

MODE OF ACTION STUDY OF *PARA*-AMINOSALICYLIC ACID AND  
STRUCTURE, FUNCTION AND INHIBITOR STUDY OF THE ISOCITRATE  
DEHYDROGENASE-2 IN *MYCOBACTERIUM TUBERCULOSIS*

A Dissertation

by

YU-SHAN CHENG

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

DOCTOR OF PHILOSOPHY

Chair of Committee,	James C. Sacchetti
Committee Members,	Tadhg P. Begley
	David P. Barondeau
	Paul D. Straight
Head of Department,	Simon W. North

August 2016

Major Subject: Chemistry

Copyright 2016 Yu-Shan Cheng

## ABSTRACT

Tuberculosis (TB) killed 1.5 million people and rivaled AIDS, becoming the leading cause of death from infectious disease in 2014. The prevalence of multidrug resistant TB has intensified the current therapeutic procedure, making it urgent to find novel anti-tubercular agents and to come up with solutions to retard the emergence of the drug resistance. This dissertation focuses on the identification of drug targets, the exploration of drug resistance mechanisms, and the identification of novel inhibitors.

In the first part, the mechanism of action of the classic anti-tubercular drug, *para*-aminosalicylic acid (PAS), was explored through genetic, cell viability and molecular modeling studies. Dihydrofolate reductase (DHFR) was identified to be the putative intracellular target of PAS. In addition, the molecular mechanism of PAS resistance was intensively investigated for the clinically relevant *Rv2671* up-regulation mutant. Biochemical assays showed that *Rv2671* exhibited a low DHFR activity with a high  $K_m$  for the substrate, 7,8-dihydrofolate. X-ray crystal structure of the *Rv2671* in complex with  $\text{NADP}^+$  and tetrahydrofolate (THF) further confirmed the structural similarity between *Rv2671* and DHFR. These studies together suggested that PAS resistance of this mutant is derived from the ability to complement the DHFR activity with the high level of *Rv2671*.

The second part of this dissertation details the characteristics of *Mycobacterium tuberculosis* isocitrate dehydrogenase-2 (*Mtb* IDH2). The kinetic study of *Mtb* IDH2

suggested that it catalyzes an ordered sequential reaction by binding NADP<sup>+</sup> first. X-ray crystal structure revealed the fairly conserved active site and dissimilar overall structure compared to human IDHs (HIDHs), suggesting a potential for drug selectivity. A screening of known inhibitors of mutant HIDHs and a high-throughput screening of *Mtb* whole cell active compounds were further implemented to identify inhibitors for *Mtb* IDH2. Two compounds from the screenings exhibited IC<sub>50</sub>s below 10 μM. The enzyme structure and the modest potency inhibitors of *Mtb* IDH2 can serve as viable starting points for the follow-up inhibitor development of *Mtb* IDH2.

## DEDICATION

To my family

## ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Sacchetti, for giving me the opportunity to conduct this research, and my committee members, Dr. Barondeau, Dr. Begley, and Dr. Straight for their guidance and support throughout the course of this research.

My thanks also go to the members of the Structural Biology Center for their support for the X-ray crystallographic data collection and Dr. Thomas R. Ioerger, who helps to analyze the sequence of resistant mutants.

I thank my colleagues and the departmental staff and faculty for their support, feedback, and friendship.

Last but not the least, I would like to thank my friends and family for their support and encouragement. This accomplishment would not have been possible without them.

## TABLE OF CONTENTS

	Page
ABSTRACT .....	ii
DEDICATION .....	iv
ACKNOWLEDGEMENTS .....	v
TABLE OF CONTENTS .....	vi
LIST OF FIGURES.....	viii
LIST OF TABLES .....	x
CHAPTER I INTRODUCTION AND LITERATURE REVIEW .....	1
Tuberculosis Is a Public Health Issue .....	1
Persistence, Latency, and Dormancy .....	2
Persistent and Dormant <i>Mtb</i> .....	9
Strategies to Identify Novel Chemotherapeutic Agents for Multidrug Resistant <i>Mtb</i> and Persistent <i>Mtb</i> .....	21
Enzymes Involved in Central Carbon Metabolism (CCM) Are Promising Targets to Combat Persistent or Dormant <i>Mtb</i> .....	24
CHAPTER II MECHANISM ACTION STUDY OF A SECOND-LINE ANTI- TUBERCULAR DRUG <i>PARA</i> -AMINOSALICYLIC ACID (PAS).....	30
Introduction .....	30
Results and Discussion.....	37
Materials and Methods .....	53
CHAPTER III STRUCTURAL INSIGHTS INTO <i>MYCOBACTERIUM</i> <i>TUBERCULOSIS</i> RV2671 PROTEIN AS A DIHYDROFOLATE REDUCTASE FUNCTIONAL ANALOGUE CONTRIBUTING TO <i>PARA</i> -AMINOSALICYLIC ACID RESISTANCE .....	57
Introduction .....	57
Results and Discussion.....	60
Materials and Methods .....	83
CHAPTER IV KINETICS, STRUCTURE AND INHIBITORS OF <i>MTB</i> IDH2 .....	89

