

**STRUCTURE-BASED DRUG MECHANISM STUDY AND  
INHIBITOR DESIGN TARGETING TUBERCULOSIS**

A Dissertation

by

FENG WANG

Submitted to the Office of Graduate Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2007

Major Subject: Chemistry

**STRUCTURE-BASED DRUG MECHANISM STUDY AND  
INHIBITOR DESIGN TARGETING TUBERCULOSIS**

A Dissertation

by

FENG WANG

Submitted to the Office of Graduate Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,  
Committee Members,

Head of Department,

James C. Sacchetti  
Julian Leibowitz  
Kevin Burgess  
Francois P. Gabbai  
David H. Russell

August 2007

Major Subject: Chemistry

## ABSTRACT

Structure-Based Drug Mechanism Study and Inhibitor Design Targeting Tuberculosis.

(August 2007)

Feng Wang, B.S., University of Science and Technology of China;

M.S., Northwestern University

Chair of Advisory Committee: Dr. James C. Sacchettini

The increase of multi-drug resistant and extensively drug resistant tuberculosis (TB) cases makes it urgent to develop a new generation of TB drugs to counter resistance and shorten treatment. Structural biology, which allows us to “visualize” macromolecules, is now playing a key role in drug discovery. In this work, a structure-based approach was applied to the study of the mode of action of current TB chemotherapies, the identification of potential therapeutic targets, and the design of new inhibitors against TB.

Knowledge of the precise mechanisms of action of current TB chemotherapies will provide insights into designing new drugs that can overcome drug-resistant TB cases. Structural biology combined with biochemical and genetic approaches was used to elucidate the mechanisms of actions of isoniazid, ethionamide and prothionamide. The active forms of these anti-TB prodrugs were identified by protein crystallography and the target-inhibitor interactions were revealed by the complex structures. Although these drugs are activated through two completely different routes, they all inhibit InhA, an

essential enzyme in mycolic acid biosynthesis, by modification of the enzyme cofactor NAD, which unveils a novel paradigm of drug action.

Isoniazid, ethionamide and prothionamide all target InhA, which validates the enzyme as a superb drug target. A structure-based approach was adopted to design new inhibitors targeting InhA, using triclosan as the scaffold. Guided by the InhA-inhibitor complex structures, two groups of triclosan analogs were identified with dramatically increased inhibitory activity against InhA.

Structural biology has also provided fundamental knowledge of two potential therapeutic targets, *Mtb*  $\beta$ -lactamase (BlaC) and fatty-acyl-CoA thioesterase (FcoT). *Mtb*  $\beta$ -lactamase has been proposed to be the most significant reason for mycobacterial resistance to  $\beta$ -lactam antibiotics. The determination of *Mtb* BlaC structure not only demonstrates the mechanism of drug resistance but also provides a solid base for the design of new  $\beta$ -lactamase inhibitors that could be used with  $\beta$ -lactam antibiotics as a new regimen to treat tuberculosis. The crystal structure of FcoT provided crucial information in identification of the function of this previously hypothetical protein. The characterization of FcoT revealed an important pathway that is critical for *Mtb*'s survival in host macrophages.

## TABLE OF CONTENTS

	Page
ABSTRACT .....	iii
TABLE OF CONTENTS .....	v
LIST OF FIGURES .....	ix
LIST OF TABLES .....	xii
CHAPTER	
I INTRODUCTION .....	1
Tuberculosis .....	1
Development of TB Chemotherapy .....	2
Current TB Drugs .....	2
Problems of Current TB Drugs .....	7
Promising Drug Candidates .....	8
Potential Therapeutic Targets .....	10
<i>Mtb</i> Cell Envelope Biosynthesis .....	11
Overview .....	11
<i>Mtb</i> Mycolic Acid Biosynthesis Pathway .....	12
<i>Mtb</i> Arabinogalactan-Peptidoglycan Complex Biosynthesis Pathway .....	18
Structure-based Drug Discovery .....	22
Overview .....	22
Virtual Screen .....	22
Fragment-based Screen .....	26
Principle of Protein Crystallography .....	27
Overview .....	27
Protein Crystallization .....	27
Protein Crystal Diffraction and Data Collection .....	29
Phasing .....	31
Phase Improvement .....	34
Model Building and Refinement .....	34
II MECHANISM OF ACTION OF ISONIAZID AGAINST MYCOBACTERIA .....	36
Background .....	36

CHAPTER	Page
Methods .....	40
Cloning, Expression and Purification .....	40
Synthesis of the INH-NAD, PNH-NAD and BZH-NAD Adducts .....	41
InhA Activity Assay.....	41
Adduct Formation by KatG Catalysis.....	42
Reduction of KatG Compound I by INH and Its Analogs.....	42
Crystallization of InhA in Complex with Hydrazides-NAD.....	43
Crystallization of InhA S94A:INH-NAD .....	43
Data Collection and Processing .....	43
Structure Determination and Model Refinement .....	44
Minimum Inhibitory Concentration (MIC) Determination.....	46
Results .....	46
InhA(S94A) Protein Is More Resistant to INH-NAD Than the Wild Type	
InhA .....	46
Antibacterial Activity of INH Analogs against <i>Mtb</i> Strains .....	49
Antibacterial Activity of INH Analogs against <i>Mtb</i> Strains .....	49
The INH-, BZH- and PNH-NAD Adducts Are Potent Inhibitors of InhA ...	50
Structures of InhA Complexes Revealed Similar Binding Modes in the	
Active Site.....	53
KatG Mediated InhA Inhibition Assay of INH, BZH and PNH.....	54
The Difference of Adduct Production Rates Is Not due to Different	
Oxidation Rates of the Hydrazides .....	57
Discussion.....	59
Why INH Is the Superior Prodrug against TB among the Hydrazides.....	59
Other Putative Targets.....	61
 III MECHANISM OF THIOAMIDE AGAINST MYCOBACTERIA .....	 66
Background.....	66
Methods .....	68
Cloning, Expression, and Purification .....	68
Isolation and Characterization of ETH-NAD and PTH-NAD .....	69
InhA Enzymatic Activity Assay.....	69
Crystallization of InhA in Complex with ETH-NAD Adduct .....	70
Data Collection and Processing .....	70
Structure Determination and Model Refinement .....	71
Control Experiment.....	71
Results and Discussion.....	73
Activation of ETH Using Purified EthA or Chemical Agents <i>in vitro</i> .....	73
Design of a Cell-Based Activation System.....	74
Isolation and Characterization of the Active Forms of ETH and PTH.....	75
Binding Mode of the ETH-NAD Adduct.....	79

CHAPTER	Page
Mechanism of Drug Action of ETH and PTH against Leprosy.....	82
Activation Mechanism of ETH and PTH.....	83
Other Anti-TB Thioamide Drugs.....	84
Crystallization Trials of EthA.....	86
Conclusion.....	86
 IV INHIBITOR DESIGN TARGETING <i>MTB</i> INHA.....	 88
Background.....	88
Methods.....	89
Cloning, Expression and Purification.....	89
InhA Enzymatic Activity Assay.....	89
Crystallization of InhA in Complex with Inhibitors.....	89
Data Collection and Processing.....	90
Structure Determination and Model Refinement.....	92
Results and Discussion.....	92
Structure Activity Relationship (SAR) Study of Triclosan Analogs.....	92
Triclosan Derivatives with Modification at the Linker.....	93
Triclosan Derivatives with Modification at 2'-Position.....	96
Triclosan Derivatives with Modification at 4'-Position.....	100
Triclosan Derivatives with Modification at 1-Position.....	104
Triclosan Derivatives with Modification at 6-Position.....	105
Triclosan Derivatives with Modification at 5-Position.....	107
High Throughput Screen Identified a Series of New Inhibitors.....	120
Computer Program Based Virtual Screen of ZINC Chemical Database.....	121
 V STRUCTURE AND ACTIVITY STUDIES OF THE <i>MTB</i> $\beta$ -LACTAMASE.....	 124
Background.....	124
Transpeptidase, an Essential Enzyme in Cell Wall Biosynthesis.....	124
Mode of Action of $\beta$ Lactam Antibiotics.....	126
Function of $\beta$ Lactamase and $\beta$ Lactam Resistance.....	128
<i>Mtb</i> $\beta$ lactamase – a Potential Anti-TB Drug Target.....	128
Methods.....	131
Cloning, Expression and Purification.....	131
Kinetic Assays.....	132
Crystallization of BlaC (41-307).....	132
Data Collection and Processing.....	133
Structure Determination and Model Refinement.....	133
Structure Analysis.....	134
Results and Discussion.....	135
Sequence Analysis of BlaC.....	135

CHAPTER	Page
Expression, Purification and Characterization of <i>Mtb</i> BlaC .....	138
Enzyme Activity of <i>Mtb</i> BlaC .....	138
Overall Structure of BlaC .....	140
Enzyme Active Site .....	143
Structural Basis for Broad Substrate Recognition .....	151
Class A $\beta$ -lactamase Inhibitors .....	155
 VI IDENTIFICATION OF A <i>MTB</i> VIRULENCE FACTOR .....	 159
Background .....	159
A Putative Operon <i>Rv0096-Rv0101</i> Is Essential for the Virulence of <i>Mtb</i> .....	159
Function and Classification of CoA Thioesterases .....	160
Methods .....	163
Cloning, Expression, and Purification .....	163
Mutagenesis, Expression, and Purification .....	164
Acyl-CoA Thioesterase Activity Assay .....	164
Crystallization of FcoT .....	165
Data Collection and Processing .....	165
Structure Determination and Model Refinement .....	166
Structure Analysis .....	168
Results and Discussion .....	168
Sequence Analysis of FcoT .....	168
Purification and Characterization of FcoT .....	168
FcoT Is Not a Fatty-acyl CoA Dehydratase .....	169
FcoT Is a Fatty-acyl CoA Thioesterase .....	169
Overall Structure of FcoT .....	175
Structural Basis of Substrate Recognition .....	179
Catalytic Site and Mechanism .....	184
Physiological Function of FcoT .....	189
 VII CONCLUSION .....	 192
REFERENCES .....	195
VITA .....	214



## LIST OF FIGURES

	Page
Figure 1 Chemical Structures of Frontline and Second-line TB Drugs .....	4
Figure 2 Chemical Structures of Some Promising TB Drug Candidates .....	9
Figure 3 Chemical Structures of Mycolic Acid. ....	13
Figure 4 FASII Biosynthesis Pathway .....	15
Figure 5 Structure of Peptidoglycan of <i>Mtb</i> .....	20
Figure 6 Biosynthesis Pathway of MAP Complex of <i>Mtb</i> .....	21
Figure 7 Virtual Screen Procedure .....	24
Figure 8 The Picture of an InhA Single Crystal .....	28
Figure 9 An Image of X-ray Diffraction of a Single Protein Crystal .....	30
Figure 10 The Fluorescence Scan Result of Se-Met .....	33
Figure 11 Model-Building Based on the Clear Electron Density .....	35
Figure 12 MIC of INH, PNH and BZH against <i>M. bovis</i> BCG Strain .....	39
Figure 13 Crystal Structure of S94A Bound with the INH-NAD Adduct. ....	48
Figure 14 Synthesis of INH-NAD, BZH-NAD, and PNH-NAD .....	51
Figure 15 Surface Diagram of InhA in Complex with Hydrazide-NAD Adducts .....	55
Figure 16 Formation of Hydrazide-NAD Adduct Catalyzed by KatG. ....	56
Figure 17 Reduction of KatG Compound I by INH and Its Analogs .....	58
Figure 18 Chemical Structures of Ethionamide and Isoniazid .....	67
Figure 19 MALDI Mass Spectra of ETH-NAD and PTH-NAD. ....	76
Figure 20 Active Sites of <i>Mtb</i> InhA Bound to Inhibitors .....	80

	Page
Figure 21 Active Site of InhA in Complex with ETH-NAD and PTH-NAD.....	81
Figure 22 Possible Reaction Mechanisms of the Activation of ETH .....	85
Figure 23 Surface Diagram of Structure of InhA in Complex with <b>7</b> .....	113
Figure 24 The Superimposition of Structures of InhA Bound with <b>24</b> and <b>25</b> .....	117
Figure 25 Structure of InhA in Complex with Compound <b>17</b> .....	119
Figure 26 Cross-Linking of Peptidoglycan Strands by Cell Wall Transpeptidases .....	125
Figure 27 Mechanism of Action of $\beta$ -lactam Antibiotics .....	127
Figure 28 Classification of $\beta$ - lactamase .....	130
Figure 29 Sequence Alignment of Class A $\beta$ -lactamases of Different Organisms .....	137
Figure 30 Mechanism of Class A $\beta$ -lactamase.....	144
Figure 31 Overall Structure of <i>Mtb</i> BlaC.....	145
Figure 32 Stereo View of Active Site of <i>Mtb</i> BlaC .....	146
Figure 33 Comparison of Active Site of <i>Mtb</i> BlaC with Homologs in Stereo View ...	150
Figure 34 Comparison of Active Sites of BlaC Bound to Inhibitors .....	153
Figure 35 Comparison of $\Omega$ loops of BlaC .....	154
Figure 36 The Critical Role of Arg220 in Inhibitor Binding .....	158
Figure 37 The Gene Cluster of <i>Rv0097</i> – <i>Rv0101</i> .....	160
Figure 38 Catalytic Mechanisms of Thioesterases.....	162
Figure 39 The Superose-6 Size-exclusion Chromatography of FcoT.....	171
Figure 40 The Substrate Profile of <i>Mtb</i> FcoT .....	172
Figure 41 Substrate Specificity of <i>Mtb</i> FcoT .....	173

	Page
Figure 42 The Activity of FcoT for Substrate Palmitoyl CoA in the pH Range 7-9.5	174
Figure 43 The Overall Architecture of FcoT. ....	176
Figure 44 The Electron Density Indicating the Disulfide Bond in <i>Mtb</i> FcoT.....	178
Figure 45 The Substrate-binding Tunnel of <i>Mtb</i> FcoT with Bound Ligand. ....	182
Figure 46 The Active Site of <i>Mtb</i> FcoT with Bound Dodecoenoate. ....	183
Figure 47 Comparison of the Catalytic Sites of FcoT and D2.3. ....	187
Figure 48 A Catalytic Mechanism Without a General Base .....	188
Figure 49 Chemical Structure of Mycobactin.....	191
Figure 50 Chemical Structure of Fortuitin .....	191

## LIST OF TABLES

	Page
Table 1 The Frontline and Second-line Drugs and Their Targets.....	6
Table 2 Data Statistics of InhA in Complex with Hydrazide-NAD Adducts .....	45
Table 3 Inhibition of InhA and S94A by INH-NAD .....	47
Table 4 The MICs of Hydrazides against <i>Mtb</i> Wild Type and Mutant Strains.....	50
Table 5 Inhibition of InhA by Hydrazide-NAD Adducts .....	52
Table 6 Data Statistics of <i>Mtb</i> and <i>M. leprae</i> InhA Complexes .....	72
Table 7 Data Statistics of InhA in Complex with Triclosan Analogs.....	91
Table 8 Inhibition of InhA by Triclosan Analogs at 10 $\mu$ M Concentration.....	94
Table 9 Inhibition of InhA by Triclosan Analogs with Substituents on B Ring.....	97
Table 10 Inhibition of InhA by Triclosan Analogs with NH Group at 2'-Position .....	98
Table 11 Inhibition of InhA by Triclosan Analogs with Cyclic NH at 2'-Position .....	99
Table 12 Inhibition of InhA by Triclosan Analogs with Substituents at 4'-Position....	101
Table 13 Inhibition of InhA by Triclosan Analogs with (CO)R Group at 4'-Position .	102
Table 14 Inhibition of InhA by Triclosan Analogs with (NH)R Group at 4'-Position .	103
Table 15 Inhibition of InhA by Triclosan Analogs with Substituents at 1-Position .....	105
Table 16 Inhibition of InhA by Triclosan Analogs with Substituents at 6-Position .....	107
Table 17 Inhibition of InhA by Triclosan Analogs with Substituents at 5-Position .....	110
Table 18 Inhibition of InhA by Triclosan Analogs with Alkyl Groups at 5-Position....	111
Table 19 Inhibition of InhA by Triclosan Analogs with Aryl Groups at 5-Position ....	115
Table 20 The Evaluation of the “Hits” Obtained from Virtual Screen .....	123

	Page
Table 21 Kinetic Parameters of <i>Mtb</i> BlaC with Different Antibiotics as Substrates ....	139
Table 22 Data Collection, Processing and Refinement Statistics of <i>Mtb</i> BlaC .....	141
Table 23 Structure Comparison of Class A $\beta$ -lactamases from Different Species .....	142
Table 24 Data Collection, Processing, and Refinement Statistics of FcoT.....	167
Table 25 Kinetic Parameters of <i>Mtb</i> FcoT Mutants Using Palmitoyl-CoA.....	186

# CHAPTER I

## INTRODUCTION

### TUBERCULOSIS

Tuberculosis (TB) is a disease with a long history. Tubercular decay was found in Egyptian mummies, which indicates TB infections in humans since 5000 years ago (1). Into the 21st century, tuberculosis remains a serious threat to the world's public health. According to WHO, nearly two billion people have tuberculosis, which causes more than 2 million deaths annually. It is projected that in the next 15 years, nearly one billion more people will be infected if no action is taken. TB is generally defined as a developing country disease. However, it has become an urgent health problem in the United States as well, due in part to HIV co-infection (2-4).

TB is caused by the pathogen *Mycobacterium tuberculosis* (*Mtb*), which is spread through the air from one person to another (5). Not everyone infected becomes sick. Only 10% of immunocompetent people infected with *Mtb* develop active disease in their lifetime (6). The other 90% do not become ill and cannot transmit *Mtb*; instead they remain in a stage called latent infection (6). However, in the immunodeficient, the rate of who develop active TB disease is much higher (5).

---

This dissertation follows the style of the *Journal of Biological Chemistry*.

*Mtb* is a slow-growing Gram-positive mycobacterium that divides every 16-20 hours. It is a small rod-like bacillus which can survive in a dry state for weeks and can withstand weak disinfectants but can only actively grow within its host (7). *Mtb* cells can be stained after the treatment of acidic solution because of its thick waxy cell wall (7). Although *Mtb* most likely grows in tissues and organs with high oxygen content, such as the lungs, it can attack any part of the body such as the kidney, spine, and brain. Less than ten bacilli are enough to initiate an infection (8). Tubercle formation begins after the inhaled bacilli are engulfed by macrophages (8). Within the tubercles, the host immune system responds to the infection by cell-mediated immunity to the tubercular proteins (9). However, *Mtb* has the ability to modulate the host immune system, which allows it to survive inside the host cells. A dynamic balance between bacterial virulence and host immune system usually reaches within two to six weeks (9). The tubercle bacilli are under control but still persist (10). In many cases, the fibrosis and calcification causes slow death of the pathogens eventually. Otherwise, the bacilli remain viable for long periods and are a potential source of reactivation when the host immune system collapses (10-12).

## **DEVELOPMENT OF TB CHEMOTHERAPY**

### **Current TB Drugs**

The current TB chemotherapy was developed based on animal tests and clinical trials mainly between the 1950s and 1970s. Currently, WHO recommended DOTS (directly observed treatment, short-course) as the frontline TB chemotherapy (2). It is a

cocktail of drugs (normally four) with both bactericidal activity and inhibition of resistance. Drugs such as isoniazid, rifampin and streptomycin are top choices in bactericidal activity, and ethambutol, pyrazinamide and rifampin have been shown to inhibit the development of resistance. Using a combination of isoniazid, pyrazinamide, ethambutol and rifampin, most patients see improvement within the first month of treatment, and the total treatment time is between 6-12 months (2). Current drugs are mainly targeting cell wall biosynthesis, nucleic acid synthesis, and protein synthesis (Figure 1 and Table 1).

Isoniazid (INH) is a prodrug activated by *Mtb* by KatG, a catalase-peroxidase, to generate its active form INH-NAD, which inhibits enoyl acyl ACP reductase (InhA), an essential enzyme in fatty acid synthesis II (FASII) pathway, thus blocking the mycolic acid biosynthesis. This mechanism has been established by genetic, structural and biochemical studies (13-15). Mutations in *katG* and *inhA* were found in the *Mtb* clinical isolates resistant to INH (15-17). NADH/NAD ratio variation caused by mutations in NADH Dehydrogenase II has also been shown to confer resistance to INH (18).

Ethionamide (ETH) is a second line drug against TB. It is structurally similar to isoniazid (Figure 1). ETH is also a prodrug that has to be activated by EthA, a monooxygenase, to exert anti-tubercular activity (19). Its active form was proposed to inhibit InhA, the same target as INH (13,14,20). The mode of action of ETH was demonstrated in a cell-based activation system (21). Its active form, the ETH-NAD adduct, was characterized using protein crystallographic and mass spectroscopic approaches (21).



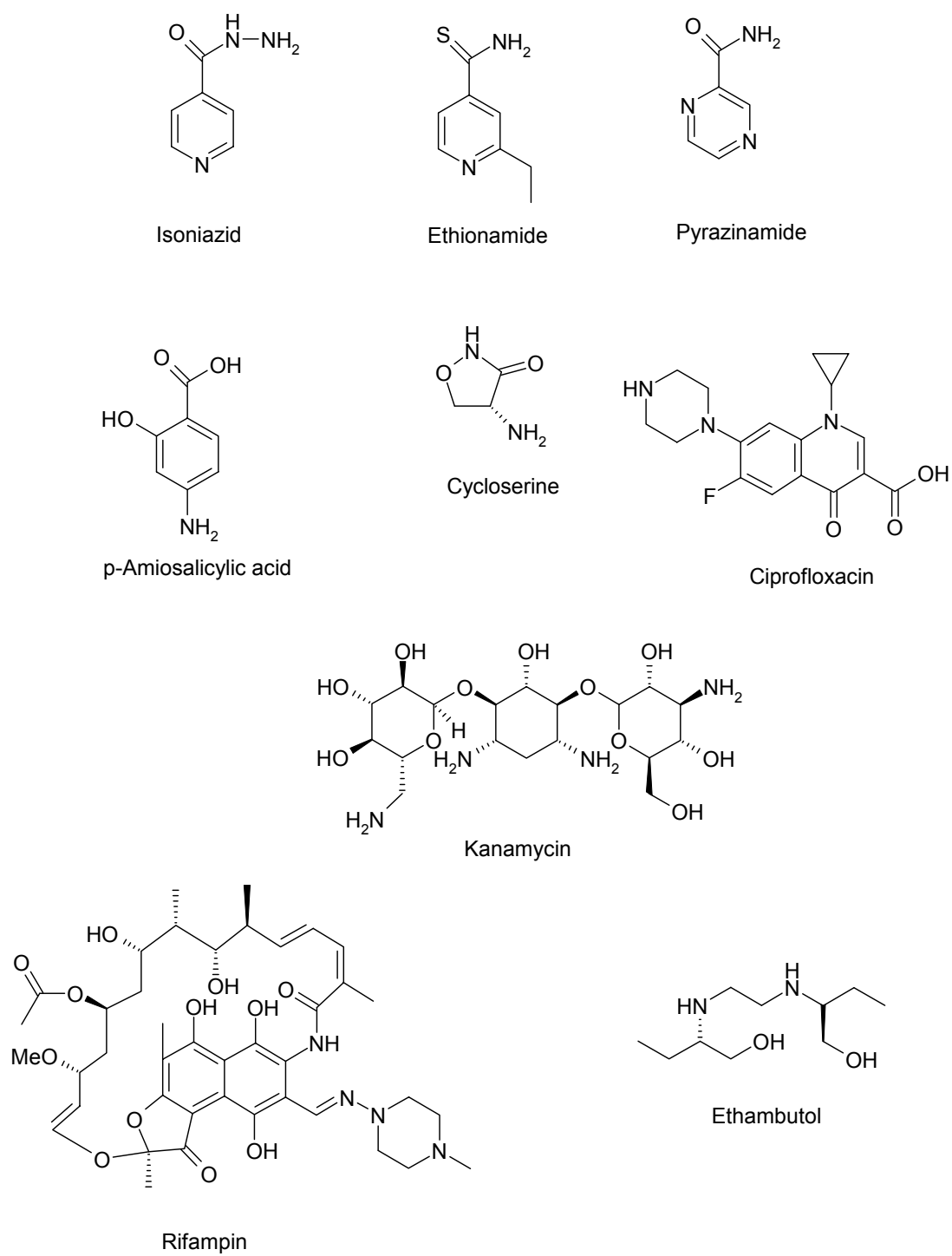


Figure 1 Chemical Structures of Frontline and Second-line TB Drugs

Cycloserine is believed to inhibit two enzymes D-alanine racemase (AlrA) and D-alanine:D-alanine ligase (DdlA), which are essential in the cell wall peptidoglycan (PG) synthesis pathway. Inactivation of *alrA* or *ddlA* gene in *M. smegmatis* caused increased sensitivity to cycloserine and overexpression of *alrA* and *ddlA* conferred 8-20 fold resistance to cycloserine (22,23). Although the genetic observations support that *alrA* and *ddlA* are possible targets of cycloserine, further biochemical studies are required to get a clear mechanism of drug action of cycloserine against *Mtb*.

Streptomycin is an aminoglycoside antibiotic, used as the first antitubercular chemotherapy. The mechanism of its drug action is to block protein synthesis by inhibiting mRNA translation by ribosomes (24). Its primary targets are protein S12 (*rpsL*) and 16S rRNA (*rrs*) in the small 30S subunit of the ribosome.

Rifampin is a series of broad-spectrum antibiotics that disrupt RNA synthesis by binding to the subunit B of bacterial DNA-dependent RNA polymerase (RpoB) (25). The crystal structure of RNA polymerase bound with rifampin, combined with biochemical studies, has demonstrated that rifampin blocked the transcription after three nucleotides (26). Rifampin is special in TB chemotherapies because it is active against both actively growing and nongrowing bacilli. Targeting the slowing or nongrowing bacilli is believed to be the key to shorten treatment length (27). Mutations in three regions of the *rpoB* are found in about 96% of rifampin -resistant *Mtb* isolates (28).

Fluoroquinolone is a class of compounds that are now used to treat multiple drug resistant (MDR) TB as second-line drugs. They are known to inhibit DNA synthesis by

targeting the type II DNA topoisomerase (GyrA and GyrB) (29). Strains that contain mutations in *gyrA* or *gyrB* were found to be resistant to fluoroquinolon (30).

It is worth noting that those chemotherapies that target protein synthesis and nucleic acid synthesis pathways are all broad-spectrum antibiotics due to these pathways being essential for bacteria. In contrast, the chemotherapies that target cell wall biosynthesis are more specific to *Mtb*, which is probably because of the uniqueness of *Mtb* cell wall.

Table 1 The Frontline and Second-line Drugs and Their Targets

Drug	Category	Target
Isoniazid	Front line	InhA
Rifampin	Front line	RpoB
Pyrazinamide	Front line	unknown
Ethionamide	Second line	InhA
Ethambutol	Second line	EmbCAB
Streptomycin	Second line	RpsL
Fluoroquinolones	Second line	GyrA, GyrB
Cycloserine	Second line	DdlA, AlrA

### **Problems of Current TB Drugs**

Although DOTS can cure TB in most of the cases, some serious problems still remain. Among those, the TB persistence and drug resistance are most challenging (31). Compared to the treatment of other infectious diseases, the treatment time for TB is long (usually 6-9 months), which makes patient compliance very difficult. The main reason for this lengthy treatment is to prevent relapse of TB. Antibiotics are usually active against growing bacteria but are less effective against nongrowing bacteria. Nongrowing bacteria includes the stationary phase bacteria, persisters not killed by antibiotic, and dormant bacteria (32). Current TB drugs such as isoniazid are mainly active against growing bacteria and are not effective against the persisters and dormant bacteria (33). The incapability of host immune system to eliminate tubercle bacilli in the lesions could be another factor that leads to the persistence of TB (33).

The first TB chemotherapy, streptomycin, came into use to treat TB in 1943 (24). And shortly after that, streptomycin resistant strains of *Mtb* were found. The following 20 years showed tremendous promise in TB chemotherapy, each time to be let down by the occurrence of *Mtb*'s resistant strains (31). Mutations have been found for almost every current antitubercular drugs (Table 1). *Mtb* has extraordinary ability to survive for long periods of time of treatment. This together with the problem of low patient compliance has led to the emergence of multiple drug resistant strains (31). The increase of multidrug-resistant strains resistant to these widely available and affordable anti-tubercular drugs is a growing clinical problem. The standard TB therapy is ineffective in

controlling TB in high MDR-TB incidence areas (34,35). Fifty million people have already been infected with drug-resistant TB (2).

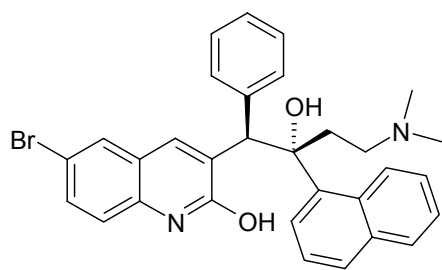
### **Promising Drug Candidates**

Facing the challenge of multiple drug resistant (MDR) and even extensively drug resistant (XDR) TB cases, it is urgent to search for new TB chemotherapies. Several new compounds that hold promise for treatment of resistant TB have been discovered by whole cell screen or optimization of previous drugs (Figure 2).

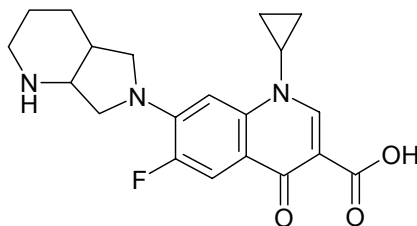
PA824, a derivative of 5-nitroimidazoles, is highly active against both *Mtb* and MDR *Mtb* with MIC as low as 0.015–0.250 µg/ml (36). PA824 is a prodrug that requires activation by a F420 dependent nitroreductase, encoded by *Rv3547* to generate an active form that could inhibit bacterial mycolic acid and protein synthesis (36). However, the activated form of PA824 and its final target remain unknown. In animal models, PA824 was active against nongrowing bacilli and its bactericidal activity is comparable to INH and Rifampin (37).

Diarylquinoline is a compound discovered by Johnson & Johnson that is highly active against mycobacteria in an *in vitro* drug screen using *M. smegmatis*. After lead optimization, one of the diarylquinolines R207910 showed a minimum inhibitory concentration (MIC) of 0.003 µg/ml for *M. smegmatis* and 0.030 µg/ml *Mtb* (38). R207910 is highly specific for mycobacteria and is much less active against other bacterial species, such as *E. coli* and *S. aureus* (MIC >32 µg/ml). Mutant screens and

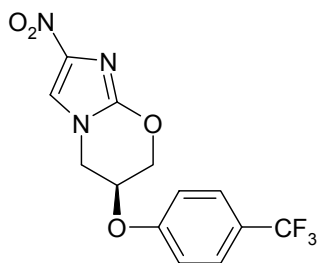
analysis have helped to identify the drug target - proton ATP synthase, which is a key enzyme for ATP synthesis (38).



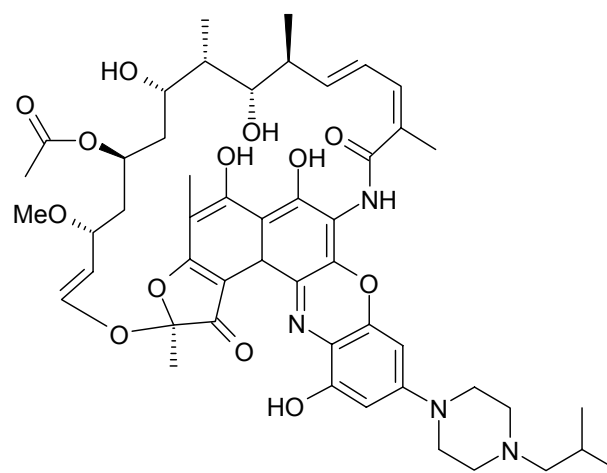
Diarylquinoline



Moxifloxacin



Niroimidazolepyran



Rifalazil

Figure 2 Chemical Structures of Some Promising TB Drug Candidates

The new moxifloxacin (MXF) and gatifloxacin (GATI) are designed based on known drugs ofloxacin and ciprofloxacin. But they are more stable and more active against *Mtb* (39). MXF has been shown to kill persisted tubercle bacilli not killed by rifampin. MXF in combination with rifampin and PZA showed higher bactericidal activity against tubercle bacilli in mice than the standard combination of INH, rifampin, and PZA and the treatment can be shortened to four months with no relapse (40). Thus, MXF has the potential to replace the standard INH, rifampin, and PZA regimen to shorten TB therapy in humans.

### **Potential Therapeutic Targets**

Transposon mutagenesis and signature-tagged mutagenesis have recently been used to identify genes essential for *Mtb* growth *in vitro* and survival *in vivo* (41,42). The identification of these essential genes provided a solid base for the search of new drug targets. Many of these essential *Mtb* genes are potentially good targets for TB drug development. However, two potential problems exist. First, the gene essentiality determined from this approach is not 100% correct. Therefore, it has to be used very carefully and most of time requires target knockout studies to verify. Secondly, the functions of a significant number of essential genes are still unknown. Further detailed biochemical characterization of target genes is necessary.

A lot of research activities are also focused on understanding the persistence mechanism of TB and developing new drugs that target the persister bacteria (43). Some genes involved in *Mtb* persistence, such as isocitrate lyase (ICL), PcaA (methyl

transferase involved in the modification of mycolic acid) and DosR (a regulator under hypoxic conditions), have been identified and could be potential targets for the development of drugs that target persistent bacilli (43-45).

## ***Mtb* CELL ENVELOPE BIOSYNTHESIS**

### **Overview**

The cell envelope of *Mtb* consists of three structural components: the inner membrane, the cell wall, and a polysaccharide-rich capsule. *Mtb*'s inner membrane, which is similar to those of other bacteria, does not have major contribution to the pathological properties. Its cell wall and outer capsule, however, are unique and are tightly associated with its pathogenicity (46). Outside the membrane, *Mtb*'s peptidoglycan (PG) layer is covalently linked to arabinogalactan (AG), which is in turn attached to the mycolic acids. This mycolate-arabinogalactan-peptidoglycan (mAGP) complex, a very hydrophobic barrier with very poor permeability, is critical to the survival of *Mtb* in macrophages and contributes significantly to the resistance against many chemotherapies (47).



### ***Mtb* Mycolic Acid Biosynthesis Pathway**

Mycolic acids are the major building blocks of the protective layer in the *Mtb* cell envelope. The cyclopropane ring structures of mycolic acid contribute to the structural integrity of the cell wall complex (48) and protect the bacillus from oxidative stress inside macrophages (49).

Mycolic acids are  $\beta$ -hydroxy fatty acids with a long  $\alpha$ -alkyl side chain (50). *Mtb* mycolic acids are formed by a saturated short fatty acyl chain of 20-26 carbon atoms and a long meromycolic acid chain of 50-60 carbon atoms. There are five forms of mycolic acids in *Mtb*,  $\alpha$ -mycolic acid, *cis* and *trans* methoxy-mycolic acid, *cis* and *trans* keto-mycolic acid (Figure 3). Among them, the  $\alpha$ -mycolic acid is the dominant species, accounting for more than 70%, whereas other forms are less than 30%. The structure and component of mycolic acids are critical to the virulence of *Mtb*. Deletion of the proximal cyclopropane ring of  $\alpha$ -mycolic acid or of methoxy- and keto-mycolates in *Mtb* was found to lead to a significant attenuation in growth of the two *Mtb* mutants in the infection studies on the mouse model (45,51), while deletion of the keto-mycolates caused restricted growth of this mutant in macrophages (52).

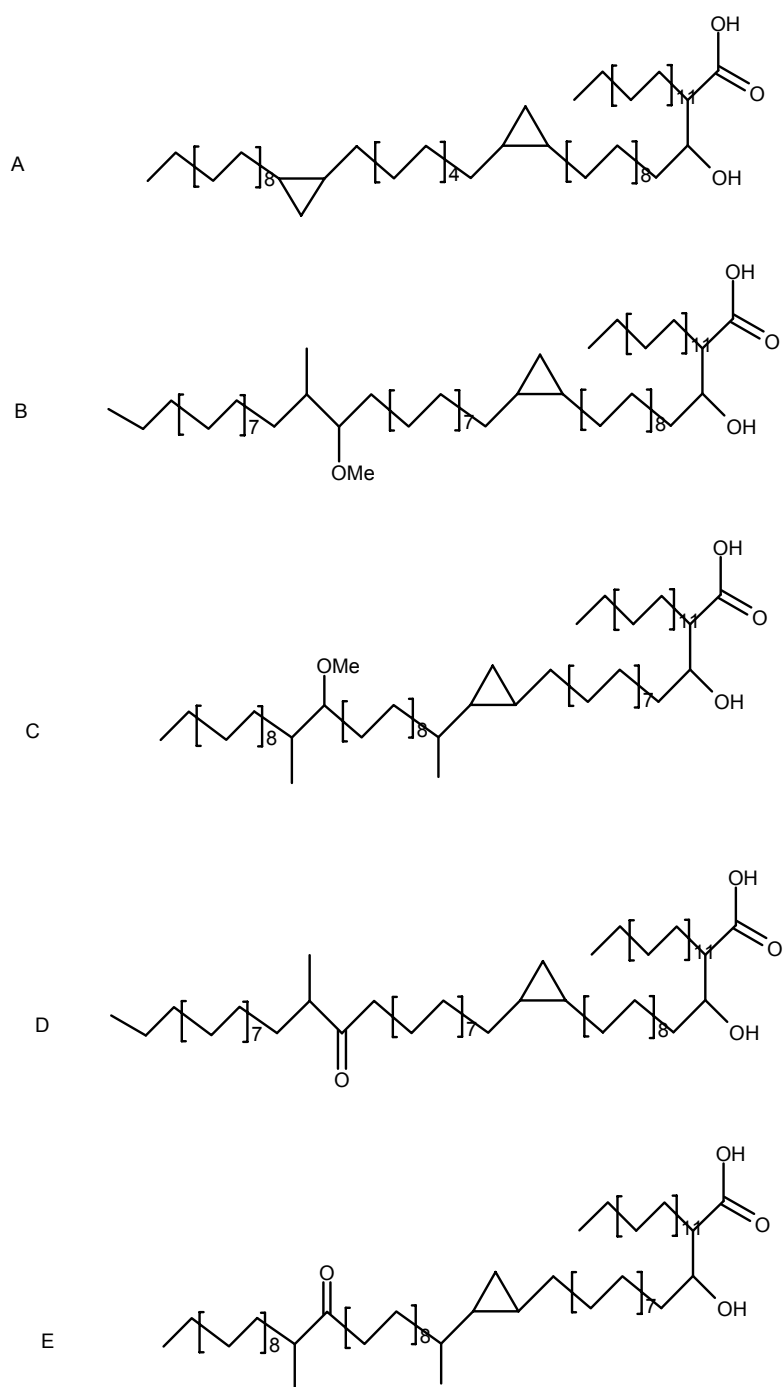


Figure 3 Chemical Structures of Mycolic Acid.

A.  $\alpha$ -mycolic acid, B. *cis* methoxy mycolic acid, C. *trans* methoxy-mycolic acid, D. *cis* keto-mycolic acid, E. *trans* keto-mycolic Acid

## 1. Biosynthesis of precursor of mycolic acid—long chain fatty acid of *Mtb*

The precursor of mycolic acid is long chain fatty acid. *Mtb* has both FAS-I and FAS-II systems. Its FAS-I, a multienzyme complex is in charge of synthesis of short chain fatty acids, and type II fatty acid synthetase (FAS-II), a discrete synthetase system continues to elongate the short chain fatty acids to long chain fatty acids.

### a) FAS-I

By sequence alignment of the *Mtb* FAS gene with other known FASs, seven functional catalytic domains have been identified in the following order: acyltransferase, enoyl reductase, dehydratase, transferase, acyl carrier protein,  $\beta$ -ketoacyl reductase, and  $\beta$ -ketoacyl synthase (53). The mycobacterial FAS-I system elongates acetyl group by two-carbon building blocks, starting with acetyl-CoA and malonyl-CoA as substrates to yield butyryl-enzyme intermediate after reaction steps of acetyltransacylation catalyzed by acyltransferase, malonyl transacylation by malonyl transferase, condensation by  $\beta$ -ketoacyl synthase,  $\beta$ -ketoacyl reduction by  $\beta$ -ketoacyl reductase, dehydration by dehydratase and enoyl reduction by enoyl reductase. After several cycles of elongation, C16- and C18- CoA products are generated and are used for the cell membrane synthesis (54). C16 and C18 products can be further elongated by *Mtb* FAS-I to generate C20 and C26 products, again as CoA derivatives. These C20 and C26 CoA products become the substrate of FAS-II systems and are further elongated to very long chain fatty acids (46).

### b) Transition steps from FAS-I to FAS-II

Malonyl-CoA has to be converted to malonyl-ACP by MtbFabD (malonyl-CoA:ACP transacylase) as the two-carbon building block for further elongation in FAS-

II system (55). MtbFabH ( $\beta$ -ketoacyl-ACP synthase-III), working as the link between the FAS-I and FAS-II systems, initiates chain elongation of C20-CoA released from FAS-I as the substrate to generate a  $\beta$ -ketoacyl-ACP product (Figure 4) (56).

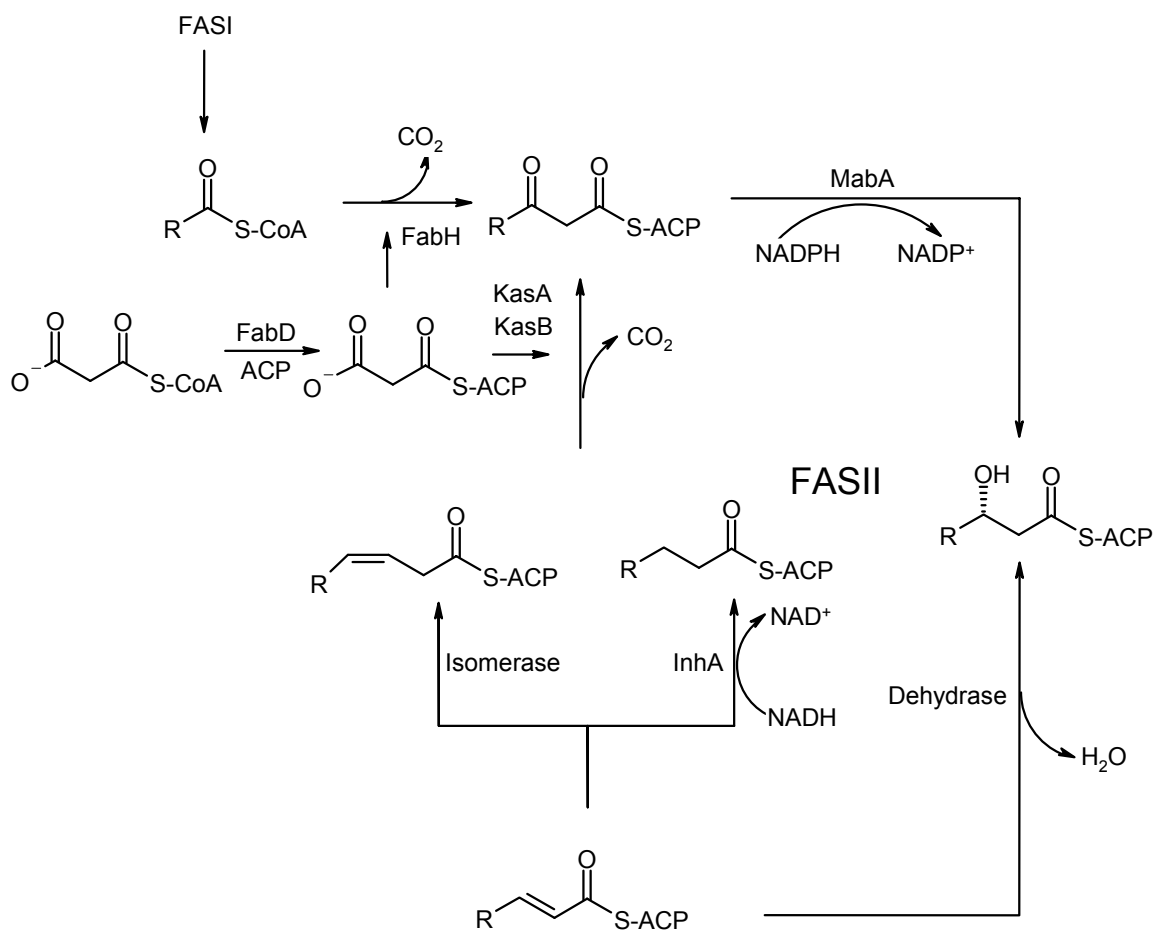


Figure 4 FASII Biosynthesis Pathway

### c) FAS-II

FAS-II of *Mtb*, a discrete synthetase system contains enzymes KasA and KasB ( $\beta$ -ketoacyl-ACP synthases), MabA ( $\beta$ -ketoacyl-ACP reductase),  $\beta$ -Hydroxyacyl-ACP dehydrase, 2-trans-Enoyl-ACP isomerase and InhA (2-trans-enoyl-ACP reductase) (46). KasA and KasB start fatty acyl elongation by catalyzing the condensation reactions (57). Although KasA and KasB share 69% sequence homology and both enzymes catalyze the condensation reaction, it is believed that they have different substrate profiles, most likely different chain length (58). MabA carries out the NADPH specific reduction of  $\beta$ -ketoacyl-ACP to  $\beta$ -hydroxyacyl-ACP.  $\beta$ -Hydroxyacyl-ACP dehydrase in *Mtb* has not been identified yet. It is supposed to dehydrate the  $\beta$ -hydroxyacyl-ACP. In *E. Coli*, two different  $\beta$ -Hydroxyacyl-ACP dehydrases *fabA* and *fabZ* have been found (59). InhA, an NADH-dependent 2-trans-enoyl-ACP reductase, catalyzes the final step of the cycle of elongation by generating saturated fatty acyl ACP. Other than dehydrase, 2-trans-enoyl-ACP isomerase has not been identified in *Mtb* either (Figure 4).

In FAS-II, KasA, KasB, MabA, dehydrase, and InhA form a cycle to catalyze the elongation of the fatty acyl chain by two carbons. 2-trans-Enoyl-ACP isomerase replaces InhA in some of the cycles as the last step enzyme to introduce *cis* double bond into the fatty acyl chain. Thus, instead of saturated fatty acid, FAS-II system generates long chain fatty acyl products with double bonds at certain positions (60). Those unsaturated long chain fatty acids will be modified into different types of meroacid in the latter modification steps.

## 2. Modification of long chain fatty acids

The functions of modification of long chain fatty acyl ACP derivatives are carried out by a group of enzymes called cyclopropane synthases and methyltransferases. Eight genes with high sequence homology *camA1*, *cmaA2*, *mmaA1-mmaA4*, *umaA1* and *pcaA* in this group have been identified. Among the enzymes, encoded by these genes, MmaA2 and PcaA convert the distal and the proximal double bonds into cyclopropane rings, respectively (Figure 3) (45). MmaA4 oxidizes the distal *cis* double bonds to a secondary alcohol with an adjacent methyl branch. MmaA1 converts the proximal *cis* double bonds to an allylic methyl branch with *trans* double bonds. SAM is the cofactor used in these methyltransferase reactions. Other modifications include the generation of *cis*-methoxy-meroacyl ACP and *trans*-methoxy-meroacyl ACP (61).

## 3. Claisen condensation to form mycolic acid

The final step in the synthesis of mycolic acid in *Mtb* is believed to be a Claisen-type condensation catalyzed by a polyketide synthase Pks13 (62).  $\alpha$ -meroacyl-AMP, produced by FadD32 and 2-carboxyl-C26-CoA derived from FAS-I are the substrates for the condensation reaction catalyzed by Pks13. Reaction starts with loading the two substrates by Pks13 and the transfer of a meroacyl group from PPB domain to the condensing domain, followed by the condensation step and reduction of the 3-carbonyl group to the secondary alcohol by a reductase to yield the mature  $\alpha$ -mycolate (63,64).

### ***Mtb* Arabinogalactan-Peptidoglycan Complex Biosynthesis Pathway**

The arabinogalactan-peptidoglycan complex biosynthesis includes the biosynthesis of peptidoglycan and arabinogalactan. Peptidoglycan, formed by polymerized  $\beta$ -1,4-linked N-acetylglucosaminyl and N-acetylmuramic acid, is the key component of a eubacterial cell wall (Figure 5) (65). The D-lactyl groups of the muramic acid are amidated with L-alanyl-D-isoglutaminyl-meso-diaminopimelyl-D-alanine stem tetrapeptides and L-alanyl-D-isoglutaminyl-meso-diaminopimelyl tripeptides. The stem peptides of adjacent glycan strands are covalently linked to form the polymeric peptidoglycan. Compared to peptidoglycan of other bacteria, Mycobacterial peptidoglycan have two distinct features; first, the muramic acid residues are N-glycolylated instead of N-acetylated as in other bacteria, and secondly, a proportion of the crosslinks in the peptidoglycan layer includes those between the diaminopimelic acid residues (so called 3-3 linkage) as well as between diaminopimelic acid and D-alanine (so called 4-3 linkage). It is thought that the N-glycolylation makes the peptidoglycan layer more rigid. Mycobacterial peptidoglycan is also highly resistant to lysozyme (47). Although, until recently, little is known about the biosynthesis of the peptidoglycan of *M. tuberculosis*, it is generally assumed that *Mtb* peptidoglycan biosynthesis is similar to that of *E. coli*, which has four sequential reaction steps (Figure 6). The initial step involves the synthesis of UDP-N-acetyl muramic acid (UDP-MurNAc) from UDP-GlcNAc by two reactions. First, enolpyruvate is added to the GlcNAc residue of the UDP-GlcNAc and subsequently the enolpyruvate moiety is reduced to form UDP-MurNAc, catalyzed by MurA and MurB, respectively. The sequential addition of L-

alanine, D-glutamate, DAP and a D-alanyl-D-alanine dipeptide to UDP-MurNAc generates UDP-MurNAc-pentapeptide. It is then linked with undecaprenyl phosphate to form Lipid I. A GlcNAc residue is subsequently added to Lipid I to produce Lipid II. MurX and MurG are believed to catalyze the formation of Lipid I and Lipid II. The Lipid II is translocated from the cytoplasm to periplasm, where it is used to assemble the peptidoglycan. The glycan chains are cross-linked by peptide bridges, catalyzed by transpeptidase. It was reported that two-thirds of the peptide bridges found in *Mtb* peptidoglycan are D, D bond between the carboxyl group of a terminal D-alanine and the amino group at the D-center of DAP (65). The rest of one third of the linkages are L, D link between the carboxyl group at the L-center of terminal DAP of one chain and the amino group at the D-center of the DAP of another peptide side chain. This suggests that there are both D, D and L, D transpeptidases in *Mtb*, however only D, D transpeptidase has been identified in *Mtb* (47). The synthesis of arabinogalactan initiated from the transfer of a GlcNAc-1-phosphate from UDP-GlcNAc to prenyl phosphate form GlcNAc-P-P-C<sub>50</sub>, which is then linked with a rhamnose (Rha). The Galf moiety is subsequently transferred to Rha-GlcNAc-P-P-C<sub>50</sub> to form (Galf)<sub>30</sub>Rha-GlcNAc-P-P-C<sub>50</sub>. Finally the Araf moiety is added to produce the (Araf)<sub>70</sub>(Galf)<sub>30</sub>Rha-GlcNAc-P-P-C<sub>50</sub>, which is used to covalently link with peptidoglycan and mycolic acid. Several genes involved in dTDP-Rha biosynthesis, such as *rmlA*, *rmlB*, *rmlC* and *rmlD*, have been identified (47).



