Nat Rev Microbiol* **5**, 39-47
6. Pandey, A. K., and Sasseti, C. M. (2008) *Proc Natl Acad Sci U S A* **105**, 4376-4380
7. Singh, A., Crossman, D. K., Mai, D., Guidry, L., Voskuil, M. I., Renfrow, M. B., and Steyn, A. J. (2009) *PLoS Pathog* **5**, e1000545
8. Griffin, J. E., Pandey, A. K., Gilmore, S. A., Mizrahi, V., McKinney, J. D., Bertozzi, C. R., and Sasseti, C. M. (2012) *Chem Biol* **19**, 218-227
9. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) *Nature* **393**, 537-544
10. Chen, L., and Vitkup, D. (2007) *Trends Biotechnol* **25**, 343-348
11. Hotter, G. S., and Collins, D. M. (2011) *Vet Microbiol* **151**, 91-98
12. Mougous, J. D., Leavell, M. D., Senaratne, R. H., Leigh, C. D., Williams, S. J., Riley, L. W., Leary, J. A., and Bertozzi, C. R. (2002) *Proc Natl Acad Sci U S A* **99**, 17037-17042
13. Mougous, J. D., Senaratne, R. H., Petzold, C. J., Jain, M., Lee, D. H., Schelle, M. W., Leavell, M. D., Cox, J. S., Leary, J. A., Riley, L. W., and Bertozzi, C. R. (2006) *Proc Natl Acad Sci U S A* **103**, 4258-4263
14. Holsclaw, C. M., Sogi, K. M., Gilmore, S. A., Schelle, M. W., Leavell, M. D., Bertozzi, C. R., and Leary, J. A. (2008) *ACS Chem Biol* **3**, 619-624
15. Glickman, M. S., Cox, J. S., and Jacobs, W. R., Jr. (2000) *Mol Cell* **5**, 717-727
16. Stover, C. K., de la Cruz, V. F., Fuerst, T. R., Burlein, J. E., Benson, L. A., Bennett, L. T., Bansal, G. P., Young, J. F., Lee, M. H., Hatfull, G. F., and et al. (1991) *Nature* **351**, 456-460

17. Hatfull, G. F., and Jacobs Jr, W. R., Jr. (eds). (2000) *Molecular Genetics of Mycobacteria*, ASM Press, Washington, D.C.
18. Akiyama, T., Kuwahara, T., Ando, T., Tada, T., and Kirikae, T. (2012) *Unpublished, Submitted to Bioproject #PRJDB66*
19. Mougous, J. D., Green, R. E., Williams, S. J., Brenner, S. E., and Bertozzi, C. R. (2002) *Chem Biol* **9**, 767-776
20. Johnston, J. B., Kells, P. M., Podust, L. M., and Ortiz de Montellano, P. R. (2009) *Proc Natl Acad Sci U S A* **106**, 20687-20692
21. Sassetti, C. M., Boyd, D. H., and Rubin, E. J. (2001) *Proc Natl Acad Sci U S A* **98**, 12712-12717
22. Sassetti, C. M., Boyd, D. H., and Rubin, E. J. (2003) *Mol Microbiol* **48**, 77-84
23. Deshayes, C., Bach, H., Euphrasie, D., Attarian, R., Coureuil, M., Sougakoff, W., Laval, F., Av-Gay, Y., Daffe, M., Etienne, G., and Reyrat, J. M. (2010) *Mol Microbiol* **78**, 989-1003

Appendix

Table 3-1: Bacterial strains used in this chapter

Strains	Genotype	Source
<i>M. smeg</i> mc ² 155	Wild type	
<i>M. smeg</i> mc ² 155 <i>rv2269c</i>	pKMS101; Kn ^r , contains <i>Rv2269c</i>	This study
<i>M. smeg</i> mc ² 155 <i>cyp128</i>	pKMS102; Kn ^r , contains <i>cyp128</i>	This study
<i>M. smeg</i> mc ² 155 <i>stf3</i>	pKMS103; Kn ^r , contains <i>stf3</i>	This study
<i>M. smeg</i> mc ² 155 <i>stf3</i> operon	pKMS104; Kn ^r , contains <i>rv2269c</i> , <i>cyp128</i> , <i>stf3</i>	This study
<i>M. smeg</i> mc ² 155 <i>cyp128</i> , <i>stf3</i>	pKMS105; Kn ^r , contains <i>cyp128</i> and <i>stf3</i>	This study
Mtb H37Rv	Wild type	
Mtb H37Rv Δ <i>cyp128</i>	Hyg ^r , hyg cassette disrupting <i>cyp128</i>	This study
Mtb H37Rv Δ <i>cyp128::cyp128</i>	Hyg ^r , Kan ^r , complemented strain of Δ <i>cyp128</i>	This study
Mtb H37Rv Δ <i>stf3</i>	Hyg ^r , <i>stf3</i> interrupted by hyg resistance cassette	Ref (1)
Mtb H37Rv Δ <i>stf3::stf3</i>	Hyg ^r , Kan ^r , complement with <i>stf3</i> under the glutamine synthase promoter modified pMV306 (3)	Ref (1)
Mtb H37Rv Δ <i>rv2269c</i>	Hyg ^r , hyg cassette disrupting <i>rv2269c</i>	This study

Table 3-2: Plasmids used in this chapter

Plasmids		
Reference name	Description	Source
pMV261	Kn ^r , pAL5000 origin, ColE1 origin, multiple cloning site, Phsp60 promoter	Ref (3)
pMV306	Kn ^r , A derivative of pMV261 lacking the Phsp60 promoter	Ref (3)
pKMS101	pMV261 derivative; contains <i>rv2269c</i>	This study
pKMS102	pMV261 derivative; contains <i>cyp128</i>	This study
pKMS103	pMV261 derivative; contains <i>stf3</i>	This study
pKMS104	pMV261 derivative; contains <i>rv2269c</i> , <i>cyp128</i> , and <i>stf3</i>	This study
pKMS105	pMV261 derivative; contains <i>cyp128</i> and <i>stf3</i>	This study
pKMS110	Plasmid used for <i>cyp128</i> disruption with hyg cassette	This study
pKMS108	Plasmid used for <i>rv2269c</i> disruption with hyg cassette	This study
pKMS133	Kn ^r , a derivative of pMV306 encoding <i>cyp128</i> with a native promoter (upstream 1 kb of the first gene in the putative operon).	This study

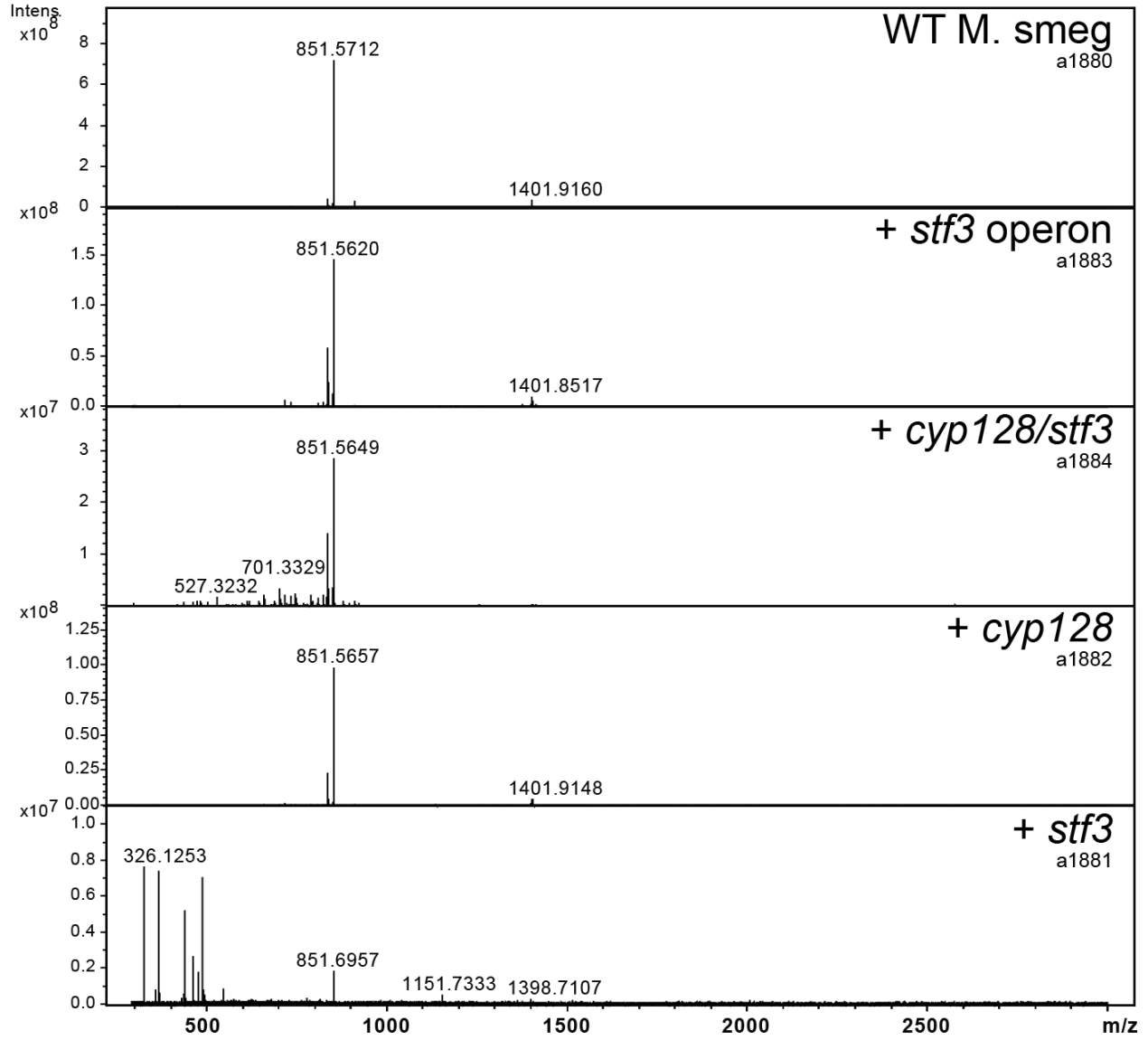
Table 3-3: Primers used in this chapter. Restriction enzymes sequences are in bold and enzyme in parentheses.

Primer name	Sequence	Description
okms102	cacttcgcaat ggcca acgatgcgaccccttagcg	5' pKMS101, pKMS104 (MscI)
okms109	actgttctacgcctctct gaatcgat agggtcatga	3' pKMS101 (ClaI)
okms100	ccagegtcagaaacaatgtg	5' pKMS102, pKMS105
okms101	cgtgacaacgggctgcttag	3' pKMS102
okms103	cacttcgcaat ggcca acgatgcgaccccttagcg	5' pKMS103 (MscI) 3' pKMS103, pKMS104, pKMS105
okms112	actgttctacgcctctct gaatcgat gtcg	(ClaI)
okms126	ccgtacgt ctc gaggtgagcaactgaccg	pKMS110 KO 5' cyp128 (XhoI)
okms127	caccat gaagctt ggtcagaccaacgtcgggc	pKMS110 KO 5' cyp128 (HindIII)
okms128	cc gggtacc gaatagaggtggtcgagc	pKMS110 KO 3' cyp128 (KpnI)
okms129	cggtact taagc gaacgtcggttgttc	pKMS110 KO 3' cyp128 (AflII)
okms213	c gcggtacc gtggccaacgatgcgcg	5' pKMS133 (KpnI)
okms196	gtcgacat cgat gcacggcgaagcggttac	3' pKMS133 (ClaI)
okms122	gtacgt ctc gagttgtaggccctcggccagcg	pKMS109 KO 5' rv2269 (XhoI)
okms123	gatcc agatatc aactgggccgactgtgtagg	pKMS109 KO 5' rv2269 (EcoRV)
okms124	gacaggact ctagac gcaattattgcgatgcccc	pKMS109 KO 3' rv2269 (XbaI)
okms125	gactagag gggtacc agcagtgtctcatag	pKMS109 KO 3' rv2269 (KpnI)