Introduction .....	89
Results and Discussion.....	95
Conclusion.....	122
Materials and Methods .....	124
CHAPTER V CONCLUSIONS AND FUTURE DIRECTIONS.....	132
REFERENCES .....	136
APPENDIX A SUPPORTING INFORMATION FOR CHAPTER III.....	168
Materials and Methods .....	168
Tables and Figures .....	170
References for Appendix Materials.....	174

## LIST OF FIGURES

	Page
Figure 1. Killing curve during standard anti-tubercular chemotherapy. ....	5
Figure 2. Regulation system of the stringent response model in <i>Mycobacterium</i> .....	17
Figure 3. Dos regulation in <i>Mtb</i> .....	19
Figure 4. Chemical structures of PAS and its structural related metabolites, <i>para</i> -aminobenzoic acid and salicylic acid .....	31
Figure 5. <i>De novo</i> folate metabolism and the sites at which antimetabolites act.....	33
Figure 6. Chemical structures of siderophores produced by <i>Mtb</i> .....	36
Figure 7. Iron acquisition system of <i>Mtb</i> .....	36
Figure 8. Thymidylate synthase reaction .....	42
Figure 9. Whole cell assay of PAS to <i>Mtb</i> mc <sup>2</sup> 7000 strain with a DHFR overexpression vector .....	44
Figure 10. Chemical structure of DHF and MTX with numbering.....	47
Figure 11. Two possible orientations of the hydroxyl groups of the hydroxy-DHFs as docking in the MTX binding site of <i>Mtb</i> DHFR .....	49
Figure 12. Detailed interactions between the top ranked poses and <i>Mtb</i> DHFR .....	50
Figure 13. Multiple sequence alignment of a collection of DHFR from different species of <i>Mycobacterium</i> .....	51
Figure 14. Characterization of Rv2671 as a DHFR .....	61
Figure 15. Crystal structure of Rv2671 and the THF binding mode.....	67
Figure 16. Crystal structure of Rv2671 in complex with TMP or TMQ .....	75
Figure 17. Residue differences between the substrate binding sites of Rv2671 and <i>Bs</i> RibG.....	80
Figure 18. Phylogenetic tree of pyrimidine reductases and DHFRs.....	82
Figure 19. Proposed catalytic mechanism of IDH .....	91



Figure 20. Double reciprocal plots of the initial velocity kinetics for <i>Mtb</i> IDH2.....	96
Figure 21. Overall structure of NADP <sup>+</sup> bound <i>Mtb</i> IDH2 and the pseudo two-fold axis in the large domain.....	100
Figure 22. Structural comparisons of small and large domains .....	101
Figure 23. Interdomain motions of monomeric IDHs.....	102
Figure 24. Detailed interactions between NADP <sup>+</sup> and <i>Mtb</i> IDH2 .....	104
Figure 25. Active site comparison between <i>Mtb</i> IDH2 and <i>Ec</i> IDH illustrate the potential catalytic residues of <i>Mtb</i> IDH2.....	105
Figure 26. Models of the proposed intradomain closure mechanism .....	106
Figure 27. Interface interaction between chain A and chain D in the NADP <sup>+</sup> bound <i>Mtb</i> IDH2 structure.....	108
Figure 28. Size exclusion chromatography of wild-type and mutants of <i>Mtb</i> IDH2s ...	110
Figure 29. Schematic representations of the NADP <sup>+</sup> binding residues of <i>Mtb</i> IDH and HIDH1/HID2 .....	112
Figure 30. Inhibitory dose-response curve to determine the IC <sub>50</sub> of SYC-3 for <i>Mtb</i> IDH2 .....	114
Figure 31. Detailed information of the hits from the HTS assay .....	117