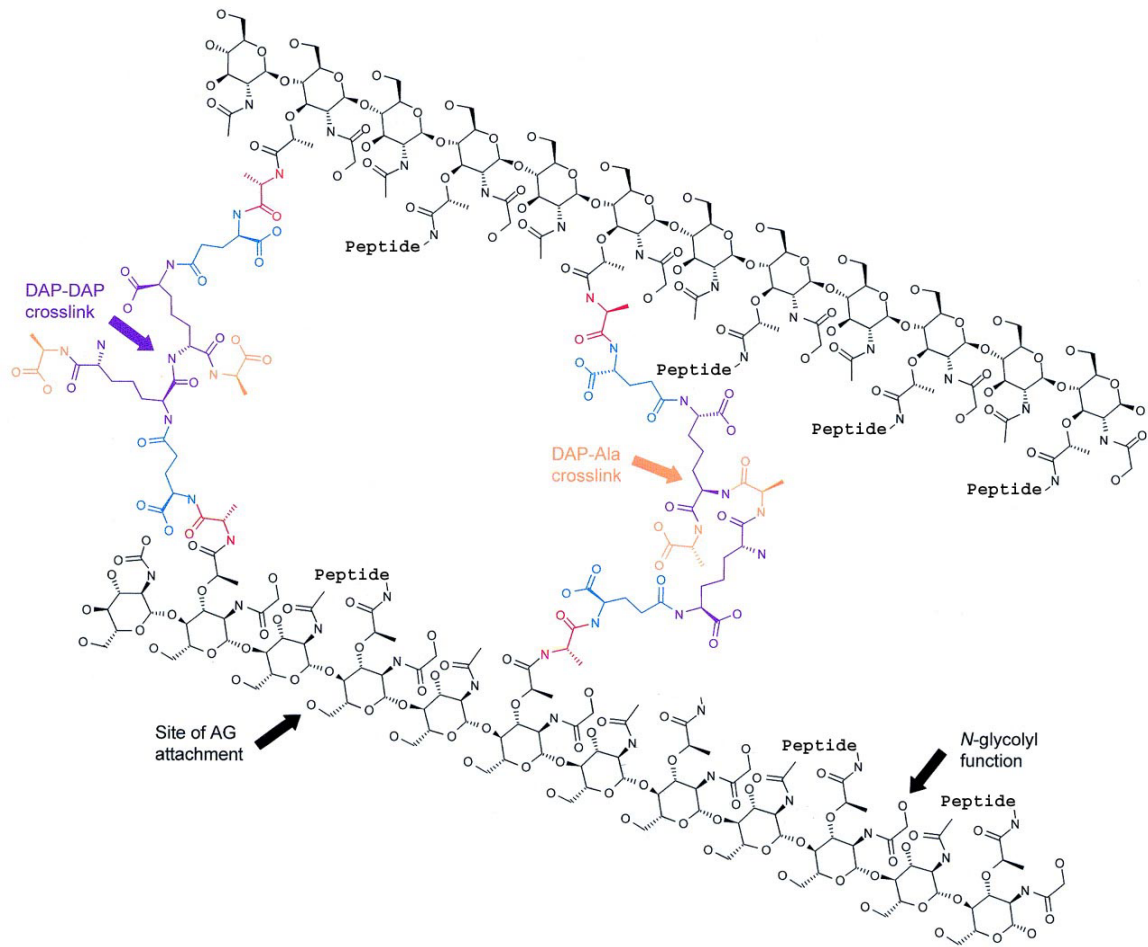


Figure 5 Structure of Peptidoglycan of *Mtb*

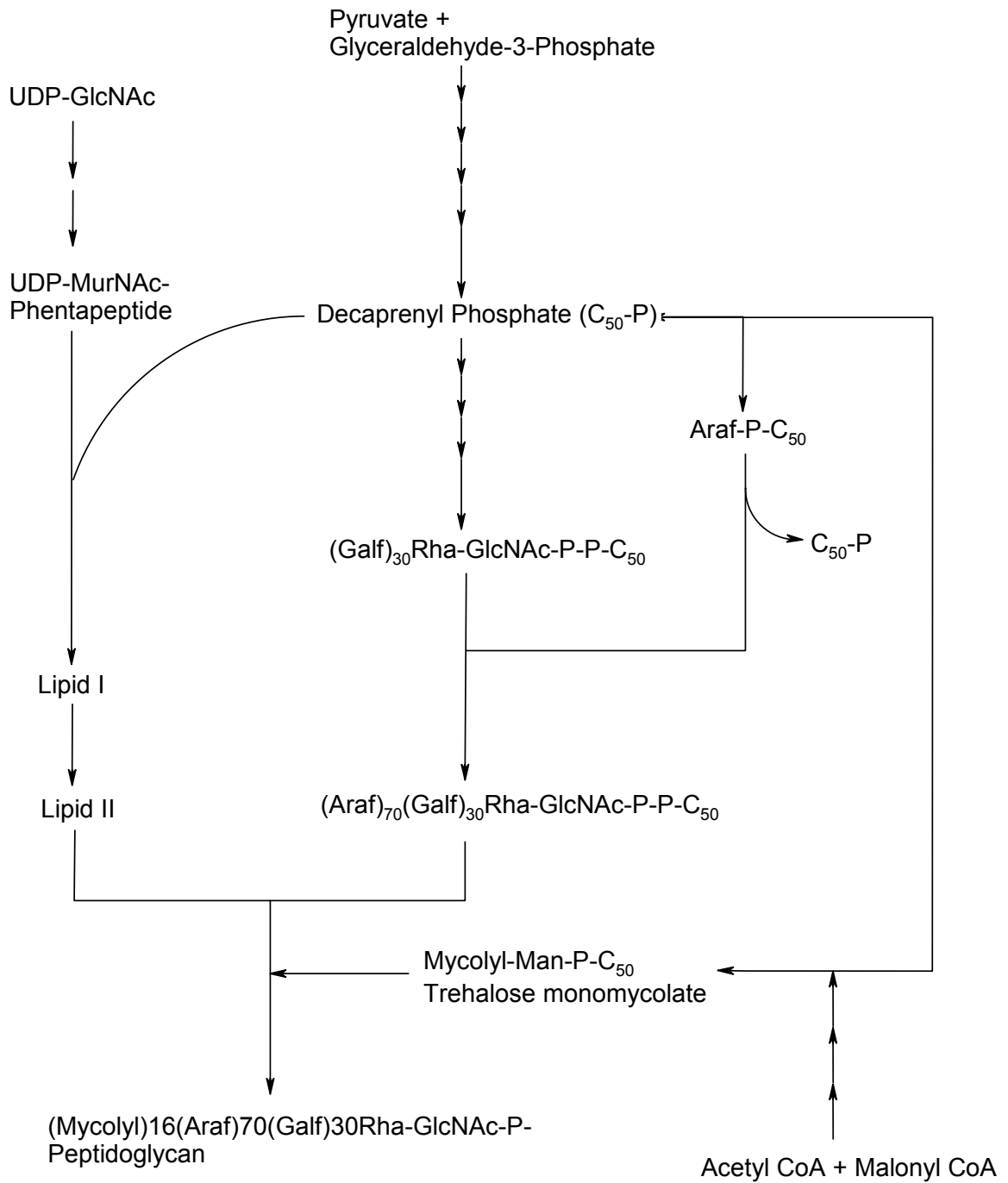


Figure 6 Biosynthesis Pathway of MAP Complex of *Mtb*

## **STRUCTURE-BASED DRUG DISCOVERY**

### **Overview**

The idea to use structural biology in drug discovery first came out in the 1970s. However, this concept became practical only after the crystal structures of real drug targets were defined. Using this method, two AIDS drugs Agenerase and Viracept were designed targeting HIV protease (66). New technical developments in protein crystallization, data collection and processing have led to a dramatic increase in the number of protein structures. There are now more than 40,000 of 3D structures in the Protein Data Bank, many of which are potential therapeutic targets ([www.rcsb.org](http://www.rcsb.org)). This provides a solid fundament of structure based drug discovery.

Before structure based drug lead search became a practical method, high-throughput screening (HTS) is the conventional approach to search for novel leads. Despite the development of HTS technology, HTS remains very costly and usually can only be afforded by big pharmaceutical companies (67). In contrast, structure based approach to lead discovery is more cost efficient and less time consuming. The structural guided lead search can either be carried out based on computational method (virtual screen) or experiment approaches (fragment based screen).

### **Virtual Screen**

Virtual screening (VS) utilizes computer programs to screen large “virtual” compound databases and selected out a small number of molecules (‘hits’) based on certain defined criteria (67). In contrast to the requirement of large real compound

libraries in the HTS, VS only needs three-dimensional coordinates of chemicals and computing power (Figure 7). Virtual screen highly relies on the development of algorithms on the calculation of receptor-ligand interactions. Many docking algorithms as well as different scoring functions have been developed (68-72). Rapidly growing computing power together with more complicated and realistic protein-ligand docking programs greatly facilitated the VS process (73,74).

A properly and accurately defined docking site of the molecular target is critical for a successful virtual screen. Once the 3D structure of the therapeutic target is available, the coordinates of the docking site (molecular receptor) have to be defined. Since most of the protein structures do not define hydrogen coordinates, the hydrogen atoms have to be added by software. After the addition of hydrogen atoms, the correct protonation and tautomeric states also need to be assigned (75). In order to remove any steric clashes introduced by the hydrogen atoms, a few cycles of energy minimization are recommended. Water molecules in the crystal structure are generally deleted, unless they have important roles in protein structures and ligand binding.

Chemical databases, such as ZINC and Chembridge chemical databases, offer atom coordinates of thousands or even millions of compounds (76). Other physical and chemical properties of the compounds provided in the database are useful to “filter” the database. The “rule of five” is the most widely used principle in compound selection (77). “Rule of five” is based on the sampling of known drugs, which claims that a drug-like molecule should have molecular weight less than 500, logP value less than 5, hydrogen donor less than 5 and hydrogen acceptor less than 10 (77). This rule provides

good prediction regarding the solubility and cell permeability of the drug. Filters are useful to exclude compounds with some functional groups that are related to poor chemical stability or known toxicity.

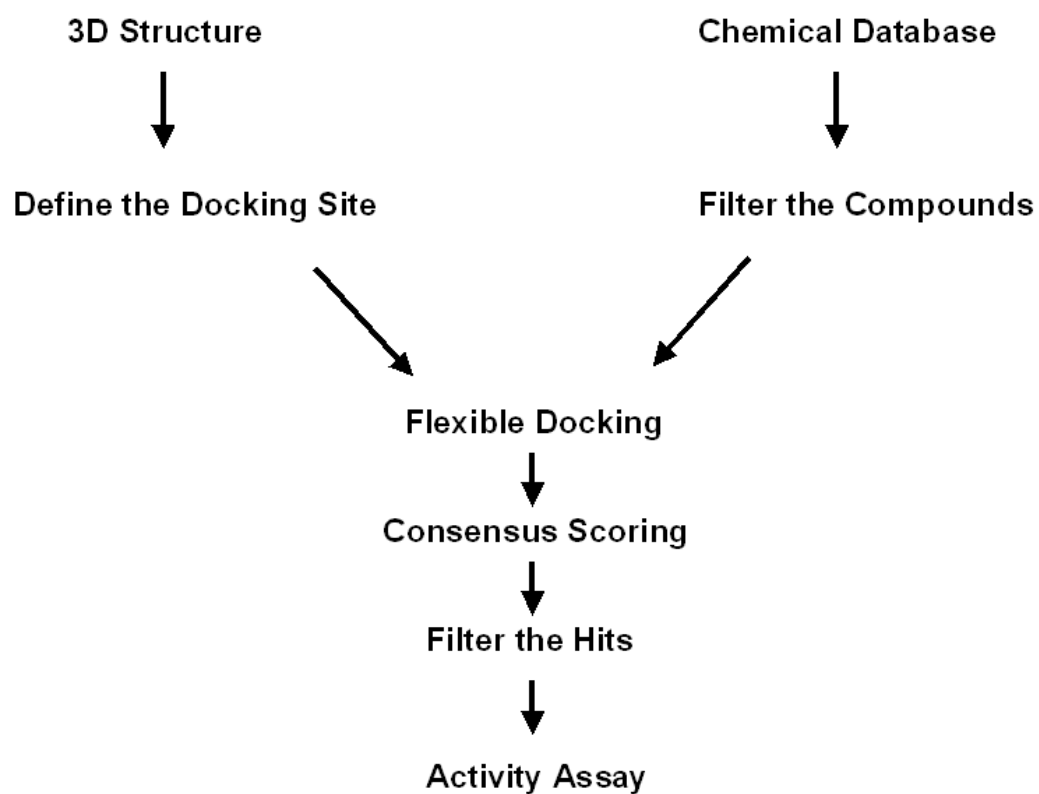


Figure 7 Virtual Screen Procedure

Due to the limitation of prediction accuracy of docking algorithms and computing power, it is impossible to put all protein and ligand conformations into consideration. Therefore, most docking programs only explore the conformation flexibility of the ligand (78). GOLD uses a genetic algorithm to generate conformers of a ligand. This algorithm is time consuming and requires a powerful grid computing system to screen large databases. FlexX uses an alternative approach called the incremental growth method, which is relatively fast and able to produce good predictions of protein-ligand complex structures. It divides the ligand into small fragments. Then an initial base fragment is selected and docked into the active site, followed by the addition of the rest of the fragments to the base fragment (79).

The virtual screen results are evaluated by scoring functions that could be generally classified as physical based, empirical or knowledge based. Physical based functions, which are used by Dock and AutoDock, calculate the free energy of binding by force fields. Empirical based functions use an additive approximation instead of a full force field to calculate energy (69). They are used in FlexX and Ludi. Empirical based score functions are faster than the physical based functions, but biased, because they are derived from binding affinity data from a small set of protein-ligand complexes. Knowledge based functions are derived from structures of protein-ligand complexes. They are fast, moreover, because they do not rely on binding affinity data and are derived from a much larger data pool, they are less biased (80). Due to the limitation of the current scoring functions, the docking result may not predict the real binding affinity accurately. Consensus score is used in order to minimize the false positive in evaluation,

in which several scoring functions are used to evaluate the docked ligands (81). Only compounds that have high scores from each scoring functions are selected out for further testing.

### **Fragment-based Screen**

In recent years, a novel approach in structure based drug design has evolved. This new approach is based on small fragment, namely fragment screening (66). Unlike typical inhibitors, fragments are typically small organic molecules (< 250 Da) and have low binding affinities (~100  $\mu$ M to 10 mM) against target proteins. Therefore, they are normally not used in HTS. The determination of interactions between these small fragments and target proteins relies on high-throughput protein crystallographic technique. Hundreds or even thousands of small molecules will be screened in order to identify fragments that bind to protein targets and then the protein-ligand complex structures will be precisely defined. Soaking method is used to quickly get suitable crystals. As long as the binding site is not blocked by the crystal packing, small molecules are usually able to diffuse into the crystals and interact as if they were in solution (82). If a fragment binds to the target, it can be identified in an electron density map. The ‘hit’ fragments will be selected out as the basis for further optimization. A synthetic chemical library will be built to explore different functional groups on the basis of the “hit” fragment. Those second generation inhibitors will then be evaluated by enzyme kinetic studies and the best inhibitor will enter into the next round of chemical modification as the new lead compound. After several cycles, a low binding affinity

initial fragment can be eventually optimized into a highly potent inhibitor. This method has been proven to be successful in the kinase drug discovery (82,83).

## **PRINCIPLE OF PROTEIN CRYSTALLOGRAPHY**

### **Overview**

X-ray crystallography is one of the major techniques developed in structural biology. Since the first crystal structure was determined in 1950s, more than 39,000 protein crystal structures have been deposited into the Protein Data Bank ([www.rcsb.org](http://www.rcsb.org)). In the past decade, the pace of protein structure determination accelerated dramatically facilitated by new techniques such as high throughput crystallization, the MAD phasing and synchrotron radiation source. Despite all the improvement in protein crystallography, the major rate limiting steps of solving protein crystal structure remains to be the growth of diffracting quality crystals and obtaining the phase.

### **Protein Crystallization**

Solving a protein crystal structure depends on the obtaining a protein crystal suitable for x-ray diffraction and data collection. Crystals are defined as solids whose components are arranged in a long-range three-dimensional order (Figure 8). In order to crystallize a protein, the purified protein undergoes slow precipitation from an aqueous solution. Individual protein molecules align themselves in a repeating series of "unit cells" periodically. The production of crystals of x-ray diffraction quality is dependent



on several factors, such as protein purity, concentration, buffer condition, temperature, and precipitants (84).

Vapor diffusion methods including hanging drop and sitting drop are often used to achieve super saturation of proteins (85). Both hanging drop and sitting drop methods use a droplet containing purified protein and crystallization solution to equilibrate with a large volume of the same crystallization solution in higher concentrations, the so called mother liquid. Initially, the droplet of protein solution contains an insufficient concentration of precipitant for crystallization, but as equilibrium is reached, the precipitant concentration increases to a level high enough for protein crystallization.

Protein crystallization usually starts from screening for initial crystallization conditions ('hits') by using random or sparse matrix screen methods. Once a hit is found, the hit will be optimized to obtain crystals with better diffraction quality and larger size. The optimization approaches usually includes varying precipitant concentration, pH, additive agents, temperature and seeding, etc.

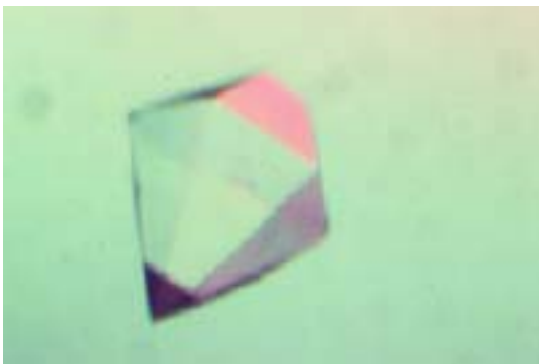


Figure 8 The Picture of an InhA Single Crystal

## **Protein Crystal Diffraction and Data Collection**

X-rays are electromagnetic radiations with wavelengths ranging from 0.1 to 1000 Å. An X-ray could be described as wave with three properties – wavelength, amplitude and phase. When incoming X-rays pass through matter, the atoms scatter the X-rays, which radiate out at the same wave as the incoming X-rays. The electrons are much more effective at scattering X-rays than protons, the total scattering from atoms is almost completely due to the electrons. In a system with multiple electrons, each electron will scatter X-rays and all these scattered X-ray waves should be summated according the vector algebra. Because crystal elements repeat periodically, only the scattered reflections that fulfill the Laue conditions are detectable in a diffraction pattern. The maxima of a diffraction pattern occur at certain positions as determined by Bragg's law.

Once suitable protein crystals become available, X-ray diffraction data are collected. The radiation source for most single crystal diffraction data collection is monochromatic X-rays from either a sealed tube generator, a rotating anode generator or from a synchrotron source. The crystal then is mounted in a loop or capillary and accurately positioned in the X-ray beam. Nowadays, data collection is typically carried out cryogenically. A cold nitrogen gas stream keeps the crystal at 100 K throughout the experiment. The incoming X-rays strike the crystal to produce a diffraction pattern, which is recorded on the X-ray detector, such as image plates or CCD area detector. During data collection three important properties: resolution, completeness, accuracy have to be optimized by adjusting the x-ray beam intensity, wavelength, detector distance, exposure time and number of images (Figure 9). Upon completing data

collection, the images with the diffraction pattern have to be processed to generate a complete reflection data file using computer program such as HKL2000, Crystal-Clear and Molsflm (86). The first step of data processing is to determine the unit cell parameters of the crystal by analyzing the diffraction pattern. Based on the unit cell parameters, the symmetry and laue group of the crystal can be obtained. All images collected can then be integrated and scaled up to produce the final scalepack data file, which contains intensity information of the diffraction in the reciprocal space.

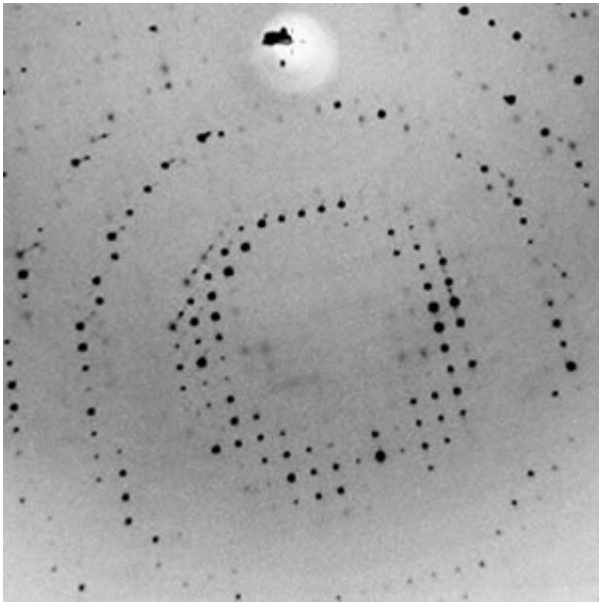


Figure 9 An Image of X-ray Diffraction of a Single Protein Crystal

## Phasing

Structure factor is a complex number, including magnitude that is proportional to the intensity measured experimentally in the X-ray diffraction experiment, and the phase angle, which cannot be obtained directly from diffraction data. Three methods, molecular replacement (MR), multiple isomorphous replacement (MIR), and multiple anomalous dispersion (MAD) are generally used to acquire the phase.

### 1. Molecular replacement

The atomic coordinates can be obtained from a known protein structure (search model) similar to the target protein. The structure factor from experimental data file can be used to calculate a Patterson map. The comparison and matching of the observed and the calculated Patterson maps is used as strategy of MR phasing. The search model is usually first rotated, then based on each rotation operation, a series of translations are performed. For each orientation, the Patterson function is calculated and compared to that from the experiment data. If a significant overlap is observed, one can assume the correct orientation of the model has been found. Once the orientation of the search model is locked, the positional parameters are determined to place the search model at the correct position in the unit cell. The correlation coefficient is used to quantify the overlap between calculated and observed Patterson maps. Molerep, Amore, EPMR and Phaser are commonly used computer programs for molecular replacement.

### 2. Multiple anomalous dispersion (MAD)

MAD method is based on one of the physical properties of elements, which is called anomalous scattering. The scattered photon has a normal component and an

imaginary component, whose phase angle is  $90^\circ$  delayed. Usually the anomalous scattering of the atom is very weak and can be ignored. However, anomalous scattering becomes significant when the incident X-ray wavelength approaches the absorption edge of the element. When anomalous scattering happens, the Friedel's law is broken, meaning the structure factor  $F(\mathbf{h}, \mathbf{k}, \mathbf{l})$  is no longer equal to its opposite phase angle structure factor  $F(-\mathbf{h}, -\mathbf{k}, -\mathbf{l})$ . MAD technique utilizes the additional information from the anomalous scattering effect to acquire an experimental phase. Because Se-Met can be easily incorporated into protein and selenium absorption edge is about  $0.9 \text{ \AA}$  which is easily available by common synchrotron sources, it is often used in the MAD experiment. Diffraction data from one single crystal under three wavelengths peak, inflection, and high remote are collected and used to determine the heavy metal sites (Figure 10). Several computer programs are currently used to determine and refine the heavy metal site, such as SOLVE, SHELXD and RANTAN (87). Once the heavy atom sites are defined, the phase of the whole protein molecule can be calculated.

### 3. Multiple isomorphous replacement (MIR)

Isomorphous replacement is the first phasing technique practical for macromolecular crystallographers. Heavy atoms such as transition metals, lanthanides, uranium, can be introduced into specific site in crystals of the native protein by soaking method. If the derivative crystal is isomorphous to the native crystal, the structure factor amplitude differences between the derivative and the native dataset can be used to locate the heavy atom sites. Similar to MAD, once the heavy atom sites are defined, the phase of the protein molecule can be deduced.

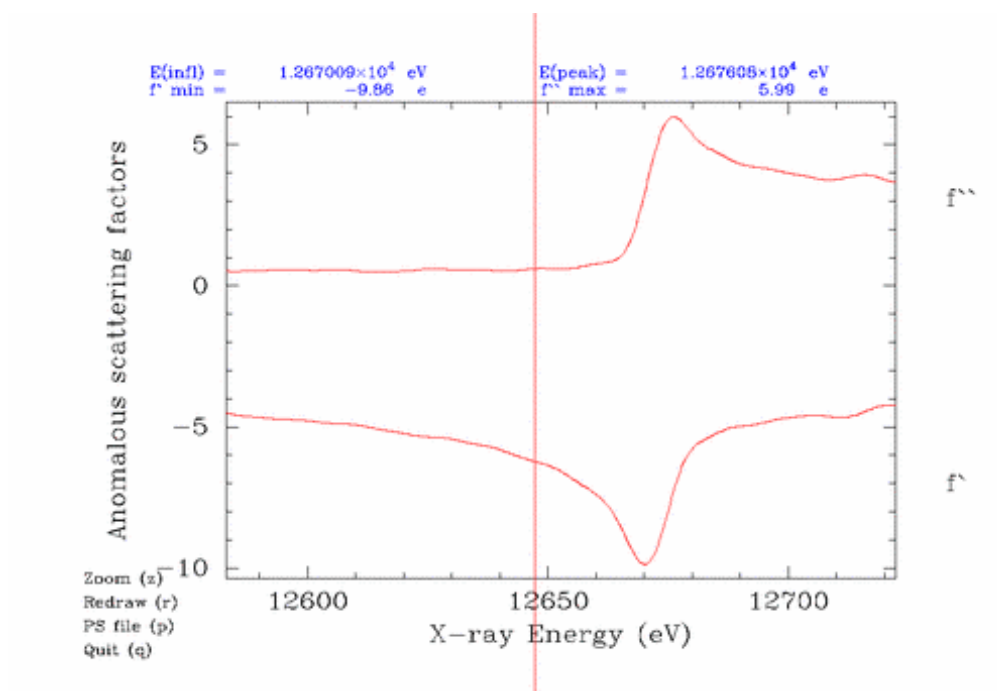


Figure 10 The Fluorescence Scan Result of Se-Met

## **Phase Improvement**

After the phase is determined, there are errors in it. Density modification is often applied to improve the phase. Solvent flattening, solvent flipping and molecular averaging are the most commonly used methods. Solvent flattening and flipping are based on the fact that in the electron density map the protein region has higher density and the solvent region is flat. Computational methods are able to define the protein-solvent boundary and flatten the density in the solvent region, which improves the density map and phase. Molecular averaging method averages the density corresponding to the symmetric molecules in one asymmetric unit to improve the phase.

Shake 'n' wARP is a computer program used to minimize the phase bias from the model in MR method by the combination of several different approaches, such as 'shaking' the model, omitting part of the structure, maximum likelihood refinement, real space placing and averaging. Shake 'n' wARP is especially useful in determination and verification of additional electron density in 2Fo-Fc map (88).

## **Model Building and Refinement**

The model building is accomplished by manually fitting molecules into the electron density with the help of computer graphic programs, such as XtalView, COOT, and O (89,90) (Figure 11). Some model-auto-building computer programs, such as ArpWarp (usually for data of resolution higher than 2.2 Å) (91) and TEXTAL (92) have also been developed and utilized in initial model building. The model is refined and

compared with the experiment diffraction data and monitored by the R factor defined as the following:

$$R = \frac{\sum ||F_{\text{obs}}| - |F_{\text{cal}}||}{\sum |F_{\text{obs}}|} \quad (\text{Eq. 1})$$

About 5-10% data are randomly selected as the free set before the first round of refinement, which is never used in the refinement, to monitor the effect of each refinement step and avoid over-refinement. The R factor calculated from free set data is defined as  $R_{\text{free}}$ , while the R factor from working data set is called  $R_{\text{factor}}$ . Only the decrease of both  $R_{\text{factor}}$  and  $R_{\text{free}}$  indicates a model improvement in refinement. The divergence of these two R factors is usually a sign of over-refinement (93).

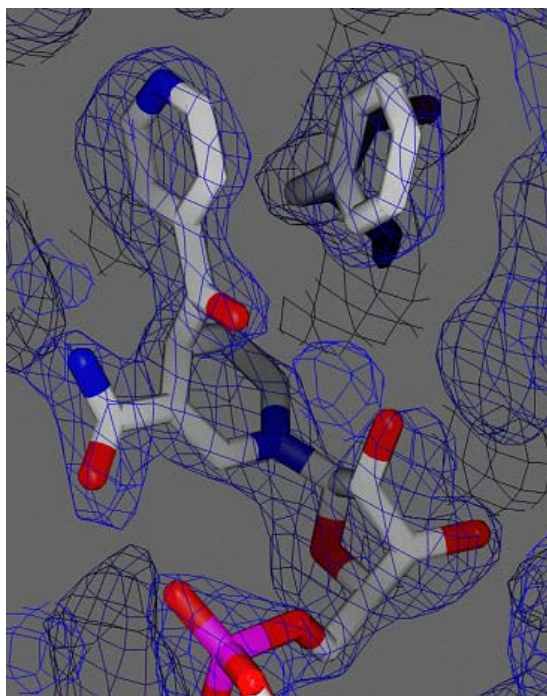


Figure 11 Model-Building Based on the Clear Electron Density



## CHAPTER II

### MECHANISM OF ACTION OF ISONIAZID AGAINST MYCOBACTERIA

#### BACKGROUND

Isoniazid (INH) was discovered to have bactericidal activity against *Mycobacterium tuberculosis* (*Mtb*) in 1952 (94). Since then, it has been used as a potent front-line drug against tuberculosis (TB) (95). The mechanism of action of INH has been studied for more than 50 years. INH was found to inhibit mycolic acid biosynthesis in *Mtb* (96). In addition, the INH-induced inhibition of mycolic acid biosynthesis was demonstrated to correlate with the bactericidal activity of INH (97). Further analysis of the lipids of INH-treated *Mtb* indicated the elongation of the fatty acid beyond C26 was inhibited, which suggested that the target of INH is an enzyme in fatty acid elongation (98). Isolated INH resistant mutants were analyzed to search for the enzymes that are related to the drug action of INH. The majority of INH-resistant strains isolated in the earlier studies were found to have lost catalase/oxidase activity (99). In 1992, *Mtb katG*, which encodes a catalase/oxidase, was shown to restore the INH susceptibility to INH resistant *M. smegmatis* mc<sup>2</sup>155 mutant strain (15). In the same study, deletion of *katG* from the chromosome was found in two INH-resistant *Mtb* clinical isolates. These studies suggested that *katG* encodes the activator of INH. Since then, 45% to 90% of the *Mtb* clinical isolates resistant to INH were found to contain mutations in the *katG* gene in different studies (16,17,100-102).

The mode of INH action remained unclear until an INH-NAD adduct was identified as the bound inhibitor in the active site of InhA, the enoyl acyl ACP reductase involved in long-chain fatty acid biosynthesis, by protein crystallography. The crystal structure of InhA bound with the adduct indicated that an isonicitinoyl moiety was covalently attached to the 4 position of the nicotinamide ring of NAD cofactor in a S configuration. The chemical structure of the INH-NAD adduct was consistent with the molecular weight obtained by the mass analysis (13). Later studies demonstrated that INH was activated by a KatG-catalyzed oxidation (103,104). One of the reaction intermediates, the isonicitinoyl free radical then attacks the nicotinamide ring of NAD to form the INH-NAD adduct, which strongly inhibits InhA ( $K_i = 5$  nM) to block mycolic acid biosynthesis (13,14,105).

Mutations within *inhA* gene or its promoter region were found in the clinical isolates resistant to INH (16,17). A S94A mutation in the *inhA* structure gene, which was originally identified in an INH-resistant *M. smegmatis* strain, was later found in three *Mtb* clinical isolates that conferred resistance to both INH and ETH (106,107). Recently, the S94A mutation allele of *inhA* has been transferred into *Mtb* by a specialized linkage transduction, which was sufficient to confer five-fold resistance to INH (14). Moreover, overexpression of *inhA* in *Mtb* was found to confer more than 10 fold resistance to INH (108). These genetic observations support that InhA is the primary target of INH.

Although the activation of INH by KatG has been studied by biochemical and structural methods (109-113), the detailed mechanism of adduct formation remains under investigation. The inhibition of InhA by INH-NAD has also been demonstrated

(105,114); however, the molecular basis of the high potency of the INH-NAD adduct has not yet been conclusively determined. Furthermore, some studies suggested that the nitric oxide released from INH oxidation is responsible for its anti-tubercular activity (115). INH, as a single probe, is not sufficient to help us fully understand and validate the mechanism of its drug action. There are many hydrazides related to INH whose antitubercular activities were studied soon after INH was discovered (94). Several of them, such as picolinic hydrazide (PNH) and benzoic hydrazide (BZH), were found to inhibit the growth of *Mycobacterium bovis* BCG strain. However, their antitubercular activity is 7.5 and 30 fold lower than that of INH, respectively, based on the MIC values (Figure 12) (94). Since those hydrazides have chemical structures similar to INH, it is very likely they share the same target and mechanism of drug action. If so, why is INH the most potent molecule to treat TB among all the hydrazides? Is there any difference in the activation step catalyzed by KatG or in the inhibitory potency to the final target, InhA?

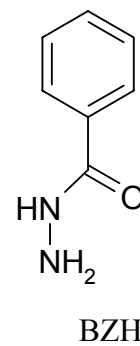
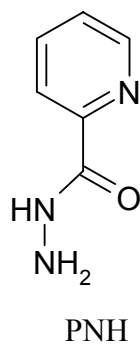
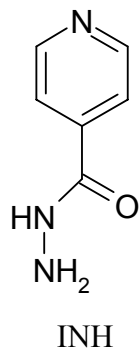
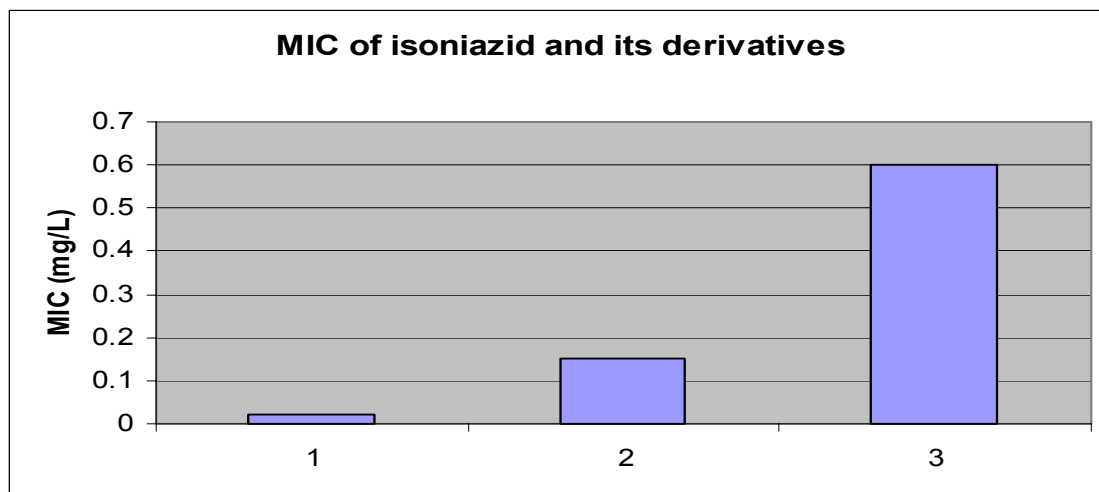


Figure 12 MIC of INH, PNH and BZH against *M. bovis* BCG Strain

## METHODS

### Cloning, Expression and Purification

*M. tuberculosis inhA* has already been cloned into *E. coli* BL21 (DE3) (116). It was cultured in Terrific Broth media containing 50 µg/ml carbenicillin at 37 °C until OD<sub>600</sub> reached 0.8. Expression of the *inhA* gene was induced for 20 h at 16 °C by addition of 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). Cells that were harvested by centrifugation were re-suspended in 50 mM PIPES (PH 6.8) and 1 mM phenylmethylsulfonyl fluoride (PMSF) and lysed with a French press. After treatment with DnaseI, the insoluble material was removed by centrifugation and the supernatant containing soluble protein was applied to HiTrap Blue Sepharose column (AP Biotech) pre-equilibrated with the same buffer, using a fast protein liquid chromatography system and eluted through a NaCl gradient (0-2 M). InhA eluted out at a NaCl concentration of 0.9 M. Fractions that contain InhA were pooled and applied to an octyl-sepharose column (AP Biotech) pre-equilibrated with 1 M NaCl and eluted out through a NaCl gradient (1-0 M). InhA fractions were pooled and a further step of gel filtration in a Superdex 200 column was performed to separate the monomeric protein from aggregated material. InhA was obtained at 40 mg/L of *E. coli* culture and appeared homogeneous by SDS-PAGE and Coomassie Blue staining.

The *M. tuberculosis* KatG was expressed and purified as previously described (110).

### **Synthesis of the INH-NAD, PNH-NAD and BZH-NAD Adducts**

Adduct of INH and NAD was synthesized using similar methods described by Meunier B. et al. (114) with some modifications. InhA (2  $\mu\text{M}$ ) was incubated with INH (500  $\mu\text{M}$ ), NAD (500  $\mu\text{M}$ ), and Mn(III) (500  $\mu\text{M}$ ) in 2 ml sodium phosphate (100 mM, pH 7.5) buffer for 1 hour. After InhA was inhibited according to the activity assay below, the enzyme solution was concentrated and applied to a Superdex 200 column pre-equilibrated with 50 mM Phosphate buffer (PH 7.5). The InhA in complex with INH-NAD was then concentrated and heated for 60 seconds under 100°C. After the heat treatment, INH-NAD was separated from denatured enzyme by filtration, using a centricon (cutoff 30 KDa). The concentration of INH-NAD was determined by its absorbance at 260 nm or 326 nm ( $\epsilon_{260} = 27 \text{ mM}^{-1}\text{cm}^{-1}$ ,  $\epsilon_{326} = 7.0 \text{ mM}^{-1}\text{cm}^{-1}$ ). BZH-NAD and PNH-NAD were synthesized and purified using the same method.

### **InhA Activity Assay**

All assays were carried out using a Cary 100 Bio Spectrophotometer at 25 °C by monitoring oxidation of NADH at 340 nm. Reactions were initiated by adding substrate, dodecenoyl-CoA (50  $\mu\text{M}$ ), to assay mixtures containing InhA (1 nM), NADH (100  $\mu\text{M}$ ) and hydrazide adduct inhibitors (3 - 2000 nM).

The  $\text{IC}_{50}$  was determined from dose-response plot of enzyme fractional activity as a function of inhibitor concentration.  $K_i$  was obtained by dividing the  $\text{IC}_{50}$  value by  $1 + [\text{S}_1]/K_{m1} + [\text{S}_2]/K_{m2}$ , where  $[\text{S}_1]$  and  $[\text{S}_2]$  are the concentration of dodecenoyl-CoA and NADH and  $K_{m1}$  and  $K_{m2}$  are their Michaelis constants.

S94A was assayed under the same condition and evaluated using the same method.

### **Adduct Formation by KatG Catalysis**

The biomimetic approach for adduct formation was carried out according to Zhao et al. (113). Briefly, INH or hydrazide analogs (50  $\mu\text{M}$ ), NAD (50 $\mu\text{M}$ ), *M. tuberculosis* KatG (0.5 $\mu\text{M}$ ) and  $\text{H}_2\text{O}_2$  (2  $\mu\text{M}/\text{min}$ , supplied by glucose/glucose oxidase system) were incubated in phosphate buffer (pH 7.2) at 25  $^\circ\text{C}$ . Adduct formation was monitored by increase in absorbance at 326 nm (14DS UV-VIS spectrophotometer, AVIV Biomedical). To eliminate background absorbance, the reference cuvette contained all reagents except NAD. The solutions were collected after 30 min incubation and were centrifuged to remove enzymes using a Centricon filter (30 kD cut-off). InhA-inhibitory activity of the product was also subsequently measured as above described. (Note: The experiment was performed by Xiangbo Zhao in Richard Magliozzo's group as a collaboration.)

### **Reduction of KatG Compound I by INH and Its Analogs**

Double mixing experiments were carried out using a stopped-flow spectrophotometer (HiTech Scientific Model SF-61DX2) according to Yu et al. (117). Briefly, resting ferric KatG was mixed with peroxyacetic acid to produce Compound I intermediate, followed by the addition of INH or its analogs after a 2 sec delay. The reduction of Compound I was monitored by the change of absorbance at 407 nm, which

returns to the value of the resting enzyme during this process. (Note: The experiment was performed by Xiangbo Zhao in Richard Magliozzo's group as a collaboration.)

### **Crystallization of InhA in Complex with Hydrazides-NAD**

Due to low yields of the adducts using the protocol with KatG, inhibitors used for co-crystallization were produced by chemical activation using manganese as described above. Crystallization was accomplished by the hanging drop vapor diffusion method (85). InhA in complex with inhibitors was co-crystallized in hanging droplets containing 2  $\mu$ l of protein solution at 10 mg/ml and 2  $\mu$ l of buffer (12% MDP, 4% DMSO, 0.1M Hepes, 0.025M Sodium Citrate) at 16 °C in Linbro plates against 1 ml of the same buffer. Diamond shaped protein crystals formed ~4 days later.

### **Crystallization of InhA S94A:INH-NAD**

Mutant S94A InhA proteins were co-crystallized with INH-NAD using the same method as wild type InhA.

### **Data Collection and Processing**

Data were collected at 121 K using cryo-protection solution containing reservoir solution with additional 30% MPD. Crystal of InhA:INH-NAD diffracted X-rays to 1.9 Å at beam line 14BMC at the Advanced Photon Source (APS), Argonne National Laboratory. Crystals of InhA:BZH-NAD and InhA:PNH-NAD diffracted X-rays to 2.6 Å and 2.2 Å respectively, using a MacScience DIP2030 image plate detector coupled to



a Rigaku X-ray generator utilizing a copper rotating anode ( $\text{CuK}_\alpha$ ,  $\lambda = 1.54 \text{ \AA}$ ). Crystals of InhA(S94A):INH-NAD diffracted to  $1.9 \text{ \AA}$  at beam line 19ID at the Advanced Photon Source (APS) in Argonne National Laboratory. Diffraction data were collected from a single crystal with one-degree oscillation widths for a range of  $180^\circ$ . The data were integrated and reduced using HKL2000 (86) (Table 2).

### **Structure Determination and Model Refinement**

Crystals produced from InhA in complex with INH-NAD, PNH-NAD and PNH-NAD were isomorphous with those of the native enzyme. Initial phases were obtained by molecular replacement using the apo-InhA structure (1ENY) and refined with CNS (118) (Table 2).  $F_o - F_c$  and  $2F_o - F_c$  electron density maps were calculated and additional density resembling the inhibitor was found. The ligand was fit into the additional density and the whole model was rebuilt using XtalView (89). During the final cycles of the refinement, water molecules were added into peaks above  $3\text{-}\sigma$  of the  $F_o - F_c$  electron density maps that were within hydrogen-bonding distances from appropriate protein atoms.

For the model of InhA(S94A):INH-NAD, the apo S94A structure (1ENZ) was used as the initial model. All other procedures of model building and refinement were the same as for InhA:INH-NAD. The final model of InhA(S94A):INH-NAD was refined to  $R_{\text{factor}} 20.1\%$ ,  $R_{\text{free}}, 22.8\%$ .

Table 2 Data Statistics of InhA in Complex with Hydrazide-NAD Adducts

Data collection				
	InhA:INH-NAD	InhA:BNH-NAD	InhA:PNH-NAD	S94A:INH-NAD
Maximum resolution (Å)	1.9	2.6	2.2	1.7
Space group	P6(2)22	P6(2)22	P6(2)22	P6(2)22
a (Å)	98.2	97.7	97.4	98.3
b (Å)	98.2	97.7	97.4	98.3
c (Å)	139.6	139.6	139.8	139.2
$\alpha$ (°)	90.0	90.0	90.0	90
$\beta$ (°)	90.0	90.0	90.0	90
$\gamma$ (°)	120.0	120.0	120.0	120
Unique reflections	31950.0	18441.0	32900.0	34383
Rsym (%)	11.4	16.0	7.6	10.1
Completeness (%)	99.8	97.7	98.0	90.5
Redundancy	13.2	11.0	7.8	14.3
I/ $\sigma$	22.4	12.5	23.8	23.7
Refinement statistics				
Resolution range (Å)	40.8-1.9	42.3-2.6	40.0-2.2	33.7-1.9
Number of reflections	25437.0	11543.0	19558.0	28957
Number of atoms / subunit				
Protein	1994.0	1994.0	1994.0	1993
Ligand (s)	52.0	52.0	52.0	52
Rcryst (%)	20.8	21.6	22.0	20.1
Rfree (%)	26.0	26.0	26.8	22.8
Average B-factors (Å <sup>2</sup> )	31.5	35.3	37.0	31.8

### **Minimum Inhibitory Concentration (MIC) Determination**

The *M. tuberculosis* H37Rv strains were obtained from laboratory stocks and grown in Middlebrook 7H9 medium (Difco) supplemented with 10% (v/v) OADC enrichment (Difco), 0.2% (v/v) glycerol, and 0.05% (v/v) Tween 80 to an  $OD_{600nm} \approx 1.0$ . The cultures were diluted 4 logs, and 0.1 ml of the diluted cultures were inoculated into 2 ml of Middlebrook 7H9 media containing the following drug concentration: 0, 0.1, 0.5, 1, 5, 10, 50, 75, 100, 200 and 500  $\mu\text{g/ml}$ . The cultures were incubated while shaking at 37°C for 4 weeks. The MIC was determined as the concentration of drug that prevented mycobacterial growth. (Note: The experiment was performed by Catherine Vilcheze in William Jacobs's group as a collaboration.)

## **RESULTS**

### **InhA(S94A) Protein Is More Resistant to INH-NAD Than the Wild Type InhA**

In order to study the molecular basis for the resistance of the *inhA* (S94A) mutant strain, the recombinant InhA (S94A) protein was expressed and purified. To determine if the InhA (S94A) enzyme is resistant to inhibition by the INH-NAD adduct, inactivation of purified InhA and S94A by this adduct was measured in a dose-dependent fashion. The InhA(S94A) enzyme was 17 times more resistant to inhibition by the INH-NAD adduct based on  $IC_{50}$  values, and showed a 30-fold increase in the  $K_i$  value for the INH-NAD adduct (Table 3). To determine the structural basis for the reduced binding of the INH-NAD adduct to the InhA(S94A) enzyme, the InhA(S94A) protein was co-crystallized in complex with the INH-NAD adduct (Figure 13a) and compared to the

wild-type protein. Although the position and orientation of the INH-NAD adduct was nearly identical in the active sites of the wild-type (Figure 13b) and InhA(S94A) proteins (Figure 13c), the loss of the hydroxyl group of serine residue resulted in the disruption of a water molecule and a perturbation in the hydrogen bonding network, which likely decreases the binding of the adduct to the InhA(S94A) protein. The biochemical and structural results clearly indicated that a single amino acid modification dramatically changed the binding of the INH-NAD adduct to InhA. This is consistent with the genetic observation that this single point mutation conferred resistance to isoniazid. Altogether, it establishes the precise molecular mechanism by which INH inhibits InhA, and supports the premise that *inhA* encodes the primary target of INH.