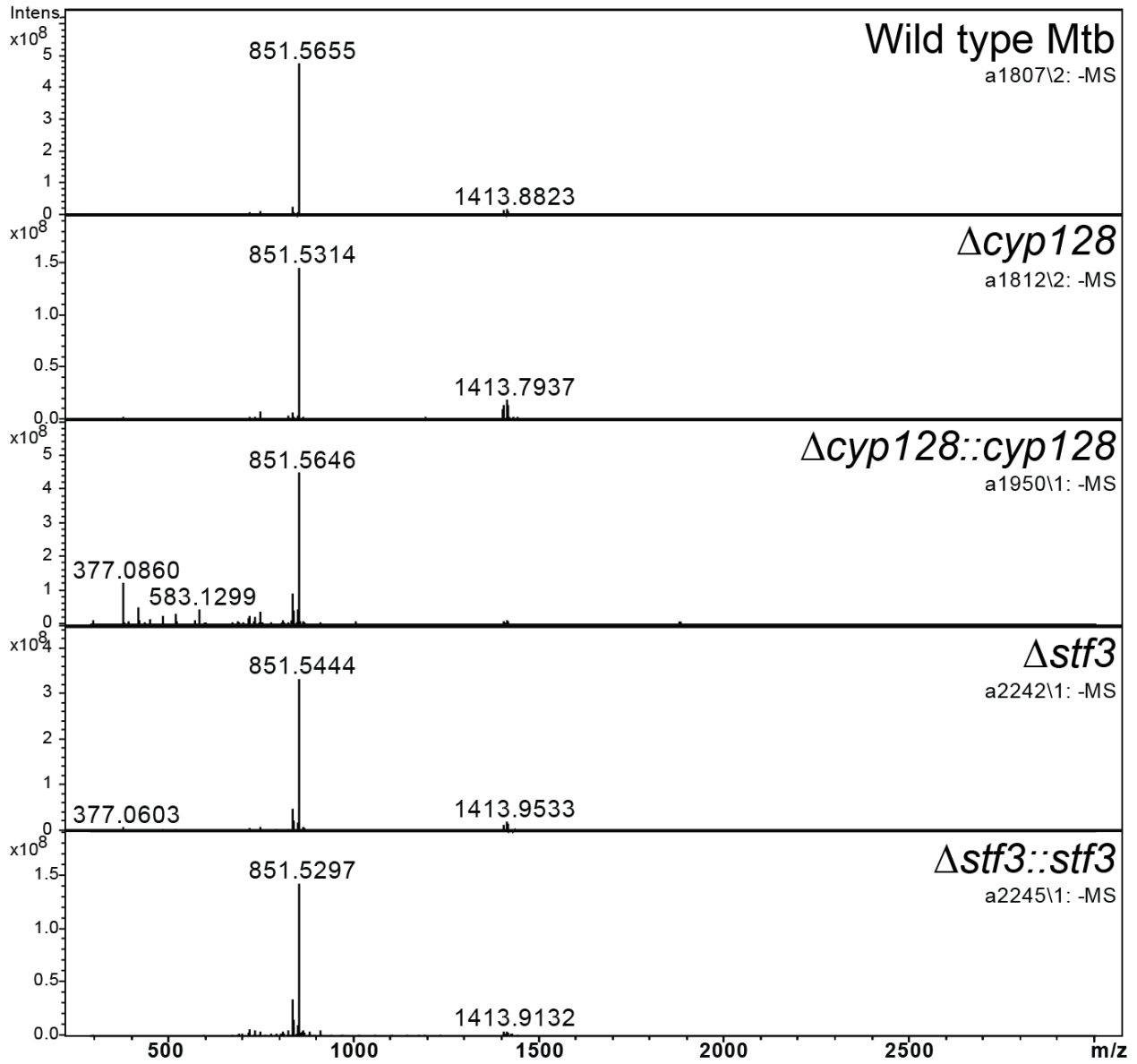
References

1. Mougous, J. D., Senaratne, R. H., Petzold, C. J., Jain, M., Lee, D. H., Schelle, M. W., Leavell, M. D., Cox, J. S., Leary, J. A., Riley, L. W., and Bertozzi, C. R. (2006) *Proc Natl Acad Sci U S A* **103**, 4258-4263
2. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) *Nature* **393**, 537-544
3. Stover, C. K., de la Cruz, V. F., Fuerst, T. R., Burlein, J. E., Benson, L. A., Bennett, L. T., Bansal, G. P., Young, J. F., Lee, M. H., Hatfull, G. F., and et al. (1991) *Nature* **351**, 456-460

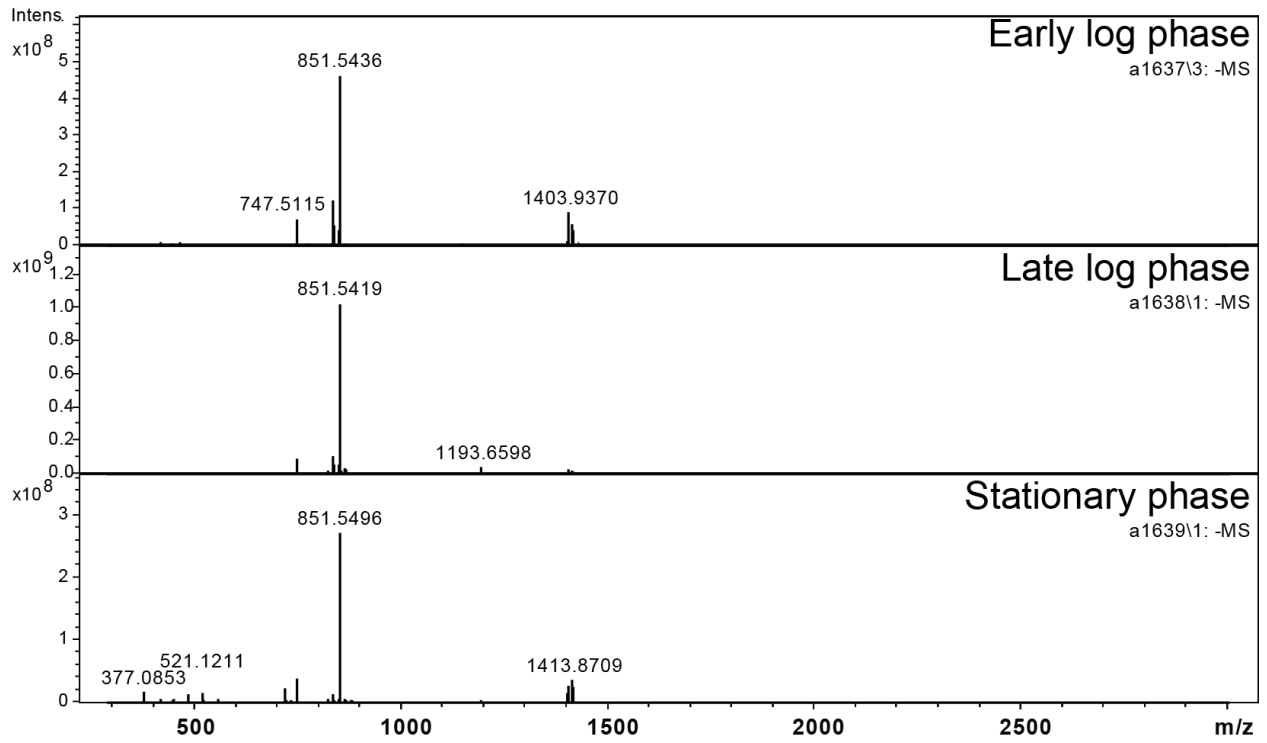
Full spectra of *M. smeg* total lipid extracts (Figure 3-2)



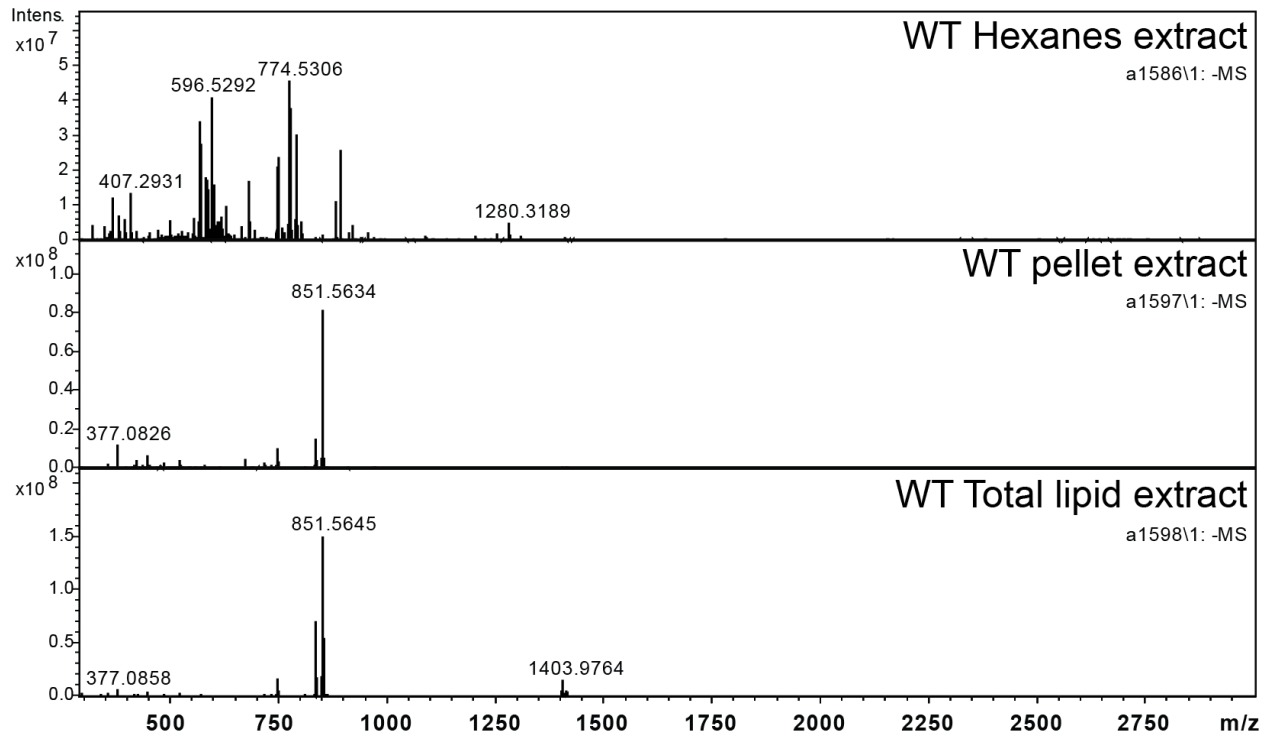
Full spectra of Mtb total lipid extracts (Figure 3-3)



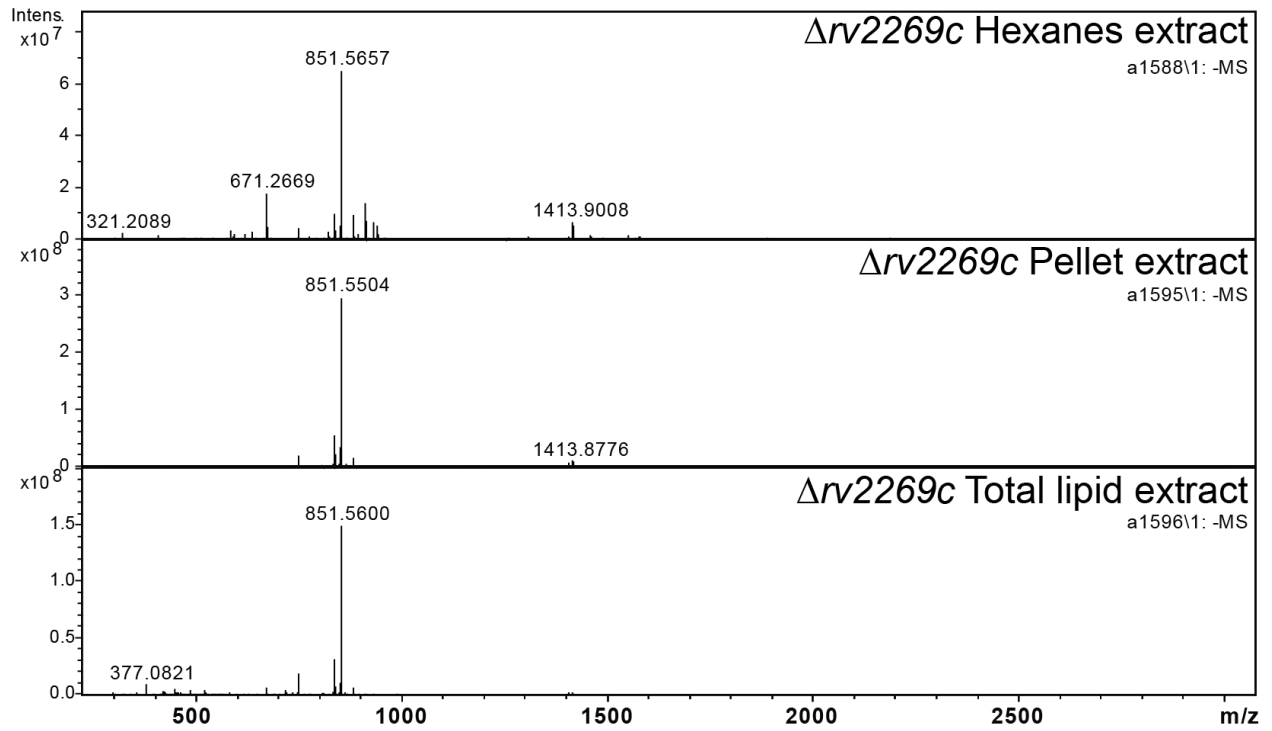
Full spectra of WT Mtb at different phases of growth (Figure 3-4)



Full spectra of WT cell wall fractionation extractions (Figure 3-5A)



Full spectra of $\Delta rv2269c$ cell wall fractionation extracts (Figure 3-5B)



Chapter 4: Using Mtb models of infection to investigate the function of S881

Introduction

In order to facilitate the study of Mtb, a number of *in vitro* models of infection have been developed over the last twenty years to study the processes necessary for Mtb to enter and exit a non-replicating state (1-3). The first and most widely used model of infection is the Wayne hypoxia model. In 1996, Wayne and his colleagues discovered that reducing the levels of oxygen in the Mtb culture caused the bacteria to enter a non-replicating state and were refractory to all antimycobacterial drugs (4). Since then, a number of other models have been developed that attempt to isolate and study one aspect of the granuloma including microaerobic/hypoxia, starvation or cell based infections to study the more complex host-pathogen interactions. More recently, models have been developed that attempt to mimic multiple aspects of the environment within a granuloma using continuous cultures systems (5).