## LIST OF TABLES

	Page
Table 1. Function and mechanism of the identified TA systems in <i>Mtb</i> .....	14
Table 2. Enzymes involved in the respiratory chains of <i>Mtb</i> .....	21
Table 3. Inhibition profile of PAS, sulfones and sulfonamides against the activity of Mtb DHPS and the growth of <i>Mtb</i> .....	34
Table 4. The sequencing results of the PAS resistant strains.....	40
Table 5. Ligand conformations derived from the Glide docking and their corresponding strain calculated by ICM-browser pro. ....	48
Table 6. Whole cell activity of PAS against different species of <i>Mycobacterium</i> and the residue of their DHFR corresponding to Gln28 in <i>Mtb</i> DHFR. ....	53
Table 7. Kinetic parameters of Rv2671 and DHFRs .....	62
Table 8. Data collection and refinement statistics.....	64
Table 9. Product inhibition patterns and kinetic parameters for <i>Mtb</i> IDH2.....	97
Table 10. Crystallographic data collection and refinement statistics of the NADP <sup>+</sup> bound <i>Mtb</i> IDH2 structure.....	99
Table 11. Oligomeric states of <i>Mtb</i> IDH2 in different buffer system.....	107
Table 12. Inhibitory activity of HIDHs' inhibitors against <i>Mtb</i> IDH2 .....	113
Table 13. Summary of the HTS results.....	116
Table 14. Structures and inhibitory activity of sub-group1 .....	119
Table 15. Structures and inhibitory activity of sub-group2 .....	120
Table 16. Structures and inhibitory activity of sub-group3 .....	121
Table 17. Structures and inhibitory activity of sub-group4 .....	122

CHAPTER I  
INTRODUCTION AND LITERATURE REVIEW

**Tuberculosis Is a Public Health Issue**

Tuberculosis (TB) is an ancient disease and was one of the top 10 leading causes of death worldwide in 2012. The emergence of HIV and the prevalence of the drug resistant strains make TB a public health issue. In 2014, 1.5 million people died because of TB and of this group, 27% were human immunodeficiency virus (HIV) positive.<sup>1</sup> In addition, 480,000 people developed multidrug-resistant TB (MDR-TB) and only 1 out of 2 MDR-TB patients can be cured.<sup>1</sup> Therefore, there is an urgent need to develop new effective anti-tubercular agents to treat the drug-resistant TB.

The occurrence of drug resistant *Mycobacterium tuberculosis* (*Mtb*) is not a new phenomenon. Indeed, *Mtb* showed a high drug resistant mutant emergence rate when patients were treated with a single anti-tubercular drug. Drug resistant *Mtb* was observed quickly following the discovery of streptomycin, one of the oldest anti-tubercular drugs. In 1942, over 80% of TB patients showed resistance in the first TB clinical trial with streptomycin.<sup>2, 3</sup> A combination of streptomycin and *para*-aminosalicylic acid effectively overcame the high resistant mutant frequency. Combination therapy then became a standard for TB treatment.

The high drug resistant characteristic of *Mtb* is associated with the high bacterial density in pulmonary cavities. The bacterial population ( $\sim 10^8$ – $10^9$ ) in the cavities of infected lungs<sup>4</sup> is higher than the typically observed resistant mutant frequency of a drug ( $\sim 10^{-6}$ – $10^{-9}$ ).<sup>5</sup> An increasing amount of literature discussed drug resistant TB after the outbreak of the MDR-TB at US hospitals in the early 1990s.<sup>6</sup> MDR-TB is resistant to two powerful anti-tubercular drugs, isoniazid and rifampicin. In addition to MDR-TB, extensively drug-resistant TB (XDR-TB) was reported.<sup>7</sup> Not only resistant to isoniazid and rifampicin, XDR-TB is also resistant to any of the fluoroquinolones and to at least one of the three injectable drugs, amikacin, capreomycin, or kanamycin. It is estimated that 9.7% of the MDR-TB cases are XDR-TB.<sup>1</sup> Moreover, cases of totally drug-resistant TB (TDR-TB) have been reported all over the world.<sup>8</sup> The first TDR-TB cases appeared in Italy in 2003.<sup>9</sup> Recently, TDR-TB was reported in Iran, India, and South Africa.<sup>10-12</sup> The emergence of TDR-TB reinforces the urgent need for new anti-tubercular drugs.

## **Persistence, Latency, and Dormancy**

### *Persistence*

Standardized TB treatment, which is suggested by WHO,<sup>13</sup> takes 6 months. It includes a combination of isoniazid, rifampicin, pyrazinamide, and ethambutol for the first 2 months, followed by a combination of isoniazid and rifampicin for the final 4 months. Most active *Mtb* are killed during the first two months, and the residual bacteria are in a persistent state (Figure 1). Persistence is a term associated with a state in which a

bacterium shows high tolerance to a sufficient amount of antibiotic that is able to kill a major bacterial population. Persistent *Mtb* shows a phenotypic resistance to most of the anti-tubercular agents except rifampicin. Therefore, isoniazid and rifampicin are implemented in the last 4 months treatment to kill the trace amount of active *Mtb* and the persistent *Mtb*, respectively. Long, multidrug therapy is often complicated by incomplete treatments for patients, which subsequently causes their strain of *Mtb* to become drug resistant.