Table 3 Inhibition of InhA and S94A by INH-NAD

	IC <sub>50</sub> (nM)	K <sub>m</sub> DD-CoA (μM)	K <sub>m</sub> NADH (μM)	K <sub>i</sub> (nM)
InhA	19 ± 10	46	66	5 ± 3
S94A	324 ± 41	104	250	172 ± 22

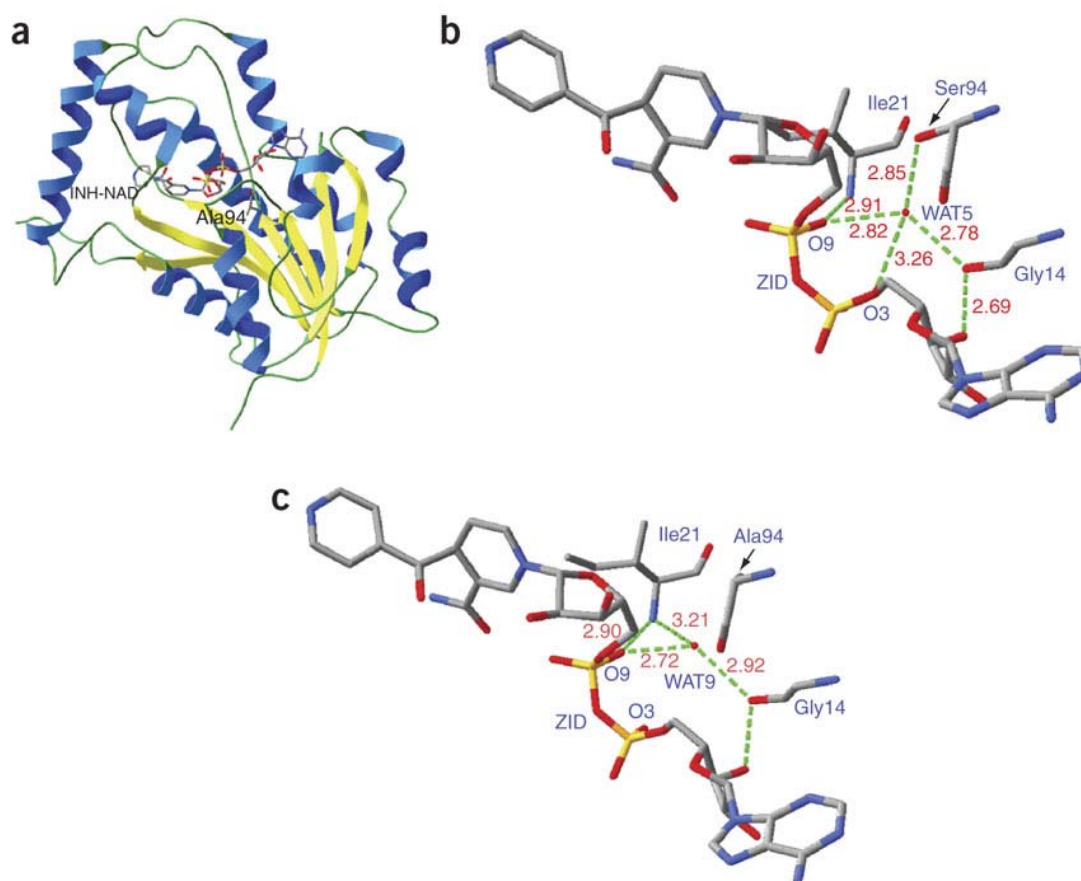


Figure 13 Crystal Structure of S94A Bound with the INH-NAD Adduct.

(a) Ribbon diagram of the S94A mutant protein with the INH-NAD adduct. (b) In the InhA:INH-NAD structure, the O9 of the phosphate of the INH-NAD adduct forms one hydrogen bond with the main-chain nitrogen atom of Ile21 and one hydrogen bond with a well-ordered water molecule. (c) In the S94A:INH-NAD structure, this hydrogen-bonding network is disrupted by the loss of the hydroxyl group in the S94A substitution

### Antibacterial Activity of INH Analogs against *Mtb* Strains

BZH, PNH, and INH were tested against three *Mtb* strains: H37Rv, a wild-type *Mtb*; H37Rv *inhA* (S94A), a *Mtb* strain containing a single point mutation in *inhA* and H37Rv *katG* (del g<sup>371</sup>), a *Mtb* strain containing a single base pair deletion in *katG* (Table 4). Mutations that prevent the activation of INH as in H37Rv *katG* (del g<sup>371</sup>) result in INH resistance while the S94A mutation in *inhA* prevents the binding drug-NAD adduct to its target, InhA, leading to INH resistance. The analogs tested are structurally similar to INH (BZH has no nitrogen in the aromatic ring while in PNH, the hydrazide group is located at the ortho position instead of the para position in INH). Each analog contains a hydrazide group that could be activated by the catalase/peroxidase KatG to form a hypothetical acyl radical. As shown in Table 4, PNH had an MIC 40 times higher in the H37Rv *inhA* (S94A) strain and 100 times higher in the H37Rv *katG* (del g<sup>371</sup>) strain suggesting that this compound is also activated by KatG and targets InhA. BZH was poorly active against wild-type *Mtb* and the *inhA* mutant strain, and slightly less active against the *katG* mutant strain suggesting that this compound might also be activated by KatG in mycobacterial cultures.

Table 4 The MICs of Hydrazides against *Mtb* Wild Type and Mutant Strains

Strains	INH (mg/ml)	PNH (mg/ml)	BZH (mg/ml)
H37Rv	0.075	5	500
H37Rv <i>inhA</i> (S94A)	0.5	200	500
H37Rv <i>katG</i> (del g <sup>371</sup> )	>25	500	>500

### The INH-, BZH- and PNH-NAD Adducts Are Potent Inhibitors of InhA

Previous studies have shown that INH is a *pro*-drug that requires activation by KatG to form an INH-NAD adduct that inhibits InhA. We hypothesized that BZH and PNH have the same mechanism of action, due to their structural similarity to INH. To test this hypothesis, the INH-NAD, BZH-NAD and PNH-NAD adducts were synthesized using Mn(III) as catalyst (Figure 14). Like INH-NAD, both BZH-NAD and PNH-NAD are very potent inhibitors of InhA. The inhibition constant of INH-NAD has been determined by several groups. In this report, the inhibition constants of INH-NAD, BZH-NAD and PNH-NAD were quantified, respectively, using the same assay conditions and analytical method (Table 5). The difference in inhibition constant for these three species was about two-fold, which is not significant compared to the experimental error. InhA:BZH-NAD and InhA:PNH-NAD complexes were isolated by size exclusion or ion exchange column chromatography as was done for the InhA: INH-NAD complex. The isolated enzyme remained inactive, suggesting that all three acyl-NAD adducts are tight binding inhibitors of InhA.

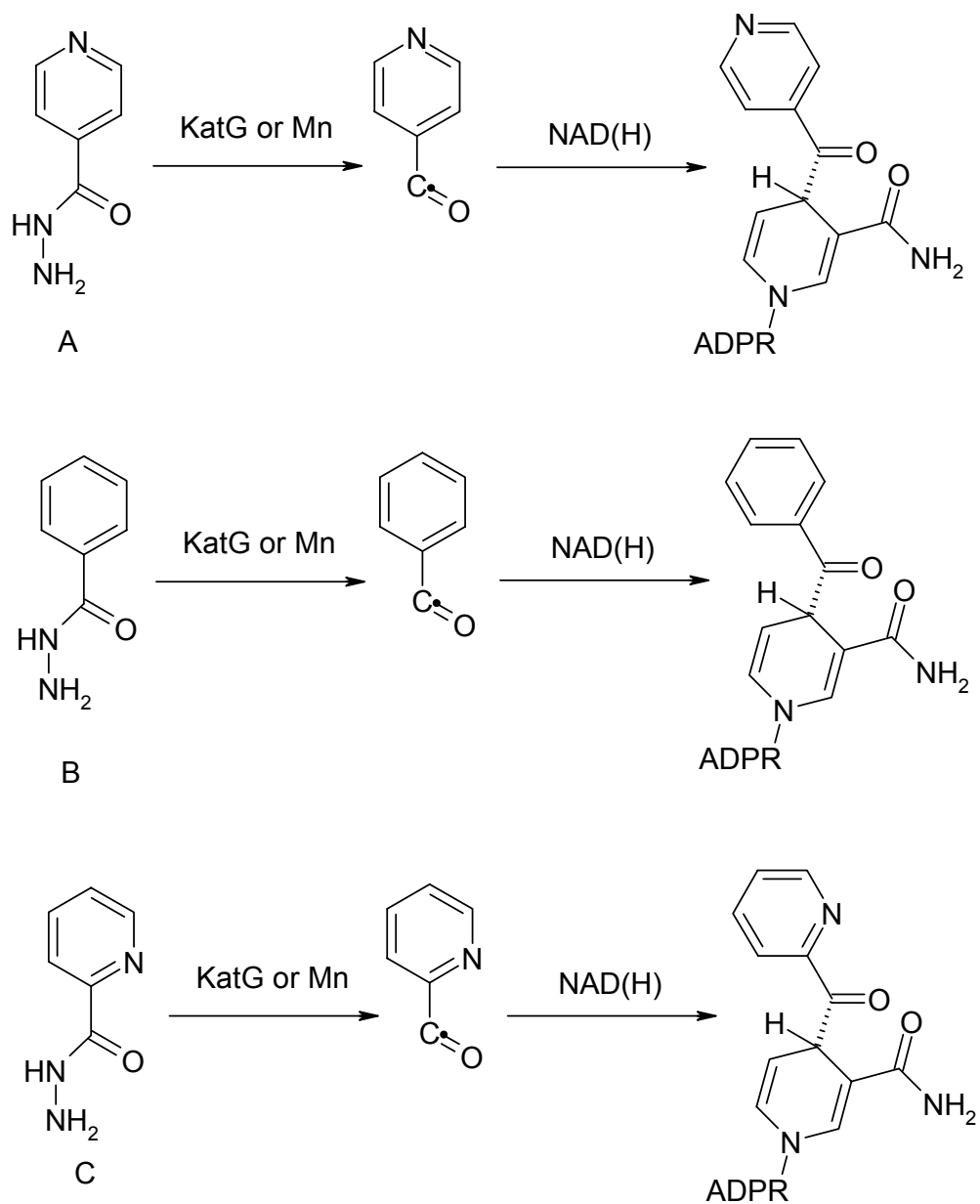


Figure 14 Synthesis of INH-NAD, BZH-NAD, and PNH-NAD



Table 5 Inhibition of InhA by Hydrazide-NAD Adducts

Inhibitors	IC <sub>50</sub> (nM)	K <sub>m</sub> DD-CoA (μM)	K <sub>m</sub> NADH (μM)	K <sub>i</sub> (nM)
INH-NAD	19 ± 10	46	66	5 ± 3
BZH-NAD	42 ± 30	46	66	11 ± 8
PNH-NAD	34 ± 19	46	66	9 ± 5

### Structures of InhA Complexes Revealed Similar Binding Modes in the Active Site

In order to find out the structural basis of the inhibition of INH-NAD, BZH-NAD and PNH-NAD against InhA, crystal structures of InhA in complex with these adducts were solved. Each structure includes 269 protein residues and one ligand.

A 2.8 Å InhA:INH-NAD structure has previously been published (13). Due to the resolution limitation, there was ambiguity of the real form of INH-NAD adduct. The 1.9 Å electron density map of the newly solved InhA:INH-NAD structure clearly showed that the open-form adduct with 4S configuration at carbon 4 of nicotinamide ring is the inhibitor bound in the enzyme active site. The INH-NAD inhibitor binding site includes the same nucleotide binding site observed in the structure of InhA:NADH, and an extended small pocket which contains the isonicotinic acyl group of the inhibitor. The small pocket, that is below and partially overlapped with the fatty acyl substrate-binding site, is composed of residues Tyr<sup>158</sup>, Phe<sup>149</sup>, Met<sup>199</sup>, Trp<sup>222</sup>, Leu<sup>218</sup>, Met<sup>199</sup>, Met<sup>155</sup>, Met<sup>161</sup>, and Pro<sup>193</sup>. Among them, Tyr<sup>158</sup>, Pro<sup>193</sup> and Gly<sup>192</sup> are conserved in other ENR homologues. The isonicotinic acyl group of the adduct points toward the entrance of the small portal. The pyridyl group of the adduct forms an aromatic stacking interaction with the flipped phenyl ring of residue Phe<sup>149</sup> and the pyridyl nitrogen atom forms a hydrogen bond with a buried water (TIP18s). The structures of the InhA:BZH-NAD and InhA:PNH-NAD complexes appear to have binding modes similar to InhA:INH-NAD based on structural alignment. The phenyl ring of BZH and the picoline ring of PNH all stack with the flipped phenyl ring of the sidechain of Phe<sup>149</sup> at a distance of 3.6 Å. Most of the interactions between ligand and protein are conserved in all three complex

structures with a few exceptions (Figure 15). For instance, in the structure of InhA:INH-NAD, the pyridyl nitrogen atom forms a hydrogen bond with a conserved water, which does not exist in InhA:BZH-NAD and InhA:PNH-NAD structures. PNH-NAD forms one unique hydrogen-bonding interaction between its pyridine nitrogen and the hydroxyl group of Tyr<sup>158</sup>. Nevertheless, the minor structural difference does not affect the potency of the inhibitors (Table 5).

### **KatG Mediated InhA Inhibition Assay of INH, BZH and PNH**

Even though hydrazide-NAD adducts could be conveniently synthesized using inorganic catalysts, the yield of adduct generated from this approach could not truly reflect an enzyme mediated process. To better mimic the *in vivo* activation of INH and its analogs, biomimetic studies were carried out to examine the KatG catalyzed adduct formation. As shown in Figure 16, the rate of INH-NAD formation is the highest and remains steady during the time period monitored, while the formation rate of PNH-NAD decreases with time and the formation of BZH-NAD is not detectable using this method. BZH was found to be activated by KatG at a rate below 2% of INH in a separate study, which is consistent with the present results (119). The reaction mixtures from the aforementioned activation were also collected, removed from KatG, and examined for their InhA-inhibitory activity. InhA was most effectively inhibited by the products derived from INH, slightly inhibited by those of PNH, and almost not affected by BZH. These observations suggest that the difference in anti-tubercular activity among INH,

PNH, and BZH results from their varied abilities to form hydrazide-NAD adduct, catalyzed by KatG.

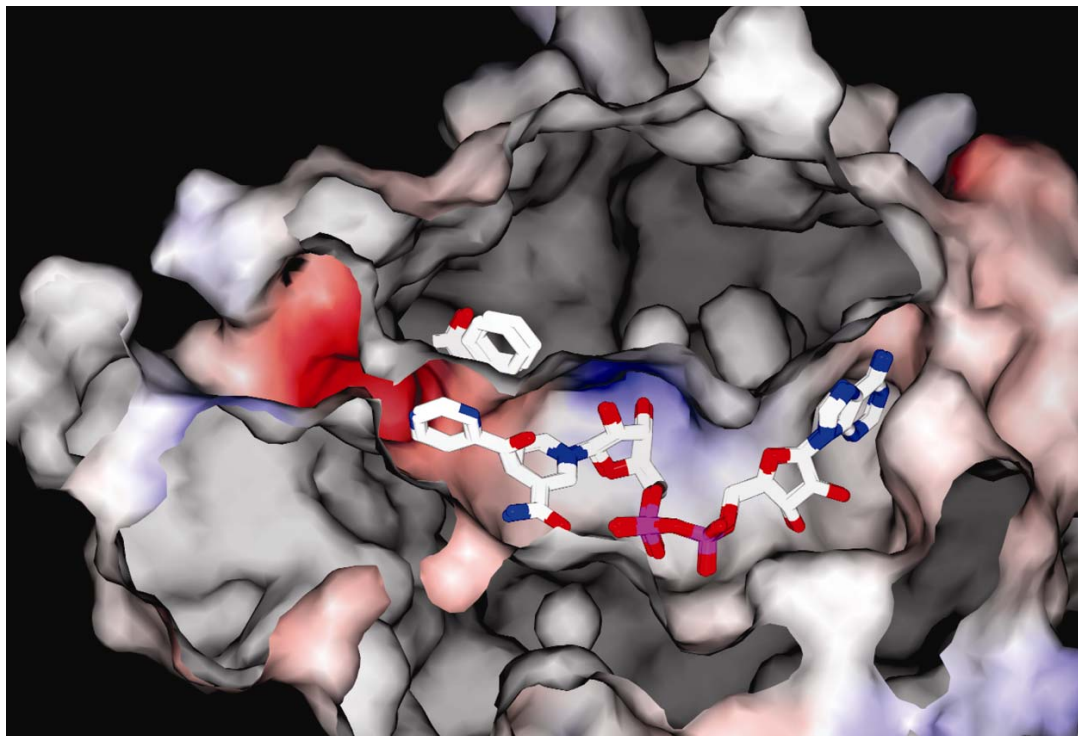


Figure 15 Surface Diagram of InhA in Complex with Hydrazide-NAD Adducts

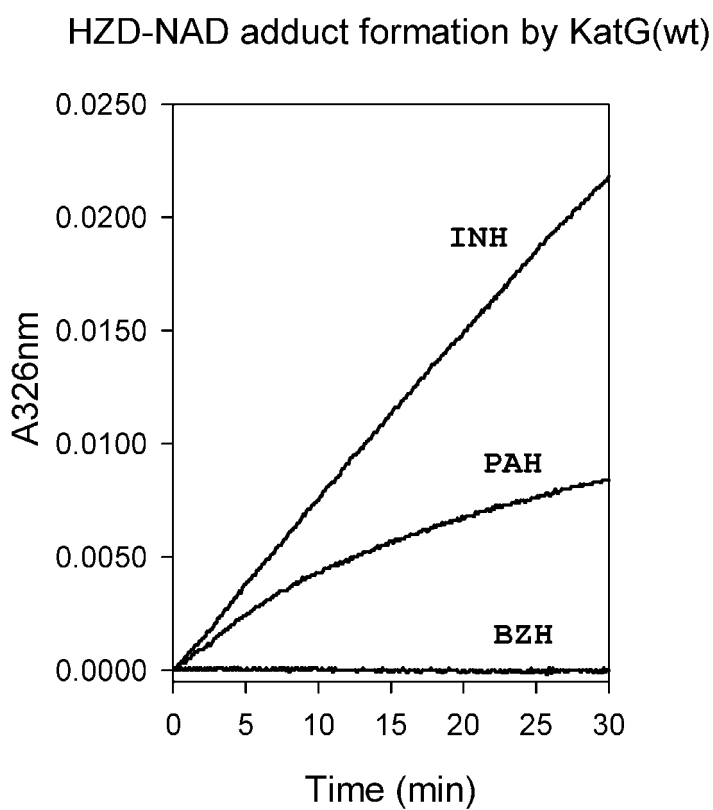


Figure 16 Formation of Hydrazide-NAD Adduct Catalyzed by KatG.

INH (50  $\mu\text{M}$ ),  $\text{NAD}^+$  (50  $\mu\text{M}$ ), KatG (0.5  $\mu\text{M}$ ) and  $\text{H}_2\text{O}_2$  (2  $\mu\text{M}/\text{min}$ , supplied by Glucose/Gox) were incubated in Phosphate buffer pH 7.2 at 25  $^\circ\text{C}$ . Absorbance at 326 nm was monitored

### **The Difference of Adduct Production Rates Is Not due to Different Oxidation Rates of the Hydrazides**

To examine in more detail the reason that INH and its analogs have different yields of adduct, the reactions between INH, BZH, PNH, and KatG Compound **I** were examined using double-mixing stopped-flow spectrophotometry, respectively. KatG Compound **I** has been demonstrated to be catalytically competent in INH activation (117). Reduction of Compound **I** by the hydrazides restored the Soret peak to its value in resting enzyme. Thus, the reaction of hydrazides with Compound **I** could be monitored by the absorbance increase at 407 nm. As shown in Figure 17, INH, PNH and BZH all react with KatG Compound **I**. Interestingly, the reaction rates for these three substrates decrease in the order BZH >> PNH > INH (Figure 17). This trend is opposite to that of their rates of adduct formation. It has been suggested that the product of the oxidation step is a radical intermediate; the subsequent reaction of this radical species with NAD is likely the rate-limiting step in adduct formation. NAD is relatively large compared to the main substrate access channel in KatG. Therefore, a radical intermediate must be stable for long enough to encounter NAD. The different aromatic structures in INH, PNH, BZH would affect the stability of the radicals.

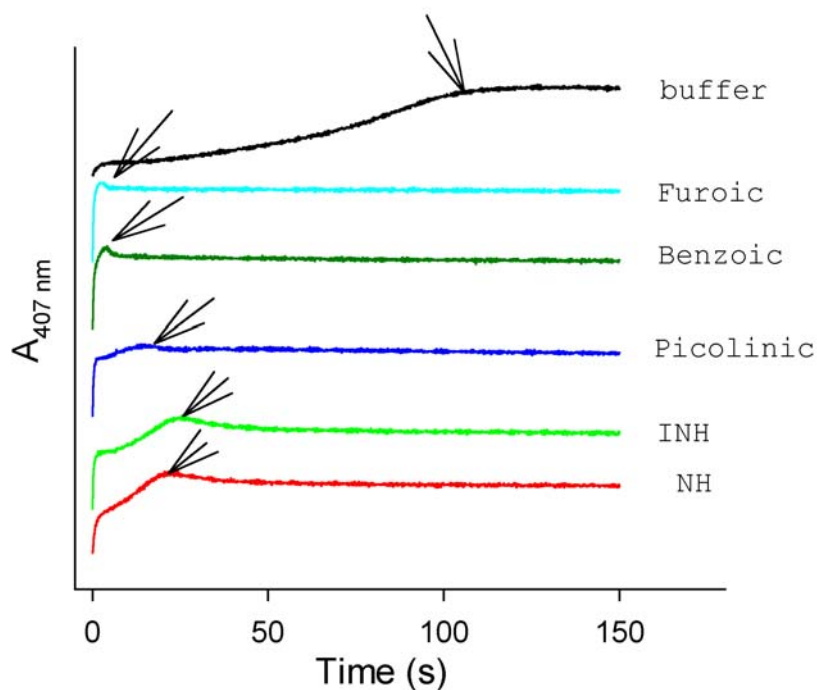


Figure 17 Reduction of KatG Compound I by INH and Its Analogs.

Resting enzymes were pre-reacted with peroxyacetic acid (PAA) to form Compound I, followed (two seconds delay) by addition of INH. The final concentrations of enzymes and PAA were 5 and 50  $\mu\text{M}$  respectively; the final concentrations of INH and analogues are 200  $\mu\text{M}$ . Arrows indicate the time points when KatG Compound I (low absorbance at 407nm) was reduced back to resting enzyme by hydrazide

## DISCUSSION

### Why INH Is the Superior Prodrug against TB among the Hydrazides

The INH analogs were investigated to probe the activation and inhibition mechanism of INH. BZH and PNH were activated by Mn(III) and formed adducts with NAD(H) similar to INH. Furthermore, the BZH-NAD and PNH-NAD were as potent as INH-NAD against InhA *in vitro*. The structures of InhA:BZH-NAD and InhA:PNH-NAD complexes were determined and compared to the InhA:INH-NAD structure. The binding modes of BZH-NAD and PNH-NAD were similar to that of INH-NAD, which correlates well with their high inhibitory activities against InhA.

However, despite all the similarities of their activated forms, the MIC values of BZH and PNH showed much less antitubercular activity than INH. The MICs of PNH and BZH on *Mtb* H37RV strain are 50 and 5000 fold higher than that of INH, respectively. To understand this tremendous difference between the *in vivo* antitubercular activity and *in vitro* inhibition activity, we have to go back to the mechanism of drug action of INH. In the cell, INH is activated by a catalase-peroxidase (KatG) and forms an INH-NAD adduct, which inhibits InhA. A Drug with this mechanism must fulfill at least two requirements to exert high antitubercular activity. First, the active form of the drug must be a potent inhibitor of InhA. Second, the formation of adducts catalyzed by KatG must have reasonable yields. Since BZH-NAD and PNH-NAD are potent inhibitors of InhA, we hypothesized that the reason for their low antitubercular activity is the low yields of formation of adducts with NAD(H) catalyzed by KatG. Formation rates of INH-NAD, BZH-NAD, and PNH-NAD obtained



in the KatG activation study strongly support this hypothesis. KatG catalyzed reaction did not yield sufficient BZH-NAD or PNH-NAD adducts to cause potent inhibition against InhA *in vitro*. The order of oxidation rates of the hydrazides by KatG Compound **I** was opposite to that of the adduct production rates. Therefore, the difference of adduct production rate was not caused by different oxidation rates of the hydrazides. Based on these results, we believe the most important factor for the adduct production yield is not the formation rate of the free radical intermediate but its stability. Since the pyridyl-acyl free radicals are more stable than phenyl-acyl free radicals, the production rate of INH-NAD is much higher than that of BZH-NAD. As the free radicals involved in the formation of INH-NAD and PNH-NAD have approximately the same stability, the difference in production rate must be attributed to a different mechanism. Based on our results, the production rate of PNH-NAD continuously decreases over time while the production of INH-NAD remains unchanged. It is likely PNH or its oxidation products inhibit KatG catalysis. This leads to the tremendous difference in the final yields of adducts between INH and PNH.

In addition, the oxidation of all three hydrazides tested in the experiments releases nitric oxide as one of the products. Since the oxidation rate of INH was the lowest compared to other hydrazides, the rate of formation and releasing of its product nitric oxide could not at least be faster than other hydrazides if not slower. This indicates that the product nitric oxide from oxidation of isoniazid should not be responsible for its anti-tubercular activity.

In summary, the high antitubercular activity of INH compared to its analogs is not due to a structurally unique moiety in the acyl-NAD inhibitor of InhA, but rather the high yield of its activated form; this is provided by two important factors: 1) long life time of the free radical intermediate; 2) no significant product inhibition to KatG.

### **Other Putative Targets**

Although genetic and biochemical studies have provided convincing evidence that InhA is the primary target of INH, other putative targets of INH have been proposed as well (120-122). Recently, an INH-NADP adduct was shown to strongly inhibit *Mtb* dihydrofolate reductase (DHFR) *in vitro* ( $K_{i,app} = 1$  nM) (123). This INH-NADP adduct was synthesized by incubating INH and  $NADP^+$  in the presence of Mn(III) catalyst. In the Mn catalyzed reaction, the product is a mixture of six adducts (two of them are acyclic and the other four are cyclic) (114). Unlike INH-NAD adduct, which can be synthesized by the KatG mediated reaction, the formation of INH-NADP adduct catalyzed by KatG *in vitro* has not been demonstrated (103). This INH-NADP adduct was found to compete with both substrates for DHFR dihydrofolate and NADPH, which is similar to the inhibition mechanism of INH-NAD against InhA. Recombinant DHFR was co-crystallized with synthesized INH-NADP adduct mixtures to find out which form of INH-NADP adduct has the highest binding affinity to the enzyme. The complex structure indicated that an acyclic 4R INH-NADP adduct was selectively bound in the active site of DHFR. The structural modeling study also suggested the 4S enantiomer would not fit in the binding pocket. To demonstrate that the inhibition of DHFR by this

4R INH-NADP adduct is relevant to the killing of *Mtb* by INH *in vivo*, it was further tested if overexpression of *Mtb dfrA* could protect *M. smegmatis* from INH inhibition. In the previous study, over-expression of *dfrA* in *M. smegmatis* caused a two fold increase of resistance to INH compared to the wild type. Based on these observations, Argyrou, et. al. claimed that *Mtb* DHFR, which is inhibited by an INH-derived 4R acyclic INH-NADP adduct, is also a target of INH.

In the previous study, a synthetic INH-NADP adduct derived from INH demonstrated strong inhibition and binding to *Mtb* DHFR *in vitro*, which is the biochemical evidence for Argyrou, et. al. to link DHFR with the anti-tubercular activity of INH (123). However, the INH-NADP adduct in the previous study was synthesized using inorganic catalysts, Mn(II) chloride. Thus, the yield of adduct generated from such approach could not truly reflect an enzyme mediated process inside the cell. To better mimic the *in vivo* activation of INH, a cell-based activation system was designed to examine the KatG catalyzed adduct formation and its binding to DHFR. This *E. coli*-based activation system is similar to the one used previously to activate the prodrug ETH and PTH (21). In this system, *katG* and *dfrA* were co-expressed in *E. coli* in the presence of INH to investigate if the activated drug would inhibit DHFR. To construct this system, *katG* and *dfrA* were co-transformed into the *E. coli* BL21(DE3) strain and selected by 50 µg/ml kanamycin and carbenicillin. The cells that exhibited high level of expressions of both KatG and InhA were selected out for further experiments. The cell growth was tested at series of concentrations of INH. 100 µg/ml of INH was found to be the highest concentration that would not cause inhibition of *E. coli* cell growth. The *E.*

*coli* strain containing *katG* and *dfrA* genes was grown and induced in the presence and absence of 100  $\mu\text{g/ml}$  INH, respectively. After the co-expressions of both genes were confirmed by SDS-PAGE, recombinant KatG and DHFR proteins were readily purified.

Mass spectroscopy method was used to investigate if DHFR purified from the experiment sample was bound with any inhibitor. Mass spectroscopy is being widely used to characterize both small compounds and macromolecules. We have applied this method to the identification of INH-NAD, ETH-NAD and PTH-NAD adducts bound to InhA (13,21). Before the mass spectroscopy experiment, the purified DHFR was concentrated to 5mg/ml. The protein was then denatured by heating at 100 °C for 60 seconds, followed by filtration (cutoff size = 3000 Da) to separate the small molecules from the denatured protein. MALDI mass spectrum of the filtrate on the range between 200 to 1200 Da was carefully analyzed. We were not able to identify any compound that has a molecular weight corresponds to the INH-NADP adduct. The peaks shown on the spectrum were most likely originated from small molecules in the buffer solution or some small fragments of the protein resulted from proteolytic activity.

An enzyme assay was performed to determine the activity of purified DHFR. DHFR isolated from the experimental sample was found to be fully active (specific activity of 12  $\mu\text{mol mg}^{-1}\text{min}^{-1}$ ) and showed no sign of inhibition compared with the enzyme purified from expression in the absence of INH. Since the 4R acyclic INH-NADP adduct is extremely potent against DHFR *in vitro*, it would tightly bind to DHFR, if the adduct is indeed generated by KatG catalysis inside the cell. However, the activity assay result and the mass analysis demonstrated that no detectable amount of INH-

NADP adduct had bound to DHFR. To confirm that the lack of the INH-NADP adduct did not result from the absence of KatG activity, the co-expressed KatG from the same experimental sample was purified and assayed for its activity *in vitro*. The specific catalase activity of isolated KatG was  $17 \text{ mol mg}^{-1}\text{min}^{-1}$ , which is comparable to the published data (activity of  $21 \text{ mol mg}^{-1}\text{min}^{-1}$ ) (113). Therefore, it is clear that 4R acyclic INH-NADP adduct is not an activated INH product generated by KatG catalysis inside the *E. coli* cell-based system.

To demonstrate a target of a bactericidal drug, it is necessary to show that binding of the drug to the putative target causes inhibition of biological activity, which leads to the death of the cell. In the case of INH, it is more complicated to elucidate its mechanism of action, since INH is a prodrug. It is necessary to not only demonstrate that the putative target is relevant to the bactericidal activity of INH, but also show how the active form of INH inhibits the target protein. The previous work postulated that the *Mtb* DHFR is the target of INH mainly based on two independent observations: 1) a 4R INH-NADP adduct synthesized from INH by non-enzymatic approach showed strong inhibition to DHFR *in vitro*; 2) Overexpression of *dfrA* in *M. smegmatis* conferred resistance to INH. However, the two observations are not relevant to each other, unless it can be demonstrated that the 4R INH-NADP adduct is a product of INH generated inside the cell by KatG catalysis, which is absent in the previous study. By coexpressing *katG* and *dfrA* genes in the *E. coli* cells in the presence of INH, we showed that the DHFR protein isolated from the experimental sample was not bound with the INH-NADP adduct. This demonstrated that the 4R INH-NADP adduct is not generated by KatG

catalysis. KatG is so far the only identified activator of INH and it is very unlikely that INH could be activated by another unknown protein to form the INH-NADP adduct in mycobacteria. Therefore we can further assert that the synthetic INH-NADP adduct is not biologically relevant to INH.

## CHAPTER III

### MECHANISM OF THIOAMIDE AGAINST MYCOBACTERIA

#### BACKGROUND

Thioamide drugs, ethionamide (ETH) and prothionamide (PTH), are clinically effective in the treatment of *Mycobacteria* infections caused by *M. tuberculosis*, *M. leprae* and *M. avium* complex (124,125). ETH and PTH are both bacteriocidal and interchangeable in a chemotherapy regimen. They are the most frequently used drugs for the treatment of drug resistant TB and therefore are becoming increasingly relevant as the number of multi-drug resistant (MDR) and extensively drug resistant (XDR) cases is increasing worldwide (126). ETH and PTH are also used in a combined chemotherapy regimen with either dapsones or rifampin to treat leprosy (127). While we have previously speculated about the mechanism of action of ETH in *M. tuberculosis* based on analogy to Isoniazid (INH)'s mode of action, definitive biochemical evidence that ETH targets InhA has not been forthcoming. ETH and PTH are structurally similar to the front line drug isoniazid (Figure 18), and it is known that all of these drugs inhibit mycolic acid biosynthesis. Banerjee et al. first demonstrated that the gene *inhA*, which encodes an essential enzyme for mycolic acid biosynthesis, was the target of INH and ETH (107). Indeed, several *M. tuberculosis* isolates resistant to INH contain mutations in the *inhA* gene and all have been found to be cross-resistant to ETH (106). An experimentally induced single amino acid mutation S94A of *inhA* was sufficient to confer resistance to both ethionamide and isoniazid in *M. tuberculosis* and *M. bovis*

BCG (128). These observations suggested that the primary target of both isoniazid and ethionamide was InhA, the enoyl-acyl ACP reductase involved in mycolic acid biosynthesis. The same enzyme has been previously identified as the target of INH (13,103,116).

Like INH, ETH is also a prodrug that requires activation to exert antitubercular activity. However, KatG mutant strains resistant to INH are still sensitive to ETH, demonstrating that ETH must have a different activator (106,129). Later on, it was found that mutations of gene *Rv3854c* (*ethA*) were repeatedly found in the clinical isolates resistant to ethionamide. The overexpression of *ethA* in *M. smegmatis* resulted in significantly increased ethionamide sensitivity (20). This evidence suggested that *ethA* is critical for the activation of ETH.

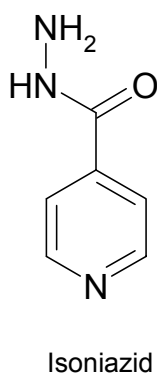
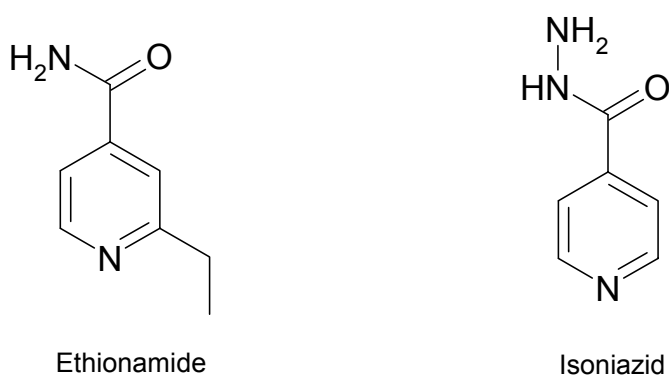


Figure 18 Chemical Structures of Ethionamide and Isoniazid



## METHODS

### Cloning, Expression, and Purification

The *Mtb inhA* has been previously cloned (116). *Mtb ethA* was cloned from genomic DNA (C.S.U. N01-AI-75320). The amplified product was inserted into pET28b using the NdeI and NotI restriction sites. *M. leprae ethA* and *inhA* were cloned from genomic DNA. The amplified product of *M. leprae ethA* was inserted into pET15b using the NdeI and BamHI restriction sites. *M. leprae inhA* was inserted into pET30b using the NdeI and HindIII restriction sites.

The plasmids of *Mtb inhA* and *ethA* were singly and doubly transformed into *E. coli* BL21 (DE3) (EMD Bioscience #69387-3). The strain containing plasmids of *inhA* and *ethA* was cultured in LB-Miller media containing 50 µg/ml kanamycin and 50 µg/ml carbanicillin at 37 °C until OD<sub>600</sub> reached 0.8. Expression of both genes was induced for 20 hours at 16 °C by addition of 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). At the same time of induction, 100 µg/ml of ethionamide or prothionamide was also added to the culture. The same protocol was used for the strain containing just the *inhA* plasmid.

Cells were harvested by centrifugation and re-suspended in 50 mM PIPES (PH 6.8) and 1 mM phenylmethylsulfonyl fluoride (PMSF) prior to lysing with a French press. After treatment with DnaseI, the insoluble material was removed by centrifugation and the supernatant containing soluble protein was applied to a HiTrap Blue Sepharose column (AP Biotech) pre-equilibrated with the same buffer, using a fast protein liquid chromatography system, and eluted through a NaCl gradient (0-2 M). Fractions that

contain InhA were pooled and applied to an octyl-sepharose column (AP Biotech) pre-equilibrated with 1 M NaCl and eluted with a NaCl gradient (1-0 M). InhA fractions were pooled and a further step of gel filtration in a Superdex 200 column was performed to separate the monomeric protein from aggregated material. The protein appeared homogeneous by SDS-PAGE and Coomassie Blue staining.

The co-expression and purification of *M. leprae ethA* and *inhA* were conducted using a protocol similar to *Mtb*.

### **Isolation and Characterization of ETH-NAD and PTH-NAD**

InhA purified from the experimental strain containing both *inhA* and *ethA* genes was then concentrated and heated for 40 seconds under 100°C. After the heat treatment, ETH-NAD or PTH-NAD was separated from denatured enzymes by filtration, using a centricon device (cutoff size 30 KDa). The concentration of ETH-NAD and PTH-NAD were determined by its absorbance at 260nm and 326nm (103). The molecular weight of both adducts was determined by MALDI, carried out on an ABI Voyager-DE STR: for ETH-NAD, calculated 797.2 and found 797.3 (negative mode); calculated 799.2 and found 799.2 (positive mode); for PTH-NAD calculated 811.2 and found 811.3 (negative mode).

### **InhA Enzymatic Activity Assay**

All assays were carried out on a Cary 100 Bio Spectrophotometer at 25 °C by monitoring oxidation of NADH at 340 nm. Reactions were initiated by adding substrate

dodecenoyl-CoA (50  $\mu$ M) to assay mixtures containing InhA (1 nM), NADH (100  $\mu$ M), and adduct inhibitors (3 - 2000 nM).

The  $IC_{50}$  was determined from the dose-response plot of enzyme fractional activity as a function of inhibitor concentration.  $K_i$  was obtained by dividing the  $IC_{50}$  value by  $1 + [S_1]/K_{m1} + [S_2]/K_{m2}$ , where  $[S_1]$  and  $[S_2]$  are the concentration of dodecenoyl-CoA and NADH and  $K_{m1}$  and  $K_{m2}$  are their Michaelis constants.

### **Crystallization of InhA in Complex with ETH-NAD Adduct**

Crystallization was accomplished by the hanging drop vapor diffusion method (85). *Mtb* InhA in complex with inhibitors were co-crystallized in hanging droplets containing 2  $\mu$ l of protein solution at 10 mg/ml and 2  $\mu$ l of buffer (12% MDP, 4% DMSO, 0.1M Hepes, 0.025M Sodium Citrate) at 16 °C in Linbro plates against 1 ml of the same buffer. Diamond shaped protein crystals formed ~4 days later.

*M. leprae* InhA in complex with inhibitors were co-crystallized in a similar manner, and the crystal has a cubic shape.

### **Data Collection and Processing**

Data were collected at 121 K using cryo-protection solution containing reservoir solution with additional 30% MPD. Crystals of *Mtb* InhA:ETH-NAD and *M. leprae* InhA:PTH-NAD diffracted X-rays to 2.2 Å and 1.8 Å using the beam line 23ID at the Advanced Photon Source (APS), Argonne National Laboratory. Diffraction data were collected from a single crystal with one-degree oscillation widths for a range of 120°.

Crystals of *Mtb* InhA:PTH-NAD diffracted to 2.5 Å using a Raxis image plate detector coupled to a Rigaku X-ray generator utilizing a copper rotating anode ( $\text{CuK}\alpha$ ,  $\lambda = 1.54$  Å). The data were integrated and reduced using HKL2000 (Table 6) (86).

### **Structure Determination and Model Refinement**

Crystals produced from InhA in complex with ETH-NAD were isomorphous to those of the native enzyme. Initial phases were obtained by molecular replacement using the apo-InhA structure (1ENY) and refined with CNS (Table 6) (118).  $F_o - F_c$  and  $2F_o - F_c$  electron density map were calculated and additional density resembling the inhibitor was found. The ligand was fit into the additional density and the whole model was rebuilt using XtalView (89). During the final cycles of the refinement, water molecules were added into peaks above  $3 \sigma$  of the  $F_o - F_c$  electron density maps that were within hydrogen-bonding distances from appropriate protein atoms. The final refinement statistics were listed in Table 6.

### **Control Experiment**

To ensure that over expression of *ethA* did not interfere with the activity of the co-expressed InhA, over expression of both *ethA* and *inhA* in the absence of ETH was used as one control (designated control 1). Furthermore, over expression of only the target protein InhA under the same concentration of ETH was used as control 2 for the experiment to verify that EthA is critical for the activation of ethionamide. InhA was

isolated from both controls and assayed for the activity. Mass analysis was performed on purified InhA to identify any bound adduct inhibitors.

Table 6 Data Statistics of *Mtb* and *M. leprae* InhA Complexes

	<i>Mtb</i>		<i>M. leprae</i>
	InhA:ETH-NAD	InhA:PTH-NAD	InhA:PTH-NAD
<b>Data Collection</b>			
Space group	P6 <sub>2</sub> 22	I2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	I2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions			
a, b, c (Å)	97.9, 97.9, 140.4	91.6, 100.5, 186.5	91.2, 100.0, 186.6
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 120	90	90
Resolution (Å)	2.2 (2.28 - 2.2)	2.5 (2.59 - 2.5)	2.1 (2.18 - 2.1)
R <sub>sym</sub> (%)	16.6 (74.3)	14.8 (92.9)	6.8 (56.9)
I/ $\sigma$ I	19.8 (3.0)	8.6 (1.3)	31.5 (2.0)
Completeness (%)	99.6 (99.9)	99.5 (100)	99.5 (95.6)
Redundancy	10.4 (9.0)	3.0 (3.0)	7.0 (4.9)
<b>Refinement</b>			
Resolution (Å)	40.85 - 2.20	19.99 - 2.5	19.99 - 2.1
No. reflections	19683	28424	47248
R <sub>work</sub> / R <sub>free</sub>	23.3 / 26.1	21.9 / 26.8	22.0 / 24.4
No. atoms / ASU			
Protein	1994	3988	4016
Ligand	54	110	110
Water	34	83	203
B-factors (Å <sup>2</sup> )			
Protein	49.1	33.8	41.3
Ligand	47.8	32.6	35.7
Water	42.5	30.3	49.5
R.m.s deviations			
Bond length (Å)	0.007	0.015	0.017
Bond angles (°)	1.3	2.1	2.1

## RESULTS AND DISCUSSION

### Activation of ETH Using Purified EthA or Chemical Agents *in vitro*

*ethA* encodes a flavin monooxygenase known to catalyze the Baeyer-Villiger reaction to detoxify aromatic and long chain ketones (130). EthA is membrane-associated and has a tendency to form large oligomers after purification, although its membrane attachment site has not been identified based on sequence prediction (130,131). The monooxygenase activity of the purified EthA is very low ( $k_{\text{cat}} = 0.00045 \text{ s}^{-1}$ ), suggesting that the enzyme may require cell membrane to be fully functional. We and other groups have tried to use purified recombinant EthA to activate ETH *in vitro*. It has been demonstrated that some polar metabolites were generated by the catalytic oxidation of ethionamide by EthA by TLC and HPLC (19). However, none of the isolated metabolites of ETH have antitubercular activity, suggesting that the active form of the ethionamide is unstable and was not isolated in previous attempts. In our experiment, we incubated EthA with ETH, NADPH, and  $\text{NAD}^+$  in phosphate buffer pH 7.5 for two hours. After incubation the small molecule was separated from EthA, and incubated with purified InhA for one hour. The enzyme assay results indicated that the products from a direct activation of ethionamide by purified recombinant EthA did not show any measurable inhibition of InhA. All these results suggested that the activation of ETH could not be reconstituted *in vitro*. Non-enzymatic approaches were also tried to activate ETH. Since Mn (II) and Mn(III) have been shown to activate INH, these two metal complexes were used to activate ETH. ETH and  $\text{NAD}^+$  were incubated with Mn (II) or Mn(III) for 30 minutes. The product mixture was then filtered and incubated with

purified InhA. The activity of InhA was assayed and compared with the native enzyme. It was found that InhA remained active even after two-hour incubation with the product mixture, which indicated that ETH was not activated by the Mn catalysts. Other oxidants, such as hydrogen peroxide, were tested using the aforementioned methods. None of them could yield any products that showed inhibition to InhA.

### **Design of a Cell-Based Activation System**

Since *in vitro* activation of ethionamide has not been possible by either chemical or enzymatic approaches, we developed a cell-based activation method. In this system, InhA and EthA were co-expressed in *E. coli* BL21(DE3) in the presence of 100 µg/ml of ethionamide. Overexpressed recombinant EthA was used to activate the thioamide drugs ETH or PTH and InhA to scavenge the activated species in *E. coli*. While ETH or PTH is a potent drug against *M. tuberculosis* (MIC = ~ 0.5 µg/ml) (132), it does not affect *E. coli* growth even at very high concentrations (100 µg/ml) primarily due to the absence of EthA, the activator of the drug. Co-expression of multiple recombinant proteins in *E. coli* has been used to get protein complexes. However, to our knowledge it is the first time that the activator and target genes have been co-expressed to isolate active species of a prodrug.

### **Isolation and Characterization of the Active Forms of ETH and PTH**

InhA and EthA were co-expressed in *E. coli* BL21 (DE3) in the presence of 100 µg/ml of ETH. The recombinant InhA was rapidly purified, and an *in vitro* enzyme assay was performed. InhA isolated from the experimental sample had less than 1% of the specific activity of InhA purified without addition of ETH under the same assay condition. Mass analysis of denatured InhA from the experimental sample indicated the presence of a small molecule with a molecular weight of 798.2 (Figure 19). This mass corresponds to the exact molecular weight of an ethyl-isonicotinic-acyl-NAD covalent adduct. Moreover, purified fractions of this small molecule showed strong inhibition to native InhA *in vitro* ( $K_i = 7 \pm 5$  nM), which is as potent as the INH-NAD adduct, the active form of isoniazid ( $K_i = 5$  nM). When prothionamide was used in the same co-expression experiment, a compound was identified with a molecular weight of 812.2 that corresponds to the mass of a propyl-isonicotinic-acyl-NAD adduct (Figure 19). This compound is also very potent against InhA *in vitro* ( $K_i = 2 \pm 0.8$  nM).



A.

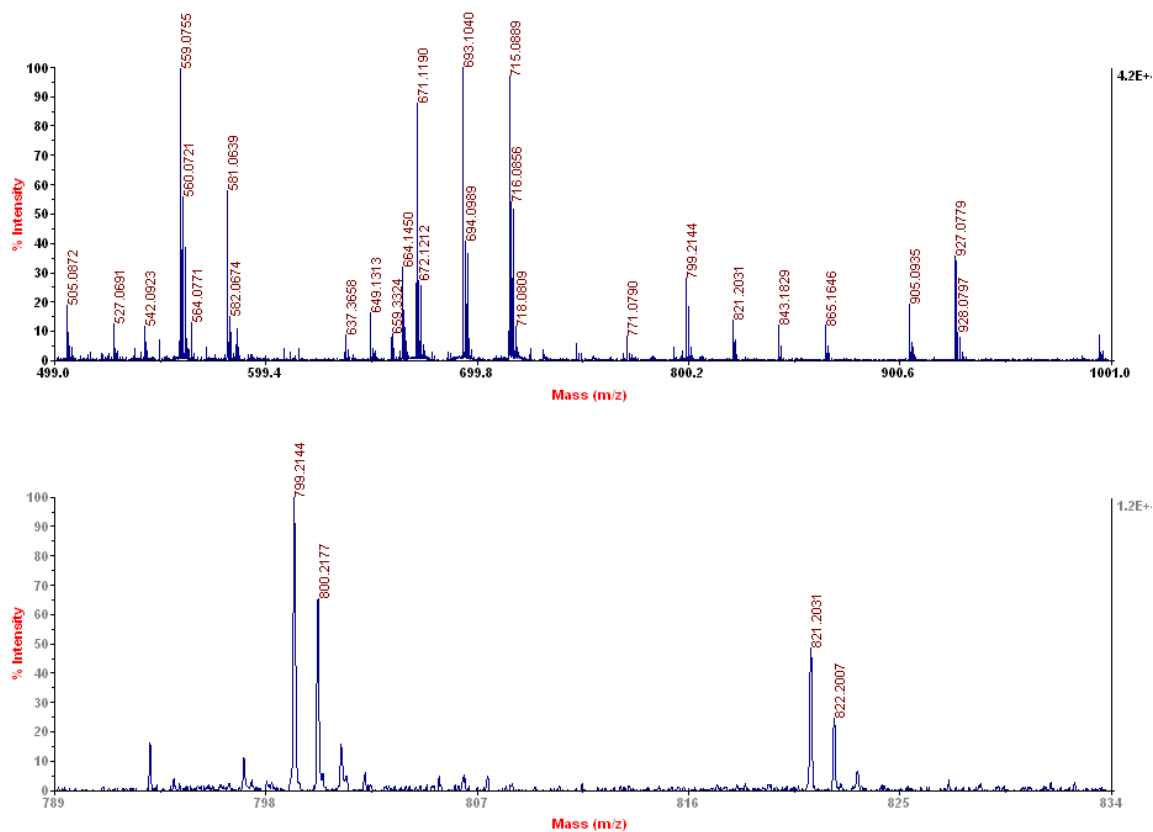


Figure 19 MALDI Mass Spectra of ETH-NAD and PTH-NAD.

A. MALDI mass spectra, carried out on an ABI Voyager-DE STR, showing that two inhibitors bound to InhA are compounds with apparent mass of 798 and 812, respectively. A. The positive mode mass spectra showing a peak at  $m/z = 799$  ( $[M + H]^+$ ). Peaks corresponding to  $[M + Na]^+$  ( $m/z = 821$ ),  $[M - H + 2Na]^+$  ( $m/z = 843$ ), and  $[M - 2H + 3Na]^+$  ( $m/z = 865$ ) can be observed as well.

B.