Through the use of microarray data analysis, we are beginning to understand how Mtb adapts to *in vitro* models of infection. For example, during the transition from aerobic to microaerobic conditions, Mtb transitions its terminal cytochrome (cyt) oxidases from the more efficient cyt *bc₁* and *aa₃* cyt c oxidase complexes to the less efficient *bd* oxidase (6). This is supported in mutant studies showing that the *bd* oxidase is required for viability under microaerobic conditions (7). As the cultures become hypoxic, some studies suggest that NO₃⁻ becomes the primary terminal electron acceptor as a result of the nitrate transporter (*narK*) being up regulated (8).

Because Mtb must adapt to changes in the oxygen tension and variability in nutrient levels, the electron transport chain in Mtb is highly modular; it can utilize a range of enzymes to fulfill each step. Of all the components in the electron transport chain, menaquinone is the only small molecule required under all conditions. Therefore all of the other components, mostly proteins, can be regulated by changing transcription and degradation rates. As a small molecule, menaquinone requires at least eight enzymes for its biosynthesis. Therefore, traditional protein regulation of menaquinone levels would be very inefficient and slow. Currently, the method of menaquinone regulation is unknown. We hypothesize that the conversion of menaquinone to S881 is a method of removing menaquinone from the electron transport chain.

Menaquinone functions by transporting electrons and protons from the cytoplasmic side of the membrane to the periplasmic side. Sulfation of menaquinone would make flipping in the membrane energetically unfavorable and may be the signal for transport to the outer cell envelope removing menaquinone from the electron transport chain. One study has characterized the changes in the menaquinone pool in Mtb during adaptation to changes in the electron transport chain (9). Honaker et al. used a hypoxia model or a strong cyt c reductant, ascorbate, to alter the electron transport chain. The ratio of menaquinone-9(H₂) (MK-9(H₂)) to menaquinone-9 (MK-9) changes in Mtb from hypoxic cultures or when treated with ascorbate. This study does not address how either form of menaquinone is removed from the system, but it is the only study that has identified a change in the menaquinone pool. S881 may be the trigger for menaquinone degradation and/or transport to the outer cell envelope.

Hypoxia studies show a slow, sequential decrease in the respiration and metabolism of Mtb. Conversely, the Loebel model of total starvation causes a rapid decrease in both the respiration and metabolism of Mtb (10-12). Similar to hypoxic studies, the nutrient starved bacteria are also resistant to traditional front line Tuberculosis drugs. Microarray data showed

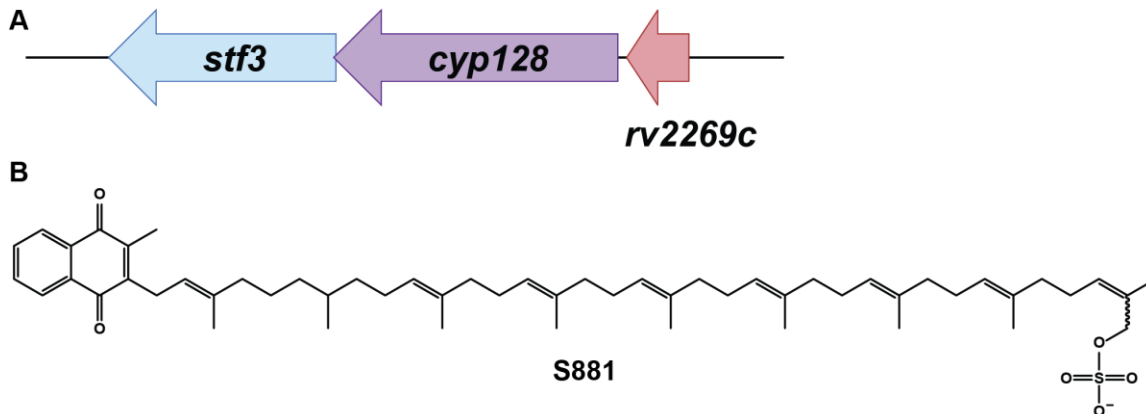


Figure 4-1: (A) The *stf3* operon. (B) The chemical structure of S881.

genes involved respiration including ATP synthase, NADH dehydrogenase and lipid biosynthesis are down regulated after 96 hours of starvation. Total starvation conditions show a small but significant induction of S881 biosynthetic genes, unlike studies of hypoxia.

We plan to use a variety of infection models to study the function of S881. To focus our initial studies, we mined the available microarray literature to identify conditions that caused the largest change in S881 gene transcription. Overall the regulation of S881 biosynthetic genes was minimal across all conditions as analyzed by microarray. Our initial studies focused on the starvation model because it showed the largest change in S881 biosynthetic genes by microarray than any other *in vitro* model. After determining that the S881 mutants did not show a phenotype in the starvation model, we used the macrophage infection model to determine if S881 function is a result of host-pathogen interactions. For the macrophage infection model, both human and murine macrophages were used and the S881 mutants showed no phenotype in either of the cell lines as measured by colony forming units (CFU). Next, we took a chemical genetic approach by screening around 45 compounds that target either Mtb cell wall biosynthesis or each step of the electron transport chain for Mtb during *in vitro* growth. While the mutants did not show a difference in survival with treatment of these inhibitors, many of the inhibitors altered the levels of S881.

Results

S881 mutants showed no phenotype in the starvation model of infection

The S881 genes, *cyp128* and *stf3*, were shown to be induced under conditions of starvation as shown by microarray analysis (11). To confirm that the S881 genes were induced, mRNA was isolated from wild type (WT) Mtb after 96 hours of incubation in phosphate buffered saline with tyloxopol (PBST) or 7H9 media supplemented with OADC and tween-80. Tyloxopol is a non-hydrolyzable detergent used to prevent clumping of the bacteria and cannot be used as a carbon source for Mtb. Transcript levels of *cyp128* and *stf3* were up regulated over cultures grown in rich media (Figure 4-2D). The up regulation of the genes was not dependent on whether the cultures were agitated or left standing during the 96 hour experiment.

To elucidate the role of S881 in relation to the electron transport chain, we measured two parameters of the electron transport chain: ability to maintain reducing equivalents and ATP

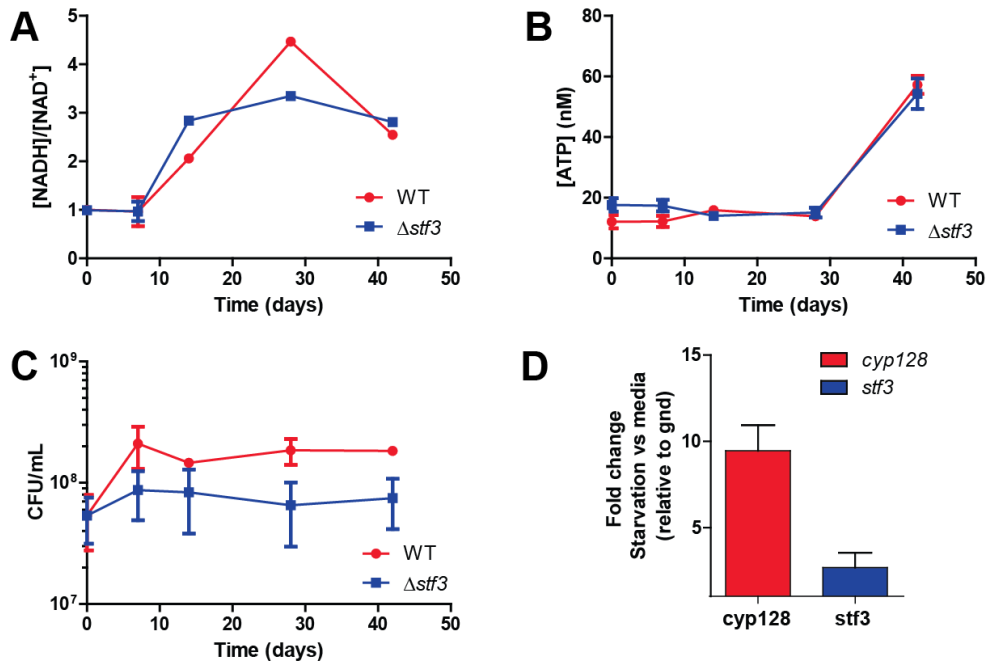


Figure 4-2: Starvation of WT Mtb and $\Delta stf3$. (A) Ratio of [NADH]/[NAD⁺]. (B) Levels of ATP. (C) Survival of WT and $\Delta stf3$ during starvation. All data from a single experiment and each experiment was done in biological replicates. The difference in CFU was not reproducible across experiments. (D) Fold change of each S881 gene relative to cultures grown in rich media. Transcripts were normalized to *gnd*.

production. During starvation in WT Mtb, the levels of ATP decrease initially and are maintained at a lower level during starvation (12). To assess survival, we enumerated the colony forming units (CFU) at 5 time points up to 42 days. Both WT and the S881 mutants showed comparable levels of ATP, reducing equivalents and showed no difference in survival by CFU (Figure 4-2). The increase in ATP content at 42 days may be due to an increase in bacterial sample as the bacteria became clumpier at the end of the experiment.

S881 mutants show no phenotype in the macrophage infection model

In addition to studying the effects S881 has on the electron transport chain of Mtb, its presence in the outer cell envelope of Mtb suggests a role in host-pathogen interactions. Another sulfated metabolite Sulfolipid-1 (SL-1) has been implicated in host-pathogen interactions (13,14), including modulating phagosome maturation (15). To further study the role S881 could be playing in the context of infection, we used the macrophage infection model. Because Mtb is strictly a human pathogen, we wanted to compare the strains survival in both human (THP-1) and murine macrophages (RAW264.7). The macrophages were infected at a multiplicity of infection (MOI) of 0.5 (one bacterium for two macrophages). Bacterial growth was measured by enumerating CFUs over six days. Figure 4-3 shows that the S881 mutants had no difference in survival in either THP-1 cells or RAW264.7 cells.

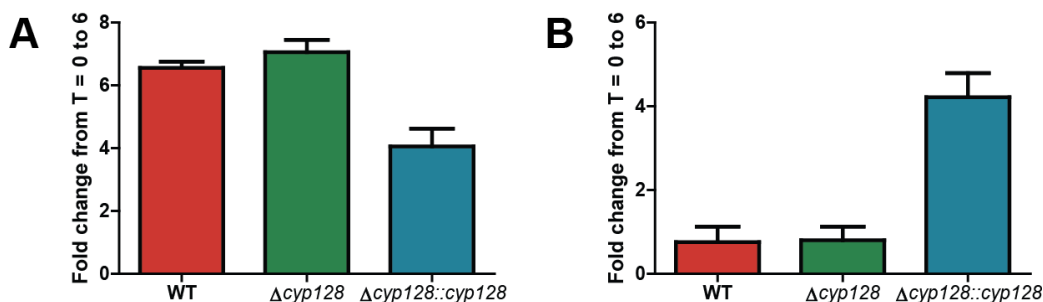


Figure 4-3: Survival of $\Delta cyp128$ mutants in macrophage infections expressed as fold change from day 0 to day 6 in (A) THP-1 cells or (B) RAW264.7 cells.