The ability of *Mtb* to persist in the host is thought to be responsible for the need of long treatment duration; hence, identifying a drug for persistent *Mtb* is another important mission for TB control. In 1994, Bigger used the term persisters to describe a subgroup of *Staphylococcus* spp. that was phenotypically insensitive to a high dose of penicillin.<sup>14</sup> He believed that the high tolerance was because these *Staphylococcus* spp. were in a dormant and non-replicating state.

Persisters were reported in almost all organisms, and recently they were observed in cancer cells.<sup>15</sup> The population of persisters is influenced by several factors such as types of antibiotics, age of bacterial culture, and environmental stresses. After a decade of study, persisters now are generally divided into two types: continuously generated persisters (Type I persisters) and stress induced persisters (Type II persisters). In general, bacteria evolve two ways to adapt the environmental change: generating different phenotypes by randomly fluctuating the gene expression or possessing a constitutively expressed sensor. If the change is a result of an unpredictable perturbation, the former is preferred to save the energy consumption. This stochastically phenotypic switch within

an exponentially growing culture leads to the formation of Type I persisters and can be considered as a “bet-hedge: strategy of bacteria.<sup>16</sup> Type II persisters, however, are induced by environmental stresses, such as starvation, heat shock, and pH change, with the assistance of the sensors. For example, the levels of persisters are significantly higher when a culture is in a stationary phase.<sup>17</sup> Bacteria usually face a challenge of low nutrient supply in this stage that was demonstrated to induce the transition from a normal growing state into a slow or non-growing state. Most antibiotics are not effective against these slow or non-growing bacteria since their molecular targets are no longer required for bacteria. The actual mechanism which triggers the formation of Type II persisters is still unclear, and may vary in different organisms or the environmental stresses. Nevertheless, mechanisms, such as toxin-antitoxin (TA) modules and stringent response, are found to participate in the regulation of Type II persisters.

### *Latency*

Latency or latent infection with *Mtb* (LTBI) is another important topic for TB especially with the current prevalence of HIV. LTBI is a clinical phenomenon of infection in which patients show positive reaction to the tuberculin skin test or the IFN- $\gamma$  release assay but without any symptoms. WHO estimated that one-third of the world’s populations carried LTBI.<sup>19</sup> It is unrealistic to treat patients with LTBI as they stay healthy, and LTBI is unable to transmit the disease to other people.

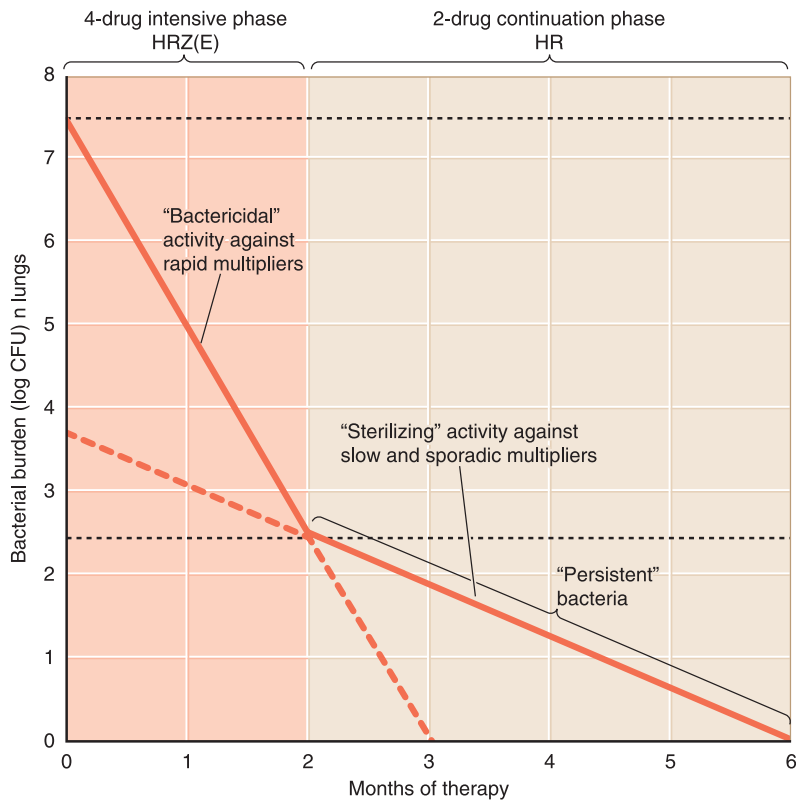


Figure 1. Killing curve during standard anti-tubercular chemotherapy.<sup>18</sup> The x-axis represents the time of the therapy and the y-axis represents the total counts of bacteria in lungs. The amount of bacteria drops rapidly at the first two month while treated with rifampin (R), pyrazinamide (Z), isoniazid (H), and ethambutol (E). Reprinted with permission from Jain, S. K., Lamichhane, G., Nimmagadda, S., Pomper, M. G., and Bishai, W. R. (2008) Antibiotic treatment of tuberculosis: old problems, new solutions, *Microbe* 3, 285. Copyright 2008 by American Society for Microbiology.