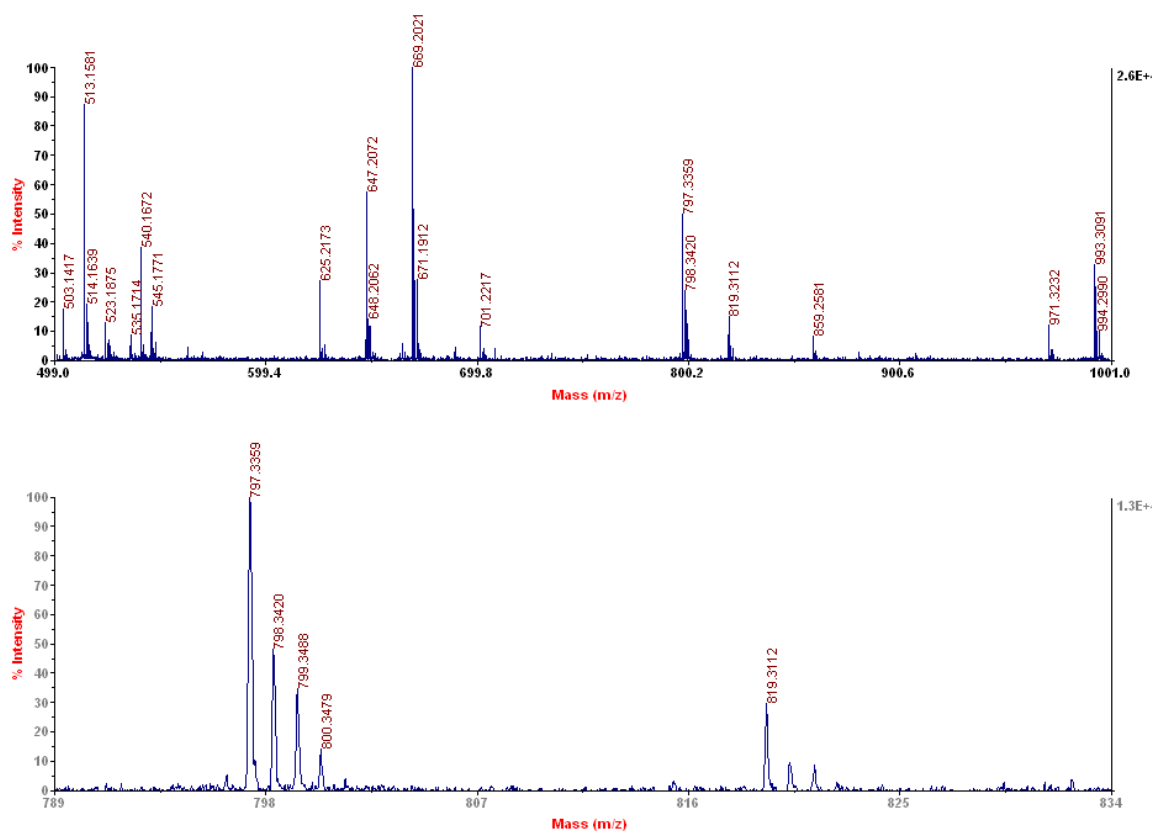


Figure 19 continued

B. The negative mode mass spectra showing a peak at  $m/z = 797$  ( $[M - H]^-$ ). Peaks corresponding to  $[M - 2H + Na]^-$  ( $m/z = 819$ ), and  $[M - 3H + Na + K]^-$  ( $m/z = 859$ ) can be observed as well. Therefore, the apparent mass of the compound is 798, which is in agreement with the chemical structure of ETH-NAD.

C.

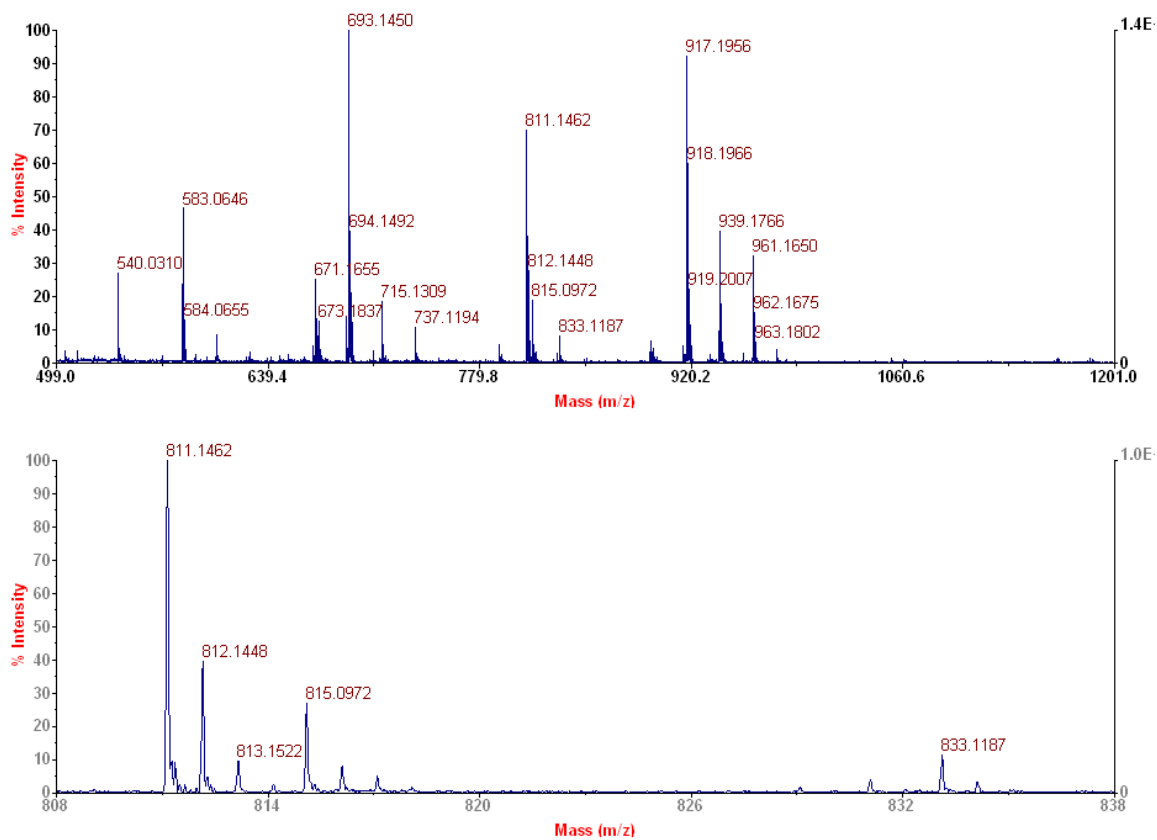


Figure 19. continued

C. The negative mode mass spectra showing peaks at  $m/z = 811$  ( $[M - H]^-$ ) and  $833$  ( $[M - 2H + Na]^-$ ), indicating an apparent mass of 812 that is in agreement with the chemical structure of PTH-NAD.

### Binding Mode of the ETH-NAD Adduct

*Mtb* InhA in complex with both inhibitors was crystallized. Dataset to 2.2 and 2.5 Å resolution were collected from single crystals of both complexes (Table 6). Unbiased electron density maps of each complex clearly indicated the presence of a modified NAD with an ethyl-isonicotinic-acyl or propyl-isonicotinic-acyl group at the 4-position of the nicotinamide ring in a 4S configuration (Figure 20A). The chemical structures of both inhibitors are consistent with molecular weights obtained by the mass analysis. Similar to the structure of InhA bound with the adduct INH-NAD (13), the ethyl-isonicotinic-acyl, or propyl-isonicotinic-acyl, moiety is found in a hydrophobic pocket that was formed by the rearrangement of the side chain of Phe<sup>149</sup> (Figure 20B, C, D). The ethyl-isonicotinic acyl, or propyl-isonicotinic-acyl, group also forces the side chain of Phe<sup>149</sup> to rotate  $\sim 90^\circ$  forming an aromatic ring-stacking interaction with the pyridine ring (Figure 21A). The pocket is lined predominantly by hydrophobic groups from the side-chains of Tyr<sup>158</sup>, Phe<sup>149</sup>, Met<sup>199</sup>, Trp<sup>222</sup>, Leu<sup>218</sup>, Met<sup>155</sup>, Met<sup>161</sup>, and Pro<sup>193</sup>, and is adjacent and partly overlapped with the fatty acyl substrate-binding site. Indeed, the atoms common to ETH-NAD, PTH-NAD, and INH-NAD are in identical positions. The only difference is the extra ethyl or propyl group at the 2 position of the pyridine ring of ETH or PTH. The ethyl group contributes to the binding of ETH-NAD adduct by forming  $\pi$ -stacking interactions with the aromatic side chain of Tyr<sup>158</sup> at a distance of  $\sim 3.3$  Å. It is also within hydrophobic interaction distances with side chains of Leu<sup>218</sup> (3.3 Å) and Met<sup>155</sup> (3.2 Å). The hydrogen-bonding interactions on the nucleotide-binding site are well conserved. Therefore, it is very likely that mutations, such as S94A, that

decrease the binding of NAD(H) and the INH-NAD adduct would also weaken the binding of ETH-NAD and PTH-NAD (Figure 21B). This explains why the S94A mutant strain of *Mtb* is co-resistant to both isoniazid and ethionamide.

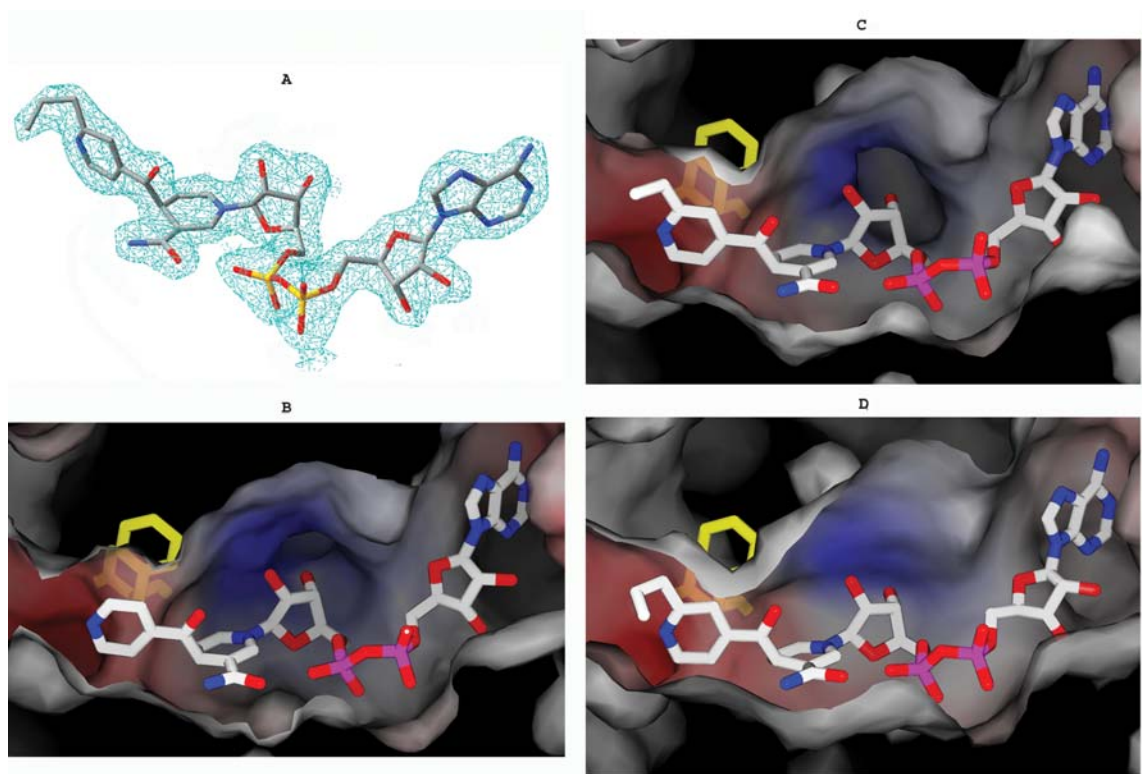


Figure 20 Active Sites of *Mtb* InhA Bound to Inhibitors

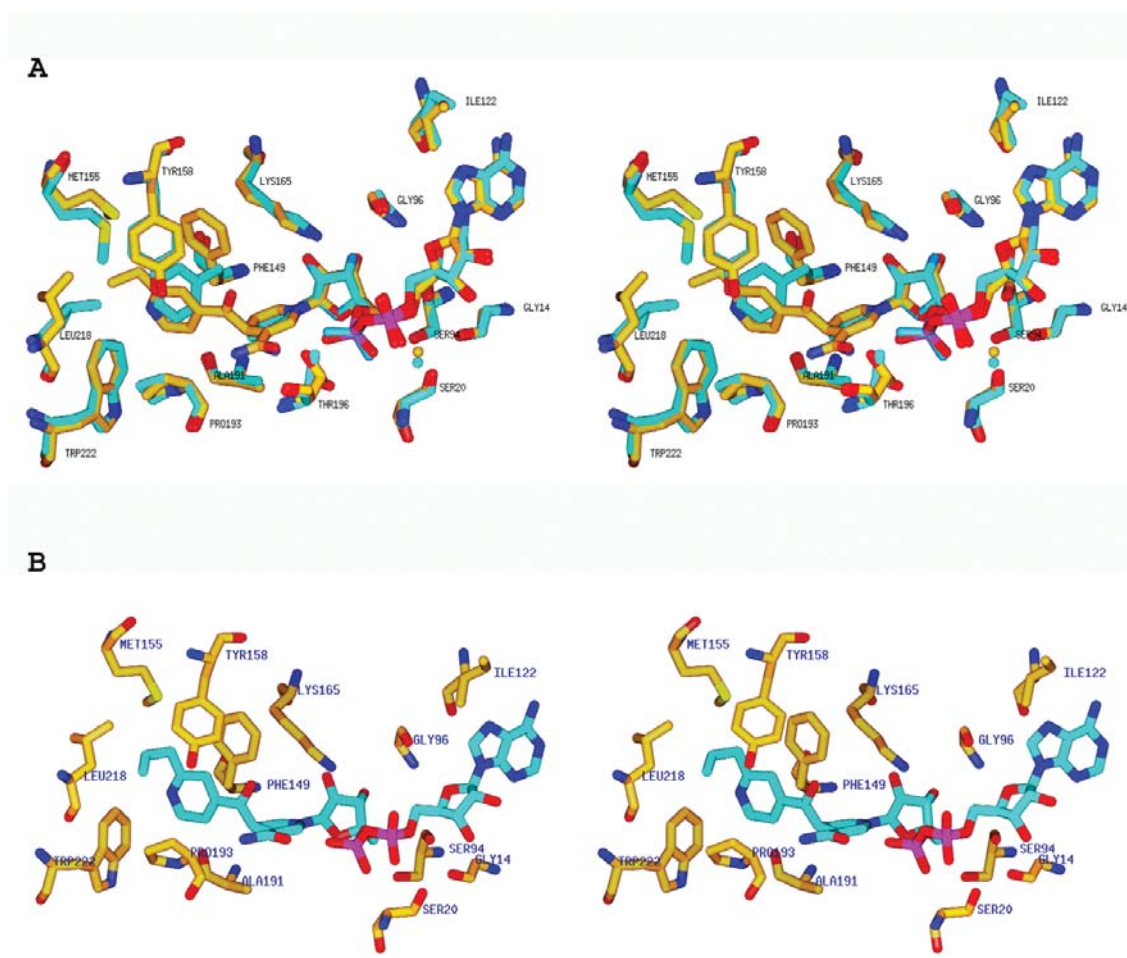


Figure 21 Active Site of InhA in Complex with ETH-NAD and PTH-NAD.

A. The stereo view of the superposition of active sites of *Mtb* InhA:NADH structure and InhA:ETH-NAD structure, showing the side chain of Phe<sup>149</sup> rotated  $\sim 90^\circ$  once the ETH-NAD adduct binds to the enzyme. B. The stereo view of the active sites of *M. leprae* InhA:PTH-NAD structure

### Mechanism of Drug Action of ETH and PTH against Leprosy

Unlike Isoniazid, the front line anti-TB drug, which is not effective against *M. leprae*, ethionamide and prothionamide are used in the treatment of leprosy. Genome analysis indicated that both *M. leprae* and *M. avium* have *ethA* and *inhA* homologs with high sequence similarity to *Mtb*. The *M. leprae ethA* and *inhA* were cloned, co-expressed, and purified in the presence of the prothionamide using a cell based activation method similar to *Mtb*. A compound with same molecular weight as PTH-NAD adduct was identified, which also showed strong inhibition to *M. leprae* InhA ( $K_i = 11 \pm 6$  nM) *in vitro*.

The crystal structure of *M. leprae* InhA in complex with PTH-NAD was solved to 2.1 Å resolution, in order to compare the binding mode of the inhibitor to the enzyme with *Mtb* InhA. The overall structure of *M. leprae* InhA is similar to *Mtb InhA* (r.m.s.d = 1.3), and the active site residues are conserved. The propyl-isonicotinic-acyl moiety of the adduct is observed inside a hydrophobic pocket. The pyridyl ring forms  $\pi$ -stacking interaction with the aromatic side chain of Phe<sup>149</sup> and the propyl group is  $\pi$ -stacking with the aromatic side chain of Tyr<sup>158</sup>. Residues including Ser<sup>94</sup> in nucleotide binding site forms hydrogen-bonding interactions with NAD moiety of the adduct in a similar manner as *Mtb* InhA (Figure 21B). These results supported our hypothesis that the active form of prothionamide inhibits *M. leprae* InhA in a similar way to *Mtb* InhA. Although no clinical or experimental mutant of *M. leprae* InhA has been reported yet, based on the binding mode of the PTH-NAD adduct, it is legitimate to predict that those InhA

mutations in ETH resistant *Mtb* mutant strains, such as S94A, would also confer resistance of *M. leprae* to ETH and PTH.

### **Activation Mechanism of ETH and PTH**

Although ethionamide, prothionamide, and isoniazid NAD adducts are similar, their activation mechanisms are very different. Isoniazid, a hydrazid, is activated by the heme-utilizing catalase-peroxidase, KatG (15,103). We proposed that isoniazid was oxidized to generate an isonicotinic acyl free radical, which subsequently attacked the NAD<sup>+</sup> to form INH-NAD (13). Ethionamide, a thioamide, has been shown to be metabolized by EthA, a FAD enzyme found in *Mtb*, however, none of the final metabolites isolated were active either *in vitro* or *in vivo* (19). It has been proposed that a free radical metabolite intermediate could be generated through EthA oxidation of ETH, in analogy to the activation of INH. There is a possibility that other unknown cell component is required for the formation of the adduct by the free radical intermediate. It is also possible that the activation of ETH itself may happen inside cell membrane, which creates a hydrophobic environment. This is usually critical for a free radical type reaction. These may explain why no active species were isolated *in vitro*. We believe those inactive metabolites isolated in previous attempts could result from the side reactions and quench of the free radical intermediate in solution. It is still not clear how the thioamide is oxidized inside the active site of EthA, although it is common for a flavin monooxygenase that a peroxide free radical is generated during oxidation, assisted by the cofactor FAD or FMN. Tokuyama et al. demonstrated that a thioamide could be used as a precursor of a synthon equivalent to an imidoyl radical in converting



thioamides to corresponding indole derivatives (133). The  $\text{Bu}_3\text{SnH}/\text{Et}_3\text{B}$  was used as a free radical initiator in pure organic solvent. Similarly, we postulate that ETH is converted to an imidoyl radical inside the active site of EthA using  $\text{O}_2/\text{FAD}$  as a radical initiator. The imidoyl radical subsequently forms an adduct with NAD, which is then converted to ethyl-isonicotinic-acyl-NAD adduct after hydrolysis to release the amine group. It is also possible that the imidoyl anion is the intermediate before forming the adduct with NAD (Figure 22).

### **Other Anti-TB Thioamide Drugs**

Other than ETH and PTH, there are several anti-TB thioamides drugs currently used as second line drugs to cure TB. Among them, the best known are thioacetazone and isoxyl. These thioamides are cheap and very effective against TB and even some drug resistant strains. Thioacetazone and isoxyl are prodrugs that also have to be activated by EthA. However, their final targets are still not clear. We grew the *E. coli* cell with both EthA and InhA co-expressed in the presence of 100  $\mu\text{g}/\text{ml}$  of thioacetazone. However, the isolated InhA are fully active. Mass and crystallographic analysis indicated that there is no small molecule tightly bound to InhA, which suggested that unlike ETH the final target of thioacetazone is not InhA. Later on, the genetic studies indicated that both thioacetazone and isoxyl are more likely targeting the mycolic acid modification pathways, not fatty acid biosynthesis, which is consistent with our results.

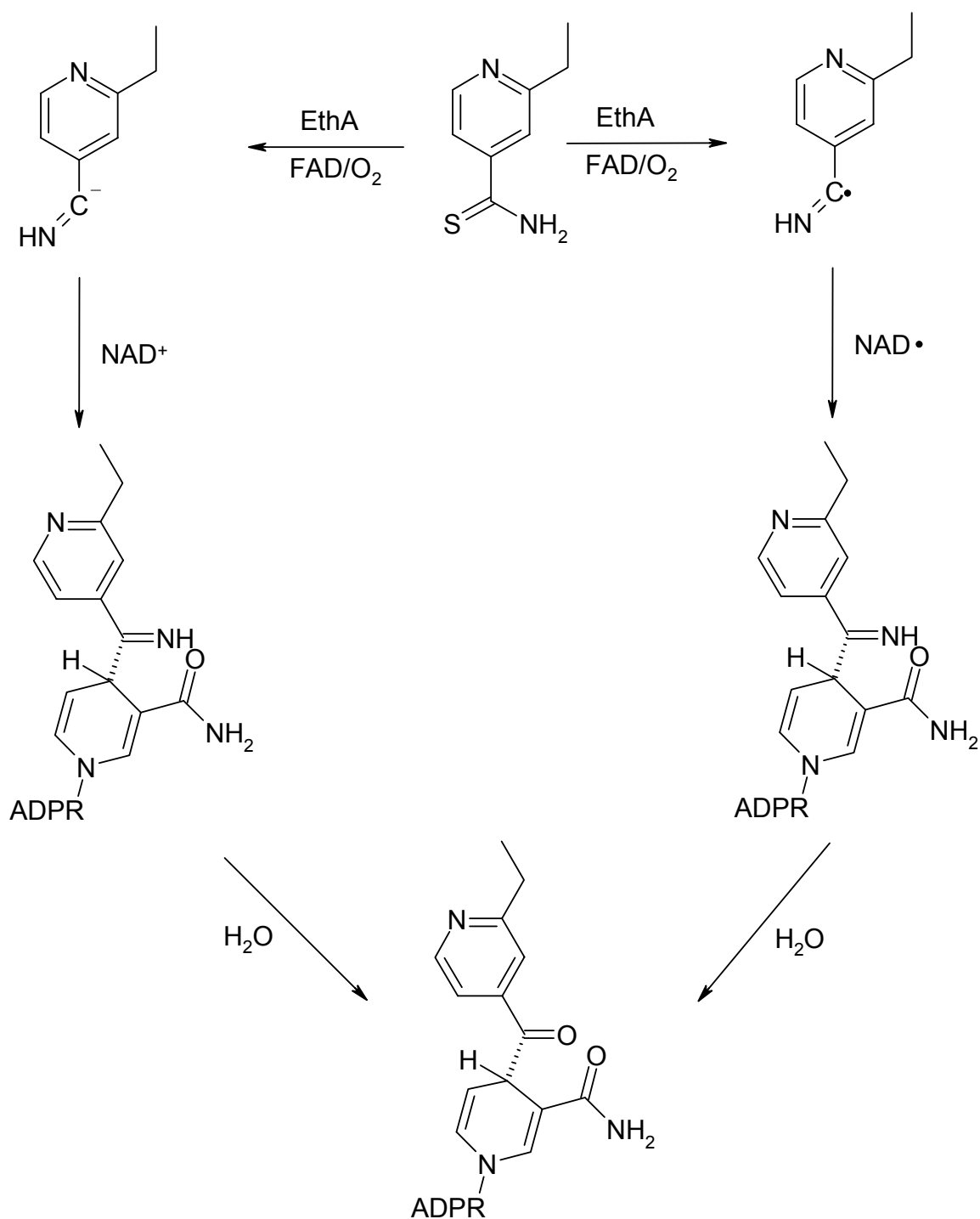


Figure 22 Possible Reaction Mechanisms of the Activation of ETH

### **Crystallization Trials of EthA**

Since EthA plays a key role in the activation of several anti-mycobacteria thioamides, it is important to understand the reaction mechanism in the active site of EthA. The structure of EthA would greatly benefit the studies of the activation mechanism. As mentioned previously, EthA is a membrane associated protein, although it is still not clear about the lipid attachment site based on sequence analysis. More than 95% of the recombinant EthA is attached to membrane after overexpression in *E. coli*. Several detergents have been tried to extract the protein from the membrane. Triton X-100 showed good yield (>90%) of extraction of the recombinant EthA. Interestingly, once the EthA is extracted from membrane, it does not require detergent in the purification buffer. However, the protein formed large oligomer during purification. Several detergents were screened to stabilize the protein monitored by dynamic light scattering method. None of the detergents tried could turn the protein into smaller oligomer state. The purified protein was tried for crystallization. More than 400 crystallization conditions were screened, however no protein crystal was observed.

### **Conclusion**

These studies describe the molecular mechanism of the drug action of ethionamide and prothionamide against *M. tuberculosis* and *M. leprae*. This information is important for the optimization of drug activity and the understanding of drug resistance. As the molecular target of ethionamide, prothionamide, and isoniazid is the same, it validates InhA as an outstanding anti-tuberculosis and anti-leprosy drug target.

Since most of the clinical strains resistant to ethionamide and prothionamide contain mutations in the *ethA* gene, it is clear that agents that inhibit InhA without the need for EthA activation will be effective against resistant bacteria.

## CHAPTER IV

### INHIBITOR DESIGN TARGETING *MTB* INHA

#### BACKGROUND

The emergence of strains resistant to INH and other widely used antitubercular drugs is a global health problem, making it critical to search for a new generation of antitubercular agents (126,134). The activities of INH and ETH are dependent on their activation in *Mtb* by KatG and EthA, respectively (19,106). Mutations in activator genes *katG* and *ethA* have been found in most of the clinical isolates resistant to these two drugs (106).

Compounds that do not require activation and directly target InhA represent a promising approach to circumvent this resistance mechanism. Aiding such an effort are several crystal structures, such as those of InhA:NADH, InhA:NAD<sup>+</sup>:triclosan, InhA:NAD:Genz10850, InhA:NAD<sup>+</sup>:C16-substrate and InhA:INH-NAD. The complex structures indicate two hydrophobic cavities capable of being filled, the substrate binding site and the pocket where the isonicotinyl group protrudes. There are also several possibilities for hydrogen-bonding, aromatic stacking, van der Waal, and hydrophobic interactions distributed throughout the active site. The precisely defined active site of InhA, by these structures, has facilitated a thorough understanding of the ligand-enzyme interactions that render potent enzyme inhibition. The design of novel inhibitors also requires the discovery of new lead compounds.

## METHODS

### Cloning, Expression and Purification

*M. tuberculosis inhA* has already been cloned into *E. coli* BL21 (DE3) (116). It was expressed and purified as described in chapter II.

### InhA Enzymatic Activity Assay

All assays were carried out using a Cary 100 Bio Spectrophotometer at 25 °C by monitoring oxidation of NADH at 340 nm. Reactions were initiated by adding substrate, dodecenoyl-CoA (50 μM), to assay mixtures containing InhA (5 nM), NADH (100 μM) and inhibitors (1 - 10000 nM).

The IC<sub>50</sub> was determined from dose-response plot of enzyme fractional activity as a function of inhibitor concentration. K<sub>i</sub> was obtained by dividing the IC<sub>50</sub> value by  $1 + [S_1]/K_{m1} + [S_2]/K_{m2}$ , where [S<sub>1</sub>] and [S<sub>2</sub>] are the concentration of dodecenoyl-CoA and NADH and K<sub>m1</sub> and K<sub>m2</sub> are their Michaelis constants.

InhA *in vitro* enzyme assay were conducted in the presence of 10 μM inhibitor as an initial screen, and those compounds with > 60% inhibition were further subjected to IC<sub>50</sub> determination, that will be used to compare the inhibition potency quantitatively.

### Crystallization of InhA in Complex with Inhibitors

Crystallization was accomplished by the hanging drop vapor diffusion method (85). InhA was incubated with inhibitors and NAD<sup>+</sup> at molar ration 1:2:100 for two hours and was then co-crystallized in hanging droplets containing 2 μl of protein

solution at 10 mg/ml and 2  $\mu$ l of buffer (20% PEG 3350, 6% DMSO, 0.1M ADA pH 6.8, 0.08 M ammonium acetate) at 16 °C in Linbro plates against 0.5 ml of the same buffer. Protein crystals formed ~4 days later.

### **Data Collection and Processing**

Data were collected at 121 K using cryo-protection solution containing reservoir solution with additional 30% ethylene glycol. Crystal of InhA:JPC5394 diffracted X-rays to 1.98 Å at beam line 23ID at the Advanced Photon Source (APS), Argonne National Laboratory. Crystals of InhA:JPC5274 and InhA:JPC5734 diffracted X-rays to 2.8 Å and 1.98 Å respectively, using a Raxis detector coupled to a Rigaku X-ray generator utilizing a copper rotating anode ( $\text{CuK}_{\alpha}$ ,  $\lambda = 1.54$  Å). Diffraction data were collected from a single crystal with 0.5 degree oscillation widths for a range of 180°. The data were integrated and reduced using HKL2000 (Table 7) (86).

Table 7 Data Statistics of InhA in Complex with Triclosan Analogs

Data collection			
	InhA:JPC5394	InhA:JPC5274	InhA:JPC5734
Maximum resolution (Å)	1.98	2.80	1.97
Space group	C2	I4(1)22	I4(1)22
a (Å)	125.6	90.0	90.0
b (Å)	92.3	90.0	90.0
c (Å)	103.0	183.1	183.9
$\alpha$ (°)	90.0	90.0	90.0
$\beta$ (°)	106.4	90.0	90.0
$\gamma$ (°)	90.0	90.0	90.0
Unique reflections	71784 (6260)	9708 (1063)	27204 (2964)
R <sub>sym</sub> (%)	9.9 (70.5)	12.8 (82.3)	6.0 (74.3)
Completeness (%)	87.3 (65.5)	99.8 (100)	99.8 (100)
Redundancy	3.7 (2.1)	8.5 (8.6)	10.2 (10.0)
I/ $\sigma$	15.7 (1.3)	22.3 (3.7)	42.5 (4.3)
Refinement statistics			
Resolution range (Å)	99.0-1.98	80.9-2.8	80.9-1.97
Number of reflections	68950	9182	25756
Number of atoms / subunit			
Protein	7806	1994	1994
Cofactor (NAD)	176	52	52
Ligand	88	23	22
R <sub>cryst</sub> (%)	24.4	23.9	25.4
R <sub>free</sub> (%)	28.5	25.8	28.0
Average B-factors (Å <sup>2</sup> )	32.2	52.3	32.4



## Structure Determination and Model Refinement

Initial phases of InhA in complex with JPC5394, JPC5274 and JPC5734 were obtained by molecular replacement using the apo-InhA structure (1ENY) and refined with CCP4 (Table 7).  $F_o - F_c$  and  $2F_o - F_c$  electron density maps were calculated and additional density resembling the inhibitor was found. The ligand was fit into the additional density and the whole model was rebuilt using XtalView (89). During the final cycles of the refinement, water molecules were added into peaks above 3- $\sigma$  of the  $F_o - F_c$  electron density maps that were within hydrogen-bonding distances from appropriate protein atoms.

## RESULTS AND DISCUSSION

### Structure Activity Relationship (SAR) Study of Triclosan Analogs

5-chloro-2 (2,4-dichlorophenoxy) phenol, known as triclosan, has been reported to potentially inhibit the enoyl acyl reductases (ENR) from several species, such as *Plasmodium falciparum*, *Staphylococcus aureus*, *Escherichia coli*, *Bacillus subtilis*, *Brassica napus* and *Pseudomonas aeruginosa*. As an antimicrobial agent incorporated into many household items, triclosan has been utilized in the United States for over 40 years.

The structures of ENR from *Escherichia coli* (FabI), *Plasmodium falciparum* (PfENR), and *Mycobacterium tuberculosis* (InhA) bound with triclosan have been defined. Although triclosan is a potent inhibitor to all these ENRs, its binding affinities to the aforementioned target enzymes are dramatically different (ranging from picomolar

to micromolar). Sequence alignment indicated low conservation in substrate binding region that is related to the triclosan binding site. Therefore using triclosan as a scaffold, it is possible to design triclosan derivatives specific to InhA, the ENR of *Mtb* to optimize potency. A library (JPC) of triclosan derivatives was built to search for inhibitors with higher potency to InhA. Those triclosan derivatives were synthesized with a variety of substituents at the 1, 5, 6, 2' and 4' positions, as well as the linker. InhA *in vitro* activity assay was conducted in the presence of 10  $\mu$ M inhibitor as an initial screen, and those hits with 50% inhibition were further subjected to IC<sub>50</sub> determination, that will be used to compare the inhibition potency quantitatively.

### **Triclosan Derivatives with Modification at the Linker**

The ether oxygen which links ring A and ring B plays an important role in the triclosan binding. It was found that replacement of the linker oxygen with carbon, sulfur or nitrogen led to more than 20 fold decrease in binding affinity. It is possible that the linker atom could have some interaction with the adjacent hydroxyl group. Replacing the linker oxygen may perturb this interaction and consequently weaken the hydrogen bonds of the hydroxyl group.

Triclosan analogs with several different linkers were synthesized and their inhibitory potency was assayed. Some of them were listed in Table 8. A major effort put on trials was to use a urea group as a new linker. Similar to the other linkers tested, the triclosan derivatives with urea linker did not show any inhibitory potency comparable to

triclosan. The urea group may be too flexible to maintain the same geometry of the ring A and ring B of triclosan.

Table 8 Inhibition of InhA by Triclosan Analogs at 10  $\mu$ M Concentration

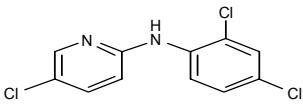
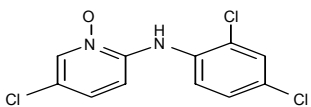
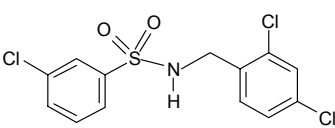
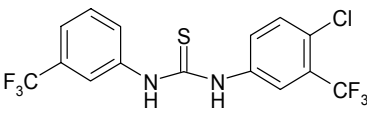
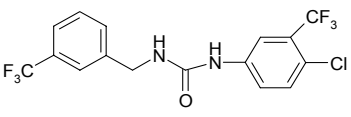
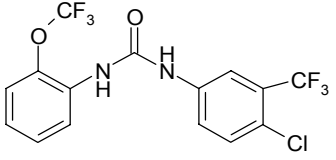
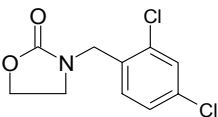
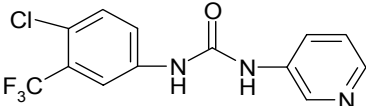
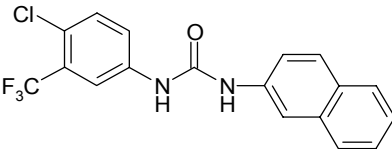
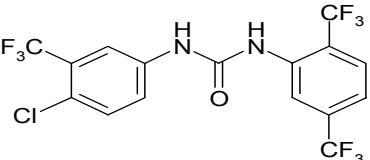
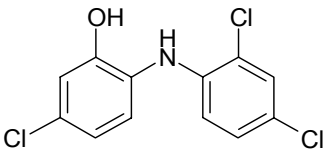
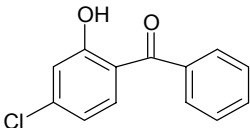
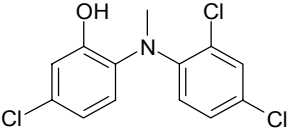
Compound	Structure	Inhibition
JPC2118		0.85
JPC2119		0.93
JPC2145		0.98
JPC2295		1.07
JPC2298		1.21
JPC2312		0.93

Table 8. Continued

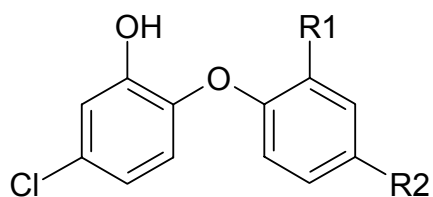
Compound	Structure	$V_i/V_0$
JPC2328		1.08
JPC2357		0.91
		0.87
JPC2449		0.89
JPC2576		0.75
JPC2596		0.78
JPC2600		1.00

### **Triclosan Derivatives with Modification at 2'-Position**

The ring B of triclosan, oriented orthogonal to the ring A, was projected along the substrate binding site toward the major portal. The chlorine atom at 2'-position was located on top of the cofactor NAD<sup>+</sup> and within 3.5 Å to its phosphate moiety. The extra space between the chlorine and the cofactor made us hypothesize that a hydrophilic substituent potentially could increase binding affinity through some new hydrogen bonding interactions with NAD<sup>+</sup> or residues nearby.

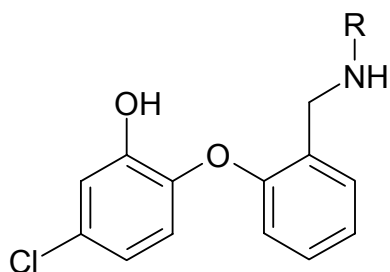
Triclosan derivatives with 2' substituent were synthesized and their activities were assayed (Table 9). It is quite clear that a polar group such as chlorine increased the inhibitory potency. For example, abolishing the chlorine atom of triclosan resulted in a dramatic decrease in potency of JPC2274 (23 fold less than triclosan). An amine group at the same position somewhat restored the potency of JPC2205 (only 4.0 fold less than triclosan). However, a carbon linker between the amine group and the ring B led to more than three fold decrease in activities, which suggested that the size of the substituent is an important factor for the potency. This is confirmed by the assay results of a series of triclosan derivatives with long or bulky substituents at 2'-position (Table 10 and 11). The complex structure of InhA:triclosan suggested that there is not enough space between the substituent on ring B and cofactor NAD<sup>+</sup> to accommodate a bulky group.

Table 9 Inhibition of InhA by Triclosan Analogs with Substituents on B Ring



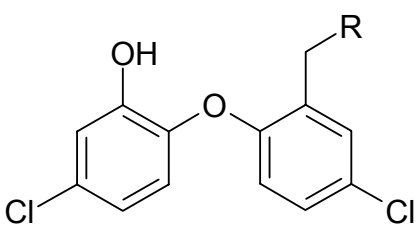
Compound	R <sub>1</sub>	R <sub>2</sub>	IC <sub>50</sub> (μM)	%
Triclosan	Cl	Cl	1.1	
JPC-2274-A-1	H	H	23	0.61
JPC-2205-A-1	NH <sub>2</sub>	H	4.2	0.25
JPC-2210-A-1	CH <sub>2</sub> NH <sub>2</sub>	H	>10	0.94


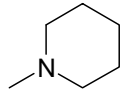
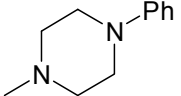
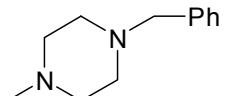
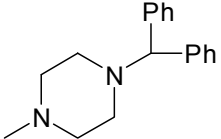
Table 10 Inhibition of InhA by Triclosan Analogs with NH Group at 2'-Position



Compound	R	IC <sub>50</sub> (μM)	V <sub>i</sub> /V <sub>0</sub>
JPC-2210-A-1	H	>10	0.94
JPC-2239-A-1	CH <sub>2</sub> Ph	>10	0.57
JPC-2240-A-1	(CH <sub>2</sub> ) <sub>3</sub> Ph	>10	0.55
JPC-2241-A-1	(CH <sub>2</sub> ) <sub>2</sub> Ph	>10	0.64
JPC-2242-A-1	1-naphthyl	>10	0.69
JPC-2243-A-1	2-naphthyl	>10	0.91

Table 11 Inhibition of InhA by Triclosan Analogs with Cyclic NH at 2'-Position



Compound	R	IC <sub>50</sub> (μM)
JPC-2468-A-1		>10
JPC-2469-A-1		>10
JPC-2442-A-1		>10
JPC-2444-A-1		>10
JPC-2470-A-1		>10

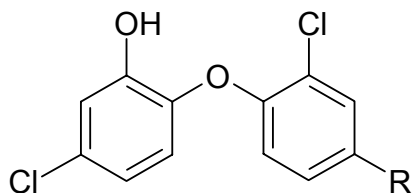


### **Triclosan Derivatives with Modification at 4'-Position**

The chlorine atom at 4'-position of triclosan was projected toward the major portal of the substrate-binding site. It is surrounded by the backbones of residues Met98, Ala198, and Ile202. Since the chlorine atom is pointing toward the major portal of the substrate-binding site, there is large space open to the substituent at 4'-position. Two strategies were applied on the design of new inhibitors to optimize the potency. First approach is to replace the chlorine atom with some other polar groups that potentially could form strong hydrogen bonding interactions with those residues surrounded. Second approach is to utilize large substituents to fit into the cavity to create hydrophobic interactions.

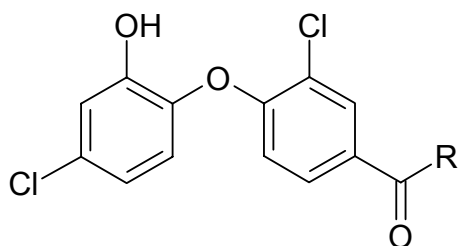
As showed in Tables 12, 13, 14, most of the triclosan derivatives with polar substituents have high inhibitory potency ( $IC_{50} < 2 \mu M$ ), except JPC2197 ( $IC_{50} > 10 \mu M$ ). It is probably due to its hydrophobic carbon chain at the end of the substituent. All compounds with hydrophilic terminals at 4'-position have good activities. When the substituents become too large, the inhibitory potency starts to decrease, most likely due to steric hindrance.

Table 12 Inhibition of InhA by Triclosan Analogs with Substituents at 4'-Position



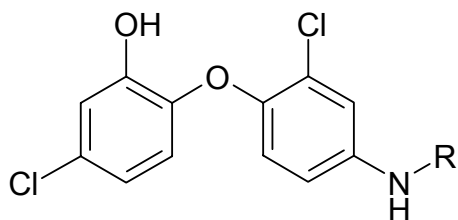
Compound	R	IC <sub>50</sub> (nM)	V <sub>i</sub> /V <sub>0</sub>
Triclosan	Cl	1100	
JPC-2136-A-1	NH <sub>2</sub>	662	0.12
JPC-2137-A-1	NO <sub>2</sub>	805	0.12
JPC-2149-A-1	CN	599	0.08
JPC-2150-A-1	tetrazol	2390	0.30
JPC-2153-A-1	COOH	2503	0.26
JPC-2181-A-1	OH	972	0.16
JPC-2197-A-1	CH(OH)CH <sub>2</sub> CH <sub>3</sub>	>10000	0.67
JPC-2211-A-1	C(O)NHOH	7833	0.40
JPC-2212-A-1	C(O)NH <sub>2</sub>	851	0.15

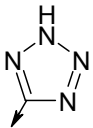
Table 13 Inhibition of InhA by Triclosan Analogs with (CO)R Group at 4'-Position



Compound	R	IC <sub>50</sub> (nM)	V <sub>i</sub> /V <sub>0</sub>
JPC-2212-A-1	NH <sub>2</sub>	851	0.15
JPC-2211-A-1	NHOH	7830	0.40
JPC-2166-A-1	N(H)Me	6510	0.47
JPC-2167-A-1	NMe <sub>2</sub>	7250	0.43
JPC-2171-A-1	N-pyrrolidine	>10000	0.55
JPC-2168-A-1	N-piperidine	5710	0.31
JPC-2169-A-1	N-morpholine	5300	0.39

Table 14 Inhibition of InhA by Triclosan Analogs with (NH)R Group at 4'-Position



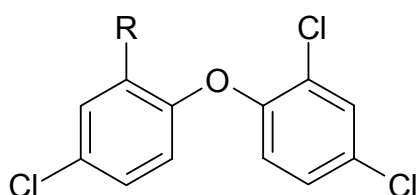
Compound	R	IC <sub>50</sub> (nM)	V <sub>i</sub> /V <sub>0</sub>
JPC-2136-A-1	H	662	0.12
JPC-2159-A-1	C(O)NH <sub>2</sub>	577	0.15
JPC-2160-A-1	SO <sub>2</sub> Ph	>10000	0.79
JPC-2162-A-1	C(O)Ph	412	0.02
JPC-2182-A-1	C(O)CH <sub>3</sub>	6460	0.44
JPC-2190-A-1	SO <sub>2</sub> (1-naphthyl)	>10000	0.53
JPC-2191-A-1	SO <sub>2</sub> (2-naphthyl)	5900	0.31
JPC-2192-A-1	CH <sub>2</sub> (2-CNPh)	2730	0.29
JPC-2193-A-1	C(NH)NHCN	3130	0.27
JPC-2194-A-1	2-pyridyl	1940	0.21
JPC-2200-A-1	SO <sub>2</sub> CF <sub>3</sub>	>10000	0.50
JPC-2263-A-B1	(C(NH)NH) <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>	>10000	0.75
JPC-2556-A-1	CH <sub>2</sub> Ph	2260	0.19
JPC-2202-A-1		3190	0.19

### Triclosan Derivatives with Modification at 1-Position

The hydroxyl group at 1-position is critical to the binding affinity and antibacterial activity. The replacement of this hydroxyl group by methoxy or sulfur derivatives resulted in a more than 10000-fold reduced affinity for *E. coli* ENR. The hydroxyl group forms hydrogen bonds to the side chain of Tyr<sup>158</sup> of InhA and the co-factor NAD<sup>+</sup> at distance of 2.2 Å and 2.8 Å, respectively. These hydrogen-bonding interactions are conserved in all complex structures of ENRs bound to triclosan. Despite the significant role of the hydroxyl group at 1-position of triclosan in the binding to the enzyme, it is rapidly glucuronidated or sulfonated *in vivo*, which causes the ineffectiveness of triclosan. Therefore, it is necessary to remove the hydroxyl group at the 1-position or replace it with other functional group that cannot be modified *in vivo*. Meanwhile, the new inhibitors should have comparable inhibitory potency as triclosan. 36 triclosan derivatives with 1-position substituents were synthesized and subsequently assayed for inhibitory potency at 10 µM concentration. The substitution of the hydroxyl group at 1-position with all 36 functional groups resulted in a significant loss of inhibitory potency (IC<sub>50</sub> > 10 µM). Six of these triclosan derivatives were listed in Table 15. Apparently, all the efforts to search for a substitute of hydroxyl group did not successfully maintain the inhibitory potency. Modeling studies indicated that it is difficult for the enzyme to accommodate those new substituents larger than hydroxyl group without major conformational changes. Slight changes in geometry could perturb the hydrogen bonding interaction. That is why some substituents very similar to the

hydroxyl group, such as an amine group that is also a hydrogen bond acceptor, still do not have a significant inhibitory potency.

Table 15 Inhibition of InhA by Triclosan Analogs with Substituents at 1-Position



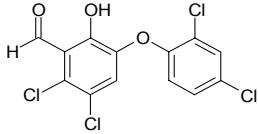
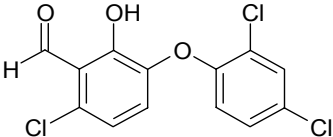
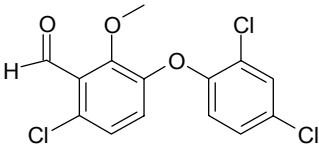
Compound	R	IC <sub>50</sub> (μM)
Triclosan	OH	1.1
JPC-1095-A-1	CONH <sub>2</sub>	>10
JPC-2025-A-1	OCH <sub>2</sub> CN	>10
JPC-2063-A-1	NH <sub>2</sub>	>10
JPC-2064-A-1	NO <sub>2</sub>	>10
JPC-2548-A-1	CF <sub>3</sub>	>10
JPC-2642-A-1	OMe	>10

### Triclosan Derivatives with Modification at 6-Position

There are two rationales behind designing the triclosan derivatives with substituents at 6-position. First of all, this design is a new trial to overcome the *in vivo*

modification of the hydroxyl group at 1-position of triclosan indirectly. As mentioned before, the trial to replace the hydroxyl group at 1-position with other substituents lead to dramatic loss of inhibitory potency. We hypothesized that placing a bulky substituent at the 6-position could potentially protect the hydroxyl group at 1-position from enzyme-catalyzed *in vivo* modification through steric hindrance. Secondly, superimposition of the structures of InhA in complex with triclosan and the substrate showed that the substrate-binding site is on top of the ring A of triclosan. Therefore, we proposed to utilize a substituent at 6-position of triclosan to project into the substrate-binding site and acquire some additional binding affinity. Thus far, three analogs were synthesized. None of them showed high inhibitory activity against InhA (Table 16). Based on the structure of InhA:triclosan, there is not too much space to contain a large substituent at 6-position of the A-ring. It is possible that the aldehyde group used in the inhibitors might caused unfavorable clash with side chains of Phe149 and Tyr158, which are only 3.6 Å and 4 Å away, respectively.

Table 16 Inhibition of InhA by Triclosan Analogs with Substituents at 6-Position

Compound	Structure	$V_i/V_0$
JPC-2605-A-1		0.73
JPC-2615-A-2		0.94
JPC-2616-A-1		0.82

### Triclosan Derivatives with Modification at 5-Position

In the structure of InhA:INH-NAD, the isonicotinyl moiety of INH-NAD adduct is found in a hydrophobic pocket that was formed by the rearrangement of the side chain of Phe149. The isonicotinic-acyl group forces the side chain of Phe149 to rotate  $\sim 90^\circ$  forming an aromatic ring-stacking interaction with the pyridine ring. The pocket is lined predominantly by hydrophobic groups from the side-chains of Tyr158, Phe149, Met199, Trp222, Leu218, Met155, Met161, and Pro193, and is adjacent and partly overlapped with the fatty acyl substrate-binding site. Among them, Tyr158, Pro193 and Gly192 are conserved in other ENR homologues. This pocket also servers as a small portal to the

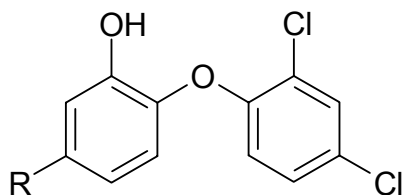


solvent at one side of the active site. The superimposition of the structures of InhA:INH-NAD and InhA:triclosan indicated that the chlorine atom at the 5-position of triclosan A ring is inside the binding pocket of isonicotinoyl moiety of INH-NAD adduct and is in van der Waals contact with Pro193, Met199, and Phe149. While triclosan is extremely potent against ecFabI ( $K_i = 7 \text{ pM}$ ), it has much lower activity against *Mtb* InhA ( $K_i = 200\text{nM}$ ). Based on the structural information, we hypothesized that it is possible to utilize this hydrophobic pocket with various hydrophobic substituents at the 5-position of triclosan to increase the *in vitro* activity against *Mtb* InhA.