S881 mutants do not show a defect in cell wall integrity

As shown previously, S881 is found in the outer cell envelope of Mtb. While not as abundant as SL-1 (16,17), S881 may have a role in maintenance or integrity of the cell envelope. Chemical oxidative stress (H_2O_2 , $HClO_3$, NO), detergent stress (SDS), and cell wall biosynthesis inhibitors (isoniazid, ethambutol, ethionamide) were screened for MIC in 96 well plates with the Mtb strains. The S881 mutants showed no difference in survival from WT as determined by OD_{600} in 96 well plates (Table 4-1).

Chemical stress	MIC	Cell wall inhibitors	MIC
H_2O_2	110 mM	INH	0.06 $\mu g/ml$
$NaNO_3$, pH 5.5	5 mM	ETA	5 μM
SDS	0.025%	ETH	6 μM

Table 4-1: Minimum inhibitory concentrations for WT, S881 mutants and complements treated with oxidative stress and inhibitors of cell wall biosynthesis. H_2O_2 hydrogen peroxide; HNO_3 , SDS sodium dodecyl sulfate, INH isoniazid, ETA ethionamide, ETH ethambutol.

S881 mutants showed no phenotype when treated with inhibitors of the electron transport chain

After establishing that S881 is not involved in the maintenance of the Mtb cell wall, we used a chemical genetic approach to elucidate the role S881 may play in the electron transport chain. To do this, we screened 35 small molecule compounds that have known inhibitory effects on the enzymes in the electron transport chain (Figure 4-4). We selected commercially available compounds that targeted each step in the electron transport chain to determine how the presence or absence of S881 would change bacterial viability. Each compound was serially diluted and inoculated with WT H37Rv, the S881 mutants or the complements. Cells were resuspended and fixed with formalin before growth was determined by optical density at 600 nm. Surprisingly, compounds known to alter levels of menaquinone showed no phenotype (full table of compounds tested in Appendix). Some inhibitors, such as metronidazole, showed no growth

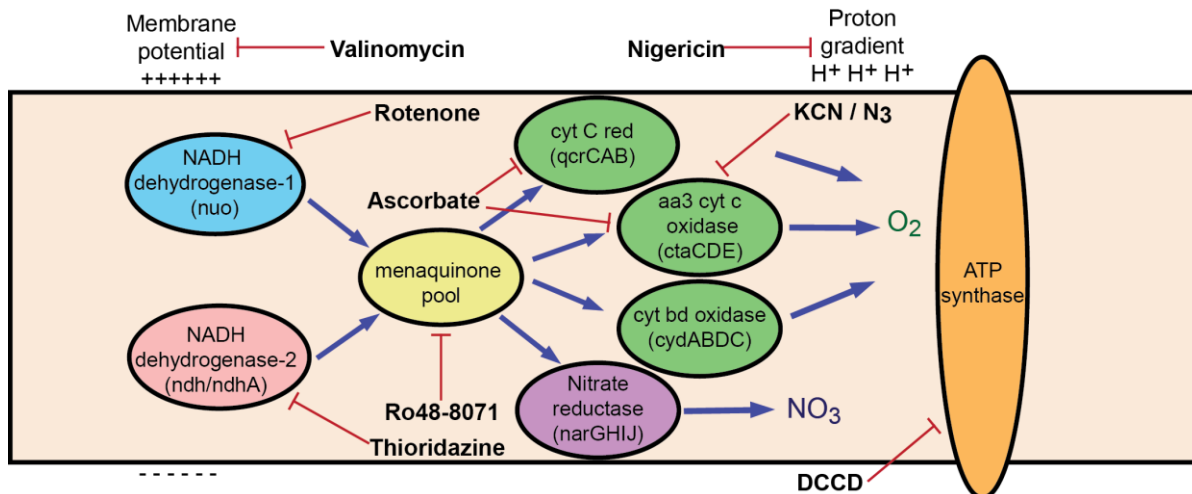


Figure 4-4: A simplified electron transport chain from mycobacteria with a selection of inhibitors tested.

inhibitory effects even at high doses due to lack of activity in aerobically grown bacteria. No reproducible phenotype was seen between S881 mutants and WT for any compound screens in this assay.

Menaquinone has been shown to be involved in the active transport of aminoglycosides in *Bacillus subtilis* and *Staphylococcus aureus* (18). A series of aminoglycosides (amikacin, gentamicin, and streptomycin) were tested against WT *Mtb* and S881 mutants to determine if S881 was involved in a similar pathway in *Mtb*. None showed a difference in survival as determined by OD₆₀₀ (data not shown). All compounds tested gave MIC values that corresponded with values reported in the literature.

S881 mutants show no phenotype when treated with human antimicrobial peptides

In addition to peptide based antibiotic, we tested the positively charged human antimicrobial peptide LL-37. LL-37 has previously been shown to have a phenotype in mutants lacking cell surface Sulfolipid-1 (19). The human antimicrobial peptide did not cause a difference in strain survival in the assays (Figure 4-5) suggesting that a loss of S881 does not change the binding of antimicrobial peptides to *Mtb*.

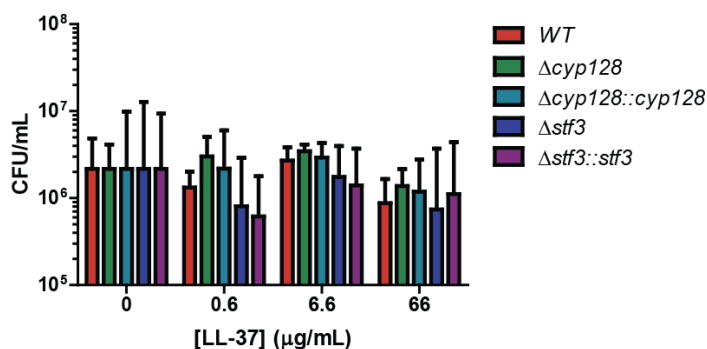


Figure 4-5: Treatment of S881 mutants with the human antimicrobial peptide LL-37. Survival of strains in CFU/ml after treatment for 3 days.

Treatment of Mtb with inhibitors of respiration may cause a change in S881 levels

To determine if inhibitors of respiration effect levels of S881, WT Mtb was treated with Valinomycin (causes dissipation of the membrane potential), Ro48-8071 (menaquinone biosynthesis inhibitor), Thioridazine (TRZ; NADH dehydrogenase-2 inhibitor) or INH (20). Levels of S881 were measured by the amount of ³⁵S-sulfate incorporation into S881. Initial results suggest that the levels of S881, normalized to total ³⁵S incorporation, in WT Mtb are altered with treatment of respiration inhibitors (Figure 4-6).

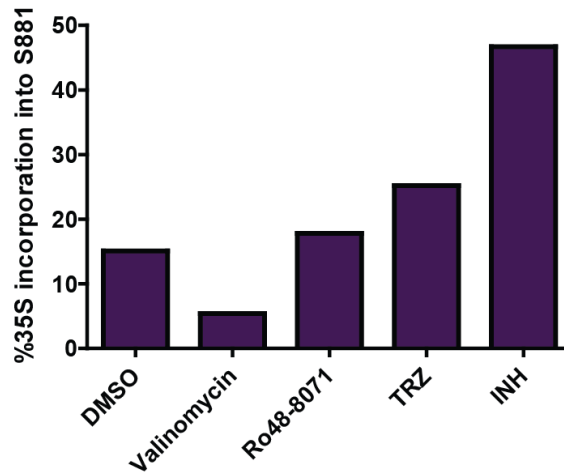


Figure 4-6: Levels of S881 from WT Mtb after treatment with inhibitors or respiration. Data shown is from a single experiment.

Discussion

Despite the wide range of assays performed, we were unable to identify a function or role of S881 that would help to understand the unusual phenotype in mice. In the starvation experiment, it is possible that the effects of S881 were lost in the large transcriptional change as shown by Betts et al. (11). Further studies using carbon starvation or sulfate starvation may be better suited to determining a phenotype with the S881 mutants. Alternatively, using a hypoxia model may be more effective for determining the role of S881 during infection.

The lack of a phenotype in the macrophage infection model was surprising. All studies were done using unactivated macrophages. It is possible that the activation of the macrophages prior to infection would cause different results. If S881 is involved in the host immune response, one possibility is that S881 alters a signaling pathway in macrophages but does not alter bacterial survival. Further studies into the cytokine profile of macrophages infected with S881 mutants are needed to determine if S881 causes a difference in macrophage signaling or activation of the innate immune response. Additionally, detailed cytokine analysis and cell composition studies of mouse lungs infected with S881 mutants compared to WT would be needed to identify a role of S881 involved in modulating the innate immune response.

The use of small molecule inhibitors of the Mtb cell wall and respiration, allowed a more targeted approach to studying S881's function. We have clearly established that S881 is not involved in maintenance of cell wall integrity. Recent studies from the Dutton and coworkers has identified a Vitamin K epoxide reductase (VKOR) homologue in Mtb that utilizes MK as a cofactor in a system that aids in disulfide bond formation in secreted proteins (21). They have shown that deletion of the VKOR homologue or inhibition with a small molecule leads to growth attenuation (21,22). It is unlikely S881 is involved in this pathway, since a lack of S881 does not change the integrity of the cell wall or bacterial survival.

We were very disappointed that no phenotype was discovered from treatment of the Mtb strains with inhibitors of the electron transport chain. Despite no change in survival, S881 levels do change with similar treatments suggesting a role for S881 in adapting to these stresses. It may be possible that the assay used to screen compounds was not sensitive enough to identify differences in survival between strains. Future studies will focus on using isolated membranes from bacteria with or without S881 to measure oxygen consumption and rate of electron transfer. In addition, measurements of oxygen consumption from isolated membranes and from whole bacteria may be sensitive enough to find differences in the Mtb strains.

In conclusion, we have yet to identify a function for S881 to explain the results from the mouse model of infection. More sensitive and targeted experiments are needed to elucidate the biology behind S881.

Materials and Methods

Bacterial strains and growth media

Growth media 7H9 and 7H11 and OADC supplement were obtained from BD Biosciences. All experiments in Mtb were done in H37Rv strain (gift from the Riley Lab, UC Berkeley). Mtb was grown in liquid 7H9 or solid 7H11 with 0.5% glycerol, 0.05% Tween-80, 10% OADC and the 7H11 plates contained 10 µg/mL cyclohexamide. For starvation studies, 7H9 media without any additive except 0.025% Tyloxopol, a non-hydrolyzable detergent was used (referred to as Starvation media). All strains used in the following studies were discussed in Chapter 3. THP-1 (ATCC TIB-202) cells were cultured in RPMI media (Hyclone) supplemented with 10% FBS (Omega Scientific), 10 mM HEPES (Sigma), 0.05 mM β-mercaptoethanol (Sigma) and penicillin/streptomycin (Hyclone). RAW264.7 cells (ATCC TIB-71) were cultured in DMEM media (Hyclone) supplemented with 10% FBS and penicillin/streptomycin.

Survival of S881 mutants during total starvation

Cultures were grown to mid-log phase in 50 mL media in roller bottles. Clumps were removed by centrifugation and washed three times in starvation media. Strains were diluted to OD₆₀₀ 0.5 in 50 mL in roller bottles and incubated at 37 °C. Aliquots were removed from the cultures at the indicated time points and serial dilutions were plated on 7H11 plates for CFU enumeration. For ATP analysis, 1 mL aliquots were transferred to screw top tubes and heat inactivated for 30 min. For NADH/NAD⁺ samples, 1 mL was transferred to a screw cap tube and pelleted. The pellet was resuspended in 0.3 mL 0.2 M NaOH (NADH) or 0.2 M HCl (NAD⁺). The sample was heat inactivated for 20 min. For qPCR, WT Mtb was grown in 10 mL

of media in 50 mL conical tubes or shaking flasks. Cultures were pelleted after 4 days and flash frozen until further use.

RT-PCR

RNA isolation and qPCR were done as previously described (23). Briefly, cell pellets were resuspended in TRIzol reagent (Invitrogen), disrupted by bead-beating, and centrifuged to clarify the cellular lysates, which were subsequently transferred to tubes containing Heavy Phase Lock Gel (5 PRIME) and chloroform. Each sample was inverted for 2 min prior to centrifugation and removal of the aqueous layer, which was then precipitated by treatment with isopropanol and high salt. RNA purification and DNase treatment were performed using the RNeasy kit (Qiagen) according to the manufacturer's instructions. Before reverse transcription, 2 µg of RNA were treated with DNase (Invitrogen) a second time, then reverse transcribed using Superscript III (Invitrogen). The cDNA was diluted to 10 ng/µL. A standard curve was made with 10 fold dilutions from 10 – 0.001 ng/µL. qPCR was performed using DyNAmo HS SYBRgreen (Finzymes) following the manufactures instructions using 1 ng/µL cDNA with primers specific for *cyp128* and *stf3* (Appendix Table 5-3). qPCR and analysis was done on a iCycler iQ5 PCR thermal cycler (Bio-Rad).