However, LTBI carriers with a higher chance of progression to active disease such as HIV infected patients or those obtaining a treatment against tumor necrosis factor (TNF) are recommended to receive a treatment to decrease the reactivation risk. Five types of treatment for LTBI are suggested by WHO: (i) 6 months isoniazid daily, (ii) 9 months isoniazid daily, (iii) 3 months rifapentine and isoniazid combination

weekly, (iv) 3–4 months isoniazid plus rifampicin daily, or (v) 3- or 4-month rifampicin alone daily.<sup>20</sup>

### *Dormancy*

*Mtb* in dormancy is thought to be responsible for the LTBI. Before moving into a review of dormancy, it is necessary to point out that there is a degree of uncertainty around the use of the terms dormancy and persistence in scientific literature. Often times, persistence is misused to describe dormancy. In this dissertation, dormancy will be used solely when referring to viable but non-culturable (VBNC) cells.<sup>21</sup> However, there is no standard definition for non-culturable, such as what kinds of media or stimuli should be tried or how long incubation should occur before concluding a cell is non-culturable. Therefore, a term “non-growing but metabolically active (NGMA)” is introduced to avoid the confusion about culturability.<sup>22</sup> The NGMA state is a VBNC-like state in which a cell is slow or non-replicating, but with detectable metabolic activities.

Dormancy is a mechanism for organisms to adapt to environmental stresses. Organisms stop growth and division during dormancy and are in a minimal metabolic state to reduce energy consumption. Dormancy is observed in various organisms such as hibernation in animals, or bud and seed dormancy in plants. Dormancy in bacteria was usually linked to a special morphology (e.g. spores or cysts). However, this concept changed when the dormancy of non-sporulating bacteria was discovered in 1982.<sup>23</sup> According to Kaprelyants *et al.* (1993)<sup>24</sup>, dormancy in non-sporulating bacteria such as



mycobacteria can be defined as a reversible state of low metabolic activity and non-replication. In addition, a dormant bacterium is unable to form a colony in a routine solid media but can resume division after a proper resuscitation from dormancy. Within this publication, dormant bacteria were defined as viable but non-culturable (VBNC) bacteria.

The existence of VBNC state was first observed on *E. coli* and *V. cholera* by Xu *et al.*<sup>23</sup> and was reported on over 85 different bacteria.<sup>25</sup> Before the observation of VBNC state, the viability of a cell equaled the ability of cell proliferation in the classical microbiology. Culturing the cells on a defined solid plate was a widely accepted method to examine the capability of cell proliferation and subsequently to estimate the cell viability. Therefore, new detection methods are required to evaluate the cell viability of the VBNC cells. The *BacLight* Live/Dead assay and RT-PCR are commonly used to quantify the total viable cells. The *BacLight* Live/Dead assay uses red fluorescing propidium iodide (PI) and green fluorescing SYTO9 to distinguish the viable and dead cells. SYTO9 can enter all cells but PI only penetrates cells with damaged cytoplasmic membrane.<sup>26</sup> A combination of the *BacLight* Live/Dead assay with flow cytometry provides an even better way to quantify the result.<sup>27</sup> However, this method is not perfect since this assay is based on the assumption that cells with damaged membrane are dead cells. Alternatively, cell viability may be determined by the metabolic activity in an RT-PCR assay. A short half-life mRNA (~ a few minutes<sup>28, 29</sup>) is measured and correlated with metabolic activity. Detectable mRNA indicates living or the recently dead cells at

the sampling time.<sup>30</sup> The results can be misleading, as the measurement is an ensemble average that may not represent the condition of a single cell or a minor subpopulation.

Another key aspect of dormancy is resuscitation. The success of resuscitation depends on organisms, dormancy inducing conditions, and the duration within the dormant state. To date, only half of the VBNC genera have been successfully resuscitated and returned to an active state. The failure of resuscitation suggested either this group of bacteria could not resume division or the currently tested methods were unable to resuscitate these bacteria. Different environmental stresses have been used to induce bacteria into a VBNC state such as starvation, unfavorable growth temperature or pH.<sup>31-33</sup> Thus, removal of the environmental stresses was often the simplest resuscitation process.<sup>34, 35</sup> However, the longer the bacterial cells are in the VBNC state, the lower population of the VBNC cells can be resuscitated.<sup>36, 37</sup> Other than changing the cultural conditions, the cell proliferation can be reactivated by compounds including amino acids<sup>38</sup>, resuscitation-promoting factors (Rpfs<sup>39</sup>) and autoinducers.<sup>40</sup>

A much debated issue is whether persistence equals dormancy. Some research groups believe they are the same, especially for Type II persisters, as they share several commonalities such as in an energy generating state and with a high tolerance to environmental stresses and antibiotics. The only confirmed difference is the duration for resuscitation from these two states. Persisters can resume replication in a short period of time after the antibiotics or the environmental stresses are removed, but the resuscitation time varied for different VBNC cells. Experimental results from early studies were insufficient to draw a conclusion whether they are the same. However, some persisters

were demonstrated to conduct translation to some degrees in a single cell flow cytometry experiment. In addition, persisters with normal replicating ability were observed in recent studies. Inability to generate a toxic radical, or the induction of the expression of efflux pumps was found to be associated with these bacterial persisters.