A series of triclosan derivatives with modifications at the 5-position of triclosan were evaluated for their inhibition of purified InhA. Inhibitors with hydrophobic substituents, such as alkyl (compounds **2** and **7**), are much more potent than those with hydrophilic substituents (compounds **3**, **4** and **5**) (Table 17). This result is consistent with our proposal that the 5-substituent of triclosan projects into the hydrophobic cavity of the enzyme.

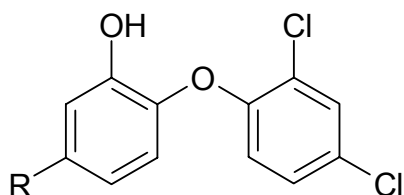
The hydrophobic substituents at the 5-position can be divided into alkyl and aryl subgroups. For the alkyl-substituted inhibitors, potency versus InhA appears to increase with the chain length (Table 18). Highest inhibitory potency was observed for compounds **10** and **12** with C4 carbon chains. Sullivan et al. observed the same trend with 5-alkyl substituted triclosan derivatives where the two chlorines of the B-ring were excised. It was found that the *in vitro* activity of those inhibitors was optimal at a carbon chain length of eight. For those inhibitors with 5-substituents of the same carbon chain length, the inhibitors studied in the present report showed better *in vitro* activity than their counterparts in Sullivan's work ( $IC_{50} = 120$  versus 2000 nM for derivatives with an ethyl substituent and  $IC_{50} = 55$  nM versus 80 nM for derivatives with a butyl substituent), suggesting that the two chlorine atoms on the B-ring contributed to the binding of the inhibitor to the enzyme. This contradicts Sullivan's assertion that the B-ring chlorines are not essential for enzyme inhibition. Indeed, as indicated in the crystal structure of InhA:NAD<sup>+</sup>:triclosan, the 4' chlorine atom engages the backbone amide nitrogen N-H of Met98 at a distance of 3.2 Å.

Table 17 Inhibition of InhA by Triclosan Analogs with Substituents at 5-Position



Compound	R	IC <sub>50</sub> (nM)
<b>Triclosan</b>	Cl	1100
<b>2</b>	Me	799
<b>3</b>	tetrazol	>10000
<b>4</b>	COOH	>10000
<b>5</b>	C(O)NH <sub>2</sub>	>10000
<b>6</b>	Ph	>10000
<b>7</b>	CH <sub>2</sub> (C <sub>6</sub> H <sub>11</sub> )	108

Table 18 Inhibition of InhA by Triclosan Analogs with Alkyl Groups at 5-Position



Compound	R	IC <sub>50</sub> (nM)
Triclosan	Cl	1100
<b>2</b>	CH <sub>3</sub>	799
<b>8</b>	CH <sub>2</sub> CH <sub>3</sub>	125
<b>9</b>	(CH <sub>2</sub> ) <sub>2</sub> CH <sub>3</sub>	91
<b>10</b>	(CH <sub>2</sub> ) <sub>3</sub> CH <sub>3</sub>	55
<b>11</b>	CH <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>	96
<b>12</b>	(CH <sub>2</sub> ) <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>	63
<b>13</b>	CH <sub>2</sub> CH(CH <sub>3</sub> )CH <sub>2</sub> CH <sub>3</sub>	129
<b>7</b>	CH <sub>2</sub> (C <sub>6</sub> H <sub>11</sub> )	108

The alkyl substituents in previous studies were all unbranched. In the present report, branched alkyl substituents (compounds **11**, **12** and **13**) are slightly less potent than their straight-chain counterparts (compounds **9** and **10**) of the same chain length. However, the active site seems to be tolerant of the bulky cyclohexyl moiety in **7**, which

demonstrated reasonable inhibitory activity against InhA *in vitro*. The crystal structure of InhA bound with Sullivan's unbranched 5-pentyl triclosan derivative showed the pentyl chain occupies nearly the same binding site as the acyl chain of the substrate. The 2.8 Å crystal structure of InhA in complex with compound **7** was solved to provide a structural basis for the activity difference amongst the alkyl substituents. The crystal belongs to  $I4_122$ , a space group for InhA that has not been reported previously. The structure was readily superimposed on the structures of InhA in complex with the 5-pentyl triclosan derivative and the C16 substrate. The 5-cyclohexylmethyl group occupied a position similar to that of the 5-pentyl triclosan derivative and formed predominantly hydrophobic interactions with the sidechains of Phe149, Leu218, Met155, and Met199 (Figure 23). The most obvious difference in overall structure is that the substrate binding loop (residues 195-205) is ordered in the structure of InhA bound with compound **7**, while it is always disordered in the structures of InhA bound with substrate, triclosan, and all other triclosan derivatives solved to date. Sullivan suggested that the ordering of the substrate-binding loop is related to the binding of the ligand. The bulky 5-cyclohexylmethyl group of **7** may have more extended hydrophobic interactions with the substrate-binding pocket, helping to order this flexible loop. Indeed, the sidechain of Met199 on the substrate-binding loop flips approximately 100° to form a hydrophobic interaction with the cyclohexyl group at a distance of 3.6 Å. Compared to the inhibitor 5-pentyl-2-phenoxyphenol, the B-ring of compound **7** also rotates about 30° to allow the 4'-chloride to form a weak hydrogen bonding interaction with the amide N-H of Met98 at distance of 3.2 Å. It is not so obvious from a structural point of view why

the branched inhibitors are less potent than the unbranched inhibitors of the same chain length. A possible explanation could be steric clashes of the sidechains of Phe149 and Met199 with the branched alkyl groups.

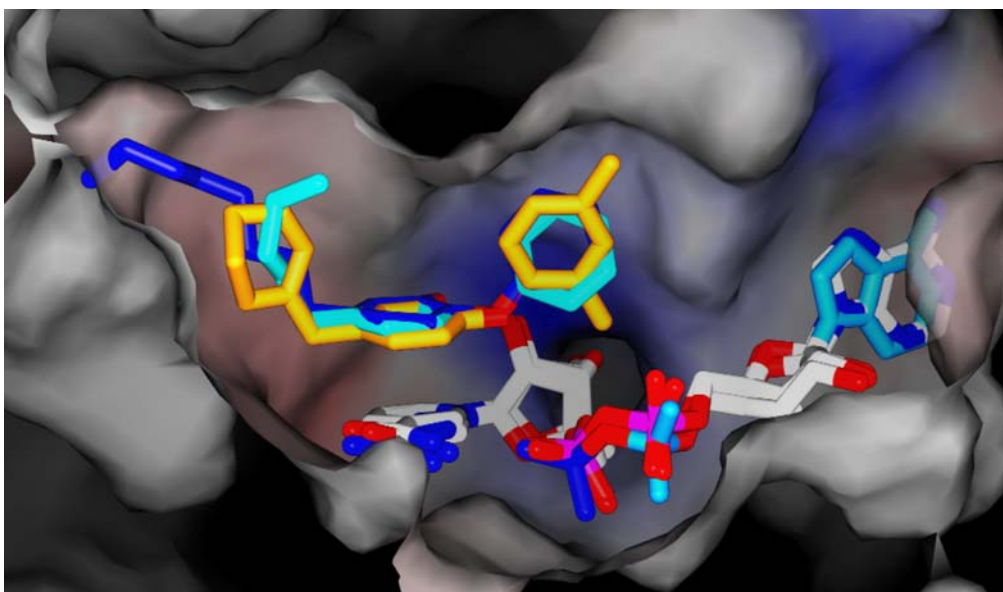
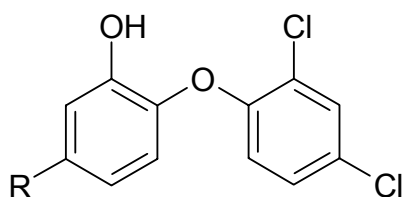


Figure 23 Surface Diagram of Structure of InhA in Complex with 7.

The structure of InhA in complex with compound 7 is superimposed with structures of InhA bound to two other triclosan analogs with linear alkyl substituents

Given the fact that the inhibitors with bulky alkyl substituent showed enhanced *in vitro* activity against InhA than triclosan, we were interested to determine if aryl and heteroaryl groups would also be accommodated in the pocket. A linker varying from 0 to 3 carbons was designed between aryl/heteroaryl groups and the 5-position carbon on the A-ring to control the flexibility and placement of the aryl/heteroaryl moiety (Table 19). A relationship between the length of the linker and the inhibitory potency is observed. Inhibitors with a phenyl or pyridyl group directly attached at the 5-position of the A-ring have very low potency against InhA. In addition, the potency of those inhibitors correlated to the position and type of substituents on the phenyl ring. Although the crystal structures for InhA bound with 5-phenyl or 5-pyridyl triclosan derivatives have not been obtainable to date, most likely due to their respective low binding affinities, the inhibitors were readily modeled into the active site based on the structure of InhA:triclosan. It is clear that the hydrophobic binding pocket is not directly adjacent to the A-ring of triclosan. Instead, the A-ring and the hydrophobic pocket are separated by a distance of 2.5 Å. A directly connected bulky group appears to be too close to the sidechains of Phe149 (shortest distance is about 1.7 Å between the aryl substituent and the phenyl group of Phe149), leading to steric clashes. Within the 5-phenyl series of triclosan derivatives, the *ortho*- substituent on the phenyl ring may “clash” with the sidechain of Leu218, while the *para*-substituent may “bump” into Tyr158. In contrast, the *meta*- methyl on the phenyl ring may be in favorable van der Waals contact with the sidechains of Met155 and Tyr158. Indeed, *m*-tolyl **22** is more potent than the corresponding *o*- and *p*- derivatives **20**, **21**, and **23** by 1.5, 12, and 12 fold respectively.

Table 19 Inhibition of InhA by Triclosan Analogs with Aryl Groups at 5-Position



Compound	R <sub>1</sub>	R <sub>2</sub>	IC <sub>50</sub> (nM)
<b>Triclosan</b>	Cl	Cl	1100
<b>14</b>	2-pyridyl	CN	>10000
<b>15</b>	3-pyridyl	Cl	>10000
<b>16</b>	4-pyridyl	CN	>10000
<b>17</b>	CH <sub>2</sub> (2-pyridyl)	Cl	29
<b>18</b>	CH <sub>2</sub> (3-pyridyl)	Cl	42
<b>19</b>	CH <sub>2</sub> (4-pyridyl)	CN	75
<b>6</b>	Ph	Cl	>10000
<b>20</b>	<i>o</i> -CH <sub>3</sub> -Ph	Cl	1334
<b>21</b>	<i>o</i> -CH <sub>3</sub> -Ph	CN	>10000
<b>22</b>	<i>m</i> -CH <sub>3</sub> -Ph	Cl	871
<b>23</b>	<i>p</i> -F-Ph	Cl	>10000
<b>24</b>	CH <sub>2</sub> Ph	Cl	51
<b>25</b>	(CH <sub>2</sub> ) <sub>2</sub> Ph	Cl	21
<b>26</b>	(CH <sub>2</sub> ) <sub>3</sub> Ph	Cl	50



Therefore, instead of a directly connected aryl/heteroaryl group, molecular modeling suggests a short carbon linker to better position the hydrophobic aryl group into the target pocket.

Consistent with molecular modeling, an evaluation of 5-position substituents of the type  $(\text{CH}_2)_n\text{Ar}$  demonstrates that carbon linkers (C1 to C3) dramatically increased the inhibitory potency versus purified enzyme. Among them, compound **25** has the highest potency of  $\text{IC}_{50} = 21 \text{ nM}$ , representing a 500-fold increase compared to **6**, a phenyl derivative. It is the most potent triclosan derivative against purified InhA we have studied to date. Compounds **24** and **26**, with similar activities against InhA ( $\text{IC}_{50}$  equal to 51 nM and 50 nM, respectively), are also potent inhibitors. The crystal structures of InhA bound with **24** and **25** showed that the  $(\text{CH}_2)_n\text{Ph}$  group extended into the pocket and formed predominantly hydrophobic interactions with residues Leu218, Ile215, Phe149, Met199, and Pro193 (Figure 24). The triclosan backbone atoms of **24** and **25** are in nearly same positions, although their 5-substituents are clearly different. The phenyl group of **24** is positioned in the center of the binding pocket, while the phenyl group of **25** protrudes about 2 Å deeper into, and is closer to the end of, the binding pocket due to the longer carbon linker. The sidechains of Leu218 and Ile215 rotate 30° and shift 1.5 Å to accommodate the phenyl ring of **25**, which also flips ca. 70° from the position of the phenyl group of **24**. All of these conformational changes suggest that there are more hydrophobic interactions between the active site residues and the 5-substituent of **25** than that of **24**, which may explain why **25** is twice as potent than **24** against InhA.

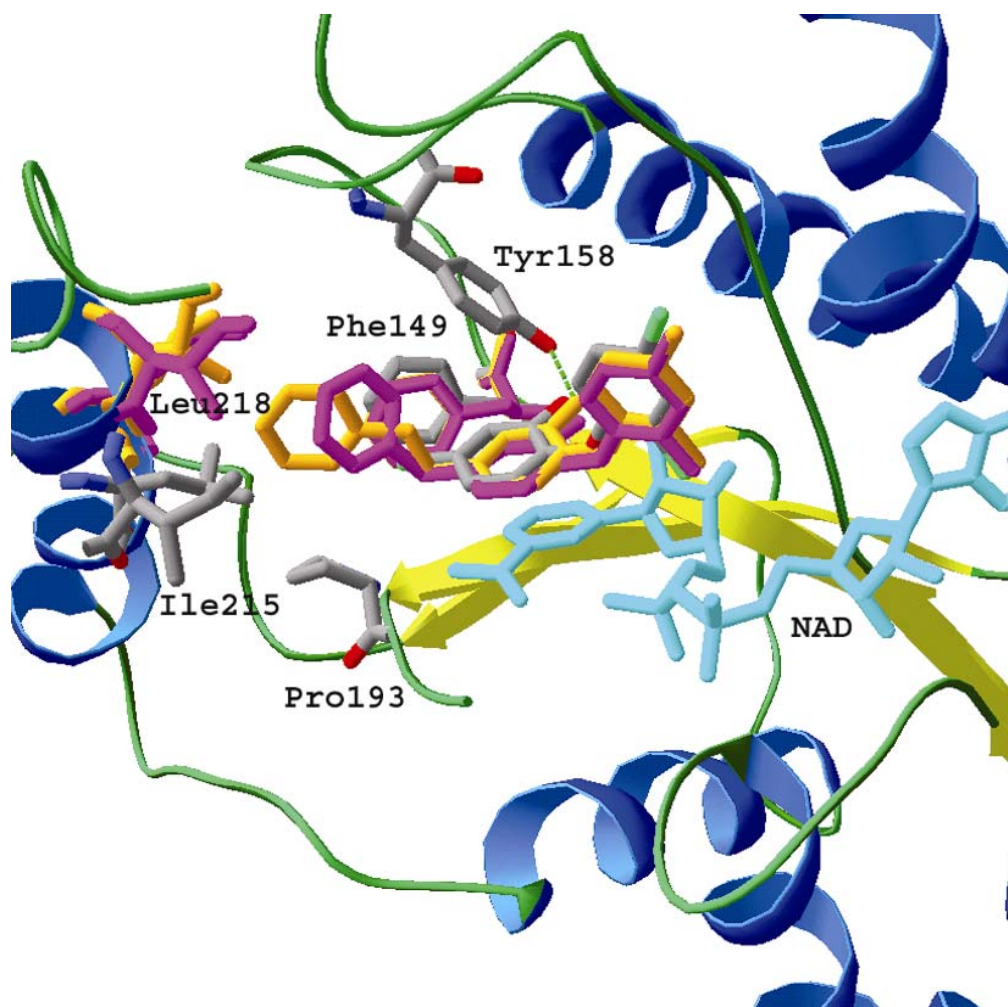


Figure 24 The Superimposition of Structures of InhA Bound with **24** and **25**

To further increase the potency of the 5- (CH<sub>2</sub>)<sub>n</sub>Ph series, derivatives with 5-CH<sub>2</sub>(n-pyridyl) (n=2, 3, 4) substituents (compounds **17** - **19**) were synthesized and examined for *in vitro* activity. The compounds were designed in light of our knowledge of the structure of InhA:INH-NAD. In this crystal structure, the pyridyl group of the INH-NAD adduct formed a hydrogen bonding interaction with a buried water molecule. In addition, some residues with charged sidechains, such as Glu219 and Arg195, are observed proximal to the binding site of the triclosan derivatives. Pyridyl analogs **17** - **19** are all potent inhibitors (IC<sub>50</sub> < 80 nM), and their IC<sub>50</sub> values are in the order **17** < **18** < **19**. Compared to the benzyl derivative **24**, the activity of **17** increases twofold while the activity of **18** increases only slightly, and **19** is less potent. Superimposition of the structures of InhA:NAD<sup>+</sup>:**17** and InhA: NAD<sup>+</sup>:**24** demonstrates that the pyridyl ring of **24** extends into the substrate-binding site and forms hydrophobic interactions with the sidechains of Tyr158 and Phe149 at a distance of 3.6 Å (Figure 25). These interactions are similar to those of the phenyl group in InhA:NAD<sup>+</sup>:**24**. However, the nitrogen atom on the 2-position of the pyridyl ring also forms a hydrogen-bonding interaction with the sidechain carboxylate oxygen of Glu219 through a water molecule. Modeling studies indicate that similar interactions could also exist for **18**, but are not likely for **19** based on the distance between the pyridyl nitrogen atom and the sidechain oxygen of Glu219. The nitrogen atom of **19** potentially points to the hydrophobic sidechains of residues Met155 and Leu218, which may not be favored energetically.

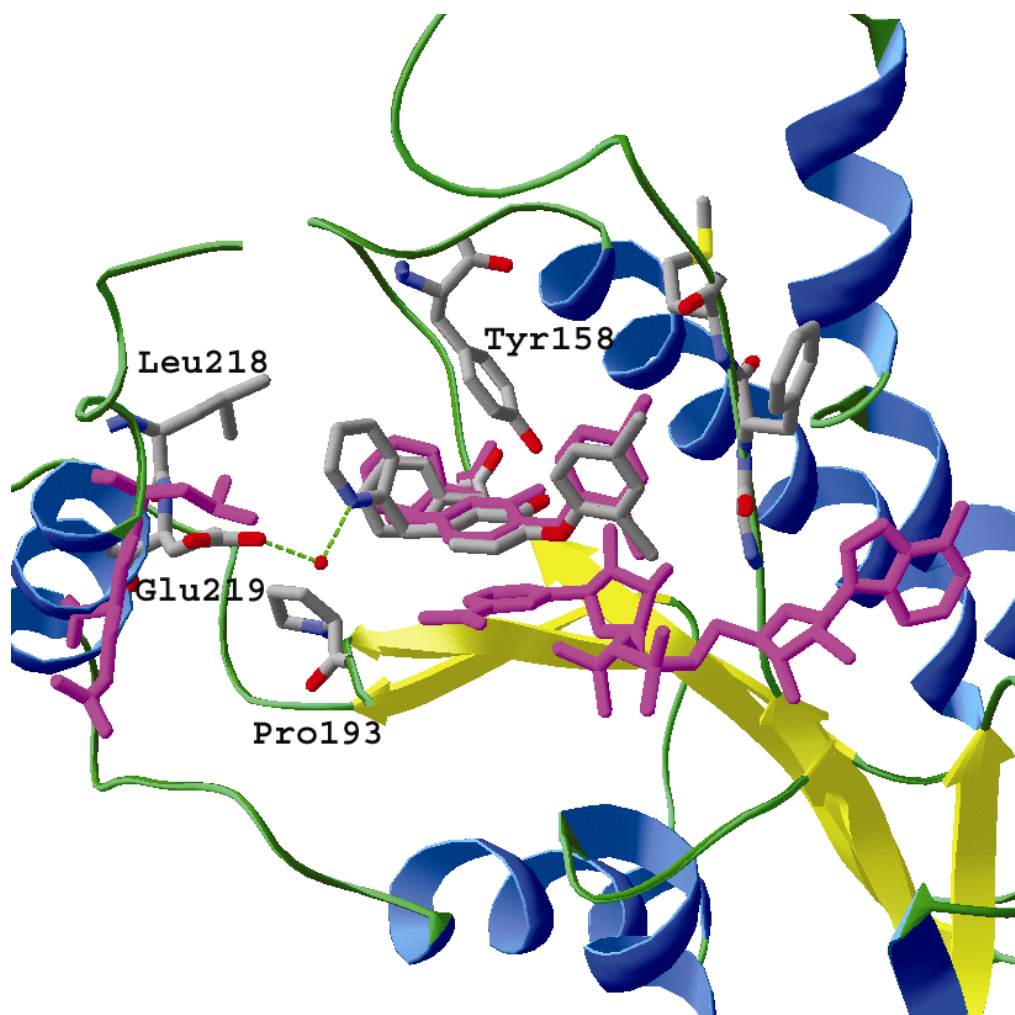


Figure 25 Structure of InhA in Complex with Compound 17

### **High Throughput Screen Identified a Series of New Inhibitors**

The compounds SB363360, SB643152 and SB346352 were found potent against *Staphylococcus aureus* and *Haemophilus influenzae* enoyl acyl reductases by high throughput screening. Interestingly, those aminopyridine derivatives showed even higher inhibitory potency against InhA *in vitro*. Therefore those compounds were utilized as new leads for the design of new inhibitors of InhA. Since it was previously reported that the indole moiety was necessary for inhibitory potency, this moiety was retained and further modifications were done on the rest of the moiety. A series of compounds with similar structures were synthesized and assayed to look for improved potency. Among them, SB643152, SB363360 and SB389656 showed high potency ( $IC_{50} < 100$  nM). The crystal structures of InhA in complex with SB643152, SB363360, SB533394 and SB389656 were determined to establish the structure-activity relationship. The structural basis for the high potency was revealed by the complex structures. The carbonyl oxygen of the carbamide group forms hydrogen bonds with Tyr158 of InhA as well as the ribose 2'-hydroxyl of  $NAD^+$ . This carbonyl oxygen occupies the same position as the hydroxyl group of triclosan and the carbonyl group of fatty acyl substrate. The indol group of the inhibitor protrudes into a hydrophobic pocket that accommodates the fatty acyl chain as showed in the structure of InhA in complex with the C16 substrate. The superimposition of the structures of InhA in complex with the inhibitor and the C16 substrate indicated that the rigid indol group of the inhibitor overlaps with the turning point of the flexible fatty acyl chain. To accommodate the indol group, the side chain of Tyr158 is forced to flip about  $90^\circ$  and in a position where its hydroxyl group can form hydrogen-bonding

interaction with the carbonyl group of the inhibitor. The interactions in this region are conserved in all four inhibitors, which is believed to be the molecular basis of the high potency. Some additional interactions could be found in the rest of the moiety of the inhibitor. For example, the backbone oxygen of Met98 forms hydrogen bonds with hydroxyl group of the inhibitor SB363360 at distance 2.8 Å. Similarly, the amine groups of the SB533394 and SB643152 are also within hydrogen bonding distances to the backbone oxygen of Met98. The amine group of SB389656 forms hydrogen bond with the oxygen of the phosphate group of  $\text{NAD}^+$ . In addition, the pyridine group of SB643152 and the bulky alkyl group of SB363360 are found to stick out of the major portal of the substrate-binding site. Since these groups are in the binding interface of InhA and acyl carrier protein, they could potentially interact with the acyl carrier protein and affect the substrate binding to InhA.

### **Computer Program Based Virtual Screen of ZINC Chemical Database**

Computer program based virtual screen was conducted using the docking program FlexX to search for new leads of InhA from the ZINC 3.3 million compound database. The free ZINC database contains commercial available chemicals annotated with molecular weight, calculated LogP, and some other physical properties. The crystal structure of InhA in complex with SB643152 was used to generate the docking site. The docking site was defined as the area that is within 7 Å distance to at least one of the atoms of inhibitor SB643152. The inhibitor SB643152 and all water molecules were removed from the coordinate file, while the cofactor  $\text{NAD}^+$  remained intact. Charges

were assigned to both protein and ligand atoms as described by Gasteiger and Marsili. FlexX programs runs on the GRID system using the default parameters provided in SYBYL 7.0. The results were retrieved and analyzed using score functions in five separate categories, FlexX, DRUG, PFM, GOLD and CHEM scores. 35 top hits were selected out as the intersection of the top 2% of each of the five categories. 11 compounds of the 35 hits were commercial available and were subsequently evaluated by enzyme assay. Only three of them showed decent inhibitory activity ( $IC_{50} < 10 \mu M$ ) *in vitro* (Table 20). The low hit rate of the docking program is possibly caused by the following reasons: 1) the inaccuracy of the energy calculation, especially for the hydrogen bonding interactions. Hydrogen bonding interactions are usually more sensitive to the distance and geometry. 2) The docking site is viewed as a rigid body in the energy minimization. 3) The charge assignment of the ligand may not be correct, due the unknown chemical environment inside the docking site.

Table 20 The Evaluation of the “Hits” Obtained from Virtual Screen

ZINC ID	IC50 ( $\mu$ M)
876044	>10
987104	7.8
1059860	>10
1154675	>10
1212620	5.5
1954491	>10
1954493	>10
2293938	>10
2863142	>10
3620257	9.1
3672291	>10



## CHAPTER V

### STRUCTURE AND ACTIVITY STUDIES OF THE *MTB* $\beta$ - LACTAMASE

#### BACKGROUND

For over 40 years, drugs that target cell wall biogenesis and chromosomal replication have been utilized as frontline agents in drug mixtures to treat *Mycobacterium tuberculosis* (102). The emergence of multidrug-resistant strains that resistant to these widely available and affordable anti-tubercular drugs becomes a growing clinical problem. New anti-tubercular drugs are needed to treat infections caused by these strains.

#### **Transpeptidase, an Essential Enzyme in Cell Wall Biosynthesis**

DD-transpeptidases, also known as penicillin-binding proteins (PBPs) catalyze the cross linking of two peptidoglycan strands, which is an essential step in bacterial cell wall biosynthesis. The reaction is initiated by the cleavage of the D-alanyl-D-alanine bond of one strand of the peptidoglycan, followed by the linkage of the second strand of the peptidoglycan (Figure 26). DD-transpeptidases (PBPs) belong to the group I of the -

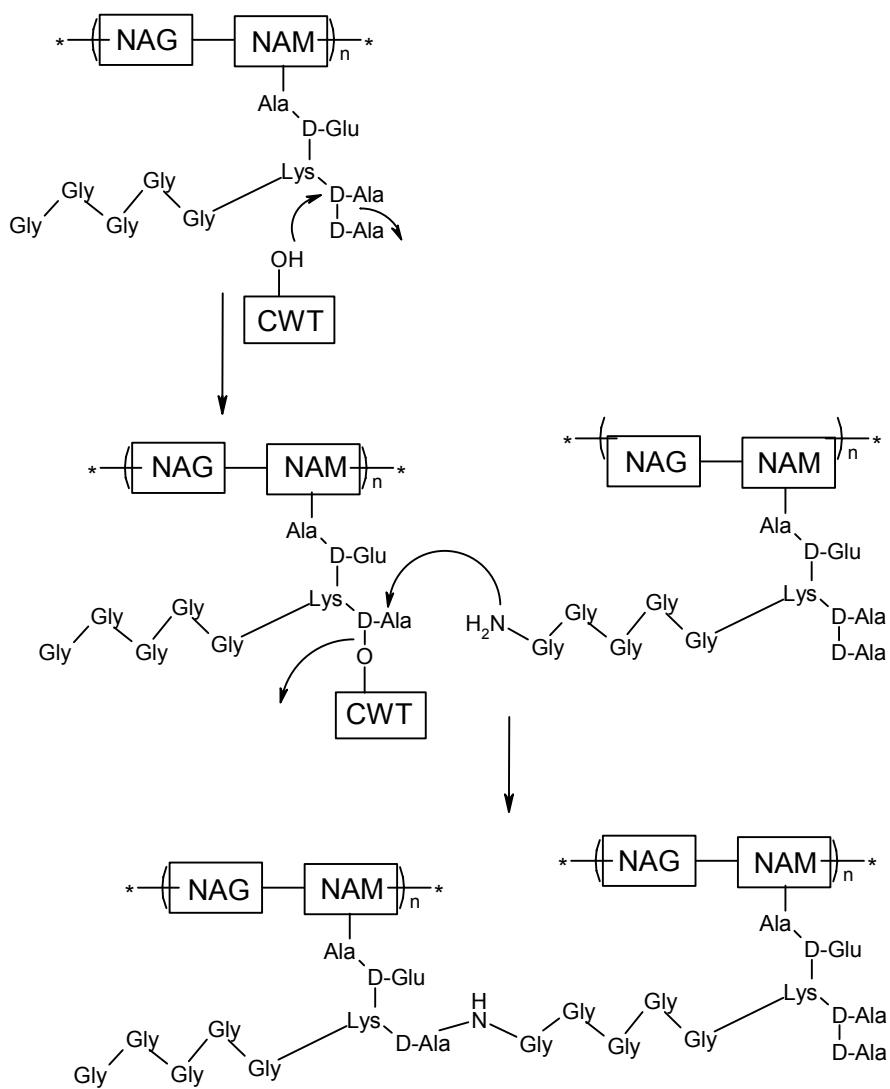


Figure 26 Cross-Linking of Peptidoglycan Strands by Cell Wall Transpeptidases

SxxK acyltransferase family. The PBPs are usually classified based on the molecular weight. The low molecular weight PBPs are single catalytic entities possessing DD-peptidase, esterase and thiol esterase activities. The high molecular mass (HMM) PBPs are multi-domain enzymes with a non-penicillin-binding domain at the N-termini and a penicillin-binding domain at the C-termini.

### **Mode of Action of $\beta$ Lactam Antibiotics**

Penicillin was the first  $\beta$ -lactam antibiotic discovered and it was used extensively to treat bacterial infections from 1940's to 1960's (135). Over the past three decades, many other potent  $\beta$ -lactam antibiotics, such as cepheems, monobactams and carbapenems have been developed and are currently being utilized to treat a wide range of bacterial infections. In fact, the high efficacy, specificity, and low rate of adverse reactions make  $\beta$ -lactam antibiotics the most prescribed antimicrobial agents worldwide (136). All  $\beta$ -lactam antibiotics show structural similarity to the natural substrates of the bacterial DD-transpeptidases. Inhibition of DD-transpeptidases by  $\beta$ -lactam antibiotics leads to disruption of the synthesis of the bacteria's cell wall (Figure 27) (137,138).

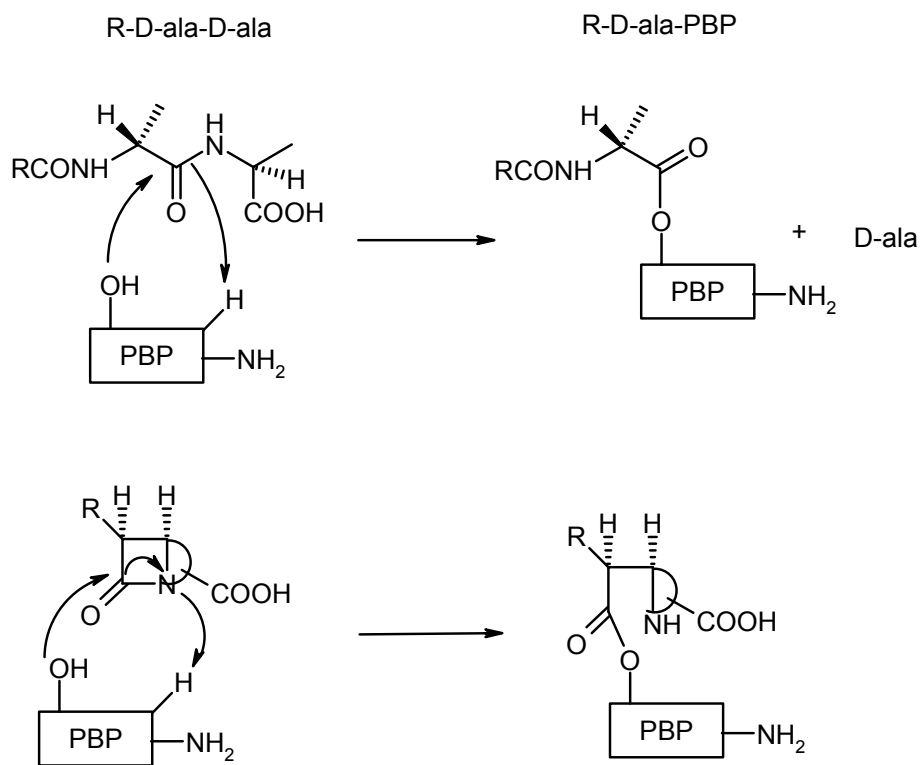


Figure 27 Mechanism of Action of  $\beta$ -lactam Antibiotics

## **Function of $\beta$ Lactamase and $\beta$ Lactam Resistance**

A significant increase in the number of  $\beta$ -lactam resistant strains of pathogenic bacterial has been reported over the past few years (139). Resistance to the  $\beta$ -lactams primarily occurs through the horizontal transfer of  $\beta$ -lactamase genes contained on plasmids (140).  $\beta$ -lactamases inactivate  $\beta$ -lactam antibiotics by efficiently hydrolyzing the amide group of the  $\beta$ -lactam ring, which is the prevalent mechanism of bacterial resistance to  $\beta$ -lactam antibiotics (141).  $\beta$ -lactamases are grouped into four classes, A, B, C, and D based on their amino acid sequence homology and general enzyme catalytic properties (Figure 29) (for a review of  $\beta$ -lactamases see (142)). Historically, class A  $\beta$ -lactamases were found to have much greater penicillinase activity than cephalosporinase activity. However, several new class A  $\beta$ -lactamases that are active against new generations of cephalosporins and carbapenems have been discovered. They are typically grouped as Extended Spectrum  $\beta$ -lactamase (ESBL) to differentiate from the non- Extended Spectrum  $\beta$ -lactamase (non-ESBL).

## ***Mtb* $\beta$ Lactamase – a Potential Anti-TB Drug Target**

Although it was found in the early studies that penicillin inhibited the growth of *Mycobacterium tuberculosis* under certain conditions, there have been few studies on the antitubercular effect of  $\beta$ -lactam antibiotics, probably due to the early discovery that *Mtb* possesses intrinsic resistance to these antibiotics (143). The production of  $\beta$ -lactamase has been proposed to be the most significant reason for mycobacterial resistance (144,145). Other factors that were believed to contribute to the ineffectiveness of  $\beta$ -

lactam in *Mtb*, include cell envelope permeability and low PBP binding affinity for  $\beta$ -lactams (146,147). Although *Mtb*'s outer cell wall, formed by waxy mycolic acid layer, was thought to make a virtually impenetrable barrier, recently,  $\beta$ -lactams have been shown to readily permeate the *Mtb* cell wall and bind the penicillin-binding proteins (144).

One approach that holds promise to counter *Mtb*'s inherent resistance is to utilize  $\beta$ -lactamase inhibitors in combination with  $\beta$ -lactam antibiotics to increase their efficacy (144). For example, class A  $\beta$ -lactamases are generally sensitive to serine  $\beta$ -lactamase inhibitors such as clavulanate and sulbactam (148,149). These mechanism-based inhibitors covalently cross link two serines in the active site that are essential to the hydrolysis of  $\beta$ -lactams and permanently inactivate  $\beta$ -lactamase. The combinations of  $\beta$ -lactam antibiotics and  $\beta$ -lactamase inhibitors have been used clinically to treat a wide range of bacterial infections (139). While they are not being used clinically, these combinations have also been shown to be somewhat effective against most strains of mycobacteria, including MDR strains *in vitro* (150-152). In fact, minimum inhibition concentrations (MICs) of  $\beta$ -lactams are less than 0.5  $\mu\text{g/ml}$  to MDR strains of *Mtb*, when used in combination with both  $\beta$ -lactamase inhibitors and anti-tubercular drug ethambutol (153). These results offer the possibility for treatment of TB with highly potent inhibitors in combination with  $\beta$ -lactam antibiotics.

*blaC* (Rv2068c) is the only gene in the *Mtb* genome that shows any detectable homology to any of the known  $\beta$ -lactamases. Recently, the Pavelka lab showed that the deletion of the *blaC* gene resulted in an increase susceptibility of *Mtb* to  $\beta$ -lactam

antibiotics by 8-256 fold (154,155), which confirmed the primary role of *blaC* in *Mtb* resistance. In previous studies, partially purified *Mtb* BlaC was classified as class A  $\beta$ -lactamase type 2b based on its relatively high level of penicillinases activity (Figure 28) (156,157).

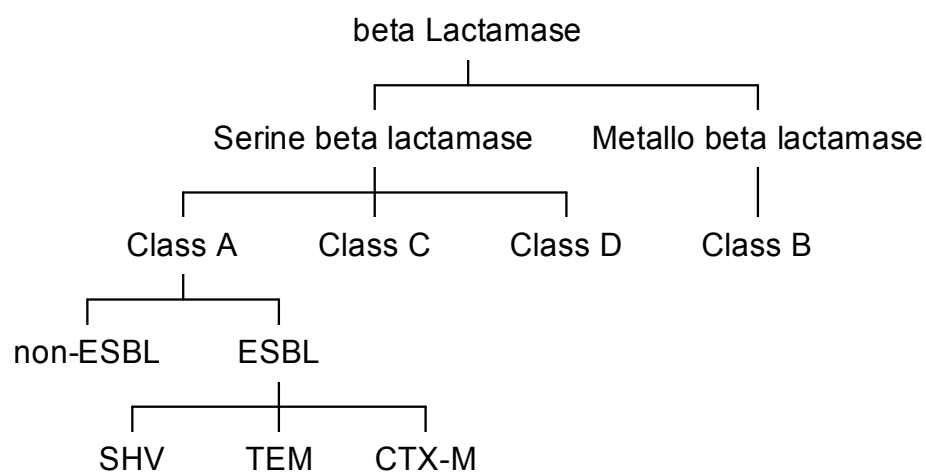


Figure 28 Classification of  $\beta$ - lactamase

## METHODS

### Cloning, Expression and Purification

*E. coli* expression plasmids of the wild type and a truncated form (41-307) of *Mtb blaC* were cloned from genomic DNA (C.S.U. N01-AI-75320). The amplified product was inserted into pET28b (EMD Biosciences #69865-3) using the NdeI and Hind-III restriction sites. This plasmid was subsequently transformed into *E. coli* BL21 (DE3) (EMD Bioscience #69387-3) and cultured in LB-Miller media containing 50 µg/ml kanamycin at 37 °C until OD<sub>600</sub> reached 0.8. Expression of *blaC* gene was induced for 20 hours at 16 °C by addition of 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). Cells were harvested by centrifugation and re-suspended in 25 mM Tris-HCl (pH 8.0), 500 mM NaCl, 2 mM β-mercaptoethanol and were lysed by French Press. After treatment with 1mM DNaseI, the insoluble material was removed by centrifugation. By SDS PAGE, full-length BlaC goes into the insoluble pellet, while ~90% of truncated BlaC (41-307) remained soluble. Therefore, further purification was performed only on truncated BlaC (41-307). The supernatant containing soluble protein was applied to a Ni<sup>2+</sup>-loaded HiTrap chelating column (AP Biotech) using a fast protein liquid chromatography system and eluted through an imidazole gradient (0-0.5M). BlaC (41-307) was eluted at imidazole concentration of 120 mM. Fractions that contain the BlaC (41-307) were pooled and dialyzed against 25 mM Tris-HCl (pH 8.0), 40 mM NaCl, 1 mM DTT. The samples were further purified by gel-filtration on a Superdex 200 column to separate the monomeric protein from aggregated material using dialysis buffer. The protein was concentrated to approximately 10 mg/ml. BlaC (41-307) was



obtained at 45 mg/L of *E. coli* culture and appeared homogeneous by SDS-PAGE and Coomassie Blue staining.

### **Kinetic Assays**

Enzyme kinetic evaluation was completed for the substrates ampicillin ( $\Delta\epsilon_{235} = -820 \text{ M}^{-1} \text{ cm}^{-1}$ ), cephalothin ( $\Delta\epsilon_{262} = -7660 \text{ M}^{-1} \text{ cm}^{-1}$ ), cefoxitin ( $\Delta\epsilon_{270} = -8380 \text{ M}^{-1} \text{ cm}^{-1}$ ), ceftazidime ( $\Delta\epsilon_{260} = -10200 \text{ M}^{-1} \text{ cm}^{-1}$ ) and meropenem ( $\Delta\epsilon_{299} = -2500 \text{ M}^{-1} \text{ cm}^{-1}$ ) (158). All assays were repeated in triplicate and carried out on a Cary 100 Bio Spectrophotometer at room temperature by monitoring the UV absorbance change of  $\beta$ -lactams (158). The *Mtb* BlaC activity assays were performed in 100 mM phosphate buffer (pH 7.5) by adding 10 nM enzyme into substrates solution.  $K_m$  and  $V_{max}$  were determined by analyzing Hanes Plot (159). Inhibition studies of clavulanate were performed by adding enzyme into solutions containing the substrate ampicillin and inhibitor clavulanate with increasing concentrations (0.1-600  $\mu\text{M}$ ). During the measurement time, no time-dependent loss of enzyme activity could be detected. Value for  $K_i$  was determined from x-intercept of a Dixon plot, assuming competitive inhibition.

### **Crystallization of BlaC (41-307)**

Crystallization of the truncated enzyme, missing the first 40 amino acids, was accomplished by the hanging drop vapor diffusion method (85). Initial crystallization conditions were discovered by sparse matrix screen using Crystal Screens from Hampton Research. Hanging drops containing 2  $\mu\text{l}$  of protein solution at 10 mg/ml and 2  $\mu\text{l}$  of

buffer (0.1 M Tris-HCl pH 8.0, 2.0 M  $\text{NH}_4\text{H}_2\text{PO}_4$ ) were equilibrated at 16 °C in Linbro plates against 1 ml of the same buffer. Protein crystals appeared in 4 days as thin plate-shaped clusters and were improved to larger rod-like clusters by co-crystallization with its substrate ampicillin. Macro seeding was then used to obtain high quality single crystals.

### **Data Collection and Processing**

A nearly complete diffraction dataset to 1.7Å resolution was collected from a single crystal at 121 K using a cryo-protection solution consisting of reservoir solution (0.1 M Tris-HCl pH 8.0, 2.0 M  $\text{NH}_4\text{H}_2\text{PO}_4$ ) with the addition of 20% glycerol. The beam line 14BMC at the Advanced Photon Source (APS), Argonne National Laboratory was used to collect the dataset using one-degree oscillation widths over a range of 180°. The data were integrated and reduced using HKL2000 (86). The completeness of low-resolution data is slightly lower, due to the high rejection rate caused ice rings. Crystals of BlaC belongs to the orthorhombic space group  $\text{P}2_12_12_1$  with cell dimensions  $a = 42.74$  Å,  $b = 71.28$  Å,  $c = 85.17$  Å,  $\alpha = 90^\circ$ ,  $\beta = 90^\circ$ ,  $\gamma = 90^\circ$ , with one molecule in each asymmetric unit.

### **Structure Determination and Model Refinement**

Initial phases were obtained by molecular replacement with the computer program MolRep, part of the CCP4 package (160), using the related structure of  $\beta$ -lactamase from *Streptomyces albus* G (PDB 1BSG) (161) as a search model. The best

solution gave a correlation coefficient of 34.2% and a  $R$ -factor of 50.9% using data between 25 and 3 Å resolution. The structure from the molecular replacement solution was refined with REFMAC, again in the CCP4 suite (160). The first step in the refinement was to use rigid body optimization, and this resulted in a slightly improved model ( $R_{\text{factor}}$ , 49.9%;  $R_{\text{free}}$ , 51.8%; 54.6-1.7 Å). Electron density maps were produced and manual rebuilding was carried out, using XtalView (89). The model was further refined with cycles of model building and REFMAC restrained refinement (160). During the final cycles of the refinement, water molecules were added based on peaks above 3- $\sigma$  in the  $F_o-F_c$  electron density maps that were within hydrogen-bonding distances from appropriate protein atom. The resulting structure had good geometry (88% of the non-glycine residues were in most favorable conformation and none in disallowed conformation) and  $R_{\text{factor}}$  of 17.9% and  $R_{\text{free}}$ , 21.4%. PROCHECK (160) revealed no disallowed  $\phi$ ,  $\psi$  and the overall structure  $G$  factor, a measure of a given stereochemical property, was consistently better than expected for this resolution.

### Structure Analysis

DALI (162) was used to search the PDB for proteins having folds similar to *Mtb*, in the order of structural similarity. BlaC and SwissPDB viewer was used to make structural alignments (163). The model was evaluated and analyzed using computer graphic program SPOCK (90).

## RESULTS AND DISCUSSION

### Sequence Analysis of BlaC

*Mtb* BlaC is a 307 amino acid protein that is predicted to have a signal peptide for translocation into the periplasm and a lipid attachment site (residue 14-24), using the computer program ScanProsite (164). This analysis suggests that the mature BlaC is linked via the covalently attached lipid to the outer leaflet of the inner membrane, after translocation to the periplasmic space.  $\beta$ -lactamase lipoproteins have been reported for several Gram-positive bacteria (165,166). However, the enzyme was also predicted, using SignalP program (167), to have a signal peptidase II cleavage site at residues 28, which would eliminate the lipid attachment site in the mature protein. It is therefore, not clear if BlaC exists free or as membrane bound protein in the periplasm.

*Mtb* BlaC shares approximately 40% amino acid sequence identity with other known class A  $\beta$ -lactamases. Interestingly, it has the highest sequence identity to the class A  $\beta$ -lactamase of *Streptomyces clavuligerus* (54%), the bacteria that produces the antibiotic cephamycin and the  $\beta$ -lactamase inhibitor clavulanate.

Three highly conserved regions of the active site, designated as motif I, II, and III, have been identified by multiple sequence alignment of class A  $\beta$ -lactamases from different bacteria. Motif I has a consensus sequence of Arg<sup>61</sup>-(Xaa)<sub>2</sub>-Glu-Xaa-Phe-(Xaa)<sub>3</sub>-Ser-Xaa-Xaa-Lys<sup>73</sup> that extends from the end of the  $\beta$ -strand B2 through a segment of the  $\alpha$ -helix H2. Motif II corresponds to sequence Ser<sup>130</sup>-Asp-Asn<sup>132</sup> and is located between  $\alpha$ -helix H6 and H7. Motif III, located on  $\beta$ -strand B3 and has consensus sequence of Asp<sup>233</sup>-Lys-Thr-Gly<sup>236</sup>. Additionally, all class A  $\beta$ -lactamases have a  $\Omega$  loop that in BlaC contains residues Val<sup>159</sup> – Thr<sup>180</sup>. While these regions are highly conserved, BlaC from *Mtb* has several notable amino acid substitutions. The substitution of a Gly for Asn at residue 132 only occurs in class A  $\beta$ -lactamase from *Mtb* and its closest homolog *S. clavuligerus* (Figure 29). The substitutions of R244A and R169A in *Mtb* are also rare in class A  $\beta$ -lactamases. Another important difference identified from amino acid sequence alignment is a four-residue insertion Gly<sup>145A</sup> to Ala<sup>150</sup> and a two-residue insertion Gly<sup>271</sup> to Tyr<sup>272</sup> in *Mtb* BlaC sequence.