ATP measurements for Mtb strains

ATP levels were determined using BacTiter Glo assay (Promega). Equal volumes of sample and BacTiterGlo reagent were added to a 96 well plate containing a standard curve ([ATP] = 10-100 nM) and read on a SpetraMax M5 microplate reader (Molecular Devices). Each well was integrated over 50 reads.

NADH/ NAD⁺ cycling assay

Measurement of the NADH/NAD⁺ ratio was done as previously described (24,25). Briefly, 250 µL of extract, 250 µL of a neutralizing acid (0.1 M HCl for NADH or 0.1 M NaOH for NAD⁺), 100 µL of 16.6 mM PES, 100 µL 4.2 mM MTT, 100 µL of absolute ethanol, and 100 µL of 40 mM EDTA (pH 8.0) were incubated for 10 min at RT. ADHII (500 U/mL; 20 µL) in 0.1 M bicine (pH 8.0) was added to the reaction mixture to begin the assay. Increase in absorbance (A_{570}) was recorded for 10 min. The rate of MTT reduction is proportional to the concentration of the cofactors. Cofactor standards were used to calibrate the assay. UV/Vis/NIR spectra were acquired on a Spectramax 190 absorbance plate reader (Molecular Devices). Reactions lacking extract or ADHII were used as negative controls.

Macrophage infections

THP-1 cells were seeded at 4×10^5 cells/well in 12 well plates and differentiated into adherent, macrophage-like cells with 10 ng/mL of phorbol myristate acetate (PMA; Sigma) for three days prior to infection. One day prior to infection, RAW cells were seeded at 1×10^5 cells/well in 12 well plates. Bacteria were grown to mid-log phase, pelleted by centrifugation at low speed to remove clumps and washed twice in PBS before determining OD₆₀₀ (OD₆₀₀ of 1.0 = 3×10^8 bacteria/mL). THP-1 and RAW cells were infected with the indicated strain of Mtb at a

multiplicity of infection (MOI) of 0.5 using infection media: (THP-1) RPMI with all additives listed above, except 10% heat killed horse serum (Sigma) instead of FBS or (RAW) DMEM with 10% heat killed horse serum instead of FBS and no antibiotic. After 4 h, extracellular bacteria were removed by washing cells three times with warm PBS. To determine viable bacteria, CFU counts of intracellular bacteria were enumerated at the indicated time points by lysing adherent cells in PBS + 0.1% TritonX-100 and plating serial dilutions on 7H11 agar. Colonies were counted weeks 2 through 4.

Minimum inhibitor concentration (MIC) determination using chemical inhibitors

Mtb was grown to mid-log phase in 50 mL media in roller bottles. Clumps were removed by centrifugation and diluted to an OD₆₀₀ of 0.2. Two fold serial dilutions of each chemical inhibitor were set up in 96 well plates from 4x literature MIC to media only in 90 µL. Each well was inoculated with 10 µL of the designated strain (in triplicate). Sterile PBS was added to remaining wells. Plates were parafilmmed and placed into a sealed Tupperware container containing damp towels to prevent evaporation. Plates were incubated for 7 days at 37 °C. Plates were fixed with 100 µL 10 % formalin in buffered PBS and read on a VersaMax UV/Vis plate reader (Molecular devices).

Antimicrobial peptide sensitivity

Sensitivity of Mtb to antimicrobial peptides was assayed as previously described (19,26). Briefly, Mtb was grown to mid-log phase, washed twice in PBS, and pelleted by centrifugation at low speed to remove clumps. 2×10^6 bacteria of each strain were exposed to the indicated concentration of LL-37 (37 amino acids; AnaSpec) dissolved in RPMI/water (1:4, v/v) at 37 °C. After three days, bacteria were plated on solid agar to enumerate the number of viable bacteria by CFU counts.

S881 quantification by ³⁵S-sulfate incorporation

WT Mtb was grown in 7H9 to late-log phase. Clumps were removed by centrifugation and washed in Sulfate free Sauton media. Ten mL cultures of WT or $\Delta cyp128$ were put in to inkwells at an OD₆₀₀ of 0.6. Drugs and 100 µCi ³⁵S-sulfate (PerkinElmer) were added to the inkwells and they were incubated for 3 days at 37 °C. WT Mtb was treated with ½ MIC of each compound: 0.25 µM Valinomycin, 5 µM Ro48-8071, 4 µM Thioridazine, 0.5 µg/ml isoniazid (20). Bacterial pellets were extracted in 1:1 chloroform/methanol as described previously. Solvent was removed by evaporation and the extracts resuspended at 1/10 or 1/20 the original volume of 1:1 chloroform/methanol. An equal volume of each fraction was spotted on silica plates (HPTLC Silica Gel 60, EMD Chemicals) and developed in 60:12:1 chloroform/methanol/water. Plates were analyzed by phosphorimaging (GE Biosciences Typhoon).

References

1. Russell, D. G. (2007) *Nat Rev Microbiol* **5**, 39-47
2. Barry, C. E., 3rd, Boshoff, H. I., Dartois, V., Dick, T., Ehrt, S., Flynn, J., Schnappinger, D., Wilkinson, R. J., and Young, D. (2009) *Nat Rev Microbiol* **7**, 845-855
3. Davis, J. M., and Ramakrishnan, L. (2009) *Cell* **136**, 37-49
4. Wayne, L. G., and Hayes, L. G. (1996) *Infect Immun* **64**, 2062-2069
5. Bacon, J., Hatch, K. A., and Allnut, J. (2010) *Methods Mol Biol* **642**, 123-140
6. Matsoso, L. G., Kana, B. D., Crellin, P. K., Lea-Smith, D. J., Pelosi, A., Powell, D., Dawes, S. S., Rubin, H., Coppel, R. L., and Mizrahi, V. (2005) *J Bacteriol* **187**, 6300-6308
7. Kana, B. D., Weinstein, E. A., Avarbock, D., Dawes, S. S., Rubin, H., and Mizrahi, V. (2001) *J Bacteriol* **183**, 7076-7086
8. Wayne, L. G., and Hayes, L. G. (1998) *Tuber Lung Dis* **79**, 127-132
9. Honaker, R. W., Dhiman, R. K., Narayanasamy, P., Crick, D. C., and Voskuil, M. I. (2010) *J Bacteriol* **192**, 6447-6455
10. Loebel, R. O., Shorr, E., and Richardson, H. B. (1933) *J Bacteriol* **26**, 139-166
11. Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A., and Duncan, K. (2002) *Mol Microbiol* **43**, 717-731
12. Gengenbacher, M., Rao, S. P., Pethe, K., and Dick, T. (2010) *Microbiology* **156**, 81-87
13. Goren, M. B., Brokl, O., and Schaefer, W. B. (1974) *Infect Immun* **9**, 142-149
14. Kato, M., and Goren, M. B. (1974) *Infect Immun* **10**, 733-741
15. Goren, M. B., D'Arcy Hart, P., Young, M. R., and Armstrong, J. A. (1976) *Proc Natl Acad Sci U S A* **73**, 2510-2514
16. Goren, M. B. (1970) *Biochim Biophys Acta* **210**, 116-126
17. Goren, M. B. (1970) *Biochim Biophys Acta* **210**, 127-138
18. Taber, H. W., Sugarman, B. J., and Halfenger, G. M. (1981) *J Gen Microbiol* **123**, 143-149
19. Gilmore, S. A., Schelle, M. W., Holsclaw, C. M., Leigh, C. D., Jain, M., Cox, J. S., Leary, J. A., and Bertozzi, C. R. (2012) *ACS Chem Biol*
20. Boshoff, H. I., Myers, T. G., Copp, B. R., McNeil, M. R., Wilson, M. A., and Barry, C. E., 3rd. (2004) *J Biol Chem* **279**, 40174-40184

21. Dutton, R. J., Boyd, D., Berkmen, M., and Beckwith, J. (2008) *Proc Natl Acad Sci U S A* **105**, 11933-11938
22. Dutton, R. J., Wayman, A., Wei, J. R., Rubin, E. J., Beckwith, J., and Boyd, D. (2010) *Proc Natl Acad Sci U S A* **107**, 297-301
23. Rustad, T. R., Harrell, M. I., Liao, R., and Sherman, D. R. (2008) *Plos One* **3**, e1502
24. Bernofsky, C., and Swan, M. (1973) *Anal Biochem* **53**, 452-458
25. Leonardo, M. R., Dailly, Y., and Clark, D. P. (1996) *J Bacteriol* **178**, 6013-6018
26. Liu, P. T., Stenger, S., Li, H., Wenzel, L., Tan, B. H., Krutzik, S. R., Ochoa, M. T., Schaubert, J., Wu, K., Meinken, C., Kamen, D. L., Wagner, M., Bals, R., Steinmeyer, A., Zugel, U., Gallo, R. L., Eisenberg, D., Hewison, M., Hollis, B. W., Adams, J. S., Bloom, B. R., and Modlin, R. L. (2006) *Science* **311**, 1770-1773

Appendix

Table 4-2: Strains used in this chapter

Strains	Genotype	Source
<i>Mtb</i> H37Rv	Wild type	
<i>Mtb</i> H37Rv $\Delta cyp128$	Hyg ^r , hyg cassette disrupting <i>cyp128</i>	Ch 3
<i>Mtb</i> H37Rv $\Delta cyp128::cyp128$	Hyg ^r , Kan ^r , complemented strain of $\Delta cyp128$	Ch 3
<i>Mtb</i> H37Rv $\Delta stf3$	Hyg ^r , <i>stf3</i> interrupted by hyg resistance cassette	Ref (1)
<i>Mtb</i> H37Rv $\Delta stf3::stf3$	Hyg ^r , Kan ^r , complement with <i>stf3</i> under the glutamine synthase promoter modified pMV306 (3)	Ref (1)

References

1. Mougous, J. D., Senaratne, R. H., Petzold, C. J., Jain, M., Lee, D. H., Schelle, M. W., Leavell, M. D., Cox, J. S., Leary, J. A., Riley, L. W., and Bertozzi, C. R. (2006) *Proc Natl Acad Sci U S A* **103**, 4258-4263
2. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) *Nature* **393**, 537-544
3. Stover, C. K., de la Cruz, V. F., Fuerst, T. R., Burlein, J. E., Benson, L. A., Bennett, L. T., Bansal, G. P., Young, J. F., Lee, M. H., Hatfull, G. F., and et al. (1991) *Nature* **351**, 456-460

Chapter 5: Characterization of Cyp128 transcriptional regulation

Introduction

As genome sequencing becomes more accessible, the number of genes of unknown function is increasing rapidly. Identification of natural products produced by polyketide synthetases or non-ribosomal pathways has been facilitated by bioinformatics and mass spectrometry. Additionally, the products of polyketide synthetases and non-ribosomal peptides can be hypothesized based on the number and type of domains contained in the protein (1). Unfortunately, these technologies do not allow us to identify the functions or substrates for many families of proteins, including cytochrome P450 enzymes which are found in both prokaryotes and eukaryotes. Cytochrome P450s (cyt P450) are easily identified by the heme binding domain, but their substrates cannot be predicted *a priori* from enzyme homology. The first cyt P450s identified and characterized were involved in scavaging carbon sources from the environment or catabolism and therefore thought to be dispensable to the organism (2). As more prokaryotic genomes were sequenced, the high abundance of cyt P450 in some organisms suggested a more complex function and a possible role in the biosynthesis of secondary metabolites. Two examples are *Mycobacterium tuberculosis* (Mtb), the causative agent of tuberculosis, and *Streptomyces coelicolor*, also a human pathogen, which encode 20 and 18 cyt P450 respectively, drastically higher than most organisms (2). The relative cyt P450 density of Mtb (genome 4.4 Mb) is almost 250 fold higher than humans (57 cyt P450 in ~3,000 Mb) highlighting the importance of cyt P450s to Mtb biology.

To date, 6 cyt P450's from Mtb have been expressed and characterized (Cyp142, Cyp51, Cyp124, Cyp125, Cyp144, Cyp121) (2-8). With the exception of Cyp51, low sequence homology between Mtb cyt P450 and previously characterized cyt P450s has precluded *a priori* assignment of the native substrates of these enzymes. Often using the resource intensive expression and biochemical characterization is not sufficient to identify the native substrates. Further analysis using genomic context may aid in the identification of substrates and functions for some cyt P450.