Taken together, Type II persisters can be predecessors of dormant cells. They are in a transition state, on the way to dormancy. Moreover, it may be necessary to introduce another category, Type III persisters, for the persisters with a regular metabolic activity and replicating state or to expand the definition of Type I persisters to refer not only to the low metabolic bacteria but all heterogenetic bacteria, which have phenotypic tolerance to antibiotics.

### **Persistent and Dormant *Mtb***

Persistence and dormancy make the TB treatment more difficult. In fact, it is hard to separate persistence and dormancy in the study of *Mtb* since the multiple environmental stresses in its life cycle within the host often prompt the formation of Type II persisters. Type II persisters, i.e., the stress induced persisters, are indeed highly correlated with dormancy. Therefore, understanding the environmental stresses *Mtb* encounters *in vivo* is required to study *Mtb* persistence or dormancy. A brief summary of the *Mtb*'s life cycle helps to rationalize the study of the models for the Type II persisters and dormant *Mtb*.

*Mtb* exists with different physiological states in response to different infection stages. The infection cycle starts from the inhalation of aerosolized *Mtb*. The aerosolized *Mtb* is taken up through phagocytosis and resides in the phagosomes of the lung macrophage. Around 5–10% of the infected individuals received will progressive to active TB, while others lapse into a dormant state resulting in a latent infection.

In the primary infection, *Mtb* arrests the phagosomes containing bacteria in an immature stage that prevents the phagosomes from fusing with a lysosome. However, some phagosomes overcome the *Mtb*'s survival strategy by the stimulation of the host immune system. More specifically, the infected phagosome becomes a phagolysosome by the stimulation of IFN- $\gamma$ .<sup>41</sup> The phagolysosome then kills some bacteria. In mature macrophages, *Mtb* needs to tolerate the low pH and oxidative stress such as NO and CO.

Some activated macrophages induce the formation of a granuloma, the hallmark of *Mtb*. For a healthy individual, TB infection is usually controlled at this stage. The granuloma is an organized structure composed of T-cells, B-cells, dendritic cells, and a fibrous capsule. A caseous granuloma is the most common type in which cell debris accumulates and forms a necrotic area in the center. In addition to caseous granuloma, various other types of granulomas are observed in humans, including non-necrotic, suppurative, fibrotic, and mineralized. In active TB, all types of granulomas coexist. However, caseous, mineralized, or fibrotic granulomas are predominant in latent TB.<sup>42</sup> Cavity formation is a key event in severe TB. Some groups believe that the cavity is produced by the expansion of the caseous granuloma. The caseous center will be softened and becomes liquefied. Eventually, a cavity forms along with overloaded

extracellular bacterial growth. However, other groups believe that the caseous granuloma plays no role in the formation of cavities. They suggest that the cavity is produced through the localized caseous pneumonia in the progressive primary TB. Massive caseation occurs, which blocks the bronchus, and subsequently results in the formation of cavities.<sup>43</sup> Nevertheless, *Mtb* grows on the surface of the cavities. When the cavities connect with the airways, *Mtb* grows exponentially. In the end of the infection, *Mtb* is released from airways and causes the infection to move to other hosts.

The granuloma is thought to be a product of coevolution in which *Mtb* and the host immunologic system reach a balance. In the granuloma, some induced or pre-existing NGMA *Mtb* survive and cause the LTBI. The center of the granuloma is thought to be low in oxygen and nutrient limited due to a dysfunctional vasculature.<sup>44</sup> In summary, *Mtb* endures starvation and low pH within macrophages, and it suffers from low oxygen tension and nutrient limitation in the granuloma.

Several *in vitro* and *in vivo* models have been designed to study Type II persisters and dormant *Mtb*. Notably, even though the original intention of these models is to study dormant *Mtb*, NGMA *Mtb* is a better term to describe the *Mtb* in these models, not only because the definition of culturable is vague, but also because of the difficulty of examining the real-time viability of a cell. The models can be mainly divided into four groups: hypoxia, starvation, macrophage, and murine infection. These models are useful to understand the NGMA *Mtb*, but none of them can fully represent the state of human infection.