70

```

BLAC  --MGADLADRFAELERRYDARLVVYPATGTAAIE-YRADERFAFCSTFKAPLVAAVLH
BS3    -----DDFAKLEEQFDAKLGIFALDTGTNRTVT-YRPDERFAFASTIKALTGVVLLQ
Toho-1 --M-NSVQQQLEALEKSSGRLGVALINTADNSQIL-YRADERFAMCSTSKVMAAAAVLK
ALBS   ---SDAERRLAGLERASGARLVYAYDTGSGRTVA-YRADELFPMCVFKTLSSAAVLR
MYC    ---APIDDQLAELERRDNVLIGLYAANLQSGRRIT-HRPDEMAMCSTFKGYVAARVLQ
K1     DNN-NTIEEQLNTLEKYSQGRGLGVALINTEDNSQIT-YRGEERFAMASTSKVMAAAVLR
TEM-1  --H-PETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSFKVLLCGAVLS
NMC-A  --N-TKGIDEIKNLETFDNGRIGVYALDTGSGKSFY-YRANERFPLCSSFKGLAAAVLK
SHV-1  --S-PQPLEQIKLSESQLSGRVGMIEMDLASGRTLTAWRADERFPMMSFKVLLCGAVLA
SCLA   -----ADLAALEREHGARLVYALETGTGAEVA-HRADERFAFCSTFKALAAAVLH

```

\* \*

132

```

BLAC  QNPLTH--LDKLIITYTSDDIRSISPVAQ----QHVQTGMTIGQLCDAAIRYSDGTAANLL
BS3    QKSIED--LNQRITYTRDDLNVNYPITE----KHVDTGMTLKELDADASLRYSDNTAQNLI
Toho-1 QSESDKHLNQRVEIKKSDLVYNPIAE----KHVNGTMTLAEELGAAALQYSDNTAMNKL
ALBS   DLDRNGEFLSRRILYTDQDVEQADGAPETGKQPONLANGMTVEELCEVSTASDNCAANLM
MYC    MAEHGEISLDNRVFDADALVPNSPVTE----ARAGAEMTLAELCQAAALQSDNTAANLL
K1     ASEKQAGLLDKNITIKKSDLVAYSPITE----KHLTTGMTLAELSAAATLQYSDNTAMNKI
TEM-1  RIDAGQEQLGRRIHYSQNDLVEYSPVTE----KHLTDGMTVRELCSSAATMSDNTAANLL
NMC-A  GSQDNRLNLNQIVNYNTRSLEFHSPITT----KYKDNMGMSLGDMAAAALQYSDNGATNII
SHV-1  RVDAGDEQLERKIHRYQQDLVDYSPVSE----KHLADGMTVGEELCAAATMSDNTAANLL
SCLA   HHPIRH--LERRVTWTRADVDSISPVTE----DHIATGLTVGQLCDAAIRHSDGTAGNLL

```

\*\*\*

164

```

BLAC  LA-DLGGPGGGTAAFTGYLRS LGD TVSRLDAEEPELNRDPPGDERDTTTPHAIALVLQQL
BS3    LK-QIGGP----ESLKKELRKIGDEVNPERFEPELNEVNPGETQDTSARALATSLQAF
Toho-1 LA-HLGGP----DKVTAFAFARSLGDETFRDRTEPTLNTAIPGDPRDITTPPLAMAQTLKLN
ALBS   LR-ELGGP----AAVTRFVRS LGDRVTRLDRWEPELNSAEPGRVDTTSPRAITRTYGRLL
MYC    LK-TIGGP----AAVTAFAFARSVGERTRLDRWEVELNSAIPGDPRDITTPAALAVGYRAI
K1     LD-YLGGP----AKVTQFARSINDVYRRLDRKEPELNTAIGHDPRDITTSPIAMAKSLQAL
TEM-1  LT-TIGGP----KELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTMPVAMATTLRKL
NMC-A  LERYIGGP----EGMTKFMRSIGDEDFRDLRWELDLNTAIPGDERDITTPAAVAKSLKTL
SHV-1  LA-TVGGP----AGLTAFLRQIGDNVTRLDRWETELNEALPGDARDITTPASMAATLRKL
SCLA   MR-DLGGP----SRLTAYLRGLGDSVSRMDQYEPPELNDHPPHDPDRDITTPRAIASDYRKL

```

\* \*

220

244

```

BLAC  VLGNALPPDKRALLTDWMARNTGAKRIRAGFPADWKVIDKTGT-GDYGRANDIAV-VWS
BS3    ALEDKLPSEKRELLIDWMKRNTG DALIRAGVPEGWEVADKTGA-GSYGRNDIAI-IWP
Toho-1 TLGKALAETQRAQLVTWLKGNTTGSASIRAGLPKSWVVGDKTGS-GDYGTNDIAV-IWP
ALBS   VLGDALNPRDRRLTSSWLLANTTS GDRFRAGLPDDWTLGDKTGA-GRYGTNNDAGV-TWP
MYC    LAGDALSPQQRGLLEDWMRANQTSS--MRAGLPEGWTTADKTGS-GDYGTNDAGIAFGP
K1     TLGDALGQSQRQQLVTWLKGNTTGDNSIKAGLPKHVVVGDKTGS-GDYGTNDIAV-IWP
TEM-1  LTGELLTTLASRQQLIDWMEADK VAGP LLRSALPAGWFIADKSGA-GERGSRGIIAA-LGP
NMC-A  ALGNILSEHEKETYQTLWKGNTTGAARIRASVPSDWVVGDKTGS CGAYGTANDYAV-VWP
SHV-1  LTSQRLSARSQRQLLQWVDDRVAGPLRSVLPAGWFIADKTGA-GERGARGIVAL-LGP
SCLA   VLGDALTPDRRALLTDWLVRNTTGRRIRAGVPSGWRVADKTGT-GNYGRANDIAV-LWP

```

\*

BLAC PTGVPPVVAVMSDR-AGGGYDAEPREALLAEAAATCVAGVLA-----

BS3 PKGDPVVLAVLSSR-DKK--DAKYDDKLI AEATKVVVKALNMNGK-----

Toho-1 ENHAPLVLVTYFTQ-PEQ--KAERRRDILAAAAKIVTHGF-----

ALBS PGRAPIVLTVLTAK-TEQ--DAARDGLVADAARVLAETLG-----

MYC DGQRLLLVMMTRSQAHPD--KAENLRPLIGELTALVLPSSL-----

K1 ENHAPLILVVYFTQ-QEQ--NAKYRKDI IAKAAEIVTKEISNSPQTK-----

TEM-1 DGKPSRIVVIYTTGSQAT--MDERNRQIAEIGASLIKHW-----

NMC-A KNRAPLISVYTTK-NEK--EAKHEDKVI AEASRIAIDNLK-----

SHV-1 NNKAERIVVIYLRDTPAS--MAERNQQIAGIGAALIEHWQR-----

SCLA PRSSPLVVAVMSDR-PGF--RTPPSERLIAEDRGAIVAGLVRHPPTHASGAGAPSRHPGP

Figure 29 Sequence Alignment of Class A  $\beta$ -lactamases of Different Organisms

### **Expression, Purification and Characterization of *Mtb* BlaC**

Initial attempts to express the full-length protein were unsuccessful, primarily due to the N-terminal signal peptide. By truncating the codons for the first 40 amino acids in our expression construct, we could produce large quantities of active enzyme. The truncation and the N-terminal His-tag were not predicted to interfere with the enzymes catalytic function based on what was known about homologous enzymes. The molecular weight of the truncated protein was confirmed by MALDI (Actual mass 30629, calculated mass 30666).

### **Enzyme Activity of *Mtb* BlaC**

Kinetic measurements of the hydrolysis of the substrates ampicillin, cephalothin, cefoxitin, ceftazidime, meropenem as well as the inhibition of the enzyme activity with clavulanate were carried out in order to compare the enzyme activity to other class A  $\beta$ -lactamases (Table 21). Overall, the results show that *Mtb* BlaC has a broad-spectrum substrate profile compared to other class A  $\beta$ -lactamases. While the penicillinase activity of BlaC ( $k_{\text{cat}}/K_m = 293 \mu\text{M}^{-1}\text{s}^{-1}$ ) is about 10-fold lower than most class A  $\beta$ -lactamases, the enzyme showed relatively good activity against cephalothin ( $k_{\text{cat}}/K_m = 103 \mu\text{M}^{-1}\text{s}^{-1}$ ), in contrast to previous studies where *Mtb* was shown to possess only minor cephalothin activity (168). Indeed, we found that the cephalosporinase activity of *Mtb* BlaC is comparable to its penicillinase activity. Furthermore, BlaC shows modest activity ( $k_{\text{cat}}/K_m \sim 10 \text{mM}^{-1}\text{s}^{-1}$ ) against cefoxitin, ceftazidime and meropenem. Inhibition studies indicated that clavulanate, a very potent inhibitor of most of class A  $\beta$ -lactamases ( $K_i <$

0.1  $\mu\text{M}$ ), is a relatively poor inhibitor to *Mtb* BlaC ( $K_i = 2.4 \mu\text{M}$ ). The substrate profile of BlaC is quite similar to that of *M. fortuitum*  $\beta$ -lactamase, which also has relatively high cephalosporinase activity (169).

Table 21 Kinetic Parameters of *Mtb* BlaC with Different Antibiotics as Substrates

Class	Antibiotics	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$K_m$ ( $\mu\text{M}$ )	$k_{\text{cat}}/K_m$ ( $\mu\text{M}^{-1}\text{s}^{-1}$ )	$K_i$ ( $\mu\text{M}$ )
penicillin	Ampicillin	18.5	63	293.65	
cephalosporin	Cephalothin	12.1	117	103.42	
	Cefoxitin	1.1	195	5.64	
	Ceftazidime	0.2	593	0.37	
carbapenem	Meropenem	0.9	279	3.19	
inhibitor	Clavulanate				2.4



## Overall Structure of BlaC

The crystal structure of BlaC was solved using molecular replacement and refined to 1.7 Å resolution. The final  $R_{\text{factor}}$  and  $R_{\text{free}}$  values for all reflections were 0.18 and 0.21, respectively (Table 22). BlaC crystals are orthorhombic, belonging to space group  $P2_12_12_1$ , with one molecule in the asymmetric unit. The refined structure of BlaC includes 265 of the 267 residues in the truncated form of BlaC (41-307) and 248 ordered water molecules. The structure of the N-terminal six histidine tag along with the first two residues Gly-Ala were not defined due to disorder of this region in the electron density map. Like other class A  $\beta$ -lactamases, the BlaC monomer contains an  $\alpha$  domain and an  $\alpha/\beta$  domain. The  $\alpha$  domain consists of six long  $\alpha$  helices (H2, H5, H6, H7, H9 and H10) and three short  $\alpha$  helices (H3, H4 and H8). The  $\alpha/\beta$  domain includes five anti-parallel  $\beta$ -strands (B1-B5), the N- and C- termini  $\alpha$  helices (H1 and H12) and one short  $\alpha$  helix H11. The overall structure of BlaC is very similar to other class A  $\beta$ -lactamase structures. The root mean square difference (RMSD) of the  $C_{\alpha}$  atoms of several class A  $\beta$ -lactamases range between 0.95 – 3.66 Å and are listed in Table 23. The  $C_{\alpha}$  alignments of BlaC with other class A  $\beta$ -lactamases of known structure indicates *Mtb* BlaC is more similar to K1 (rmsd = 0.95 Å) and NMC-A (rmsd = 1 Å) (enzymes from *P. vulgaris* and *E. cloacae*, respectively) compared to SHV-1 and TEM-1 (rmsd > 2.5 Å) (enzymes from *K. pneumoniae* and *E. coli*). ALBS (PDBID 1BSG) was used as search model because it showed significant sequence homology to BlaC. Surprisingly, it showed a fairly large rmsd value (3.66 Å) compared to BlaC. Indeed, several other class A  $\beta$ -lactamases with lower sequence homology are structurally more similar to BlaC.

Table 22 Data Collection, Processing and Refinement Statistics of *Mtb* BlaC

Data collection	
	BlaC
Maximum resolution (Å)	1.72
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
a (Å)	42.7
b (Å)	71.3
c (Å)	85.2
Unique reflections	25908 (2808)
R <sub>sym</sub> (%)	11.4 (30.6)
Completeness (%)	91.1 (100)
Redundancy	6.8 (6.9)
I/σ	23.2 (6.8)
Refinement statistics	
Resolution range (Å)	54.64-1.72
Number of reflections	24570
Number of atoms / subunit	
Protein	2230
Solvent	248
R <sub>cryst</sub> (%)	17.9
R <sub>free</sub> (%)	21.4
Average B-factors (Å <sup>2</sup> )	
Protein	20.7
Solvent	32.8
rmsd from ideal values	
bonds (Å)	0.011
angles (deg)	1.44

$R_{\text{sym}} = \frac{\sum_{\text{hkl}} \sum_i |I_i(\text{hkl}) - \langle I(\text{hkl}) \rangle|}{\sum_{\text{hkl}} \sum_i I_i(\text{hkl})}$ , where  $I_i$  is the intensity of observation and  $\langle I \rangle$  is the mean intensity of reflection.

The numbers given in parentheses denote the respective values of the highest resolution shell.

Table 23 Structure Comparison of Class A  $\beta$ -lactamases from Different Species

Enzyme	Source	PDB_ID	Sequence identity (%)	Rmsd (Å)
BlaC	<i>M. tuberculosis</i>			
BS3	<i>B. licheniformis</i>	1I2S	47.3	1.57
ALBS	<i>S. albus</i>	1BSG	45.0	3.66
Toho-1	<i>E. coli</i>	1IYS	46.5	1.87
SME-1	<i>S. marcescens</i>	1DY6	42.5	0.97
NMC-A	<i>E. cloacae</i>	1BUE	41.1	1.00
K1	<i>P. vulgaris</i>	1HZO	42.9	0.95
MYC	<i>M. fortuitum</i>	1MFO	40.7	1.54
SHV-1	<i>K. pneumoniae</i>	1SHV	40.0	2.74
TEM-1	<i>E. coli</i>	1BT5	38.0	2.85

## Enzyme Active Site

In all class A  $\beta$ -lactamases, the hydrolysis of the  $\beta$ -lactam substrates is accomplished by a nucleophilic attack initiated by an active site serine residue Ser<sup>70</sup>. It has been proposed that Glu<sup>166</sup>, a general base in the active site is the serine70-activating residue, based on the ultrahigh resolution crystal structure (0.91 Å SHV-2 and 0.85 Å TEM-1) studies (170,171). However, others believe that Lys<sup>73</sup> is involved in the activation (172). Although the mechanism of Ser<sup>70</sup> activation is contentious (173), the deacylation mechanism has been established. Glu<sup>166</sup> has been demonstrated to be critical for the deacylation of the acyl-enzyme intermediate based on site-directed mutagenesis studies (Figure 30) (174). The enzyme active site is located at the interface between the  $\alpha$  and the  $\alpha/\beta$  domains (Figure 31). Ser<sup>70</sup> and Lys<sup>73</sup>, located on  $\alpha$ -helix H2 in the center of active site (Figure 32), are completely conserved in all class A  $\beta$ -lactamases and are surrounded by other key residues on  $\beta$ -strand B3 (Lys<sup>234</sup>, Thr<sup>235</sup>, Thr<sup>237</sup>), and the loop region between H5 and H6 (Ser<sup>130</sup>, Gly<sup>132</sup>) as well as the  $\Omega$  loop (Glu<sup>166</sup>). These eight residues are all involved in direct hydrogen bonding interactions with  $\beta$ -lactam substrates. Two bound water molecules, WAT<sup>36</sup> and WAT<sup>65</sup>, are also highly conserved in the structures of all class A  $\beta$ -lactamases, determined to date. WAT<sup>65</sup> is part of a hydrogen-bond network with the side chain oxygen of Ser<sup>70</sup>, as well as the side chain oxygen of Glu<sup>166</sup> at distances of 2.5 Å and 2.5 Å, respectively. This hydrolytic water functions to deacylate Ser<sup>70</sup> from the product after hydrolysis of  $\beta$ -lactam substrates (175-177). WAT<sup>36</sup> is in the oxyanion hole, forming hydrogen bonds with the backbone nitrogen of Ser<sup>70</sup> and the backbone nitrogen of Thr<sup>237</sup> at distance of 2.6 Å and 2.7 Å,

respectively. Although ampicillin improved the diffraction quality of the crystals, neither the substrate nor any products could be identified in the active site.

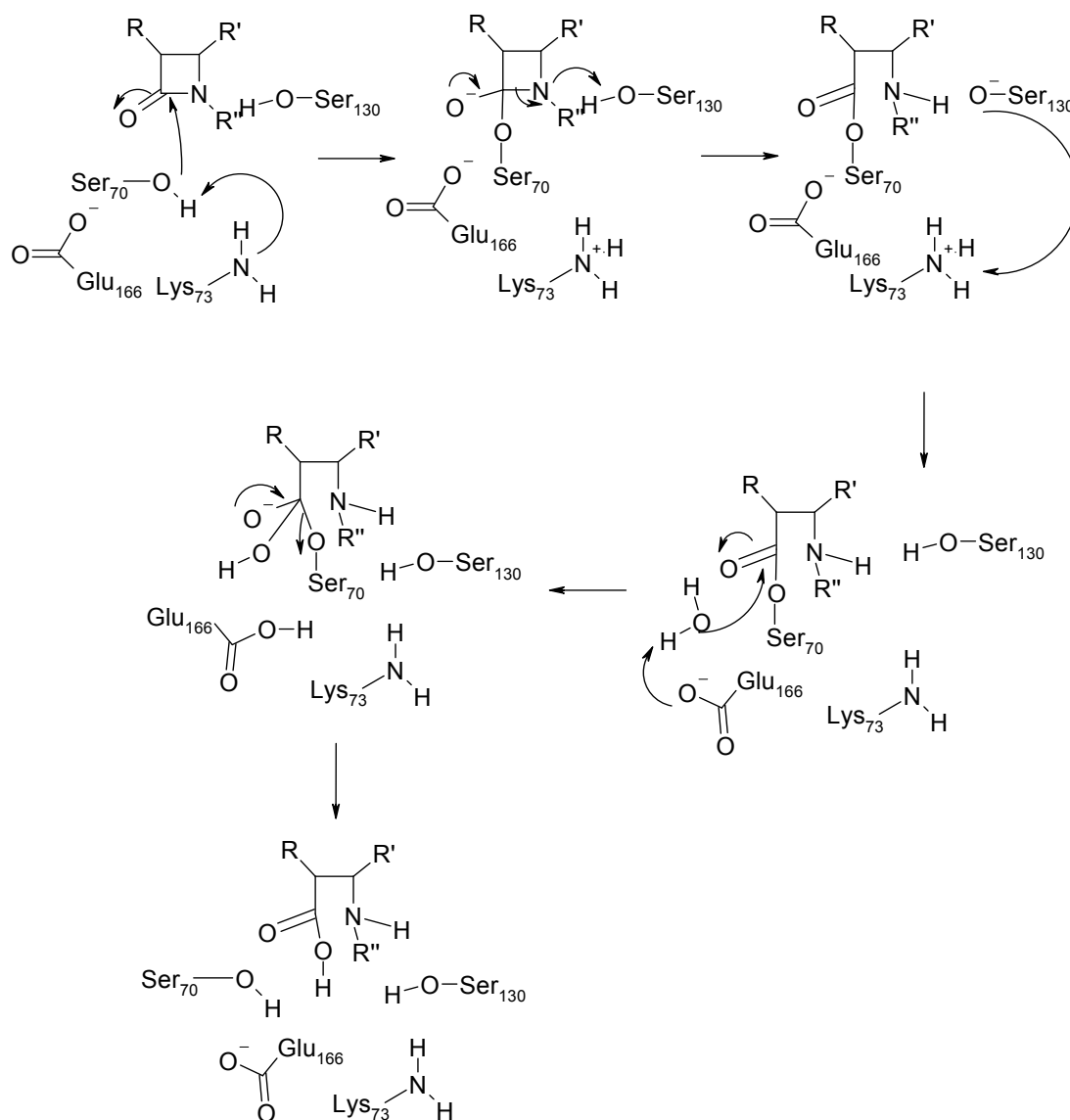
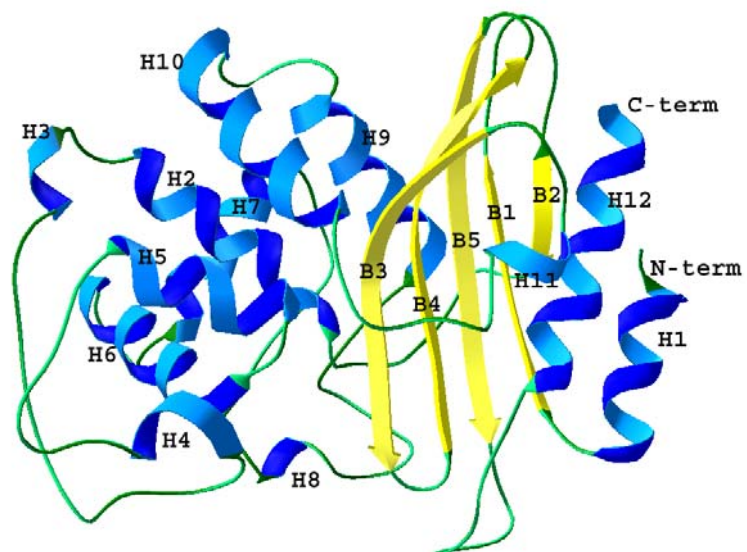
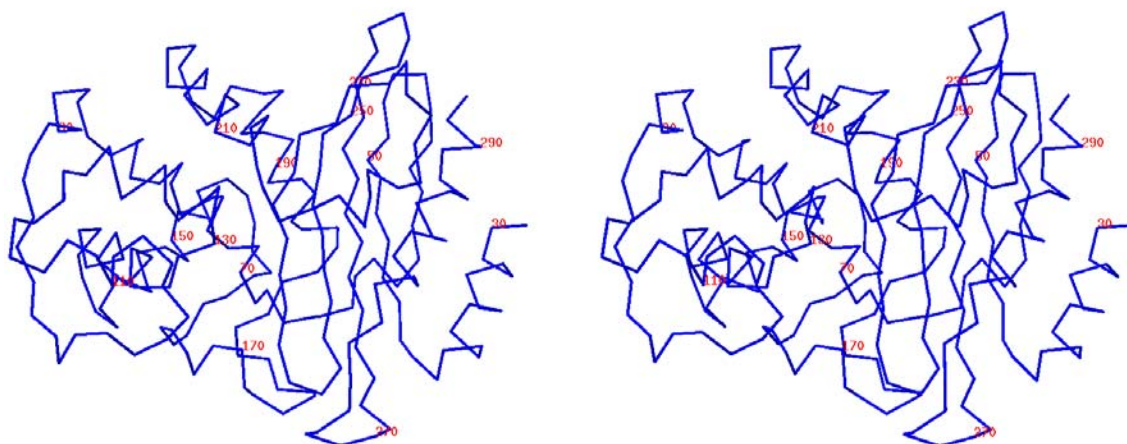


Figure 30 Mechanism of Class A  $\beta$ -lactamase

A.



B.

Figure 31 Overall Structure of *Mtb* BlaC

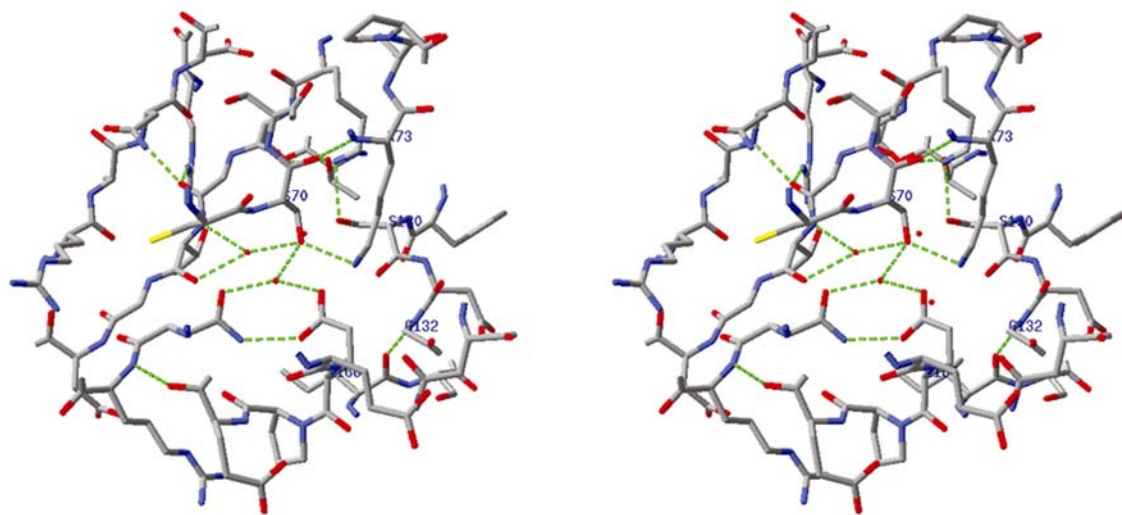


Figure 32 Stereo View of Active Site of *Mtb* BlaC

Although the overall structure of *Mtb* BlaC is similar to other class A  $\beta$ -lactamases, some differences are clear, the most obvious difference being the size of the active site. The active site of *Mtb* BlaC is significantly larger than the other class A  $\beta$ -lactamases with known structures. Three factors that contribute to the greater size of the active site are: 1) amino acid substitution of several key residues, 2) rearrangement of conserved side chains, and 3) a four residue long insertion in the active site region. The  $\beta$ -lactam-binding motif Ser-Asp-Asn is highly conserved among the class A  $\beta$ -lactamases. In *Mtb* BlaC, Glycine occupies the position of asparagine. In the class A  $\beta$ -lactamase with structures determined thus far, the side chain of this asparagine points towards the center of active site, forming a hydrogen bond with the side chain nitrogen of Lys<sup>73</sup>. The carbamide group of the asparagine is in close contact with the ester carbonyl group of penicillin and cepheems in the complex structures published (178). The 6 $\alpha$ -1R-hydroxyl group of imipenem was also observed to hydrogen bond with the side chain oxygen of Asn<sup>132</sup> in the structure of *E. coli* TEM-1 (PDBID: 1BT5) (179) and *E. cloacae* NMC-A (PDBID: 1BUE) (180). It was proposed that extra space at this location, created by about a 1 Å shift of Asn<sup>132</sup> (equivalent to Gly<sup>132</sup> in *Mtb* BlaC) and its side chain pointing away from catalytic center, would accommodate the substituents at the R6 position of carbapenem and therefore increase carbapenemase activity (180,181). In fact, this is true for all the carbapenemases of known structures. Having a glycine instead of asparagine in this position makes the BlaC substrate-binding pocket wider and likely more flexible. Therefore, this substitution in BlaC offers an alternative approach to enlarge the active site and could allow *Mtb* BlaC to specifically bind



carbapenems or cepheems with large  $\alpha$  substituents. Furthermore, on the same side of the substrate-binding site as Gly<sup>132</sup>, the side chain of Ser<sup>104</sup> is unusual in that it is extended away from the substrate-binding site. In other class A  $\beta$ -lactamases, this position is typically an asparagine or an aspartate, and the side chain points into the active site. Like the Asn<sup>132</sup> to Gly<sup>132</sup> substitution, the conformation of the Ser<sup>104</sup> side chain creates more space in the active site and reduces the potential hydrogen bonding interactions with the substrate. On  $\beta$ -strand B3 of BlaC, Thr<sup>237</sup> is unusual for class A  $\beta$ -lactamases. This location is most often occupied by an alanine, as found in non- Extended Spectrum  $\beta$ -lactamase (non-ESBL) or a serine in the ESBLs. The hydroxyl group of Thr<sup>237</sup> points to the same direction as the comparable serine in the ESBLs. Ser<sup>237</sup> (equivalent to Thr<sup>237</sup> in *Mtb* BlaC) is believed to play a critical role in extended spectrum activity of K1 (*P. vulgaris*)  $\beta$ -lactamase. Superimposition of BlaC structure with substrate bound complex forms of Bs3 from *B. licheniformis* (PDBID: 1I2W) and mutant of the RTEM-1 from *E. coli* (PDBID: 1FQG) (172,182) indicates that Thr<sup>237</sup> potentially forms three hydrogen bonding interactions with the substrate: backbone oxygen to carbamide nitrogen of  $\alpha$  substituent of  $\beta$ -lactam ( $\sim 3.5$  Å), backbone nitrogen to ester carbonyl oxygen of  $\beta$ -lactam ( $\sim 2.8$  Å), and side chain oxygen to carboxyl oxygen of  $\beta$  substituent of  $\beta$ -lactam ( $\sim 3$  Å). The side chain of Arg<sup>220</sup> extends into the top part of the active site and it forms hydrogen bonds with the side chain oxygen of Thr<sup>237</sup>, the backbone oxygen of Gly<sup>236</sup> and the ester carboxyl oxygen of bound  $\beta$ -lactam. A similar hydrogen bond is only seen in the carbapenemase NCM-A (*E. cloacae*). The position of Arg<sup>220</sup> is usually occupied by a serine in ESBL type enzymes. The hydrogen bonding between Thr<sup>237</sup> and Arg<sup>220</sup> of *Mtb*

BlaC pulls the side chain of Thr<sup>237</sup> about 2.5 Å away from the active site. This conformation helps to contain the relatively large substituents, such as 6 $\alpha$ -1R-hydroxyl group of carbapenems, which could be critical for their binding (Figure 33). On the other side of the  $\beta$  sheets, Arg<sup>243</sup> is unusual for class A  $\beta$ -lactamases. This position is typically occupied by a threonine. The two nitrogens of the guanidinyll group of Arg<sup>243</sup> form hydrogen bonds with the backbone CO group of Asp<sup>172</sup> and Asn<sup>170</sup> at distance of 2.8 Å and 3.1 Å, respectively, and serve to link B4 with  $\Omega$  loop. Another observed difference of potential importance is that Arg<sup>171</sup> in *Mtb* BlaC is usually occupied by either a threonine or a serine in other class A  $\beta$ -lactamases. The long side chain of Arg<sup>171</sup> extends toward the side of the entrance to the active site. On the opposite site of the  $\Omega$  loop, the region Ile<sup>102</sup>-Ile<sup>105</sup> of BlaC is quite different from other class A  $\beta$ -lactamases as well. While all other class A  $\beta$ -lactamases have either tyrosine or histidine with their aromatic side chains covering the entrance, Ile<sup>105</sup> at the same position makes BlaC's active site  $\sim$ 3 Å wider than other class A  $\beta$ -lactamases. A four residue insertion, consisting of Gly<sup>145A</sup> to Thr<sup>145D</sup> at the beginning of  $\alpha$  helix H7 makes H7  $\sim$ 6 Å longer than that of other class A  $\beta$ -lactamases.

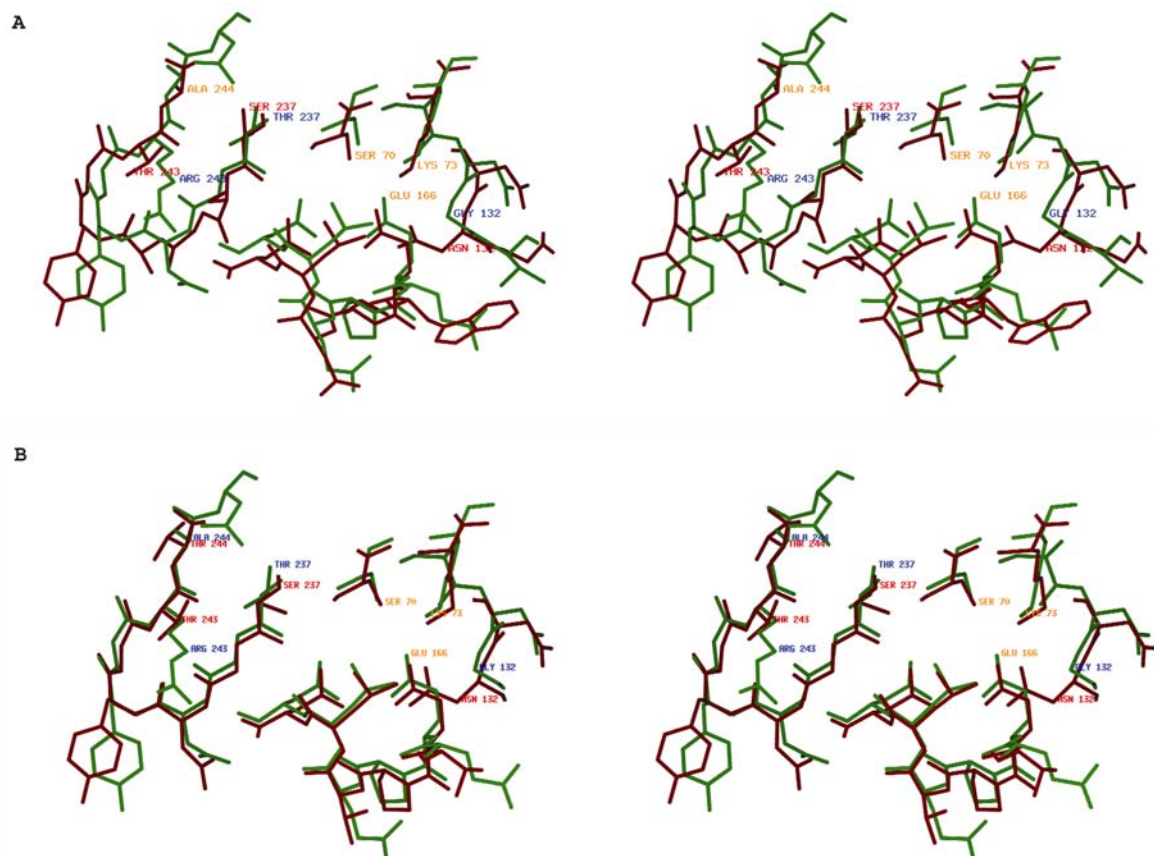


Figure 33 Comparison of Active Site of *Mtb* BlaC with Homologs in Stereo View

### Structural Basis for Broad Substrate Recognition

*Mtb* BlaC has a broad-spectrum substrate profile. The penicillinase activity of BlaC is about 10 fold lower than most of class A  $\beta$ -lactamases. However, the enzyme showed moderate activity against cephalothin and some activity against cefoxitin, ceftazidime and meropenem. The broad substrate profile is due to the enzyme's large and flexible substrate-binding site that allows it to accommodate different types of substrates. While a large active site can provide broader substrate specificity, it can also lead to weak enzyme-ligand interactions. This could explain the generally low activity of *Mtb* BlaC in hydrolyzing substrates and inhibitors. The substrate profile of *Mtb* BlaC is quite similar to the  $\beta$ -lactamase of *S. clavuligerus*, however the *Mtb* enzyme has better activity against all  $\beta$ -lactams we have tested.

We have tried to crystallize BlaC bound with substrates and inhibitors. However, no clear density of substrates or inhibitors could be identified after numerous trials. This is primarily due to the high turnover rate of hydrolysis by BlaC. Nevertheless, structure alignments suggest that the reason for the low penicillinase activity is in part due to the Asn to Gly substitution at position 132. As a conserved residue in other class A  $\beta$ -lactamase, the nitrogen atom of Asn<sup>132</sup> side chain forms a hydrogen bond with the carbonyl oxygen of the  $\alpha$  substituent of  $\beta$ -lactam substrates. Glycine at this position would result in a loss of hydrogen bonding to the substrate as shown by the superimposition of Toho-1-benzylpenicillin (178) and *Mtb* BlaC structures. The hydrogen bond is apparently important for stabilizing the substrate in the active site. Therefore, the loss of hydrogen bonding interactions could explain the low activity of

*Mtb* BlaC, especially to penicillin type substrates (Figure 34A and 34B). The critical role of Asn<sup>132</sup> to the activity of class A  $\beta$ -lactamase has been proven by the study of a N132A mutant of *S. albus* G  $\beta$ -lactamase (183), which showed a 100 to 200 fold decrease of penicillinase and cephalosporinase activities. Unlike the N132A mutant of *S. albus* G  $\beta$ -lactamase, *Mtb* BlaC has relatively good activity against cephalothin. The notable R164A substitution in the neck of  $\Omega$  loop is likely the reason why BlaC maintain good cephalothin activity. In other class A  $\beta$ -lactamases with known structures, this position is always occupied by Arg<sup>164</sup>. Normally, Arg<sup>164</sup> hydrogen-bonds to another highly conserved residue Asp<sup>179</sup> to form a salt bridge in the neck of  $\Omega$  loop, which holds the  $\Omega$  loop closed and rigid (Figure 35B). Loss of this hydrogen bond could potentially destabilize  $\Omega$  loop (Figure 35A), which has been shown to play an important in the expansion of activity against cephalosporines, based on a mutagenesis study (184). Indeed, compared to many other class A  $\beta$ -lactamases, part of the  $\Omega$  loop region Asn<sup>170</sup> – Pro<sup>173</sup> of BlaC is curved away from B4 by  $\sim 2$  Å.

Several other residue substitutions could contribute to the boost of cephalosporinase activity. Superimposition to structures of ESBLs in complex with cephalosporins, indicated that the hydroxyl group of Thr<sup>237</sup> (equivalent to Ala<sup>237</sup> or Ser<sup>237</sup> in other class A  $\beta$ -lactamase) of *Mtb* BlaC occupies the same position as the side chain of Ser<sup>237</sup> in the ESBLs and the hydroxyl group is within the hydrogen bonding distance (2.5 Å) to the carboxyl group of  $\beta$  substituents of cephalosporins. This could help to stabilize the acyl-intermediate of cephalosporins and increase the activity. This is consistent with the finding that A237T mutation increased the relative cephalosporinase

activity in TEM  $\beta$ -lactamase (non-ESBL) (185). Since *S. albus* G  $\beta$ -lactamase is a non-ESBL, at the equivalent position, it has alanine not serine or threonine. The loss of the hydrogen bonding interactions at both sites could be the reason why the N132A mutant totally lost the cephalosporinase activity.

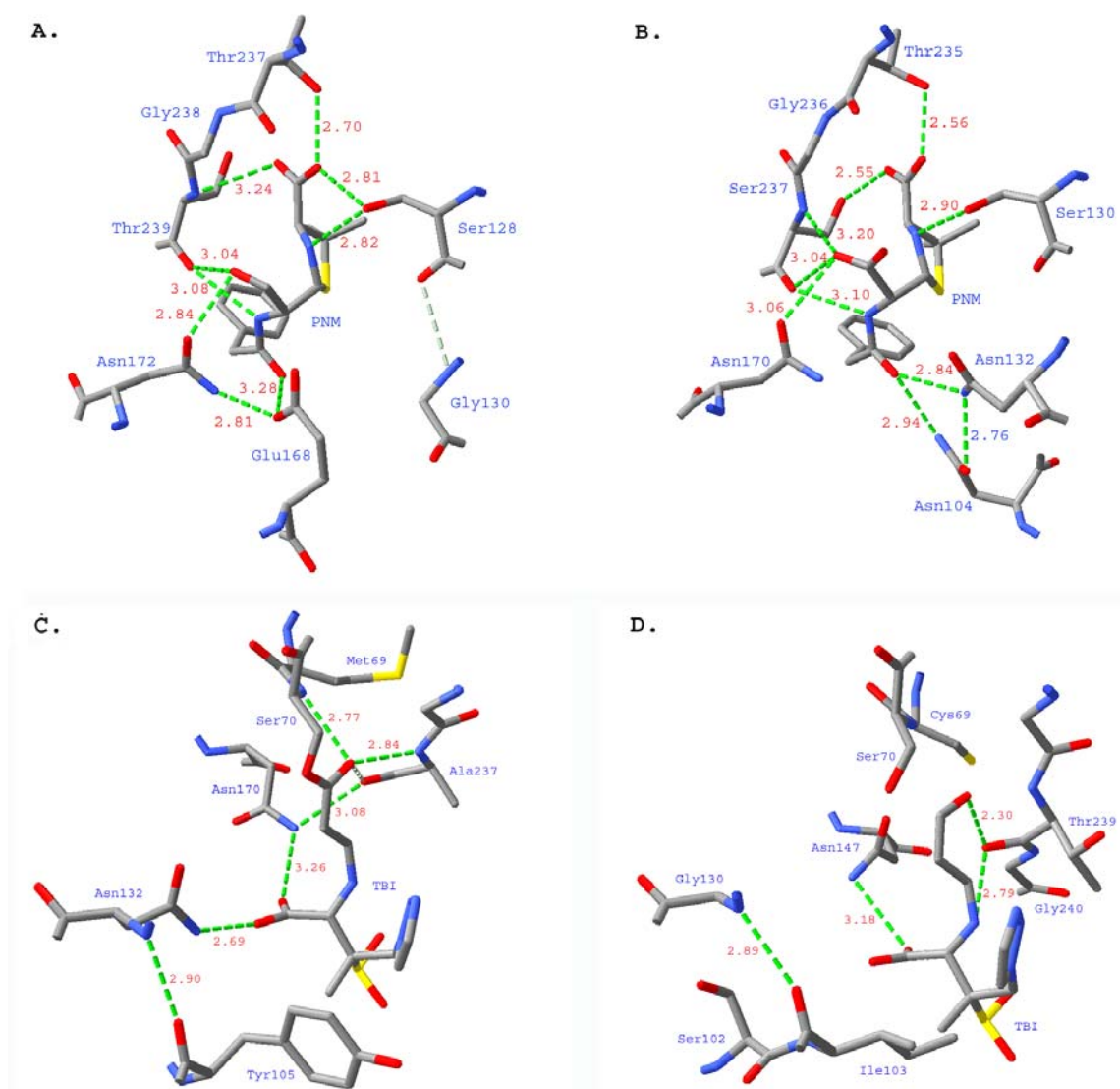
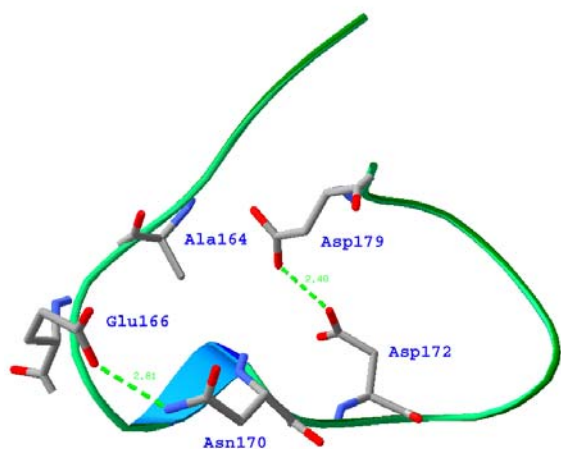
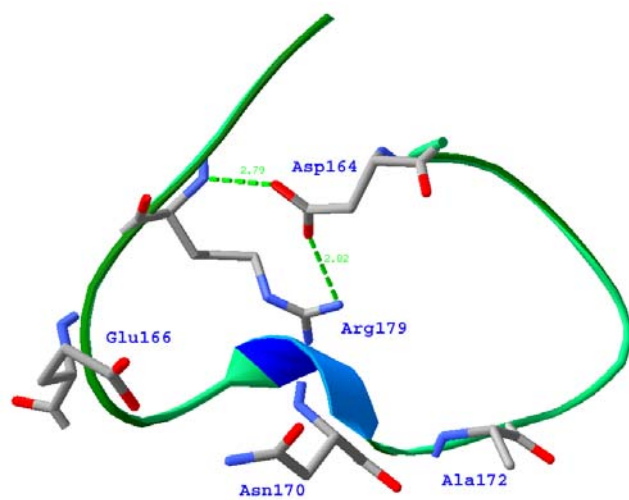


Figure 34 Comparison of Active Sites of BlaC Bound to Inhibitors

A.



B.

Figure 35 Comparison of  $\Omega$  loops of BlaC

### **Class A $\beta$ -lactamase Inhibitors**

Clavulanate is a suicide inhibitor specific for class A  $\beta$ -lactamases. This potent ( $K_i < 0.1 \mu\text{M}$ ) inhibitor forms a covalent acyl enzyme complex with class A  $\beta$ -lactamase. After acylation, the intermediate can either tautomerize to form a transient inhibitor to the enzyme or acylate another active site serine, Ser<sup>130</sup> to form a cross-linked, irreversibly inactivated enzyme (186). Our inhibition studies show that clavulanate is a relatively poor inhibitor of *Mtb* BlaC ( $K_i = 2.4 \mu\text{M}$ ) with a 24 fold higher  $K_i$  compared to TEM-1. Class A  $\beta$ -lactamases with amino acid substitutions at Arg<sup>244</sup>, Asn<sup>276</sup>, Arg<sup>275</sup>, Met<sup>69</sup>, Met<sup>182</sup> and Trp<sup>165</sup> have been reported to be resistant to clavulanate and other mechanism-based inhibitors (187). Among these residues, the substitutions of R244A, N278E, M69C and M182T are all found in BlaC. It is interesting to note that clavulanate is produced by *S. clavuligerus* whose  $\beta$ -lactamase is the closest homolog of *Mtb* BlaC. Not surprisingly, *S. clavuligerus*  $\beta$ -lactamase is resistant to clavulanate ( $K_i = 12 \mu\text{M}$ ) (188). The superimposition of a E166A variant of SHV-1 with tazobactam bound (189) and *Mtb* BlaC structures reveals that the N132G substitution would also cause the loss of a hydrogen bond between side chain of Asn<sup>132</sup> and 3- $\beta$ -carboxyl group of the inhibitor (Figure 34C and 34D). In the complex structure of wild type SHV-1 with tazobactam bound, Asn<sup>132</sup> was also observed hydrogen-bonding with the sulfinic acid group of the trans-enamine intermediate (190). Based on structure and activity relationship, it is quite clear that *Mtb* and *S. clavuligerus* share a very similar strategy in the resistance of general class A  $\beta$ -lactamase inhibitors. That is, to weaken the binding of enzyme to



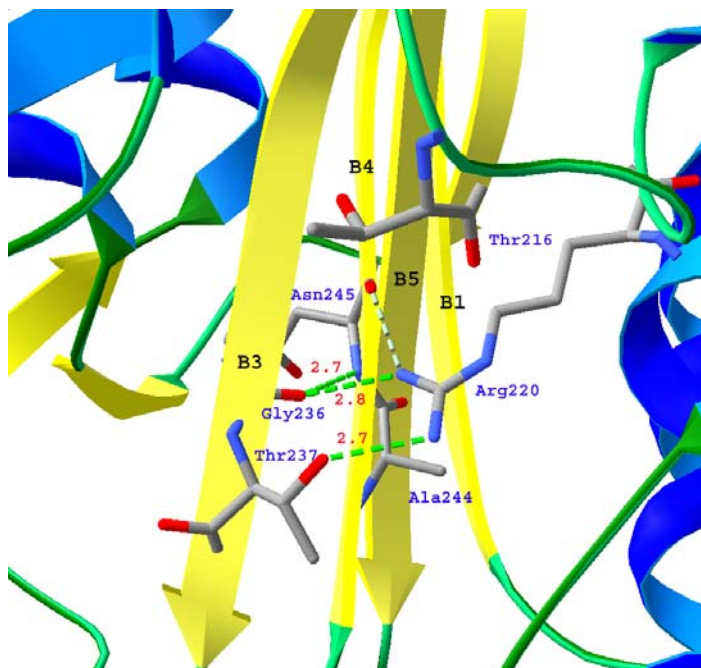
those inhibitors mainly through a single amino acid substitution N132G, even though this will also sacrifice enzyme activity to other  $\beta$ -lactam substrates.

BlaC has Ala<sup>244</sup> on  $\beta$  strand B4. This position is usually occupied by arginine in non-ESBL enzyme and threonine in ESBL enzyme. Interestingly, in *Mtb* BlaC, the side chain of Arg<sup>220</sup> is almost in the same position usually occupied by side chain of Arg<sup>244</sup> in other class A  $\beta$ -lactamases. This Arg<sup>220</sup> - Ala<sup>244</sup> residue pair only exists in *Mtb* BlaC, carbapenemase NMC-A (*E. cloacae*) and SCLA (*S. clavuligerus*). Mutation of Arg<sup>244</sup> to Serine in the TEM-1  $\beta$ -lactamase has been shown to produce resistance to inactivation by clavulanate in the mutant enzyme and resistance to ampicillin plus clavulanate in a strain of *E. coli* producing this enzyme (191). The modeling studies indicated that the guanidinyll group of Arg<sup>244</sup> and an ordered water molecule stabilize the carboxyl group of clavulanate and its acylated intermediate through hydrogen bonding (191). In the R244S mutant, these hydrogen bonds would be lost, which could potentially cause resistance to clavulanate. It is tempting to speculate that the R244A substitution in BlaC would have the similar mechanism of resistance to clavulanate. However, the position of the Arg<sup>244</sup> side chain in the TEM-1 is occupied by the guanidinyll group of Arg<sup>220</sup> in BlaC that does not exist in TEM-1 (Figure 36), which complicates the situation. Even though the water molecule in the TEM-1 structure that is believed to be important to clavulanate inhibition, is not observed in the BlaC structure and the guanidinyll group of Arg<sup>220</sup> in BlaC does not have the same orientation as that of Arg<sup>244</sup> in TEM-1, it is possible that the presence of the guanidinyll group of Arg<sup>220</sup> could to some degree restore sensitivity of BlaC to clavulanate. Indeed the resistance of BlaC to clavulanate ( $K_i = 2.4$

$\mu\text{M}$ ) is about 10 fold less than the R244S mutant of TEM-1 ( $K_i = 33 \mu\text{M}$ ) (192). However, the clavulanate inhibition is biphasic and  $K_i$  only reflects the first phase. Arg<sup>244</sup> in TEM-1 has been shown to be important to both the binding of clavulanate and the chemistry of inactivation (191). Arg<sup>220</sup> in BlaC may be able to help the binding of clavulanate by hydrogen bonding interaction in the first phase, but not likely to facilitate the chemistry of inactivation in the second phase.

Although clavulanate is not an optimal inhibitor to *Mtb* BlaC, it has been demonstrated that the combinations of  $\beta$ -lactam antibiotics and mechanism-based inhibitors such as clavulanate are effective in the treatment of *Mtb*, even MDR strains. This makes *Mtb* BlaC a promising target for the design of new inhibitors that are specific and potent to *Mtb* BlaC. Based on the structure of *Mtb* BlaC, it is now possible to either modify current inhibitors to increase inhibitory potency or design novel inhibitors that are specific to the enzyme. In combination to the highly potent *Mtb* BlaC inhibitors,  $\beta$ -lactam antibiotics could become a new regimen to treat tuberculosis.

A.



B.

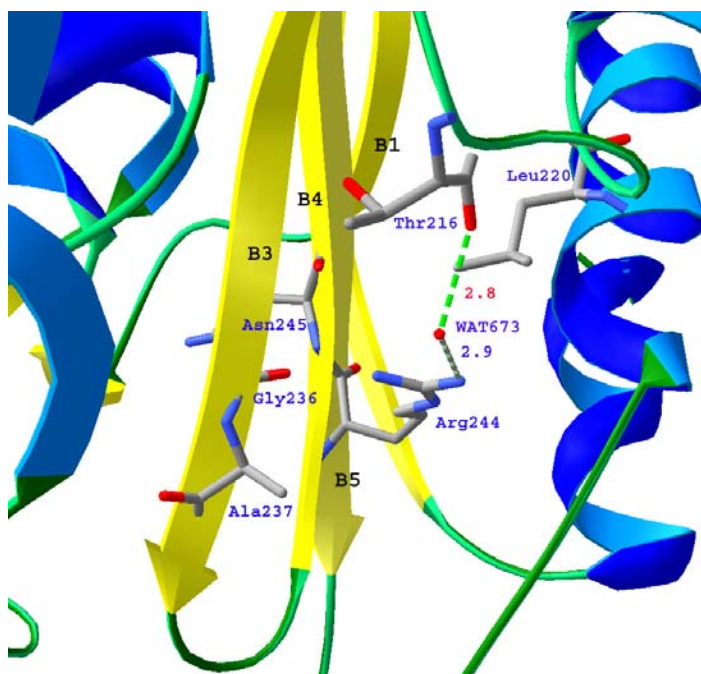


Figure 36 The Critical Role of Arg220 in Inhibitor Binding

## CHAPTER VI

### IDENTIFICATION OF A *MTB* VIRULENCE FACTOR

#### BACKGROUND

##### **A Putative Operon *Rv0096-Rv0101* Is Essential for the Virulence of *Mtb***

*Mtb Rv0098*, unique to actinomycetes, encodes a hypothetical protein that shares no homology with any protein with a defined function. *Rv0098* is found on an operon containing five other genes, *Rv0096*, a member of the PPE family, *Rv0097*, an oxidoreductase, *Rv0099*, a fatty acid AMP ligase, *Rv0100*, which encodes an acyl carrier protein, and *Rv0101*, a non-ribosomal peptide synthase. They were found to be strongly co-induced (193). While these genes are not essential for *Mtb* growth in media, *Rv0098* – *Rv0101* are required for its survival in macrophages (Figure 37) (41,42). Furthermore, the whole operon is conserved in other pathogenic mycobacteria. It was demonstrated that in *M. bovis*, this operon was involved in the biosynthesis of PDIMs, the virulence-enhancing lipids produced by *Mtb* and closely related species. The disruption of gene *Mb0100* (equivalent to *Mtb Rv0097*) caused a loss of virulence and an alteration of colony morphology. Besides, the mutations of *Mb0100* had a polar effect on the expression of the downstream genes *Mb0101* (equivalent to *Mtb Rv0098*) – *Mb0104* (equivalent to *Mtb Rv0101*), which clearly indicated that these genes are in the same operon and functionally related (194). Recently, Joshi, et al, found that the virulence of *Rv0098* is dependent on the functional interaction with the *mce* loci that was proposed to encode an ABC-like transport system (195).