Previously we had identified a sulfotransferase *stf3* that is necessary for biosynthesis of a sulfomenaquinone, named S881 (9,10). S881 was identified as a product of *stf3* using a combination genetic and mass spectrometry approach. After identifying four sulfotransferases in the Mtb genome by bioinformatics, each one was interrupted and the total lipid extracts from the Mtb mutants were analyzed by mass spectrometry and compared to wild type (WT) total lipid extracts (11). Using this approach, we identified the sulfotransferase involved in sulfolipid-1 biosynthesis (12) and *stf3* the sulfotransferase required for the biosynthesis of the previously unknown S881. Through extensive mass spectrometric fragmentation analysis, S881 was characterized as a terminally sulfated menaquinone (10). In the Mtb genome, *stf3* was found to be clustered with a putative cyt P450, *cyp128*. We have demonstrated that *cyp128* is required for S881 biosynthesis in Mtb. The clustering of *stf3* and *cyp128*, suggested that both genes were

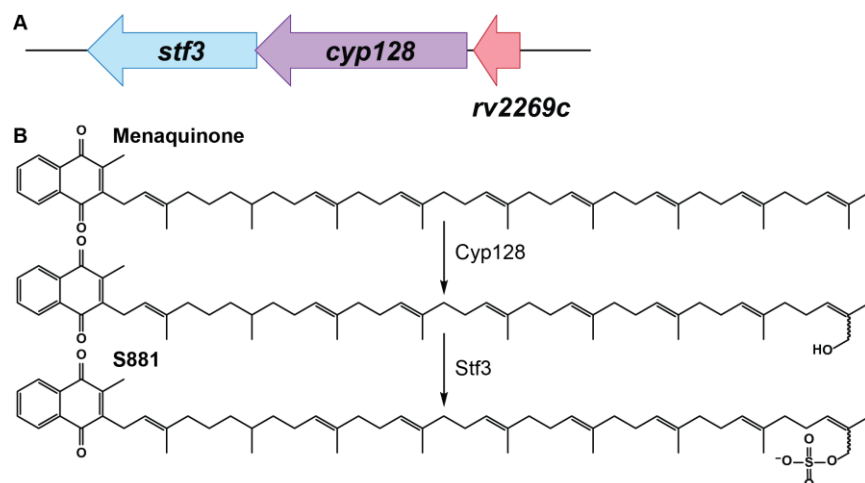


Figure 5-1: (A) The *stf3* operon shows the clustering of *stf3* and *cyp128*. (B) The proposed biosynthesis of S881 from menaquinone.

involved in the same pathway. Without characterization of neighboring genes in the same operon, the role of *cyp128* could not have been hypothesized from bioinformatics as S881 is the first reported example of a modified menaquinone.

We identified the genes necessary for S881 biosynthesis genetically and used these mutants to study the function of S881 in Mtb. Here we want to further understand how S881 biosynthesis is regulated and study the enzyme's biochemistry. We identified the native promoter of *cyp128* as contained within *rv2269c*, a gene of unknown function. Complementation studies suggest that *rv2269c* is not the promoter for *stf3* which is differentially regulated from *cyp128*. The gene *rv2269c* exhibited promoter activity in a GFP assay and that *rv2269c* promoter was only recognized in Mtb and *Mycobacterium marinum* (*M. mar*), the causative agent of tuberculosis disease in fish and amphibians (13), but was not recognized in *Mycobacterium smegmatis* (*M. smeg*). In addition, we identified an alternative transcriptional start site for *cyp128* transcript in Mtb grown under starvation conditions.

Results

Bioinformatic analysis of rv2269c and cyp128

The protein products of the genes *rv2269c* and *cyp128* were submitted to bioinformatics analysis. The gene *rv2269c* is 333 bp (111 aa) and analysis with BLAST indicated that *rv2269c* has very low homology to other proteins of its size. A protein structure prediction program, Phyre, suggests that it may have a short alpha helix fold but is mostly unstructured.

The gene *cyp128* is annotated as 1470 bp (489 aa) in the H37Rv genome on Tuberculist (<http://tuberculist.epfl.ch/>), but has been differentially annotated with an alternative start site (362 bp/residues 97-489 relative to H37Rv sequence) in more recently sequenced genomes

(SUMu, Erdman) (14). Cyp128 is not predicted to encode a signal sequence or any transmembrane domains suggesting that Cyp128 is likely cytosolic. The N-terminal 96 aa (H37Rv) is predicted to be unstructured, but the rest of the enzyme includes the conserved domains for heme binding and hydroxylase activity.

The gene *stf3* was identified as a sulfotransferase in a bioinformatics search for Mtb sulfotransferases (15). *Stf3* is predicted to have a single transmembrane domain and both PAPS binding domains.

The three genes, *rv2269c*, *cyp128*, and *stf3*, are predicted to be in a putative operon based on their organization. Conversely, most microarray data show differential regulation of each of the three genes (e.g. ref. (16,17)). RNA-seq data (unpublished from Tige Rustad, Sherman lab, Seattle Biomedical Institute) indicate that the rate of transcription of these three genes is very low, particularly *rv2269c*.

The gene rv2269c acts as the promoter for cyp128 translation

During the construction of the $\Delta cyp128$ complement plasmid, an effort was made to identify the native promoter of *cyp128*. We hypothesized that the promoter for the *stf3* operon was contained in the one kilobase of DNA upstream of *rv2269c* (referred to as P_{nat}). P_{nat} was

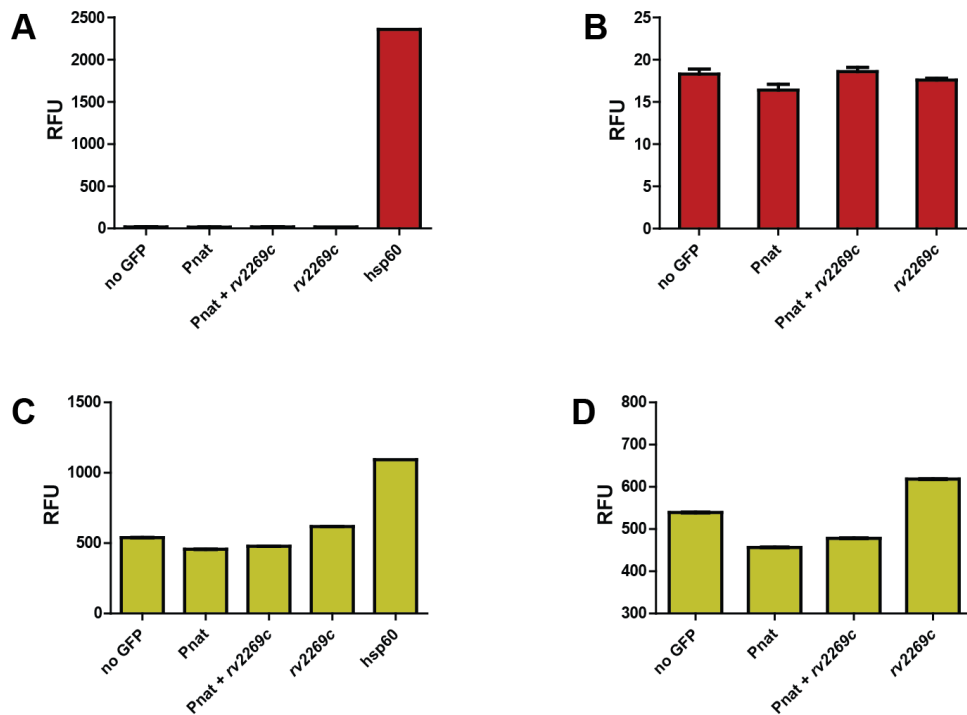


Figure 5-2: Promoter activity in *M. smeg* and *Mtb* using a GFP assay. (A) Promoter activity in *M. smeg*. (B) Promoter activity in *Mtb*. Promoter *hsp60* was used as a positive control and the wild type strain was used as a negative control. Error bars indicate SD. Data for *M. mar* are similar to *Mtb* data.

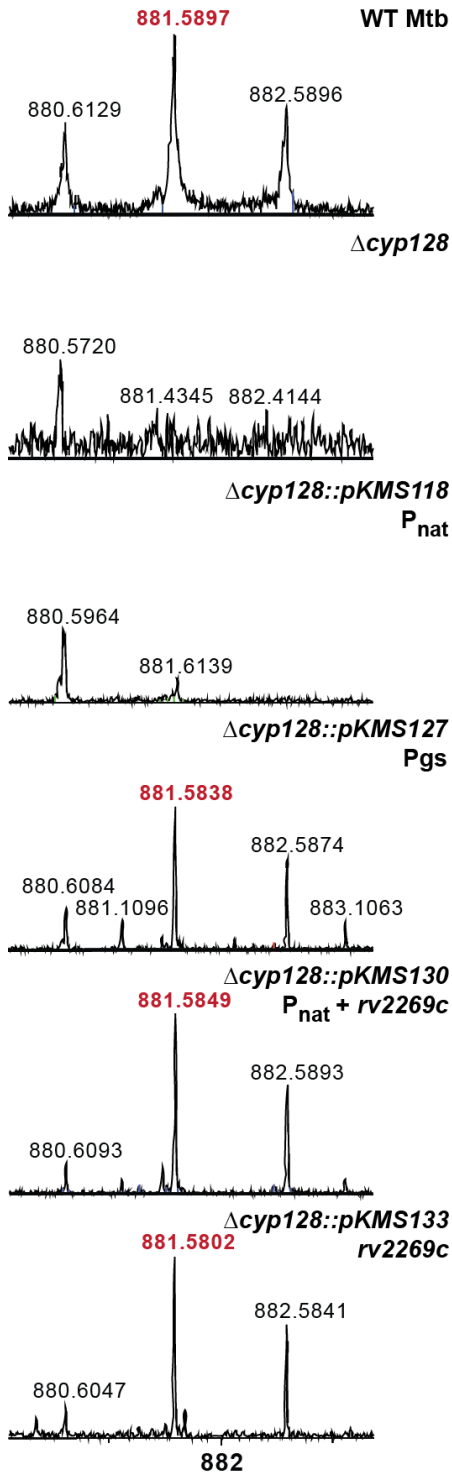


Figure 5-3: Mass spectrometry analysis of the $\Delta cyp128$ mutant and complemented strains. Peak corresponding to S881 is denoted in red. P_{gs} glutamine synthetase promoter

cloned as the promoter for *cyp128*. Total lipid extracts from the complement were analyzed by mass spectrometry to confirm complementation but no peak corresponding to S881 was observed. To identify the native promoter, a series of plasmids were constructed using either one kb of DNA upstream of *cyp128* (including the gene *Rv2269c*, $P_{nat-rv2269c}$) or just the 333 bp of *rv2269c* DNA sequence ($P_{rv2269c}$). Both plasmids containing $P_{nat-rv2269c}$ and $P_{rv2269c}$ successfully complemented S881 production in the $\Delta cyp128$ mutant (Figure 5-3).

To determine if *rv2269c* exhibited promoter activity, it was cloned upstream of the GFP gene. Promoter activity was measured by an increase in GFP fluorescence from whole cells (Figure 5-2). The plasmids were transformed into *M. smeg*, *M. mar* and Mtb for comparison of promoter activity. No promoter activity was seen in *M. smeg* with any of the promoters (Figure 5-2A). Conversely, significant promoter activity was observed with $P_{rv2269c}$ and $P_{nat-rv2269c}$ in *M. mar* and Mtb cells but not with P_{nat} (Figure 5-2B). Truncation of *rv2269c* resulted in abrogation of all promoter activity. Therefore full length *rv2269c* is necessary for promoter activity. This suggests that *rv2269c* contains all promoter activity for *cyp128* and the machinery that recognizes *rv2269c* as a promoter is found in both *M. mar* and Mtb.

To further support the differential transcriptional regulation of *cyp128* and *stf3*, a disruption of the *cyp128* gene would normally lead to a polar effect on *stf3* and stop its transcription. In fact, deletion of *cyp128* appears to have no effect on *stf3*. The $\Delta cyp128$ mutant only requires complementation with *cyp128* for restoration of S881 biosynthesis. If the deletion of $\Delta cyp128$ caused a polar effect, complementation would require both genes, but this is not the case. The native promoter of *stf3* has yet to be identified. Using DNA 1 kb upstream of *cyp128* does not restore S881 production in the $\Delta stf3$ mutant.

Mtb transcribes *cyp128* from an alternate start site when grown in starvation conditions

In 2002, Betts et al. investigated the transcriptional profile of *Mtb* grown in total starvation conditions (Phosphate buffered saline with a non-hydrolyzable detergent). Total starvation of *Mtb* showed a significant induction of the *stf3* operon after 24 hours. While starvation shows no phenotype for survival of the S881 biosynthetic mutants, the genes were confirmed to be up regulated after 24 hours of starvation (Figure 5-4A).