The hypoxia model,<sup>45</sup> a gradual O<sub>2</sub> depletion model, mimics the low oxygen situation in a granuloma but is missing the stresses generated from immune response and macrophage phagocytosis. In addition to the same limitations as the hypoxia model, the oxygen constraint does not occur in the starvation model. The macrophage model captures the situation for initial infection but not for the late infection. Even though the infection and reactivation are established, the mouse model suffers from a high bacterial load at the chronic infection and lacks a well formed granuloma, especially lacking the caseous and necrotic granulomas.

Nevertheless, studies carried out with these models reveal the physiology of NGMA *Mtb*. Transcriptomic and proteomic studies were performed to understand the molecular mechanism and metabolic realignment of NGMA *Mtb* in these models. The exact molecular mechanisms of dormancy and antibiotic persistence remained unclear, but several genes have encoded critical regulators for these NGMA bacteria. Three molecular mechanisms associated with these genes are proposed to explain the formation of the mycobacterial persistence and dormancy: the Toxin-Antitoxin (TA) model, the stringent response, and the Dos regulation.

The TA model is the most studied molecular mechanism for bacterial persistence. It is a two-component system with a toxin and an antitoxin. Toxin arrests the cell growth by interfering with either transcription or translation. Its unstable antitoxin counterpart neutralizes the activity of the toxin. The TA model was first introduced when the observation of a mutant in gene *hipA* resulted in a significant increase of persistent *E. coli*.<sup>46</sup> HipA was later proved to be a toxin with an antitoxin

counterpart, HipB. In fact, TA models can be a plasmid maintenance system or a chromosomally-encoded system, and they are highly redundant in cells. Knocking-out an individual pair does not change the subpopulation of persisters. However, overexpression of toxin usually slows down the bacterial growth. TA systems can be classified on the basis of the nature of antitoxin into five groups. In the type I TA family, the antitoxin is an antisense RNA, which regulates the toxin synthesis by binding to the toxin mRNA. In the type II TA family, the antitoxin is a toxin binding protein, inhibiting the toxin by forming a complex with it. In the type III TA family, the antitoxin is an RNA, which inhibits the toxin by forming a protein-RNA complex. In the type IV TA family, the antitoxin is protein, which inhibits the toxin by competing for the same target. In the type V TA family, the antitoxin is an endoribonuclease, which specifically cleaves the mRNA of the toxin.<sup>47-50</sup> Based on a sequence similarity search, around 80–90 TA loci were found in *Mtb*. The results of a study conducted by Sala *et al.*<sup>51</sup> with a brief introduction of each TA system are listed in Table 1.

In their report, 79 loci were assigned as TA systems including 67 from the type II TA family, 1 tripartite type II TAC system, 3 putative type IV systems, and 8 uncharacterized TA pairs. 30 showed an intracellular function in an ectopic expression assay within *M. smegmatis*. Evidence suggests that some TA pairs contributed to the antibiotic and environmental persistence in *Mtb*. The expression level of different subsets of TA loci is altered in the presence of environmental stresses such as starvation, hypoxia, and macrophage infection.<sup>52-54</sup> In addition, 10 TA pairs are up-regulated in D-cycloserine induced persisters.<sup>55</sup>

Table 1. Function and mechanism of the identified TA systems in *Mtb*<sup>50</sup>

TA family (locus)	Number of loci	Toxin	Target of toxin	Inhibitory mechanism of the toxin	Antitoxin
VapBC	50	VapC	Translation	Cleavage of RNA molecules <sup>52</sup> or RNA binding <sup>56</sup>	VapB
MazEF	10	MazF	Translation	Cleavage of a sequence-specific mRNA, <sup>57-59</sup> cleavage 23S rRNA, <sup>58</sup> or interacting with DNA topoisomerase I <sup>60</sup>	MazE
RelBE	2	RelE	Translation elongation	Binding with 16S rRNA, interacting with the 30S subunit and resulting in the cleavage of mRNA at ribosomal A site <sup>61</sup>	RelB
YefM/YoeB	1	YoeB	Translation initiation	Interaction with the 50S subunit and preventing the formation of the initiation complex on mRNAs <sup>62</sup>	YefM
HigBA	2	HigB	Translation	Binding with the ribosomal 50S subunit and cleavage of mRNA at AAA sites <sup>63</sup>	HigA
ParDE	2	ParE	Cell division	Targets DNA gyrase <sup>64</sup>	ParD
TAC system*	1	HigB1	Translation	Binding with the ribosomal 50S subunit and cleavage of mRNA at AAA sites <sup>63</sup>	HigA1
Others	11				
* In TAC system, the antitoxin is in collaboration with a SecB type chaperone					