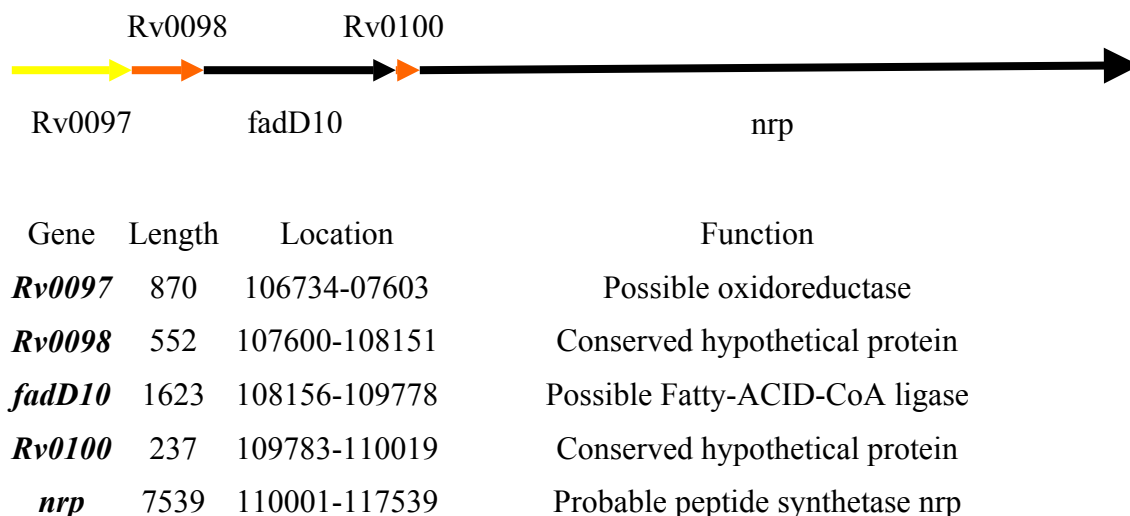


Figure 37 The Gene Cluster of *Rv0097* – *Rv0101*

### Function and Classification of CoA Thioesterases

Based on the protein fold and the structure with bound dodecenoate, we hypothesized that R0098 is a long chain acyl-CoA thioesterase. Acyl-CoA thioesterases, that hydrolyze fatty-acyl-CoA to fatty acid and coenzyme A, are found in archaean, prokaryotic, and eukaryotic organisms. Although the thioesterase activity was discovered fifty years ago, the precise physiological function of many of these enzymes is still unclear (196). Thioesterases are involved in polyketide, fatty acid, and non-ribosomal peptide biosynthesis (197-199). Some thioesterases also play an essential role in signal transduction by removing lipids from post-translationally modified proteins (200). Two families of thioesterases have been identified, thioesterase I and thioesterase

II, based on their distinct protein folds, amino acid sequence, and different catalytic mechanisms (Figure 38). Enzyme in the thioesterase I family usually exists as a fatty acid product-releasing domain of multi-domain proteins, such as polyketide synthase, FASI, and non-ribosomal peptide synthase. Thioesterase I enzymes all have an  $\alpha/\beta$  fold and use serine or cysteine as the catalytic residue, which upon activation, initiates a nucleophilic attack on the carbonyl carbon of the thioester substrate (28). Thioesterase II enzymes are discrete proteins that have been found to catalyze the hydrolysis of short or medium chain acyl-CoAs (63,196,201,202). Thioesterase II enzymes exist either as a homodimer in a single “hot dog” fold or as a monomer in a double “hot dog” fold. The active sites of thioesterase II enzymes invariably contain either an aspartate or glutamate that acts as a general base for the reaction (203). While the precise catalytic mechanism remains controversial, structural and functional studies reported for several of these enzymes have all shown that a general base is required for catalysis (196,202,204). In *Mtb*, FcoT is the only thioesterase identified thus far that is not part of a large multi-domain enzyme.

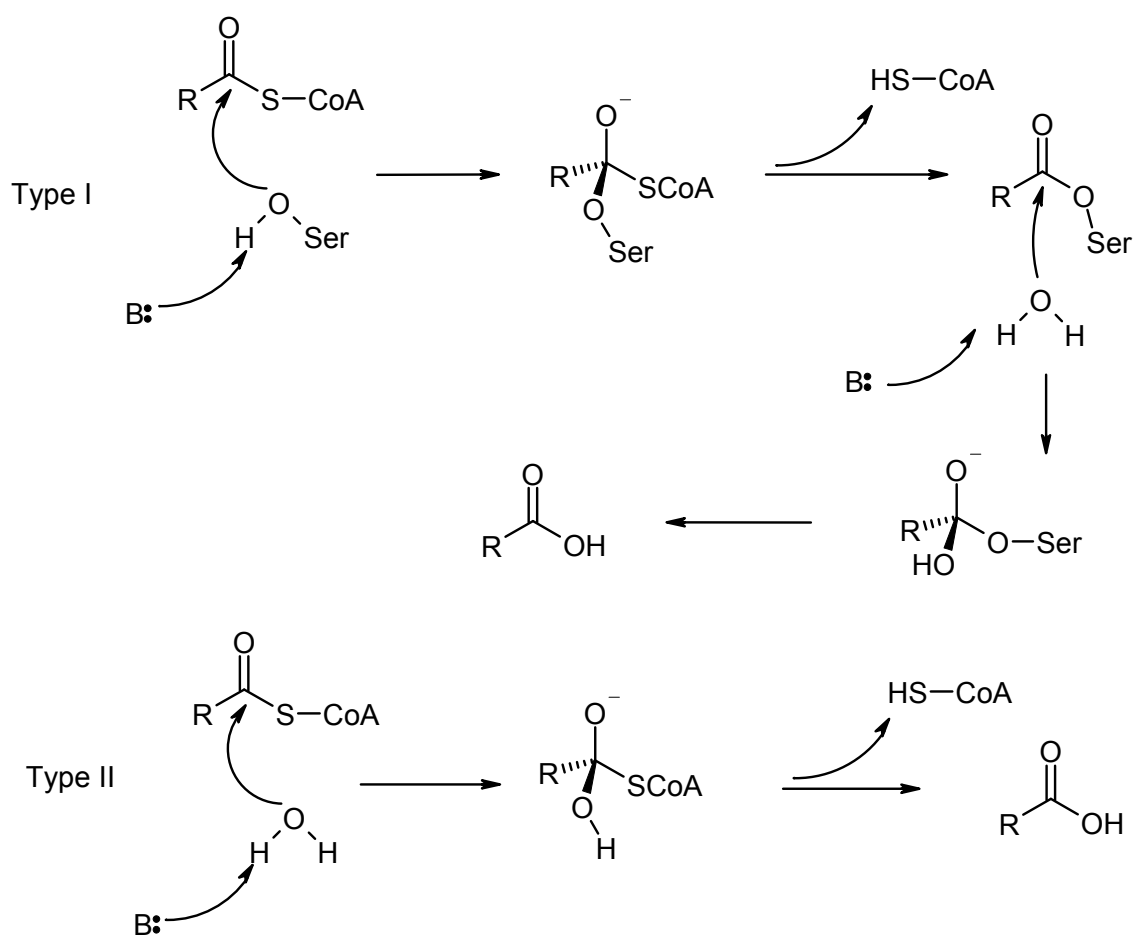


Figure 38 Catalytic Mechanisms of Thioesterases

## METHODS

### Cloning, Expression, and Purification

*E. coli* expression plasmids of *Mtb fcoT* were cloned from genomic DNA (C.S.U. N01-AI-75320). The amplified product was inserted into pET28b (EMD Biosciences #69865-3) using the NdeI and Hind-III restriction sites. This plasmid was subsequently transformed into *E. coli* BL21 (DE3) (EMD Bioscience #69387-3) and cultured in LB-Miller media containing 50 µg/ml kanamycin at 37 °C until OD<sub>600</sub> reached 0.8. Expression of the *fcoT* gene was induced for 20 hours at 16 °C by addition of 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). Cells harvested by centrifugation were re-suspended in 25 mM Tris-HCl (pH 8.0), 500 mM NaCl, and 2 mM β-mercaptoethanol, and lysed by a French press. After treatment with 1mM DNaseI, the insoluble material was removed by centrifugation. By SDS PAGE, ~ 15% of FcoT remained soluble. The supernatant, which contains soluble protein, was applied to an Ni<sup>2+</sup>-loaded HiTrap chelating column (AP Biotech) using a fast protein liquid chromatography system, and was eluted through an imidazole gradient (0-0.5M). FcoT was eluted out at an imidazole concentration of 180 mM. Fractions that contain the FcoT were pooled and dialyzed against 25 mM Tris-HCl (pH 8.0), 60 mM NaCl, and 1 mM DTT. Using the dialysis buffer, the samples were further purified by gel-filtration on a Superdex 200 column to separate the monomeric protein from aggregated material. The protein was concentrated to approximately 10 mg/ml. FcoT was obtained at 6 mg/L of *E. coli* culture, and appeared homogeneous by SDS-PAGE and Coomassie Blue staining.



In order to perform a multiwavelength anomalous diffraction (MAD) experiment, Se-Met substituted protein was overexpressed in the B834 (DE3) *E. coli* strain and purified using the same method as native FcoT.

### **Mutagenesis, Expression, and Purification**

FcoT mutants N83A, Y66F, Y33F, N74A, Y87F, E77Q, and H72A were constructed using the QuickChange site-directed mutagenesis kit (Stratagene) and confirmed by DNA sequencing. All mutant proteins were expressed and purified using the same protocol as the wild-type FcoT.

### **Acyl-CoA Thioesterase Activity Assay**

Acyl-CoA thioesterase activity assays were completed for the substrates hexanoyl-CoA (C6:0), lauroyl-CoA (C12:0), palmitoyl-CoA (C16:0), stearoyl-CoA (C18:0), dodecenoyl-CoA (C12:1 *trans*), crotonoyl-CoA (C4:1 *trans*), and benzyl-CoA. All assays were repeated in triplicate and carried out on a Cary 100 Bio Spectrophotometer at room temperature by monitoring the UV absorbance increase at 412 nm using 5, 5' dithiobis (2-nitrobenzoic acid) (DTNB), as previously described (205). The *Mtb* FcoT activity assays were performed in 100 mM phosphate buffer (pH 7.5) by adding 100 nM enzyme into solutions of acyl-CoA substrate and DTNB (100  $\mu$ M).  $K_m$  and  $V_{max}$  were determined using software KaleidaGraph (Synergy Software).

## Crystallization of FcoT

Crystallization of the native and Se-Met substituted FcoT was accomplished by the hanging-drop vapor diffusion method (85). Initial crystallization conditions were discovered by sparse matrix screen using Crystal Screens from Hampton Research. Hanging drops containing 2  $\mu$ l of protein solution at 10 mg/ml and 2  $\mu$ l of buffer (0.1 M Hepes pH 7.5, 14-16% isopropanol, 200-300mM NaCl, 10mM MgCl<sub>2</sub>) were equilibrated at 16 °C in Linbro plates against 1 ml of the same buffer in the reservoir. Protein crystals appeared in two days and were improved by co-crystallization with dodecenoyl CoA.

## Data Collection and Processing

Diffraction data from a single crystal of protein with dodecenoyl CoA was collected at 121 K using a cryo-protection solution consisting of a reservoir solution (0.1 M Hepes pH 7.5, 14-16% isopropanol, 200-300mM NaCl, 10mM MgCl<sub>2</sub>) with the addition of 30% 2-methyl-2, 4-pentanediol (MPD). Crystals diffracted to 2.3 Å at beam line 23ID at the Advanced Photon Source (APS), Argonne National Laboratory. The selenium absorption edge was determined to be at 12.6618 KeV by fluorescence scanning of a Se-Met FcoT crystal. Diffraction data of three wavelengths (peak: 0.9792 Å; inflection: 0.9794 Å; high-energy remote: 0.9611 Å) were collected from a single crystal with one-degree oscillation widths through a range of 360° for each wavelength. The data were integrated and reduced using HKL2000. Crystals of FcoT belong to the space group I2<sub>1</sub>3 with cell dimensions a = 100.18 Å, b = 100.18 Å, c = 100.18 Å,  $\alpha$  = 90°,  $\beta$  = 90°, and  $\gamma$  = 90°, with one molecule in each asymmetric unit.

### Structure Determination and Model Refinement

MAD phasing was achieved by running the program Sharp (<http://www.globalphasing.com>). Three wavelengths (peak, inflection, high energy remote) of data were included into the calculation within the resolution range 2.3 to 20 Å. One selenium site in one asymmetric unit was found. The overall figure of merit is 0.25 (Table 24). The electron density map was further improved by Sharp's solvent flipping and flattening program. A clearly interpretable electron density map was calculated using FFT in CCP4 suite (160). The initial model was built by an automatic modeling building program, TEXAL (92,206). Once again, in the CCP4 suite the structure was refined with REFMAC rigid body refinement program (160), and manual rebuilding was performed using XtalView (89). The model was further refined with cycles of model building and REFMAC restrained refinement to an  $R_{\text{factor}}$  of 23.9% and an  $R_{\text{free}}$  of 29.1%. During the final cycles of the refinement, water molecules were added into peaks above 3- $\sigma$  of the  $F_0-F_c$  electron density maps that were within hydrogen-bonding distance from appropriate protein atoms.

Table 24 Data Collection, Processing, and Refinement Statistics of FcoT

<b>SeMet FcoT</b>			
<b>Data collection</b>			
Space group	I2 <sub>3</sub>		
a (Å)	100.18		
	<i>Peak</i>	<i>Inflection</i>	<i>Remote</i>
Wavelength	0.9795 Å	0.9796 Å	0.9537 Å
Resolution (Å)	50-2.19 (2.29-2.19)	50-2.29 (2.41-2.29)	50-2.35 (2.49-2.35)
Rsym (%)	8.1 (85.9)	7.7 (65.5)	7.5 (53.6)
Completeness (%)	98.4 (96.8)	99.9 (98.6)	99.9 (100)
Redundancy	12.2 (4.8)	9.2 (5.2)	10.1 (6.7)
I / $\sigma$ I	35.6 (1.7)	30.3 (2.6)	28.2 (4.2)
FOM	0.40		
<b>Refinement statistics</b>			
Resolution range (Å)	20.0 - 2.3		
Number of reflections	7549		
Number of atoms / subunit			
Protein	1272		
Ligand	14		
Solvent	24		
R <sub>work</sub> / R <sub>free</sub>	23.3 / 27.2		
B-factors (Å <sup>2</sup> )			
Protein	54.9		
Ligand	42.6		
Solvent	60.8		
rmsd from ideal values			
bonds (Å)	0.02		
angles (°)	1.95		

The numbers given in parentheses denote the respective values of the highest resolution shell.

FOM. (Figure of merit) is the mean value of the cosine of the error in phase angles:  $\langle \cos(\alpha - \alpha_{\text{best}}) \rangle$ .

## Structure Analysis

DALI (162) was used to search the PDB for proteins having folds similar to *Mtb* FcoT. SwissPDB viewer was used to make structural alignments (163). The model was evaluated and analyzed using SPOCK (90).

## RESULTS AND DISCUSSION

### Sequence Analysis of FcoT

In *Mtb*, the *Rv0098* gene encodes a 183 amino acid protein of unknown function. A protein similarity search of all sequences, using BLAST, did not provide any homologs (>20% identity) with defined function. In fact, only two homologs could be clearly identified, ML1993 from *Mycobacterium leprae* (68% sequence identity) and SAV606 from *Streptomyces avermitilis* (36% sequence identity), and both were annotated as hypothetical proteins. *Rv0098* was previously predicted to encode a  $\beta$ -hydroxyl acyl CoA dehydratase (FabZ) based on sequence analysis (46). However, the sequence identity between *Rv0098* and any FabZ or FabA proteins was low (<20%).

### Purification and Characterization of FcoT

*Mtb* FcoT was expressed in *E. coli* BL21(DE3) cells as a fusion protein with an N-terminal His-tag. The protein was purified on a Ni-NTA affinity column to homogeneity. Purified recombinant protein, encoded by *Rv0098*, has a monomeric molecular weight of 20 kDa. The molecular weight of the native protein was determined to be approximately 110 kDa by gel filtration chromatography, suggesting that it is a

hexamer in solution (Figure 39). It is interesting to note that some FabZ dehydratases were previously reported to exist as a trimer of dimers of identical subunits in solution (207), however most type II thioesterases are dimers.

### **FcoT Is Not a Fatty-acyl CoA Dehydratase**

The purified recombinant enzyme was tested for dehydratase and hydratase activity using crotonoyl,  $\beta$ -hydroxyl butyl, and *trans*-2-dodecenoyl CoAs. The conversion of crotonoyl-CoA to  $\beta$ -hydroxybutyryl-CoA or vice versa, was monitored by monitoring the increase or decrease in absorbance at 260 nm. This protocol has been used to characterize the activity of PfFabZ (208). However, *Mtb* FcoT did not show any detectable activity using any of these substrates.

### **FcoT Is a Fatty-acyl CoA Thioesterase**

Only after the structure of the protein bound with dodecenoate was solved, as described below, was it possible for us to predict that the enzyme was a thioesterase. The thioesterase activity of the enzyme was confirmed by *in vitro* assays using a series of

different chain length (C4-C18) fatty acyl CoA substrates (Figure 40). Among them, hexanoyl-CoA (C6:0), lauroyl-CoA (C12:0), palmitoyl-CoA (C16:0), and stearoyl-CoA (C18:0) contain a saturated fatty acyl chain; dodecenoyl-CoA (C12:1 *trans*) and crotonoyl-CoA (C4:1 *trans*) have unsaturated fatty acyl chains; and benzyl-CoA has only an aryl group. The enzyme actively cleaved the thioester linkage of all of the substrates tested, but showed a marked preference for those thioester substrates with long chain fatty acyl groups. Of the substrates tested, maximal  $k_{\text{cat}}$  ( $0.037 \text{ s}^{-1}$ ) was observed for palmitoyl-CoA (C16:0) and the enzyme activity decreased for substrates with either longer or shorter acyl chains. At the same chain length, saturated substrates were marginally favored by the enzyme compared to unsaturated substrates, although the former are much more soluble for a given hydrocarbon chain length (Figure 41). Relative to the fatty acyl CoAs there was low hydrolysis activity ( $k_{\text{cat}} = 0.004 \text{ s}^{-1}$ ) and a large  $K_m$  value (272  $\mu\text{M}$ ) for benzyl-CoA as a substrate. Its  $k_{\text{cat}}/K_m$  value is  $\sim 0.3\%$  of that of palmitoyl-CoA, which suggests that the enzyme does not prefer aryl thioesters. In general, the activity of FcoT is  $10^2 - 10^4$  fold less than other thioesterases.

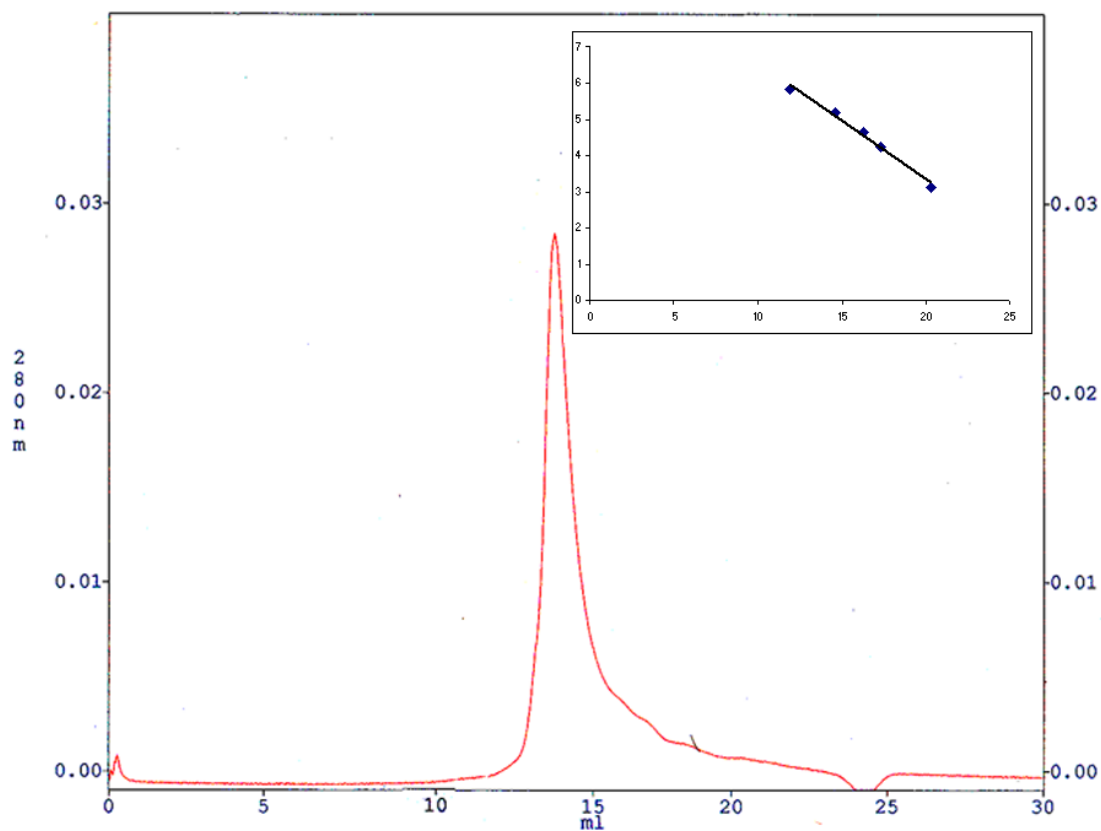


Figure 39 The Superose-6 Size-exclusion Chromatography of FcoT.

The elution profile of FcoT and calibration of the Superose-6 size-exclusion Column. Thyroglobulin, bovine gamma-globulin, chicken ovalbumin, equine myoglobin and vitamin 12 (from left to right) were used to calibrate the size-exclusion column. The logarithm of the molecular weight (MW) and the elution volume were fit with linear regression. The resulting equation was used to estimate the molecular weight of FcoT from its elution volume



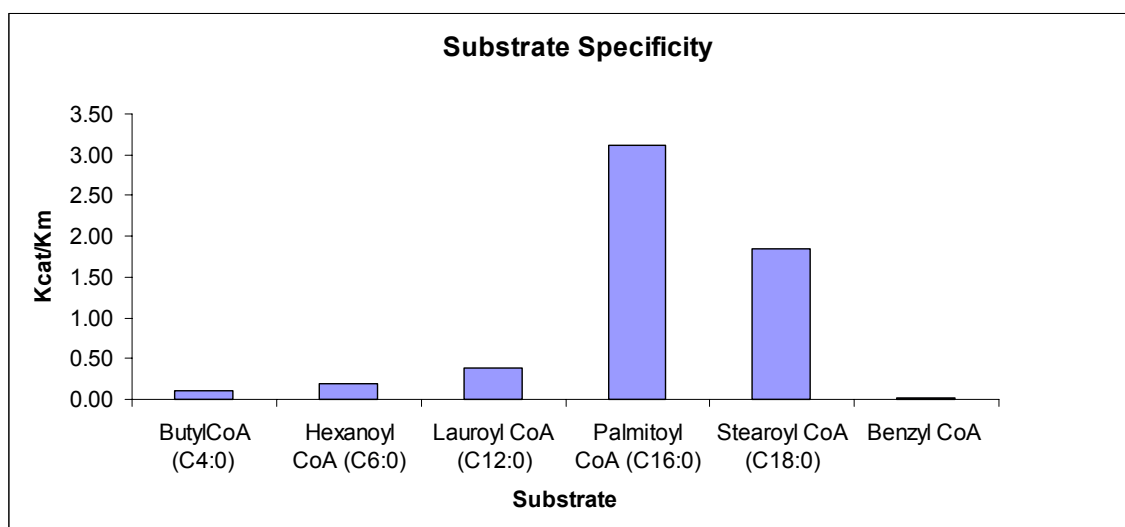


Figure 40 The Substrate Profile of *Mtb* FcoT

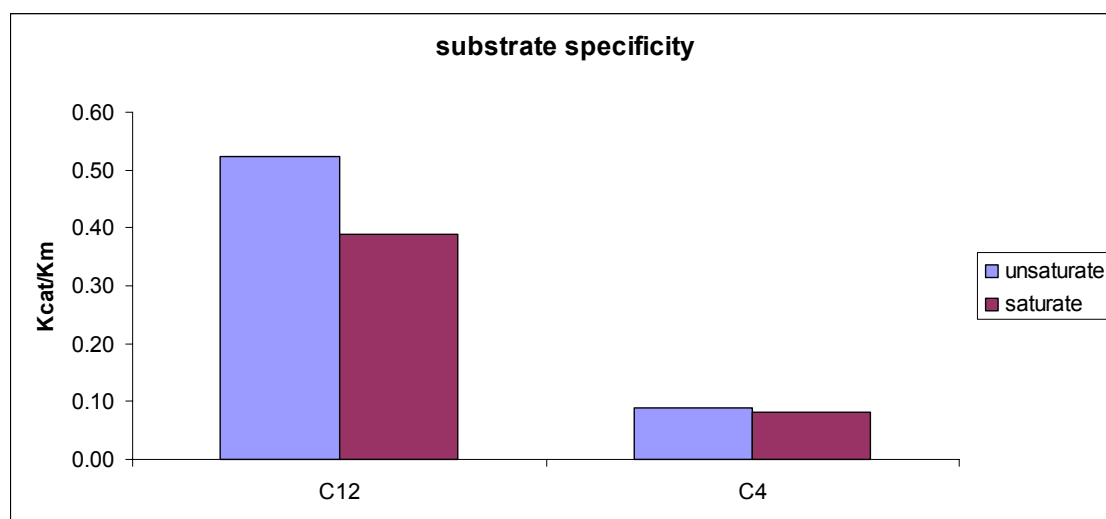


Figure 41 Substrate Specificity of *Mtb* FcoT

The rate of hydrolysis of the acyl CoA substrate by FcoT is pH dependent. In fact below pH 7, there was no detectable activity; while in the range of 7 – 8.5, the activity showed a linear phase and reached a plateau at pH 8.5 (Figure 42). These results suggest that the catalytic rate depends on the concentration of hydroxide anions. Although the pH dependence for a thioesterase reaction has not been reported before, several esterolytic antibodies showed a similar pH profile to *Mtb* FcoT (209).

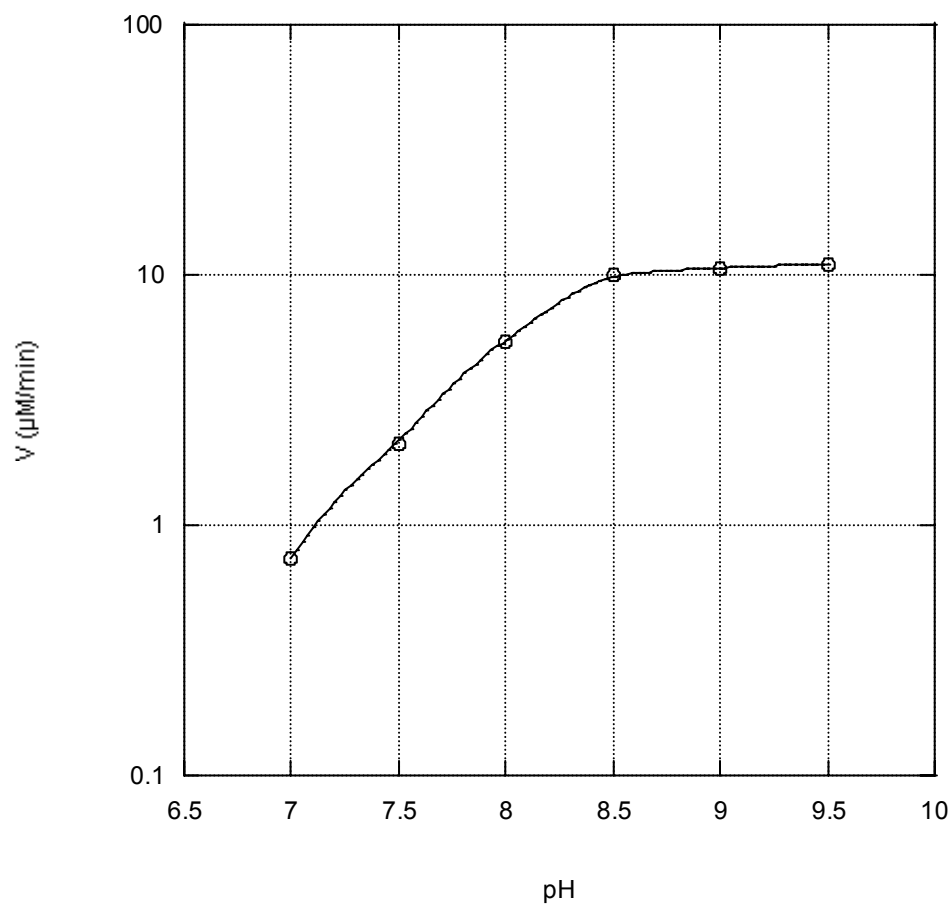


Figure 42 The Activity of FcoT for Substrate Palmitoyl CoA in the pH Range 7-9.5

## Overall Structure of FcoT

The crystal structure of FcoT was solved using MAD methods and refined to 2.3 Å resolution. The final  $R_{\text{factor}}$  and  $R_{\text{free}}$  values for all reflections were 0.23 and 0.27, respectively (Table 24). FcoT crystals, produced in the presence of the substrate dodecenoyl CoA, are cubic belonging to space group  $I2_13$ , with one molecule in the asymmetric unit. The refined model of FcoT includes 160 of the 183 total residues of the wild-type FcoT, the product dodecenoic acid, and 24 ordered water molecules. The structures of the N-terminal six-histidine tag, the first 19 residues Met<sup>1</sup>-Val<sup>19</sup>, and the four residues Glu<sup>149</sup>, Arg<sup>150</sup>, Thr<sup>151</sup>, and Glu<sup>166</sup> were not defined due to disorder of these regions in electron density maps. A “trimer of dimers” quaternary structure was identified after applying crystallographic symmetry, which is consistent with the gel filtration chromatography result that showed that the protein is a hexamer in solution (Figure 43A). The dimer is composed of two identical protomers and has overall dimensions of 32 Å x 44 Å x 67 Å (Figure 43B). Three identical dimers join together “head to tail” to form an equilateral triangle (approximately 105 Å each side) with its 3-fold rotation axis coincident to a crystallographic triad. This quaternary arrangement has not been observed in other type II thioesterases, which are usually packed as dimers or tetramers; however, a similar quaternary structure was found in FabZ with a ring-shaped packing (207).

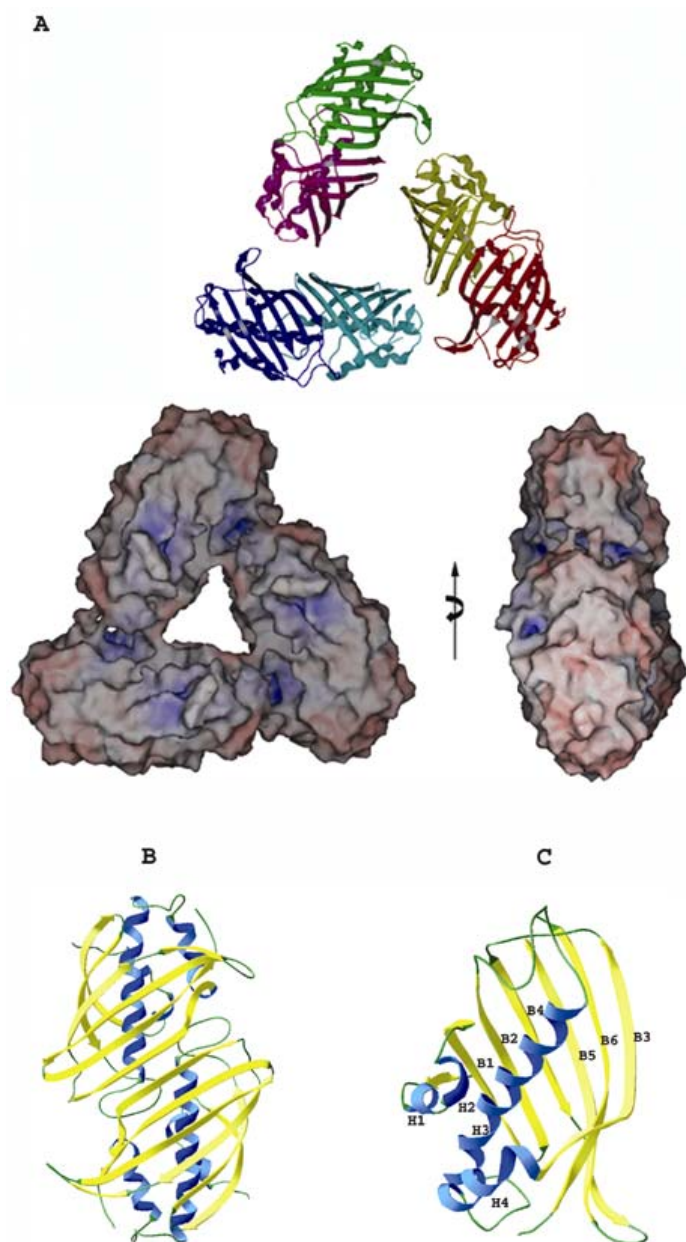


Figure 43 The Overall Architecture of FcoT.

(A) Ribbon and surface diagrams showing a “trimer of dimers” quaternary structure of *Mtb* FcoT. (B) Ribbon diagrams of a dimer showing the “hot dog” fold. (C) Subunit of FcoT. The  $\alpha$  helices are represented as H1-H4 and the strands as B1-B6

The *Mtb* FcoT protomer has an  $\alpha/\beta$  “hot dog” fold, in which a six-antiparallel  $\beta$ -sheet (B1-B6), referred to in the literature as the “bun”, wraps around a six-turn  $\alpha$  helix (H3) typically called the “sausage”, with both ends sealed, and three short  $\alpha$  helices (H1, H2 and H4) that cover the open side (Figure 43C). This “hot dog” topology is common for dehydratases and thioesterases (210).

The interface of the dimer is quite extensive. The solvent-accessible buried surface area is about 3012 Å<sup>2</sup> (18% of the surface area of one subunit), suggesting a strong interaction between the two protomers. Three major groups of interactions contribute to the dimerization of two subunits: the hydrogen bonding interaction between the anti-parallel  $\beta$  strand B3 (Arg<sup>121</sup> – Phe<sup>127</sup>) from each subunit, extensive hydrophobic interactions between side chains of residues on the two central  $\alpha$  helices (H3) (Val<sup>76</sup> – Ile<sup>79</sup>), and the hydrophobic and hydrophilic interactions between the loop region (Ala<sup>65</sup> – Phe<sup>73</sup>) of one subunit and two  $\alpha$  helices (H2, H4) from the other subunit. Specifically, most of the residues of B3 are hydrophilic, with their side chains pointing out towards the solvent, and the backbone atoms forming hydrogen bonds with their counterparts from the other subunit. This anti-parallel  $\beta$  strand interaction is conserved in all known “hot-dog” fold proteins. In the center of the dimer interface, the ends of the two “sausage”  $\alpha$  helices bundle with each other in an anti-parallel manner. A notable salt bridge, formed by side chains of Arg<sup>51</sup> and Glu<sup>14</sup>, links the loop region from one subunit to the short  $\alpha$  helix H2 from the other.

An unusual intramolecular disulfide bridge is observed between Cys<sup>142</sup> and Cys<sup>159</sup> that links B4 and B5 (Figure 44). In *Mtb* FcoT, the Cys<sup>142</sup> and Cys<sup>159</sup> disulfide

bridge is buried between the  $\beta$  sheet and the central  $\alpha$  helix, which is a probable reason why this disulfide bond remains oxidized despite the presence of the reducing agent, DTT, in the buffer.

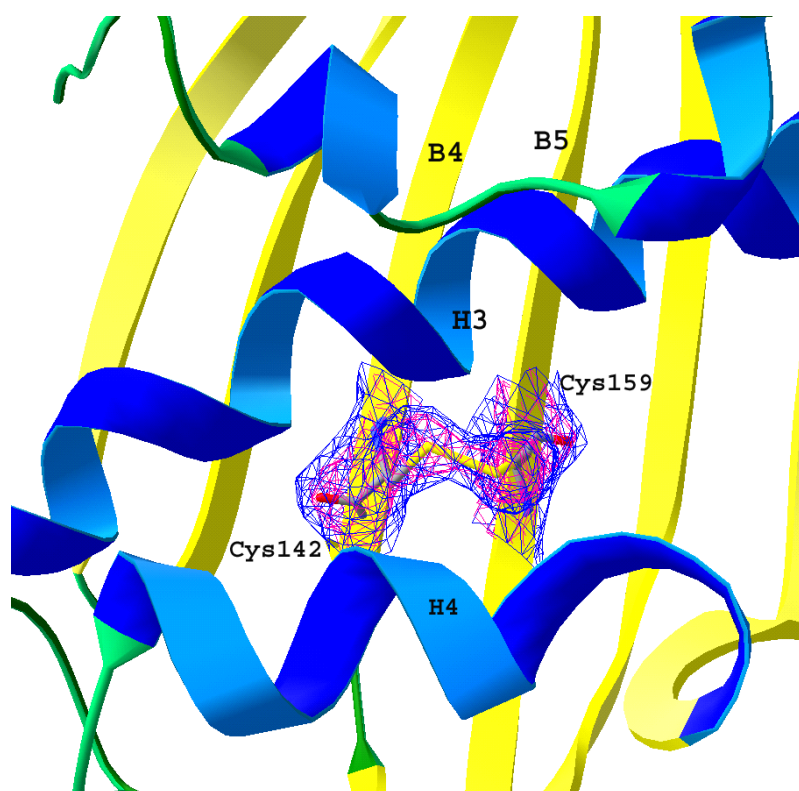


Figure 44 The Electron Density Indicating the Disulfide Bond in *Mtb* FcoT

A structure similarity search of the protomer found several proteins with modestly similar folds, most of which were dehydratases (FabZ or FabA) and thioesterases, however none of them had an RMSD less than 2.5 Å, suggesting that they were significantly different. Unlike other hotdog fold proteins, the two protomers of *Mtb* FcoT interact in such a manner that two “sausage”  $\alpha$  helices are anti-parallel to each other and on the same plane, that is perpendicular to the 2-fold rotation axis of the dimer. This unique arrangement makes the dimer of *Mtb* FcoT longer and straighter in one dimension than other proteins with a “hot dog” fold.

### **Structural Basis of Substrate Recognition**

FcoT has a long, narrow, hydrophobic substrate binding site. This “L” shaped substrate-binding tunnel is composed of two branches. The longer branch, buried inside one subunit, is formed by hydrophobic residues from two bundled  $\alpha$  helices, H3 and H4, and the  $\beta$ -sheet (B1-B6). It is approximately 24 Å deep and has a cross-sectional diameter of about 5 Å. This part of the tunnel is generally straight, extending along the “sausage”  $\alpha$  helix, H3. The bottom of the tunnel is at one end of the subunit, sealed by residues Ile<sup>99</sup> – Ile<sup>107</sup> from the loop region between H3 and H4. The other end opens to the dimer interface. The shorter branch of the tunnel is about 9 Å long and is formed by residues on the dimer interface from both subunits. It has a narrow opening (about 5 Å) to the solvent and is surrounded by residues Leu<sup>119</sup> and Pro<sup>130</sup>. Although one of the products, coenzyme A, is absent in the structure, the superimposition of the structure of FcoT and the 4-hydroxybenzoyl-CoA thioesterase with bound products clearly indicates



that FcoT can bind co-enzyme A in a similar manner, when CoA was modeled into FcoT. The pantetheine moiety of CoA fits inside the short branch of the substrate-binding tunnel of FcoT surrounded by residues His<sup>72</sup>, Arg<sup>121</sup>, Asn<sup>74</sup>, Met<sup>118</sup>, and Tyr<sup>87</sup>. The adenine ring and the phosphate group are outside of the tunnel and within hydrogen bonding distance to the residues on the surface of the protein, such as Ser<sup>116</sup>, Ser<sup>117</sup>, and Trp<sup>152</sup>.

The product of FcoT reaction, dodecenoic acid, is located inside the longer branch of the substrate tunnel (Figure 45) along with four ordered water molecules. The acyl chain portion of the product extends along the central  $\alpha$  helix, H3, into the tunnel, stabilized by van der Waals interactions with the hydrophobic residues Ala<sup>86</sup>, Tyr<sup>87</sup>, Phe<sup>90</sup>, Leu<sup>102</sup>, Tyr<sup>110</sup>, Met<sup>118</sup>, and Ile<sup>120</sup>. Its carboxyl group is located in a hydrophilic cavity created by the residues Asn<sup>83</sup> (subunit A), Asn<sup>74</sup>, and His<sup>72</sup> (subunit B) on the dimer interface (Figure 46). There is a 6 Å gap between the end of the acyl chain and the bottom of the tunnel, suggesting that acyl substrates longer than C12 could easily be accommodated. This observation correlates well with our results that FcoT has a preference for long chain fatty-acyl-CoA substrates. Although the tunnel appears to be of

optimal length for C16-18 to bind, it is possible that more space can be created to accommodate even longer substrates through conformation rearrangements of flexible residues on the bottom of the tunnel.

The narrow substrate-binding tunnel suggests that FcoT would not efficiently bind any substrates with bulky aryl groups. The superimposition of structures of *Mtb* FcoT and the *Arthrobacter* 4-hydroxybenzoyl-CoA thioesterase with bound products indicates that while the overall shape of the CoA binding site is well conserved, the binding pocket for the product, 4-hydroxybenoate, is blocked by the side chain of Tyr<sup>70</sup> in *Mtb* FcoT (Figure 45). Indeed, only negligible thioesterase activity (0.3% of palmitoyl-CoA) could be detected when benzyl-CoA was used as a substrate.

FcoT does not seem to have much selectivity between saturated and unsaturated fatty acyl substrates. With the same length of the carbon chain, the enzyme favors dodecenoyl CoA ( $k_{cat}/K_m = 0.52 \mu\text{M}^{-1}\text{s}^{-1}$ ) slightly more than lauroyl CoA ( $k_{cat}/K_m = 0.39 \mu\text{M}^{-1}\text{s}^{-1}$ ). This is probably because the hydrophobic tunnel is able to bind both saturated and unsaturated fatty acyl chains in a similar manner.

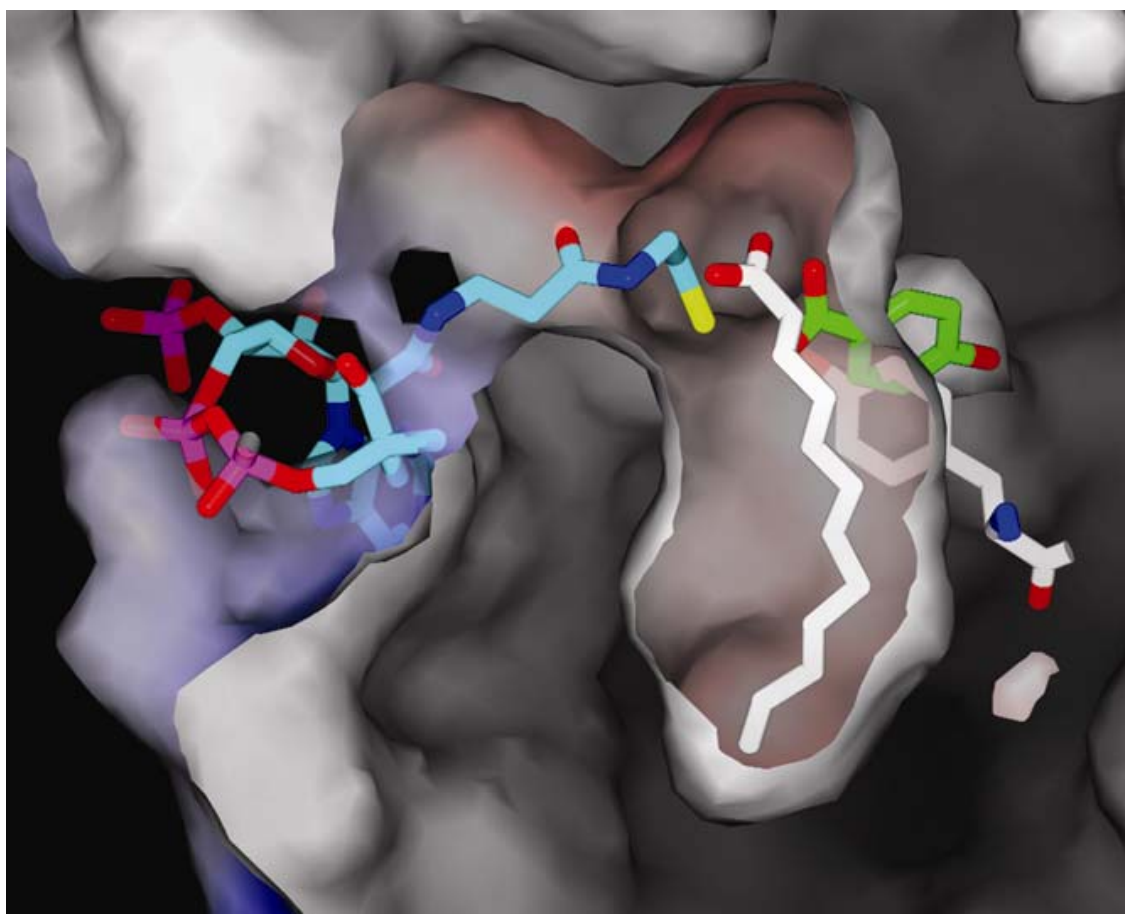


Figure 45 The Substrate-binding Tunnel of *Mtb* FcoT with Bound Ligand.

Cross section through the surface of the aligned FcoT active site with bound product dodecoenoyl acid shown to illustrate the “L” shaped substrate-binding tunnel of FcoT. Carbon atoms of 4-hydroxybenzoic acid and CoA in the structure of *Arthrobacter* 4-hydroxybenzoyl-CoA thioesterase are colored in green and cyan, respectively. The carbon atoms of dodecoenoyl acid in the structure of *M. tuberculosis* FcoT are colored in white

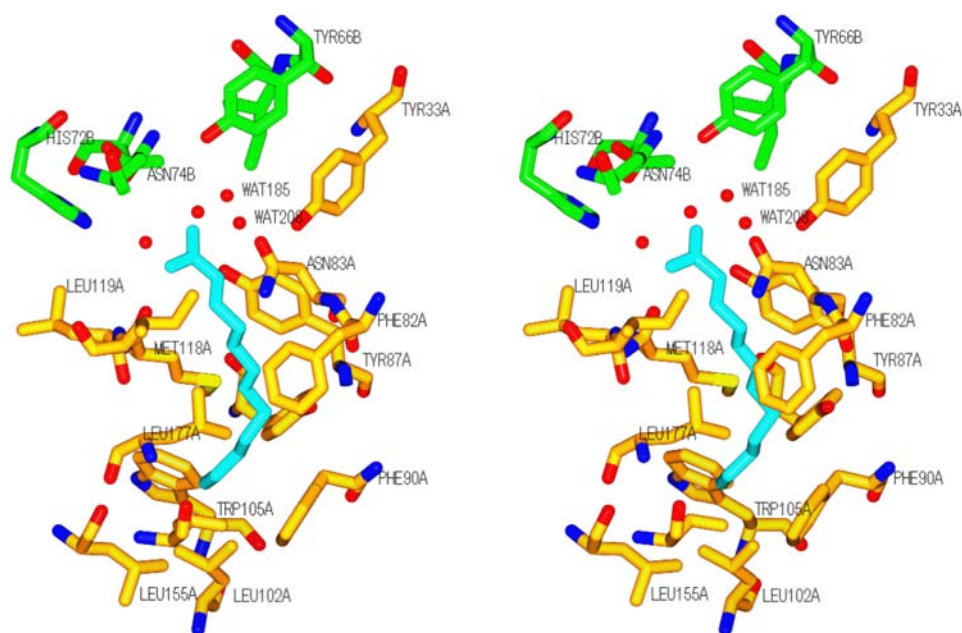


Figure 46 The Active Site of *Mtb* FcoT with Bound Dodecoenoate.

The dodecoenoyl acid is colored cyan and water is red. The residues in the active site are colored according to the atom type as following: carbon atoms are gold in subunit A and green in subunit B, oxygen atoms are red, nitrogen atoms are blue, sulfur atoms are yellow, and phosphorus atoms are orange

### Catalytic Site and Mechanism

With the help of the bound product dodecenoic acid, two symmetrically related active sites were identified at the interface of subunit A and B, ~20 Å apart. Each catalytic site is located at the intersection of the two branches of the “L” shaped substrate-binding tunnel. Upon the acyl-CoA substrate binding to the enzyme, the acyl chain portion fits into the longer hydrophobic branch of the tunnel while the pantetheine moiety extends out along the shorter hydrophilic portion (Figure 46). The thioester group is positioned in a cavity formed by Asn<sup>83</sup>, Tyr<sup>87</sup>, Tyr<sup>33</sup>, Met<sup>118</sup> (subunit A), and Tyr<sup>66</sup>, Thr<sup>70</sup>, His<sup>72</sup>, Asn<sup>74</sup> (subunit B).