After identifying *rv2269c* as the promoter for *cyp128* transcription, we wanted to investigate if the transcription seen for *rv2269c* in microarrays was due to a long 5'UTR of *cyp128*. To identify the 5' end of *cyp128* mRNA transcript, we used a kit that amplifies the 5' end of DNA or RNA (5' RACE; Invitrogen). We isolated RNA from WT *Mtb* cultures grown either in starvation media or 7H9 media and converted to cDNA. Unexpectedly, we identified a shorter version of *cyp128* mRNA isolated after total starvation of wild type *Mtb* (Figure 5-4B). The shorter version of *cyp128* corresponds to a 5' truncation of the *cyp128* transcription arising from an alternate transcriptional start site. The alternative transcriptional start site corresponds to a truncation of 171 bp from the annotated start site in the H37Rv genome (18). Annotated genomes from other strains of *Mtb* (Erdman, SUMu) have used the “alternate” start site for *cyp128*. In addition, the truncated mRNA corresponds to an unstructured portion of Cyp128 from bioinformatics analysis. The truncated mRNA was only identified in RNA isolated from starved *Mtb* cultures and only full length *cyp128* was identified in *Mtb* grown in rich media.

Expression and purification of Cyp128

We initially began by expressing and purifying a truncated form of Cyp128 (tCyp128) from *E. coli*. We determined that the C-terminal hexahis-tagged tCyp128 showed more soluble protein production than the N-terminal construct, in addition to enabling us to purify only full length protein. Using the C-terminal hexahis-tagged tCyp128, we began optimizing expression conditions, but were unable to increase yield of pure protein. All attempts to purify tCyp128 were confounded as a result of alternate products that were in a complex with tCyp128 in solution (Figure 5-5). The undesired bands were isolated and identified by Mass spectrometry

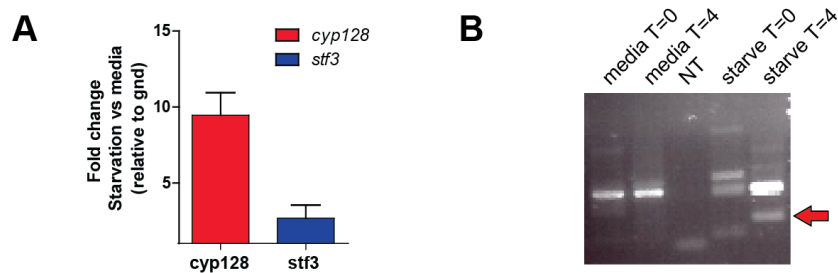


Figure 5-4: Starvation of WT *Mtb* causes changes in the *cyp128* mRNA transcript. (A) The S881 genes are induced under starvation conditions by qPCR. (B) The *cyp128* mRNA transcript is shorter when *Mtb* is starved.

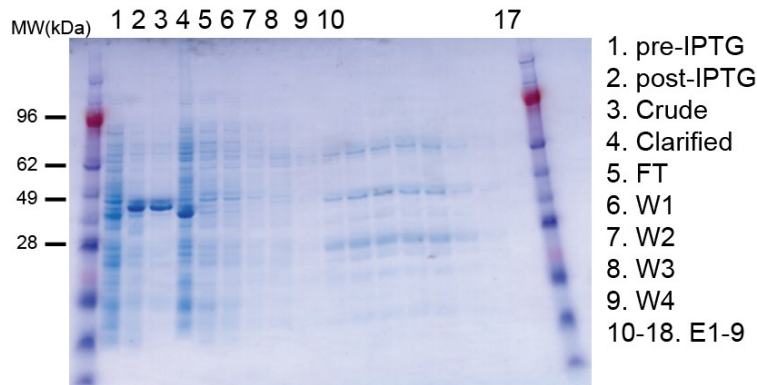


Figure 5-5: Expression of a truncated Cyp128 in *E. coli*. Coomassie stained gel from a Ni-NTA enrichment with protein loading normalized to protein concentrations. Major elution bands co-purify on a S300 size exclusion column.

(UC Davis Proteomics facility). Both upper and lower bands were identified as containing peptides from tCyp128 but the upper band also contained a high level of *E. coli* derived peptides. Due to the extremely low yields of protein and the low activity of the protein, we are currently pursuing expression in *M. smeg*.

Discussion

We have identified *rv2269c* as the native promoter for *cyp128* transcription. Initial studies suggest that *cyp128* and *stf3* are differentially regulated and *stf3* is controlled from a different promoter. All attempts to complement the Δ *stf3* mutant with a native promoter proved unsuccessful.

The promoter *rv2269c* is a weak promoter in Mtb when grown in rich aerated media. The machinery to recognize *rv2269c* as a promoter was only found in Mtb and *M. mar*, both pathogenic forms of mycobacteria that are more closely related to each other than to *M. smeg*. With the identification of *rv2269c* as a promoter, the GFP reporter strains can be used to study how *cyp128* is regulated under a variety of conditions that mimic an infection.

Since *rv2269c* is recognized as a promoter in both Mtb and *M. mar*, which does not encode orthologues of *cyp128* or *stf3*, suggests that *cyp128* is regulated by a transcription factor that may be responsible for a larger transcriptional response. We began initial studies to identify the transcription factor that recognized *rv2269c* using a gel shift assay. Unfortunately this work was never finished and the results remain inconclusive. Of the sigma factors that have been studied, none have suggested a role in regulating S881 biosynthesis as determined by microarrays. With increasing interest in large scale “omics” studies, particularly in the area of lipids, we hope that either *rv2269c* or S881 could be identified as part of a larger transcriptional response that may suggest a function for S881.

In the process of studying *rv2269c*, we discovered that *cyp128* is transcribed from an alternative start site when Mtb is grown under starvation conditions. A number of reasons could account for this. During starvation conditions, resources are low and an alternative shorter transcript is more resource efficient. Conversely, the alternate start site might remove the ability

of a second layer of regulation for protein function. When grown in rich media, Mtb may regulate the activity of Cyp128 by producing an inactive form of the enzyme that needs to be activated for activity or may be produced with a degradation tag to limit its residence time. More experiments are needed to elucidate how the alternative start site could change the role Cyp128 plays in the biology of Mtb. The alternate start sites annotated in other genomes is most likely due to differences in genome analysis. It is unlikely that the sequences are very different but further experiments in this area would be needed.

In conclusion, we have demonstrated that *rv2269c* acts as a native promoter for *cyp128* transcription and is recognized in both Mtb and *M. mar.* Under starvation conditions, *cyp128* is transcribed from a different transcriptional start site that may have implications into how *cyp128* is regulated either post-transcriptionally or post-translationally. Future studies will focus on the biochemistry of Cyp128 and Stf3 including *in vitro* biosynthesis of S881.

Materials and Methods

Reagents and Chemicals

Pfu DNA polymerase was from Stratagene (La Jolla, CA). Oligonucleotides were from Elim Biopharmaceuticals, Inc. (Hayward, CA). Restriction enzymes were from New England Biolabs (Ipswich, MA). Qiagen (Valencia, CA) kits were used for plasmid DNA purification, the extraction of DNA from agarose gels, and RNA purification. T4 DNA ligase, DNase and superscript III (reverse transcriptase) were purchased from Invitrogen. DNA sequencing was performed by Elim Biopharmaceuticals, Inc. and UC Berkeley sequencing facility. All other chemicals were purchased from Sigma, Spectrum (New Brunswick, NJ), or Fluka (St. Louis, MO) and used without further purification.

Bacterial strains and growth media

Growth media 7H9 and 7H11 and OADC supplement were obtained from BD Biosciences. Cloning and plasmid propagation were performed in *E. coli* DH5 α and XL-1Blue strains. All experiments in Mtb were done in H37Rv strain (gift from the Riley Lab, UC Berkeley) and all genes were cloned from its genomic DNA. The growth medium was 7H9 with 0.5% glycerol, 0.05% Tween-80, and 10% ADC for *M. smeg* and 10% OADC for Mtb and *M. mar.* For starvation studies, 7H9 media without any additive except 0.025% Tyloxopol, a non-hydrolyzable detergent was used. Antibiotics carbenicillin and kanamycin were obtained from Sigma. For selective media, antibiotic concentrations were 100 μ g/mL carbenicillin, 50 μ g/mL kanamycin or 100 μ g/mL hygromycin for *E. coli* and 20 μ g/mL kanamycin or 50 μ g/mL hygromycin for mycobacteria.

Sequence homology and analysis and structure prediction for rv2269c and cyp128

Amino acid sequences for Rv2269c and Cyp128 (Rv2268c) were obtained from Tuberculist (<http://tuberculist.epfl.ch/>). Both Rv2269c and Cyp128 sequences were submitted to Phyre for protein fold and structure prediction (<http://www.sbg.bio.ic.ac.uk/~phyre/>).

Promoter identification using reporter gene GFP

See appendix for a list of all plasmid, primers and strains used in the study. The reporter gene constructs were made by cloning in the putative native promoter (Pnat; 1kb upstream of *rv2269c*), Pnat + *rv2269c*, or *rv2269c* upstream of eGFP (referred to as GFP) in a modified pMV261 plasmid (19,20). Plasmids were transformed into *M. smeg*, *M. mar* and *Mtb* to compare promoter activity. GFP measurements were done as previously reported (19). Briefly, bacterial strains were grown up in 50 mL 7H9 media for 5 days. Clumps were removed by centrifugation and cell growth was determined by optical density (A_{600}). Emission from whole-cell suspensions was measured at 510 nm with excitation at 450 nm and a 495 nm high-pass cutoff filter in a Gemini XPS fluorescence microplate reader (Molecular Devices). For *Mtb*, cells were resuspended and incubated at room temperature for 1 h in 200 μ L phosphate-buffered 10% formalin prior to fluorescence measurement. Fluorescence was normalized to optical density.

Starvation assay and RNA isolation

Wild type *Mtb* was grown up to mid-log phase and clumps were removed by centrifugation. Cultures were washed three times in Starvation media and diluted to OD_{600} 0.5 in 10 mL. Cultures were incubated at 37 °C for 4 days and then pelleted and flash frozen until ready for use. RNA isolation was done as described previously (21). Briefly, frozen pellets were resuspended in 1 mL TRIzol (Invitrogen) and transferred to tubes containing 0.5 mL glass beads. Samples were lysed by bead beating three times for 30 seconds at high speed with 30 seconds of being placed on ice in between cycles. Beads were pelleted and supernatant was transferred to tubes containing 300 μ L Chloroform and Heavy Phase Lock Gel I (Eppendorf). Tubes were inverted for 2 min and then centrifuged for 5 min. The aqueous layer was transferred to a new tube containing equal volumes Isopropanol and 5 M NaCl. Samples were incubated overnight at 4 °C. RNA was pelleted and washed in 70% EtOH. RNA was purified and DNase treated using a RNeasy kit (Qiagen). Purified RNA was dissolved in 40 μ L ddH₂O and concentrations were measured by Nanodrop (Thermo).

Determination of the cyp128 transcriptional start site using 5'RACE

For determination of the *cyp128* transcriptional start site, we used the 5' Rapid Amplification of cDNA Ends (RACE) from Invitrogen. For all primers used, see appendix. Briefly, RNA was treated with DNase (Promega). PCR of the 5' end of *cyp128* was amplified using superscript II (provided in kit) and the PCR product was purified by the provided spin column followed by addition of a TdT tail. Two rounds of PCR (round 1 Taq polymerase (Promega) was used; round 2 *pfu* turbo (Stratagene) was used for TOPO cloning compatibility) were done to amplify the product. The product was purified by 2% agarose gel and was cloned into a TOPO vector for sequencing.

Cloning and expression in E. coli

A truncated Cyp128 (residues 97-489) was cloned into expression vector pET-28a (Novagen). Cyp128 with an N-terminal his-tag was cloned between NdeI and XhoI. The C-terminal his-tag version was cloned between NcoI and XhoI. Both constructs were transformed into BL21(DE3) cells. Protein expression of Cyp128 was done as described in ref (22). Briefly, a single colony was used to inoculate 500 mL of modified LB media (10 g Bacto Tryptone, 5 g Bacto yeast extract, 8 mL 50% glycerol, 50 mL K-PO₄ solution, 1 L H₂O, 250 µL microelement solution, 1 mM thiamin, 50 µg/mL Kanamycin; K-PO₄ solution 23.1 KH₂PO₄, 125.4 g K₂HPO₄, 500 mL water; microelement solution 2.7 g FeCl₃·6H₂O, 0.2 g ZnCl₂·4H₂O, 0.2 g CoCl₂·6H₂O, 0.2 g Na₂MoO₄·2H₂O, 0.1 g CaCl₂·5H₂O, 0.05 g H₃BO₃ in 90 mL water). Cultures were incubated at 37 °C until mid-log phase when the temperature was reduced to 18 °C and 300 µM IPTG, 1 mM δ-aminolevulinic acid, and 50 µg/mL kanamycin was added. The cultures were incubated overnight before pelleting. Pellets were flash frozen and stored at -20 °C until further use.