The active site of FcoT is very different from any other type II thioesterases characterized to date. In all type II thioesterases with defined structures, a residue with a carboxyl group side chain (either aspartate or glutamate) is observed in the active site and acts as general base to deprotonate the water molecule, assisted usually by either a threonine or serine. None of these residues are in the catalytic site of *Mtb* FcoT, suggesting a different catalytic mechanism. Seven FcoT mutants, N83A, Y66F, Y33F, N74A, Y87F, E77Q, and H72A were constructed to identify the residues critical to the catalysis. Mutations of Y66F, Y33F, and N74A abolished the activity of the enzyme, while Y87F, E77Q, H72A, and N83A decreased the enzyme activity to less than ten fold. Although no atom of the protein directly forms hydrogen bonds with the product, the carboxyl group of the product interacts with Tyr<sup>87</sup> and Tyr<sup>33</sup> (subunit A), as well as Tyr<sup>66</sup> (subunit B) through two ordered water molecules (Figure 46). WAT<sup>185</sup> forms hydrogen bonds with the oxygen of dodecenoic acid (2.6 Å) and the side chain of Tyr<sup>66</sup>

(3.0 Å). WAT<sup>208</sup> is in the center of a hydrogen-bonding network with the oxygen of dodecenoic acid, the side chain hydroxyl group of Tyr<sup>33</sup>, and Tyr<sup>87</sup> at distances of 2.3 Å, 2.6 Å, and 3.0 Å, respectively. Interestingly, WAT<sup>208</sup> forms a low energy barrier hydrogen bond (211) to the oxygen of the fatty acid. The electron density and the B factor of WAT<sup>208</sup> are comparable to the average of the protein, which suggests that it is not a coordinated metal. Although Tyr<sup>33</sup> (subunit A), and Tyr<sup>66</sup> (subunit B) of FcoT are critical to catalysis, they do not seem to function as general bases in the active site, because the pH profile indicated the pKa of the reaction is only 8.1, much less than tyrosine's pKa, 10.4. Moreover, there is no ionizable group within 7 Å of the tyrosine residues, so it is not likely that the environment of the tyrosines could dramatically decrease the pKa to 8.1. While the residues in the catalytic site have never been observed in any other thioesterase, either type I or type II, the catalytic site of FcoT is similar to that of hydrolytic antibody D2.3. The catalytic site of D2.3 is located in the antibody-combining site at the interface between the heavy and light chain variable regions (212). By comparing the catalytic sites of FcoT and D2.3, the residues critical for the D2.3 activity, such as TyrH100, TyrL27, TyrL96, AsnL34, and HisH35, can be observed in the catalytic site of FcoT in a similar arrangement (Figure 47). Furthermore, the activity of D2.3 is also pH dependent and was maximal at pH 9. The catalytic mechanism of D2.3 is believed to be a direct hydroxide attack on the ester, in which the catalytic rate depends on the concentration of the hydroxide anions. It is likely that FcoT adopts a similar catalytic mechanism that activates the thioester substrate and stabilizes the negatively charged oxyanion intermediate resulting from a hydroxide attack on the

substrate, by the side chains of three residues, Tyr<sup>33</sup>, Asn<sup>74</sup>, and Tyr<sup>66</sup> (Table 25). The catalytic activity of this mechanism is  $10^2$ – $10^4$  times lower than those of type I and type II thioesterases or esterases, which is seldom observed in natural enzymes. Because of the distinct active site and catalytic mechanism, FcoT should be classified as a type III thioesterase (Figure 48).

Table 25 Kinetic Parameters of *Mtb* FcoT Mutants Using Palmitoyl-CoA

	$k_{\text{cat}}$ (s <sup>-1</sup> )	$K_m$ (μM)
Y87F	1.10E-03	29.31
N83A	9.30E-04	17.50
E77Q	2.60E-02	12.19
H72A	1.70E-02	35.14
Y66F	n/a	n/a
Y33F	n/a	n/a
N74A	n/a	n/a

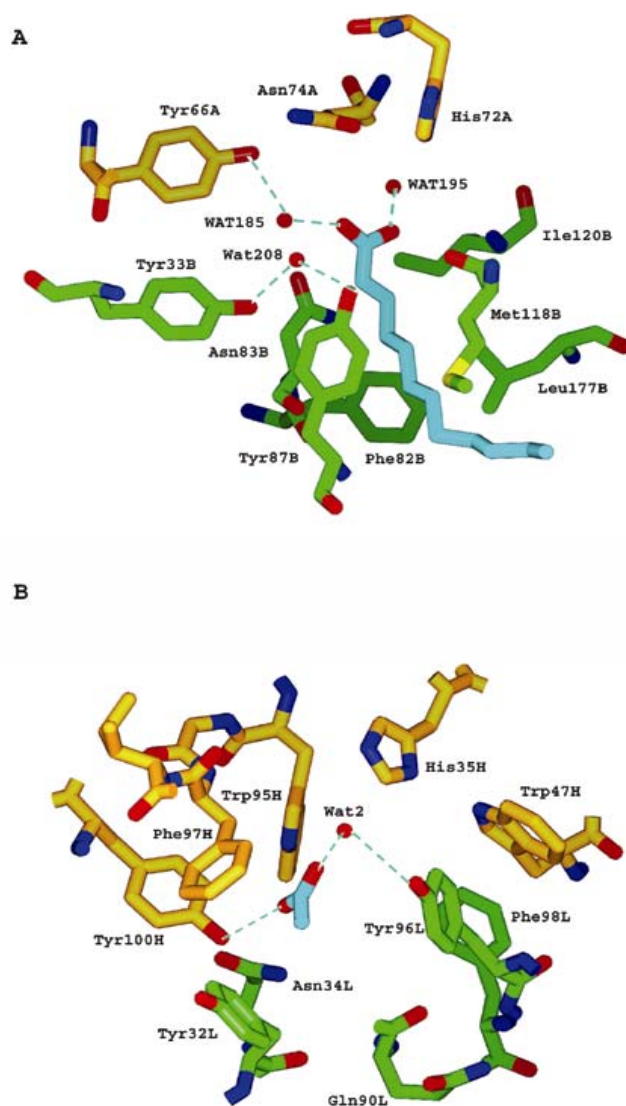


Figure 47 Comparison of the Catalytic Sites of FcoT and D2.3.

A. Catalytic site of FcoT bound to dodecoenoic acid. Carbon atoms of residues from subunit A and B are colored in gold and green, respectively. B. Catalytic site of D2.3 bound to acetic acid. Carbon atoms of residues from heavy chain and light chain are colored in gold and green, respectively



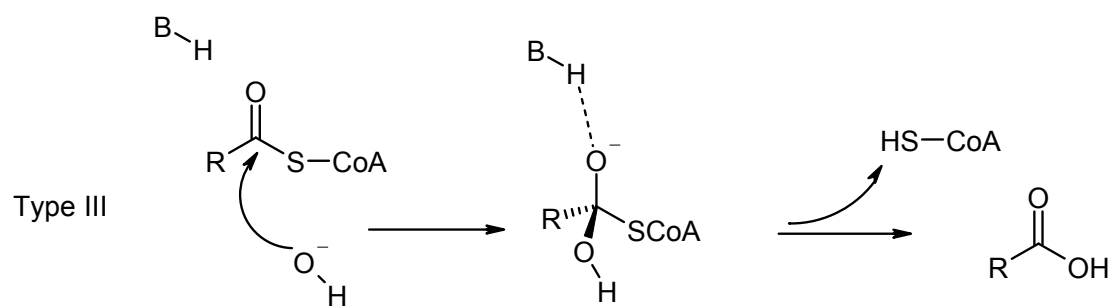


Figure 48 A Catalytic Mechanism Without a General Base

### Physiological Function of FcoT

*Rv0098* – *Rv0101* are critical for *Mtb* survival in the macrophage (41,42). Moreover, mouse infection studies have indicated that a knockout of any of the four genes leads to the attenuation of pathogenicity. Recently, it is found that the pathogenicity of *Rv0098* is highly dependent on the *mce* loci that encodes an ABC-like transport system (195). This newly identified transport system is believed to be responsible for both the export of a variety of complex lipid virulence factors and import of fatty acids as a source of nutritional carbons during *in vivo* growth. Based on sequence alignment, *Rv0099* (*fadD10*) is tentatively predicted to be a fatty acid AMP ligase, *Rv0100* encodes an acyl carrier protein, and *Rv0101* is a non-ribosomal peptide synthase (NRPS) (213). There are only two NRPS genes (or gene clusters) identified in the *Mtb* genome. The product of the other NRPS complex is known to be mycobactin (Figure 49). In other organisms, type II thioesterase genes have been found to be associated with the NRPS gene in surfactin, tyrocidine, and bacitracin biosynthesis pathways (201). Unlike the thioesterase I domain in NRPS, type II thioesterases are not involved in the termination of the synthesis. It was previously reported that the carrier domain of a NRPS might become misprimed by an acylated phosphopantetheine (PPT) instead of being the holo-enzyme. Type II thioesterases could hydrolyze acylated PPT to correct the misprimed NRPS (201); they have also been shown to catalyze deaminoacylation in order to edit the misloaded NRPS module (214). However, *Mtb* FcoT prefers long chain substrates, while the thioesterase II that functioned to regenerate

misprimed NRPS only showed activity to very short chain (up to C3) substrates (201). Furthermore, no deaminoacylation activity was detected in *Mtb* FcoT.

We hypothesize that *Mtb* FcoT hydrolyzes fatty-acyl-CoA substrates to generate free fatty acids with specific lengths (C16-C18) and the protein encoded by *fadD10* subsequently catalyzes the ligation of the fatty acid with the “specialized” acyl carrier protein encoded by *Rv0100* to produce acyl-ACP, a precursor for the non-ribosomal peptide synthase to generate a lipopeptide. Unlike the mycobactin NRPS, it is believed that the gene product of *Rv0101* does not have a cyclization domain, based on sequence analysis (213). Therefore, unlike mycobactin, the product of this pathway could be linear. No linear lipopeptide has been identified in *Mtb*, but a linear lipopeptide product, fortuitin, has been found in *M. fortuitum* (Figure 50) (215). Lipopeptides are major constituents of the cell envelopes of mycobacteria. They alter colony morphology and usually relate to virulence. In *M. bovis*, the product of this *PPE - nrp* operon is essential in the biosynthesis of PDIMs. In *Mtb*, this operon is likely related to PDIM synthesis, although it was found that *fadd26-mmpL7* gene cluster played a major role in the synthesis of PDIMs. Nevertheless, the identification of *Mtb* FcoT provides new insights into this heretofore uncharacterized operon that is crucial for *Mtb* pathogenicity.

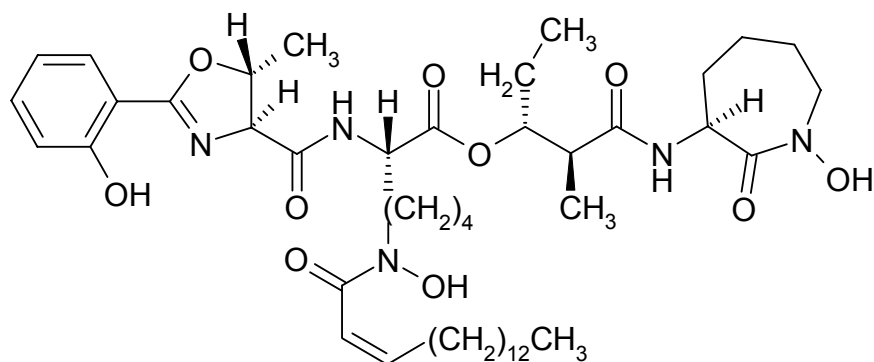


Figure 49 Chemical Structure of Mycobactin



Figure 50 Chemical Structure of Fortuitin

## CHAPTER VII

### CONCLUSION

As the world's most successful pathogen, *Mtb* has infected more than one third of the human population. The emergence of MDR and XDR TB strains and the co-infection of *Mtb* with HIV make it more deadly than ever. It is urgent to develop a new generation of TB drugs to counter resistance.

As Sir James Black said, “the most fruitful basis to develop a new drug is an old drug”. INH and ETH (and PTH) have been widely used to treat TB for more than 40 years. Knowledge of their mechanism of action is crucial to the design of new drugs that can overcome drug-resistant TB cases. INH has been demonstrated to be activated by KatG, a catalase-peroxidase, to form the INH-NAD adduct that inhibits InhA to block mycolic acid biosynthesis. *Mtb* strains harboring the *inhA*(S94A) mutation have been shown to confer resistance to INH. We purified the InhA(S94A) recombinant protein. The mutant protein showed much higher IC<sub>50</sub> (17 fold) and Ki (30 fold) values, compared to the wild type InhA, in the inhibition studies of the INH-NAD adduct. The crystal structure of InhA(S94A) in complex with INH-NAD adduct clearly showed the structural basis for the resistance. The substitution of Ser94 to alanine disrupted a hydrogen-bonding network, which weakens the binding of INH-NAD adduct. These results helped to elucidate the mode of action of INH and confirmed that InhA is the target of INH. In order to find out the basis for the extremely high activity of INH, we used three hydrazides as probes to study their difference in the activation and inhibition

steps. All three hydrazides were activated and formed adducts with NAD cofactor. The INH-NAD, PNH-NAD, and BZH-NAD adducts are potent inhibitors against InhA. However, the yield of these adducts are dramatically different. Therefore, the high antitubercular activity of INH compared to its analogs is not due to its structurally unique moiety in the INH-NAD adduct, but rather the high yield of its activated form; this is provided by two important factors: 1) high stability of the isonicotinoyl intermediate; 2) no significant product inhibition to KatG during activation.

The mechanism of action of ETH and PTH was demonstrated in a cell-based activation system. In this experiment, *Mtb ethA* and *inhA* were co-expressed inside *E. coli* cells in the presence of ETH or PTH. The ETH-NAD and PTH-NAD adducts were identified by solving the crystal structures of InhA bound with the inhibitors and confirmed by mass analysis. Similar to the INH-NAD adduct, the ETH-NAD and PTH-NAD adducts are extremely potent against purified InhA *in vitro* ( $K_i = 7$  nM and 2 nM respectively), which elucidated the mode of action of ETH and PTH against both *Mtb* and *M. leprae*.

Since INH, ETH and PTH all target InhA, this validates the enzyme as a good drug target. Using a structure-based drug design approach with triclosan as the scaffold, a series of 5-substituted derivatives of triclosan was developed targeting InhA. Two groups of triclosan derivatives with alkyl and aryl substituents, respectively, were identified with dramatically increased potency against purified InhA. Crystal structures of InhA in complex with these compounds demonstrated the structural basis for their activities. Moreover, the correlation of the activities of these triclosan derivatives against

purified enzyme and *Mtb* strains supported the proposal that InhA is the biological target.

Other than the known target InhA, we also tried to search for new potential therapeutic targets in *Mtb*. Two important *Mtb* proteins were structurally and functionally characterized. *Mtb* Class A  $\beta$ -lactamase has been proposed to be the most significant reason for mycobacterial resistance to  $\beta$ -lactam antibiotics. The determination of a high resolution structure of BlaC not only demonstrates the mechanism of drug resistance but also provides a solid base for the design of new  $\beta$ -lactamase inhibitors that could be used with  $\beta$ -lactam antibiotics as a new regimen to treat TB. *Mtb* FcoT, encoded by *Rv0098*, is functionally characterized by a combination of biochemical and structural approaches. The previously hypothetical protein is now annotated as a long chain fatty acyl CoA thioesterase. Since it lacks a general base or nucleophile in the active site, which always exists in type II and type I thioesterases, respectively, FcoT adopts a new type of catalytic mechanism. The characterization of FcoT revealed an important pathway that is critical for *Mtb*'s survival in host macrophages.

## REFERENCES

1. Vuorinen, H. S. (1997) *Hippokrates*, 74-97
2. WHO (2003) Treatment of Tuberculosis: Guidelines for National Programmes, Report ([http://www.who.int/tb/publications/cds\\_tb\\_2003\\_313](http://www.who.int/tb/publications/cds_tb_2003_313))
3. Chin, D. P., and Hopewell, P. C. (1996) *Clin Chest Med* **17**, 697-711
4. Chin, D. P., Osmond, D., Page-Shafer, K., Glassroth, J., Rosen, M. J., Reichman, L. B., Kvale, P. A., Wallace, J. M., Poole, W. K., and Hopewell, P. C. (1996) *Am J Respir Crit Care Med* **153**, 1982-1984
5. Corbett, E. L., Watt, C. J., Walker, N., Maher, D., Williams, B. G., Raviglione, M. C., and Dye, C. (2003) *Arch Intern Med* **163**, 1009-1021
6. Enarson, D. A., Murray J. F. (1996) *Global Epidemiology of Tuberculosis*, Lippincott Williams and Wilkins, Philadelphia
7. Harada, K., Gidoh, S., and Tsutsumi, S. (1976) *Microsc Acta* **78**, 21-27
8. Dannenberg, A. M., Jr. (1993) *Hosp Pract (Off Ed)* **28**, 51-58
9. Monack, D. M., Mueller, A., and Falkow, S. (2004) *Nat Rev Microbiol* **2**, 747-765
10. Fernandez, S. H. (1999) *Malays J Pathol* **21**, 111-115
11. Butler, R. N. (1993) *Geriatrics* **48**, 15-16
12. Gracey, D. R. (1988) *Mayo Clin Proc* **63**, 1251-1255
13. Rozwarski, D. A., Grant, G. A., Barton, D. H., Jacobs, W. R., Jr., and Sacchettini, J. C. (1998) *Science* **279**, 98-102



14. Vilcheze, C., Wang, F., Arai, M., Hazbon, M. H., Colangeli, R., Kremer, L., Weisbrod, T. R., Alland, D., Sacchettini, J. C., and Jacobs, W. R., Jr. (2006) *Nat. Med.* **12**, 1027-1029
15. Zhang, Y., Heym, B., Allen, B., Young, D., and Cole, S. (1992) *Nature* **358**, 591-593
16. Guo, H., Seet, Q., Denkin, S., Parsons, L., and Zhang, Y. (2006) *J Med Microbiol* **55**, 1527-1531
17. Telenti, A. (1998) *Thorax* **53**, 793-797
18. Vilcheze, C., Weisbrod, T. R., Chen, B., Kremer, L., Hazbon, M. H., Wang, F., Alland, D., Sacchettini, J. C., and Jacobs, W. R., Jr. (2005) *Antimicrob. Agents Chemother.* **49**, 708-720
19. DeBarber, A. E., Mdluli, K., Bosman, M., Bekker, L. G., and Barry, C. E., 3rd. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 9677-9682
20. Baulard, A. R., Betts, J. C., Engohang-Ndong, J., Quan, S., McAdam, R. A., Brennan, P. J., Loch, C., and Besra, G. S. (2000) *J. Biol. Chem.* **275**, 28326-28331
21. Wang, F., Langley, R., Gulten, G., Dover, L. G., Besra, G. S., Jacobs, W. R., Jr., and Sacchettini, J. C. (2007) *J Exp Med* **204**, 73-78
22. Caceres, N. E., Harris, N. B., Wellehan, J. F., Feng, Z., Kapur, V., and Barletta, R. G. (1997) *J Bacteriol* **179**, 5046-5055
23. Chacon, O., Feng, Z., Harris, N. B., Caceres, N. E., Adams, L. G., and Barletta, R. G. (2002) *Antimicrob Agents Chemother* **46**, 47-54

24. Spotts, C. R., and Stanier, R. Y. (1961) *Nature* **192**, 633-637
25. Murakami, K. S., Masuda, S., and Darst, S. A. (2002) *Science* **296**, 1280-1284
26. Campbell, E. A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., and Darst, S. A. (2001) *Cell* **104**, 901-912
27. Mitchison, D. A. (1985) *Tubercle* **66**, 219-225
28. Telenti, A., Imboden, P., Marchesi, F., Lowrie, D., Cole, S., Colston, M. J., Matter, L., Schopfer, K., and Bodmer, T. (1993) *Lancet* **341**, 647-650
29. Kocagoz, T., Hackbarth, C. J., Unsal, I., Rosenberg, E. Y., Nikaido, H., and Chambers, H. F. (1996) *Antimicrob Agents Chemother* **40**, 1768-1774
30. Takiff, H. E., Salazar, L., Guerrero, C., Philipp, W., Huang, W. M., Kreiswirth, B., Cole, S. T., Jacobs, W. R., Jr., and Telenti, A. (1994) *Antimicrob Agents Chemother* **38**, 773-780
31. Pablos-Mendez, A., Raviglione, M. C., Laszlo, A., Binkin, N., Rieder, H. L., Bustreo, F., Cohn, D. L., Lambregts-van Weezenbeek, C. S., Kim, S. J., Chaulet, P., and Nunn, P. (1998) *N Engl J Med* **338**, 1641-1649
32. Flynn, J. L., and Chan, J. (2001) *Infect Immun* **69**, 4195-4201
33. Zhang, Y. (2004) *Front Biosci* **9**, 1136-1156
34. De Cock, K. M., and Chaisson, R. E. (1999) *Int J Tuberc Lung Dis* **3**, 457-465
35. Kimerling, M. E., Kluge, H., Vezhnina, N., Iacovazzi, T., Demeulenaere, T., Portaels, F., and Matthys, F. (1999) *Int J Tuberc Lung Dis* **3**, 451-453

36. Stover, C. K., Warrenner, P., VanDevanter, D. R., Sherman, D. R., Arain, T. M., Langhorne, M. H., Anderson, S. W., Towell, J. A., Yuan, Y., McMurray, D. N., Kreiswirth, B. N., Barry, C. E., and Baker, W. R. (2000) *Nature* **405**, 962-966
37. Lenaerts, A. J., Gruppo, V., Marietta, K. S., Johnson, C. M., Driscoll, D. K., Tompkins, N. M., Rose, J. D., Reynolds, R. C., and Orme, I. M. (2005) *Antimicrob Agents Chemother* **49**, 2294-2301
38. Andries, K., Verhasselt, P., Guillemont, J., Gohlmann, H. W., Neefs, J. M., Winkler, H., Van Gestel, J., Timmerman, P., Zhu, M., Lee, E., Williams, P., de Chaffoy, D., Huitric, E., Hoffner, S., Cambau, E., Truffot-Pernot, C., Lounis, N., and Jarlier, V. (2005) *Science* **307**, 223-227
39. Ji, B., Lounis, N., Maslo, C., Truffot-Pernot, C., Bonnafous, P., and Grosset, J. (1998) *Antimicrob Agents Chemother* **42**, 2066-2069
40. Alvarez-Freites, E. J., Carter, J. L., and Cynamon, M. H. (2002) *Antimicrob Agents Chemother* **46**, 1022-1025
41. Sasseti, C. M., Boyd, D. H., and Rubin, E. J. (2003) *Mol. Microbiol.* **48**, 77-84
42. Sasseti, C. M., and Rubin, E. J. (2003) *Proc. Natl. Acad. Sci. U S A* **100**, 12989-12994
43. McKinney, J. D., Honer zu Bentrup, K., Munoz-Elias, E. J., Miczak, A., Chen, B., Chan, W. T., Swenson, D., Sacchettini, J. C., Jacobs, W. R., Jr., and Russell, D. G. (2000) *Nature* **406**, 735-738

44. Dahl, J. L., Kraus, C. N., Boshoff, H. I., Doan, B., Foley, K., Avarbock, D., Kaplan, G., Mizrahi, V., Rubin, H., and Barry, C. E., 3rd. (2003) *Proc Natl Acad Sci U S A* **100**, 10026-10031
45. Glickman, M. S., Cox, J. S., and Jacobs, W. R., Jr. (2000) *Mol Cell* **5**, 717-727
46. Takayama, K., Wang, C., and Besra, G. S. (2005) *Clin. Microbiol. Rev.* **18**, 81-101
47. Brennan, P. J. (2003) *Tuberculosis (Edinb)* **83**, 91-97
48. George, K. M., Yuan, Y., Sherman, D. R., and Barry, C. E., 3rd. (1995) *J Biol Chem* **270**, 27292-27298
49. Yuan, Y., Lee, R. E., Besra, G. S., Belisle, J. T., and Barry, C. E., 3rd. (1995) *Proc Natl Acad Sci U S A* **92**, 6630-6634
50. Asselineau, J., and Lederer, E. (1950) *Nature* **166**, 782-783
51. Dubnau, E., Chan, J., Raynaud, C., Mohan, V. P., Laneelle, M. A., Yu, K., Quemard, A., Smith, I., and Daffe, M. (2000) *Mol Microbiol* **36**, 630-637
52. Yuan, Y., Zhu, Y., Crane, D. D., and Barry, C. E., 3rd. (1998) *Mol Microbiol* **29**, 1449-1458
53. Zimhony, O., Vilcheze, C., and Jacobs, W. R., Jr. (2004) *J Bacteriol* **186**, 4051-4055
54. Schweizer, E., and Hofmann, J. (2004) *Microbiol Mol Biol Rev* **68**, 501-517
55. Kremer, L., Nampoothiri, K. M., Lesjean, S., Dover, L. G., Graham, S., Betts, J., Brennan, P. J., Minnikin, D. E., Locht, C., and Besra, G. S. (2001) *J Biol Chem* **276**, 27967-27974

56. Choi, K. H., Kremer, L., Besra, G. S., and Rock, C. O. (2000) *J Biol Chem* **275**, 28201-28207
57. 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
58. Gao, L. Y., Laval, F., Lawson, E. H., Groger, R. K., Woodruff, A., Morisaki, J. H., Cox, J. S., Daffe, M., and Brown, E. J. (2003) *Mol Microbiol* **49**, 1547-1563
59. Marrakchi, H., Choi, K. H., and Rock, C. O. (2002) *J Biol Chem* **277**, 44809-44816
60. Takayama, K., and Qureshi, N. (1978) *Lipids* **13**, 575-579
61. Glickman, M. S. (2003) *J Biol Chem* **278**, 7844-7849
62. Portevin, D., De Sousa-D'Auria, C., Houssin, C., Grimaldi, C., Chami, M., Daffe, M., and Guilhot, C. (2004) *Proc Natl Acad Sci U S A* **101**, 314-319
63. Kim, B. S., Cropp, T. A., Beck, B. J., Sherman, D. H., and Reynolds, K. A. (2002) *J. Biol. Chem.* **277**, 48028-48034
64. Schwarzer, D., Mootz, H. D., Linne, U., and Marahiel, M. A. (2002) *Proc Natl Acad Sci U S A* **99**, 14083-14088

65. Brennan, P. J., and Nikaido, H. (1995) *Annu Rev Biochem* **64**, 29-63
66. Congreve, M., Murray, C. W., and Blundell, T. L. (2005) *Drug Discov Today* **10**, 895-907
67. Bajorath, J. (2001) *Drug Discov Today* **6**, 989-995
68. Clark, R. D., Strizhev, A., Leonard, J. M., Blake, J. F., and Matthew, J. B. (2002) *J Mol Graph Model* **20**, 281-295
69. Eldridge, M. D., Murray, C. W., Auton, T. R., Paolini, G. V., and Mee, R. P. (1997) *J Comput Aided Mol Des* **11**, 425-445
70. Godden, J. W., Stahura, F., and Bajorath, J. (1998) *J Mol Graph Model* **16**, 139-143, 165
71. Gohlke, H., Hendlich, M., and Klebe, G. (2000) *J Mol Biol* **295**, 337-356
72. Gohlke, H., and Klebe, G. (2001) *Curr Opin Struct Biol* **11**, 231-235
73. Cavasotto, C. N., and Abagyan, R. A. (2004) *J Mol Biol* **337**, 209-225
74. Liu, Z., Dominy, B. N., and Shakhnovich, E. I. (2004) *J Am Chem Soc* **126**, 8515-8528
75. Xue, L., and Bajorath, J. (2000) *Comb Chem High Throughput Screen* **3**, 363-372
76. Irwin, J. J., and Shoichet, B. K. (2005) *J Chem Inf Model* **45**, 177-182
77. Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (2001) *Adv Drug Deliv Rev* **46**, 3-26
78. Lorber, D. M., and Shoichet, B. K. (1998) *Protein Sci* **7**, 938-950
79. Goodsell, D. S., Morris, G. M., and Olson, A. J. (1996) *J Mol Recognit* **9**, 1-5

80. Muegge, I., and Martin, Y. C. (1999) *J Med Chem* **42**, 791-804
81. Bissantz, C., Bernard, P., Hibert, M., and Rognan, D. (2003) *Proteins* **50**, 5-25
82. Parang, K., and Sun, G. (2004) *Curr Opin Drug Discov Devel* **7**, 617-629
83. Mol, C. D., Fabbro, D., and Hosfield, D. J. (2004) *Curr Opin Drug Discov Devel* **7**, 639-648
84. Retailleau, P., Ries-Kautt, M., and Ducruix, A. (1997) *Biophys J* **73**, 2156-2163
85. McPherson, A. (1982) *Preparation and Analysis of Protein Crystals*, Waverly, Baltimore, MD
86. Otwinowski, Z., Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode, C.W. Carter, Jr. & R. M. Sweet (eds.). *Methods in Enzymology*, **276**, pp. 307-326, Academic Press, New York
87. Schneider, T. R., and Sheldrick, G. M. (2002) *Acta Crystallogr D Biol Crystallogr* **58**, 1772-1779
88. Reddy, V., Swanson, S. M., Segelke, B., Kantardjieff, K. A., Sacchettini, J. C., and Rupp, B. (2003) *Acta Crystallogr D Biol Crystallogr* **59**, 2200-2210
89. McRee, D. E. (1999) *J. Struct. Biol.* **125**, 156-165
90. Christopher, J. A. (1998) *SPOCK: The Structural Properties Observation and Calculation Kit*. Program Manual, The Center for Macromolecular Design, Texas A&M University, College Station, TX
91. Lamzin, V. S., and Wilson, K. S. (1993) *Acta Crystallogr D Biol Crystallogr* **49**, 129-147
92. Ioerger, T. R., and Sacchettini, J. C. (2003) *Methods Enzymol* **374**, 244-270

93. Brünger, A. T. (1997) *Methods Enzymol.* **277**, 366
94. Bernstein, J., Lott, W. A., Steinberg, B. A., and Yale, H. L. (1952) *Am Rev Tuberc* **65**, 357-364
95. Bloch, K. (1977) *Adv Enzymol Relat Areas Mol Biol* **45**, 1-84
96. Winder, F. G., Collins, P., and Rooney, S. A. (1970) *Biochem J* **117**, 27P
97. Takayama, K., Wang, L., and David, H. L. (1972) *Antimicrob. Agents Chemother.* **2**, 29-35
98. Takayama, K., Schnoes, H. K., Armstrong, E. L., and Boyle, R. W. (1975) *J Lipid Res* **16**, 308-317
99. Middlebrook, G. (1954) *Am Rev Tuberc* **69**, 471-472
100. Madison, B. M., Siddiqi, S. H., Heifets, L., Gross, W., Higgins, M., Warren, N., Thompson, A., Morlock, G., and Ridderhof, J. C. (2004) *J Clin Microbiol* **42**, 1294-1295
101. Hazbon, M. H., Brimacombe, M., Bobadilla del Valle, M., Cavatore, M., Guerrero, M. I., Varma-Basil, M., Billman-Jacobe, H., Lavender, C., Fyfe, J., Garcia-Garcia, L., Leon, C. I., Bose, M., Chaves, F., Murray, M., Eisenach, K. D., Sifuentes-Osornio, J., Cave, M. D., Ponce de Leon, A., and Alland, D. (2006) *Antimicrob Agents Chemother* **50**, 2640-2649
102. Escalante, P., Ramaswamy, S., Sanabria, H., Soini, H., Pan, X., Valiente-Castillo, O., and Musser, J. M. (1998) *Tuber Lung Dis* **79**, 111-118
103. Lei, B., Wei, C. J., and Tu, S. C. (2000) *J. Biol. Chem.* **275**, 2520-2526
104. Wilming, M., and Johnsson, K. (1999) *Angew Chem Int Ed Engl* **38**, 2588-2590



105. Rawat, R., Whitty, A., and Tonge, P. J. (2003) *Proc Natl Acad Sci U S A* **100**, 13881-13886
106. Morlock, G. P., Metchock, B., Sikes, D., Crawford, J. T., and Cooksey, R. C. (2003) *Antimicrob. Agents Chemother.* **47**, 3799-3805
107. Banerjee, A., Dubnau, E., Quemard, A., Balasubramanian, V., Um, K. S., Wilson, T., Collins, D., de Lisle, G., and Jacobs, W. R., Jr. (1994) *Science* **263**, 227-230
108. Larsen, M. H., Vilcheze, C., Kremer, L., Besra, G. S., Parsons, L., Salfinger, M., Heifets, L., Hazbon, M. H., Alland, D., Sacchettini, J. C., and Jacobs, W. R., Jr. (2002) *Mol. Microbiol.* **46**, 453-466
109. Chouchane, S., Girotto, S., Yu, S., and Magliozzo, R. S. (2002) *J Biol Chem* **277**, 42633-42638
110. Chouchane, S., Lippai, I., and Magliozzo, R. S. (2000) *Biochemistry* **39**, 9975-9983
111. Ghiladi, R. A., Medzihradzsky, K. F., Rusnak, F. M., and Ortiz de Montellano, P. R. (2005) *J Am Chem Soc* **127**, 13428-13442
112. Jakopitsch, C., Droghetti, E., Schmuckenschlager, F., Furtmuller, P. G., Smulevich, G., and Obinger, C. (2005) *J Biol Chem* **280**, 42411-42422
113. Zhao, X., Yu, H., Yu, S., Wang, F., Sacchettini, J. C., and Magliozzo, R. S. (2006) *Biochemistry* **45**, 4131-4140
114. Nguyen, M., Quemard, A., Broussy, S., Bernadou, J., and Meunier, B. (2002) *Antimicrob Agents Chemother* **46**, 2137-2144

115. Timmins, G. S., Master, S., Rusnak, F., and Deretic, V. (2004) *J Bacteriol* **186**, 5427-5431
116. Dessen, A., Quemard, A., Blanchard, J. S., Jacobs, W. R., Jr., and Sacchettini, J. C. (1995) *Science* **267**, 1638-1641
117. Yu, S., Giroto, S., Lee, C., and Magliozzo, R. S. (2003) *J Biol Chem* **278**, 14769-14775
118. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D Biol. Crystallogr.* **54** (Pt 5), 905-921
119. Johnsson, K., Schultz, P. G. (1994) *J. Am. Chem. Soc.* **116**, 7425-7426
120. Mdluli, K., Slayden, R. A., Zhu, Y., Ramaswamy, S., Pan, X., Mead, D., Crane, D. D., Musser, J. M., and Barry, C. E., 3rd. (1998) *Science* **280**, 1607-1610
121. Argyrou, A., Jin, L., Siconilfi-Baez, L., Angeletti, R. H., and Blanchard, J. S. (2006) *Biochemistry* **45**, 13947-13953
122. Timmins, G. S., Master, S., Rusnak, F., and Deretic, V. (2004) *Antimicrob Agents Chemother* **48**, 3006-3009
123. Argyrou, A., Vetting, M. W., Aladegbami, B., and Blanchard, J. S. (2006) *Nat Struct Mol Biol* **13**, 408-413
124. Fajardo, T. T., Guinto, R. S., Cellona, R. V., Abalos, R. M., Dela Cruz, E. C., and Gelber, R. H. (2006) *Am. J. Trop. Med. Hyg.* **74**, 457-461

125. Yajko, D. M., Nassos, P. S., and Hadley, W. K. (1987) *Antimicrob. Agents Chemother.* **31**, 117-120
126. Crofton, J., Chaulet, P., Maher, D., Grosset, J., Harris, W., Norman, H., Iseman, M. Watt, B. (1997) *Guidelines for the management of Multidrug-resistant Tuberculosis*, Report, World Health Organization, Geneva
127. Katoch, K., Natarajan, M., Bhatia, A. S., and Yadav, V. S. (1992) *Indian J. Lepr.* **64**, 303-312
128. Bardarov, S., Bardarov Jr, S., Jr., Pavelka Jr, M. S., Jr., Sambandamurthy, V., Larsen, M., Tufariello, J., Chan, J., Hatfull, G., and Jacobs Jr, W. R., Jr. (2002) *Microbiology* **148**, 3007-3017
129. Fattorini, L., Iona, E., Ricci, M. L., Thoresen, O. F., Orru, G., Oggioni, M. R., Tortoli, E., Piersimoni, C., Chiaradonna, P., Tronci, M., Pozzi, G., and Orefici, G. (1999) *Microb. Drug Resist.* **5**, 265-270
130. Fraaije, M. W., Kamerbeek, N. M., Heidekamp, A. J., Fortin, R., and Janssen, D. B. (2004) *J. Biol. Chem.* **279**, 3354-3360
131. Vannelli, T. A., Dykman, A., and Ortiz de Montellano, P. R. (2002) *J. Biol. Chem.* **277**, 12824-12829
132. Quemard, A., Laneelle, G., and Lacave, C. (1992) *Antimicrob. Agents Chemother.* **36**, 1316-1321
133. Tokuyama, H., Yamashita, T, Reding, M. T., Kaburagi, Y., Fukuyama, T. (1999) *J. Am. Chem. Soc.* **121**, 3791-3792
134. CDC. (2006) *Morb. Mortal. Wkly. Rep.* **55**, 301-305

135. Bentley, R. (2005) *Perspect Biol. Med.* **48**, 444-452
136. Edwards, J. R., and Betts, M. J. (2000) *J. Antimicrob. Chemother.* **45**, 1-4
137. Severin, A., Severina, E., and Tomasz, A. (1997) *Antimicrob. Agents Chemother.* **41**, 504-510
138. Sugai, M., Yamada, S., Nakashima, S., Komatsuzawa, H., Matsumoto, A., Oshida, T., and Suginaka, H. (1997) *J. Bacteriol.* **179**, 2958-2962
139. Medeiros, A. A. (1997) *Clin. Infect. Dis.* **24** Suppl 1, S19-45
140. Davies, J. (1994) *Science* **264**, 375-382
141. Amyes, S. G. (1997) *J. Med. Microbiol.* **46**, 454-457
142. Fisher, J. F., Meroueh, S. O., Mobashery, S. (2005) *Chem. Rev.* **105**, 395-424
143. Iland, C. N. (1946) *J. Pathol. Bacteriol.* **58**, 495-500
144. Chambers, H. F., Moreau, D., Yajko, D., Miick, C., Wagner, C., Hackbarth, C., Kocagoz, S., Rosenberg, E., Hadley, W. K., and Nikaido, H. (1995) *Antimicrob. Agents Chemother.* **39**, 2620-2624
145. Voladri, R. K., Lakey, D. L., Hennigan, S. H., Menzies, B. E., Edwards, K. M., and Kernodle, D. S. (1998) *Antimicrob. Agents Chemother.* **42**, 1375-1381
146. Jarlier, V., Gutmann, L., and Nikaido, H. (1991) *Antimicrob. Agents Chemother.* **35**, 1937-1939
147. Trias, J., Jarlier, V., and Benz, R. (1992) *Science* **258**, 1479-1481
148. Ambler, R. P., Coulson, A. F., Frere, J. M., Ghuysen, J. M., Joris, B., Forsman, M., Levesque, R. C., Tiraby, G., and Waley, S. G. (1991) *Biochem. J.* **276**, 269-270

149. Frere, J. M. (1995) *Mol. Microbiol.* **16**, 385-395
150. Dincer, I., Ergin, A., and Kocagoz, T. (2004) *Int. J. Antimicrob. Ag.* **23**, 408-411
151. Sorg, T. B., and Cynamon, M. H. (1987) *J. Antimicrob. Chemother.* **19**, 59-64
152. Wong, C. S., Palmer, G. S., and Cynamon, M. H. (1988) *J. Antimicrob. Chemother.* **22**, 863-866
153. Abate, G., and Miorner, H. (1998) *J. Antimicrob. Chemother.* **42**, 735-740
154. Flores, A. R., Parsons, L. M., and Pavelka, M. S., Jr. (2005) *Microbiology* **151**, 521-532
155. Flores, A. R., Parsons, L. M., and Pavelka, M. S., Jr. (2005) *J. Bacteriol.* **187**, 1892-1900
156. Bush, K., Jacoby, G. A., and Medeiros, A. A. (1995) *Antimicrob. Agents Chemother.* **39**, 1211-1233
157. Segura, C., Salvado, M., Collado, I., Chaves, J., and Coira, A. (1998) *Antimicrob. Agents Chemother.* **42**, 1524-1526
158. Ibuka, A. S., Ishii, Y., Galleni, M., Ishiguro, M., Yamaguchi, K., Frere, J. M., Matsuzawa, H., and Sakai, H. (2003) *Biochemistry* **42**, 10634-10643
159. Wong, J. T. (1985) *Kinetics of Enzyme Mechanisms*, Academic Press, New York
160. Collaborative Computational Project, N. (1994) *Acta Crystallogr. D* **50**, 760-763
161. Dideberg, O., Charlier, P., Wery, J. P., Dehottay, P., Dusart, J., Erpicum, T., Frere, J. M., and Ghuysen, J. M. (1987) *Biochem. J.* **245**, 911-913
162. Holm, L., and Sander, C. (1995) *Trends in Biochemical Sciences* **20**, 478-480
163. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* **18**, 2714-2723

164. Gattiker, A., Gasteiger, E., and Bairoch, A. (2002) *Appl. Bioinformatics* **1**, 107-108
165. Nielsen, J. B., and Lampen, J. O. (1982) *J. Bacteriol.* **152**, 315-322
166. Sutcliffe, I. C., and Russell, R. R. (1995) *J. Bacteriol.* **177**, 1123-1128
167. Bendtsen, J. D., Nielsen, H., von Heijne, G., and Brunak, S. (2004) *J. Mol. Biol.* **340**, 783-795
168. Zhang, Y., Steingrube, V. A., and Wallace, R. J., Jr. (1992) *Am. Rev. Resp. Dis.* **145**, 657-660
169. Quinting, B., Reyrat, J. M., Monnaie, D., Amicosante, G., Pelicic, V., Gicquel, B., Frere, J. M., and Galleni, M. (1997) *FEBS Lett* **406**, 275-278
170. Minasov, G., Wang, X., and Shoichet, B. K. (2002) *J. Am. Chem. Soc.* **124**, 5333-5340
171. Nukaga, M., Mayama, K., Hujer, A. M., Bonomo, R. A., and Knox, J. R. (2003) *J. Mol. Biol.* **328**, 289-301
172. Strynadka, N. C., Adachi, H., Jensen, S. E., Johns, K., Sielecki, A., Betzel, C., Sutoh, K., and James, M. N. (1992) *Nature* **359**, 700-705
173. Meroueh, S. O., Fisher, J. F., Schlegel, H. B., and Mobashery, S. (2005) *J Am Chem Soc* **127**, 15397-15407
174. Ibuka, A., Taguchi, A., Ishiguro, M., Fushinobu, S., Ishii, Y., Kamitori, S., Okuyama, K., Yamaguchi, K., Konno, M., and Matsuzawa, H. (1999) *J. Mol. Biol.* **285**, 2079-2087
175. Herzberg, O., and Moulton, J. (1987) *Science* **236**, 694-701

176. Lim, D., Sanschagrin, F., Passmore, L., De Castro, L., Levesque, R. C., and Strynadka, N. C. (2001) *Biochemistry* **40**, 395-402
177. Nukaga, M., Mayama, K., Crichlow, G. V., and Knox, J. R. (2002) *J. Mol. Biol.* **317**, 109-117
178. Shimamura, T., Ibuka, A., Fushinobu, S., Wakagi, T., Ishiguro, M., Ishii, Y., and Matsuzawa, H. (2002) *J. Biol. Chem.* **277**, 46601-46608
179. Maveyraud, L., Pratt, R. F., and Samama, J. P. (1998) *Biochemistry* **37**, 2622-2628
180. Swaren, P., Maveyraud, L., Raquet, X., Cabantous, S., Duez, C., Pedelacq, J. D., Mariotte-Boyer, S., Mourey, L., Labia, R., Nicolas-Chanoine, M. H., Nordmann, P., Frere, J. M., and Samama, J. P. (1998) *J. Biol. Chem.* **273**, 26714-26721
181. Mourey, L., Miyashita, K., Swaren, P., Bulychev, A., Samama, J., Mobashery, S. (1998) *J. Am. Chem. Soc.* **120**, 9382-9383
182. Fonze, E., Vanhove, M., Dive, G., Sauvage, E., Frere, J. M., and Charlier, P. (2002) *Biochemistry* **41**, 1877-1885
183. Jacob, F., Joris, B., Lepage, S., Dusart, J., and Frere, J. M. (1990) *Biochem. J.* **271**, 399-406
184. Vakulenko, S. B., Taibi-Tronche, P., Toth, M., Massova, I., Lerner, S. A., and Mobashery, S. (1999) *J. Biol. Chem.* **274**, 23052-23060
185. Hall, A., and Knowles, J. R. (1976) *Nature* **264**, 803-804
186. Yang, Y., Rasmussen, B. A., and Shlaes, D. M. (1999) *Pharmacol Ther* **83**, 141-151

187. Therrien, C., and Levesque, R. C. (2000) *FEMS Microbiol. Rev.* **24**, 251-262
188. Perez-Llarena, F., Martin, J. F., Galleni, M., Coque, J. J., Fuente, J. L., Frere, J. M., and Liras, P. (1997) *J. Bacteriol.* **179**, 6035-6040
189. Padayatti, P. S., Helfand, M. S., Totir, M. A., Carey, M. P., Hujer, A. M., Carey, P. R., Bonomo, R. A., and van den Akker, F. (2004) *Biochemistry* **43**, 843-848
190. Kuzin, A. P., Nukaga, M., Nukaga, Y., Hujer, A., Bonomo, R. A., and Knox, J. R. (2001) *Biochemistry* **40**, 1861-1866
191. Imtiaz, U., Billings, E., Knox, J. R., Manavathu, E. K., Lerner, S. A., Mobashery, S. (1993) *J. Am. Chem. Soc.* **115**, 4435-4442
192. Imtiaz, U., Manavathu, E. K., Lerner, S. A., and Mobashery, S. (1993) *Antimicrob. Agents Chemother.* **37**, 2438-2442
193. Parish, T., Smith, D. A., Roberts, G., Betts, J., and Stoker, N. G. (2003) *Microbiology* **149**, 1423-1435
194. Hotter, G. S., Wards, B. J., Mouat, P., Besra, G. S., Gomes, J., Singh, M., Bassett, S., Kawakami, P., Wheeler, P. R., de Lisle, G. W., and Collins, D. M. (2005) *J Bacteriol* **187**, 2267-2277
195. Joshi, S. M., Pandey, A. K., Capite, N., Fortune, S. M., Rubin, E. J., and Sasseti, C. M. (2006) *Proc Natl Acad Sci U S A*, **103**, 11760-11765
196. Li, J., Derewenda, U., Dauter, Z., Smith, S., and Derewenda, Z. S. (2000) *Nat. Struct. Biol.* **7**, 555-559
197. Finking, R., and Marahiel, M. A. (2004) *Annu. Rev. Microbiol.* **58**, 453-488
198. Katz, L., and Donadio, S. (1993) *Annu. Rev. Microbiol.* **47**, 875-912



199. Smith, S. (1994) *FASEB. J.* **8**, 1248-1259
200. Hunt, M. C., and Alexson, S. E. (2002) *Prog. Lipid Res.* **41**, 99-130
201. Schwarzer, D., Mootz, H. D., Linne, U., and Marahiel, M. A. (2002) *Proc. Natl. Acad. Sci. U S A* **99**, 14083-14088
202. Thoden, J. B., Holden, H. M., Zhuang, Z., and Dunaway-Mariano, D. (2002) *J. Biol. Chem.* **277**, 27468-27476
203. Naggert, J., Narasimhan, M. L., DeVeaux, L., Cho, H., Randhawa, Z. I., Cronan, J. E., Jr., Green, B. N., and Smith, S. (1991) *J. Biol. Chem.* **266**, 11044-11050
204. Thoden, J. B., Zhuang, Z., Dunaway-Mariano, D., and Holden, H. M. (2003) *J. Biol. Chem.* **278**, 43709-43716
205. Hunt, M. C., Solaas, K., Kase, B. F., and Alexson, S. E. (2002) *J. Biol. Chem.* **277**, 1128-1138
206. Ioerger, T. R., and Sacchettini, J. C. (2002) *Acta Crystallogr. D Biol. Crystallogr.* **58**, 2043-2054
207. Kimber, M. S., Martin, F., Lu, Y., Houston, S., Vedadi, M., Dharamsi, A., Fiebig, K. M., Schmid, M., and Rock, C. O. (2004) *J. Biol. Chem.* **279**, 52593-52602
208. Sharma, S. K., Kapoor, M., Ramya, T. N., Kumar, S., Kumar, G., Modak, R., Sharma, S., Surolia, N., and Surolia, A. (2003) *J Biol Chem* **278**, 45661-45671
209. Lindner, A. B., Kim, S. H., Schindler, D. G., Eshhar, Z., and Tawfik, D. S. (2002) *J. Mol. Biol.* **320**, 559-572
210. Dillon, S. C., and Bateman, A. (2004) *BMC Bioinformatics* **5**, 109

211. Cleland, W. W., and Kreevoy, M. M. (1994) *Science* **264**, 1887-1890
212. Gigant, B., Charbonnier, J. B., Eshhar, Z., Green, B. S., and Knossow, M. (1997) *Proc. Natl. Acad. Sci. U S A* **94**, 7857-7861
213. Quadri, L. E., Sello, J., Keating, T. A., Weinreb, P. H., and Walsh, C. T. (1998) *Chem. Biol.* **5**, 631-645
214. Yeh, E., Kohli, R. M., Bruner, S. D., and Walsh, C. T. (2004) *Chembiochem* **5**, 1290-1293
215. Lederer, E. (1964) *Angew Chem Int Ed Engl* **3**, 393-452

## VITA

- NAME:** Feng Wang
- ADDRESS:** Department of Biochemistry and Biophysics  
Texas A&M University  
College Station, TX 77843
- EDUCATION:** Texas A&M University, College Station, TX  
Ph.D, Biochemistry, 2007
- Northwestern University, Evanston, IL  
M.S., Chemistry, 2000
- University of Science & Technology of China,  
Hefei, China  
B.S., Chemistry, 1998
- MEMBERSHIP:** Sigma Xi: The Scientific Research Honor Society