Frozen cell pellets were resuspended in a 1:5 (w/v) ratio of lysis buffer (50 mM KH₂PO₄, pH 7.5, 300 mM NaCl, 20 mM imidazole, 2 mM βMe, 0.5 mM EDTA, 10 µg/mL lysozyme). EDTA-free Complete Protease Inhibitor Cocktail Tablets (Roche) and 1 µM hemin were added to the lysis buffer. Cells were lysed via three passes through an Emulsiflex-C3 homogenizer (Avestin) at approximately 15,000 psi. Insoluble debris was removed by centrifugation in a SS-34 rotor (Sorvall) at 19,500 rpm for 45 min at 4 °C. Clarified lysate was incubated with Ni-NTA agarose resin (Qiagen) in batch for 1 h at 4 °C ; resin was subsequently washed with 5 column volumes of wash buffer (50 mM KH₂PO₄, pH 7.5, 300 mM NaCl, 50 mM imidazole, 2 mM βMe, 0.5 mM EDTA). Bound protein was eluted with 20 mL of elution buffer (50 mM KH₂PO₄, pH 7.5, 300 mM NaCl, 250 mM imidazole, 2 mM βMe, 0.5 mM EDTA).

References

1. Challis, G. L. (2008) *Microbiology* **154**, 1555-1569
2. McLean, K. J., Clift, D., Lewis, D. G., Sabri, M., Balding, P. R., Sutcliffe, M. J., Leys, D., and Munro, A. W. (2006) *Trends Microbiol* **14**, 220-228
3. Driscoll, M. D., McLean, K. J., Cheesman, M. R., Jowitt, T. A., Howard, M., Carroll, P., Parish, T., and Munro, A. W. (2011) *Biochim Biophys Acta* **1814**, 76-87
4. Driscoll, M. D., McLean, K. J., Levy, C., Mast, N., Pikuleva, I. A., Lafite, P., Rigby, S. E., Leys, D., and Munro, A. W. (2010) *J Biol Chem* **285**, 38270-38282
5. McLean, K. J., Lafite, P., Levy, C., Cheesman, M. R., Mast, N., Pikuleva, I. A., Leys, D., and Munro, A. W. (2009) *J Biol Chem* **284**, 35524-35533
6. Johnston, J. B., Kells, P. M., Podust, L. M., and Ortiz de Montellano, P. R. (2009) *Proc Natl Acad Sci U S A* **106**, 20687-20692
7. Johnston, J. B., Ouellet, H., and Ortiz de Montellano, P. R. (2010) *J Biol Chem* **285**, 36352-36360
8. Ouellet, H., Guan, S., Johnston, J. B., Chow, E. D., Kells, P. M., Burlingame, A. L., Cox, J. S., Podust, L. M., and de Montellano, P. R. (2010) *Mol Microbiol* **77**, 730-742
9. Mougous, J. D., Senaratne, R. H., Petzold, C. J., Jain, M., Lee, D. H., Schelle, M. W., Leavell, M. D., Cox, J. S., Leary, J. A., Riley, L. W., and Bertozzi, C. R. (2006) *Proc Natl Acad Sci U S A* **103**, 4258-4263
10. Holsclaw, C. M., Sogi, K. M., Gilmore, S. A., Schelle, M. W., Leavell, M. D., Bertozzi, C. R., and Leary, J. A. (2008) *ACS Chem Biol* **3**, 619-624
11. Mougous, J. D., Leavell, M. D., Senaratne, R. H., Leigh, C. D., Williams, S. J., Riley, L. W., Leary, J. A., and Bertozzi, C. R. (2002) *Proc Natl Acad Sci U S A* **99**, 17037-17042
12. Mougous, J. D., Petzold, C. J., Senaratne, R. H., Lee, D. H., Akey, D. L., Lin, F. L., Munchel, S. E., Pratt, M. R., Riley, L. W., Leary, J. A., Berger, J. M., and Bertozzi, C. R. (2004) *Nat Struct Mol Biol* **11**, 721-729
13. Davis, J. M., and Ramakrishnan, L. (2009) *Cell* **136**, 37-49
14. Akiyama, T., Kuwahara, T., Ando, T., Tada, T., and Kirikae, T. (2012) *Unpublished, Submitted to Bioproject #PRJDB66*
15. Mougous, J. D., Green, R. E., Williams, S. J., Brenner, S. E., and Bertozzi, C. R. (2002) *Chem Biol* **9**, 767-776
16. Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A., and Duncan, K. (2002) *Mol Microbiol* **43**, 717-731

17. Boshoff, H. I., Myers, T. G., Copp, B. R., McNeil, M. R., Wilson, M. A., and Barry, C. E., 3rd. (2004) *J Biol Chem* **279**, 40174-40184
18. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) *Nature* **393**, 537-544
19. Seeliger, J. C., Topp, S., Sogi, K. M., Previti, M. L., Gallivan, J. P., and Bertozzi, C. R. (2012) *Plos One* **7**, e29266
20. Stover, C. K., de la Cruz, V. F., Fuerst, T. R., Burlein, J. E., Benson, L. A., Bennett, L. T., Bansal, G. P., Young, J. F., Lee, M. H., Hatfull, G. F., and et al. (1991) *Nature* **351**, 456-460
21. Sherman, D. R., Voskuil, M., Schnappinger, D., Liao, R., Harrell, M. I., and Schoolnik, G. K. (2001) *Proc Natl Acad Sci U S A* **98**, 7534-7539
22. von Kries, J. P., Warriar, T., and Podust, L. M. (2010) *Curr Protoc Microbiol* **Chapter 17**, Unit17 14

Appendix

Table 5-1: Bacterial strains used in this chapter

Strains		Genotype	Source
<i>M. smeg</i> mc ² 155		Wild type	
<i>M. smeg</i> mc ² 155	<i>GFP</i>	pKMS136; Kn ^r	This study
<i>M. smeg</i> mc ² 155	<i>GFP</i>	pKMS137; Kn ^r ,	This study
<i>M. smeg</i> mc ² 155	<i>GFP</i>	pKMS138; Kn ^r ,	This study
<i>M. smeg</i> mc ² 155	<i>GFP</i>	pMWS114; Kn ^r ,	This study
<i>M. mar</i> strain M		Wild type, gift from Brown lab (Genentech)	
<i>M. mar</i> strain M	<i>GFP</i>	pKMS136; Kn ^r	This study
<i>M. mar</i> strain M	<i>GFP</i>	pKMS137; Kn ^r ,	This study
<i>M. mar</i> strain M	<i>GFP</i>	pKMS138; Kn ^r ,	This study
Mtb H37Rv		Wild type	
Mtb H37Rv	<i>GFP</i>	pKMS136; Kn ^r	This study
Mtb H37Rv	<i>GFP</i>	pKMS137; Kn ^r ,	This study
Mtb H37Rv	<i>GFP</i>	pKMS138; Kn ^r ,	This study
Mtb H37Rv	<i>GFP</i>	pMWS114; Kn ^r ,	This study
Mtb H37Rv	$\Delta cyp128$	Hyg ^r , hyg cassette disrupting <i>cyp128</i>	Ch 3
Mtb H37Rv	$\Delta cyp128::pKMS118$	Hyg ^r , Kan ^r , $\Delta cyp128$ complement with P _{nat}	This study
Mtb H37Rv	$\Delta cyp128::pKMS127$	Hyg ^r , Kan ^r , $\Delta cyp128$ complement with P _{gs}	This study
Mtb H37Rv	$\Delta cyp128::pKMS130$	Hyg ^r , Kan ^r , $\Delta cyp128$ complement with P _{nat+rv2269c}	This study
Mtb H37Rv	$\Delta cyp128::cyp128$	Hyg ^r , Kan ^r , $\Delta cyp128$ complement with P _{rv2269c}	Ch 3
<i>E. coli</i> BL-21(DE3)	<i>cyp128</i> (171-1469 bp)	<i>E. coli</i> expression system for tCyp128	This study
<i>M. smeg</i> mc ² 155	<i>cyp128</i> (171-1469 bp)	Expression system for tCyp128	This study

Table 5-2: Plasmids used in this Chapter

Plasmids	Description	Source
pMV261	Kn ^r , pAL5000 origin, ColeE1 origin, multiple cloning site, P _{hsp60} promoter	Ref (1)
pMV306	Kn ^r , A derivative of pMV261 lacking the P _{hsp60} promoter	Ref (1)
pMWS114	pMV261 derivative containing eGFP	Ref (2)
pKMS136	pMWS114 derivative replacing the promoter with P _{nat}	This study
pKMS137	pMWS114 derivative replacing the promoter with P _{nat+rv2269c}	This study
pKMS138	pMWS114 derivative replacing the promoter with P _{rv2269c}	This study
pKMS118	pMV306 derivative encoding <i>cyp128</i> with promoter P _{nat}	This study
pKMS127	pMV306 derivative encoding <i>cyp128</i> with promoter P _{glutamine synthase}	This study
pKMS130	pMV306 derivative encoding <i>cyp128</i> with P _{nat+rv2269c}	This study
pKMS133	pMV306 derivative encoding <i>cyp128</i> with P _{rv2269c}	This study
pET28a	Novagen	
pKMS156	pET28a derivative with <i>cyp128</i> (171-1469 bp) using C-terminal hexahis-tag	This study
pKMS159	pET28a derivative with <i>cyp128</i> (171-1470 bp) using N-terminal hexahis-tag	This study
pKMS160	pMV306 derivative with <i>cyp128</i> (171-1469 bp) including C-terminal hexahis-tag with the glutamine synthase promoter	This study

Table 5-3: Primers used in this Chapter. Restriction enzyme recognition sequence in bold and name in parenthesis.

Primer	Sequence	Description
okms218	gtt atc gataaattcgtgagcaagggcgag	5' pKMS136 (ClaI)
okms219	ctag ttaaca agcttttactgtacagctcgcc	3' pKMS136, pKMS137, pKMS138 (HpaI)
okms220	gc gtctaga aattcgtgagcaagggcgag	5' pKMS138, pKMS137 (XbaI)
okms177	a atctaga gtggcttgccatgctgttatgag	5' pKMS118 (XbaI)
okms178	ca tttcgaa acagggcgcgactgcg	3' pKMS118 (BstBI)
okms197	g tttctaga atgaccgcgacacagtccc	5' pKMS127 (XbaI)
okms196	gtcgac atcgat gcacggcggaagcggttac	3' pKMS127 (ClaI)
okms181	gc ggtacc gtggcttgccatgctgttatgag	5' pKMS130 (KpnI)
okms196	gtcgac atcgat gcacggcggaagcggttac	3' pKMS130, pKMS133 (ClaI)
okms213	cg cggtacc gtggccaacgatgcgcg	5' pKMS133 (KpnI)
ogkf111	gataatgcgggtaagg	5' RACE <i>cyp128</i> GSP1
ogkf112	tcgaacgggtcaaagtcggtgag	5' RACE <i>cyp128</i> GSP2
ogkf113	tcaacctggaccgagccttc	5' RACE <i>cyp128</i> GSP3
ogkf114	cgatcaacctggaccgagccttc	5' RACE <i>cyp128</i> GSP3 with NotI site
ogkf115	cgtagaacagtcacag	5' RACE <i>rv2269c</i> GSP1
ogkf116	aacgcaataagatggttaccggag	5' RACE <i>rv2269c</i> GSP2
ogkf117	cgatggttaccggagttcggac	5' RACE <i>rv2269c</i> GSP3 with NotI site
okms224	gtccacaacggcatcgagta	gnd 5' qPCR primer
okms225	gctgtccagatcgccattg	gnd 3' qPCR primer
okms230	aggagcacctccgctttatc	<i>cyp128</i> 5' qPCR primer
okms231	catccgggtcttcgtactgg	<i>cyp128</i> 3' qPCR primer
okms233	aaggccgagaactcgttgac	<i>stf3</i> 5' qPCR primer
okms234	taactcgaagtcgccag	<i>stf3</i> 3' qPCR primer

References

1. Stover, C. K., de la Cruz, V. F., Fuerst, T. R., Burlein, J. E., Benson, L. A., Bennett, L. T., Bansal, G. P., Young, J. F., Lee, M. H., Hatfull, G. F., and et al. (1991) *Nature* **351**, 456-460
2. Seeliger, J. C., Topp, S., Sogi, K. M., Previti, M. L., Gallivan, J. P., and Bertozzi, C. R. (2012) *Plos One* **7**, e29266