The stringent response model is described as a series of global physiological changes induced by the accumulated alarmone, (p)ppGpp. It occurs when a bacterium signals a nutrition deprivation, especially the amino acid depletion. The intracellular concentration of (p)ppGpp is controlled by a ATP: GTP/GDP/ITP 3'-pyrophosphoryltransferase and a (p)ppGpp 3'-pyrophosphohydrolase. In *Mycobacterium*, a bifunctional enzyme, RelA, carries out both reactions.<sup>65</sup> The regulatory system of RelA in *Mycobacterium* has been characterized in *M. smegmatis*<sup>66</sup> (Figure 2). The transcription of *relA* is mediated by a positive feedback system composed of a sigma factor ( $\sigma^E$ ), two stress response enzymes (MprAB) and a signaling molecule, polyphosphate. The stress-responsive sensor, histidine kinase (MprB), autophosphorylates using the polyphosphate as a phosphate donor and then transfers the phosphate group to the transcriptional response regulator (MprA). The phosphorylated MprA conducts tighter binding and causes up-regulations to its targets, promoters of MprA and  $\sigma^E$  in this case. The transcription of *rel* gene is induced by the up-regulated  $\sigma^E$ . The amount of polyphosphate is controlled by polyphosphate kinase (PPK)<sup>67</sup> and exopolyphosphatase (PPX)<sup>68-70</sup> that mediate the synthesis and hydrolysis of polyphosphate, respectively. Notably, the function of PPXs is inhibited by (p)ppGpp,<sup>68, 71</sup> therefore, generating another positive feedback loop. In summary, high levels of polyphosphate result in high levels of RelA. As mentioned above, RelA in *Mycobacterium* carries out two reactions. The preference of the reactions is determined by the ratio of charged and uncharged tRNA.<sup>72</sup> The (p)ppGpp synthesis activity of RelA increases in the presence of uncharged tRNA.<sup>73</sup> As a result, (p)ppGpp

accumulates and subsequently increases the intracellular concentrations of polyphosphate by inhibiting the PPXs in starvation.

In respect to the (p)ppGpp accumulation, rRNA and ribosomal protein syntheses are down-regulated, and some amino acid biosynthetic enzymes are up-regulated.<sup>74, 75</sup> The decline of rRNA and ribosome causes cell growth arrest. In *E. coli*, the transcriptional realignment is triggered by a direct interaction between (p)ppGpp and RNA polymerase in cooperation with a transcriptional factor DskA. In *B. subtilis*, the changes result from the (p)ppGpp mediated intracellular GTP depletion through inhibiting GTP biosynthesis.<sup>76</sup> Neither mechanism has been validated in *Mtb*. However, the interference of RNA polymerase activity by (p)ppGpp was observed in several studies.<sup>77</sup>

Even though some molecular mechanisms remain unclear, the stringent response model may be critical for *Mtb* survival *in vivo*. Growth attenuation was observed when RelA knocked-out *Mtb* was cultured in the long-term nutrient or oxygen depletion.<sup>78</sup> In addition, *Mtb* loses its ability to persist in mice and guinea pig when RelA is deleted.<sup>54, 79</sup>

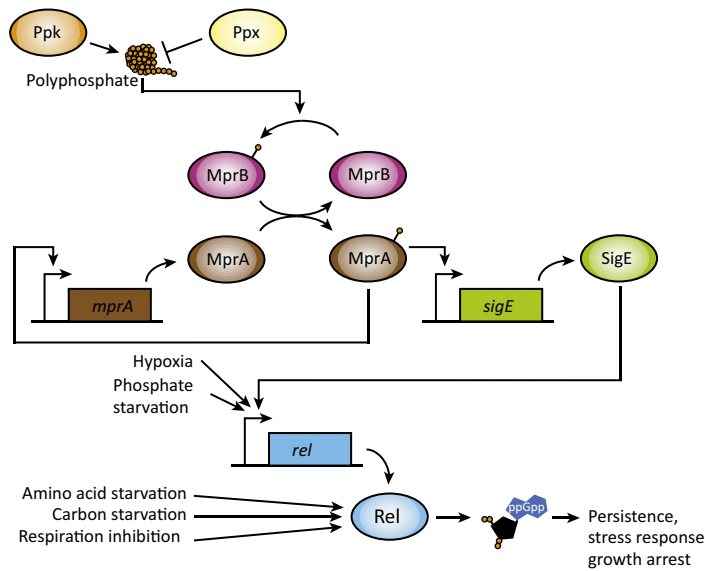


Figure 2. Regulation system of the stringent response model in *Mycobacterium*.<sup>66</sup> The core of the stringent response is the synthesis of (p)ppGpp by Rel. In *Mycobacterium*, the expression of Rel depends on the intracellular levels of polyphosphate. High levels of polyphosphate result in the increase concentrations of MprA and SigE and subsequently increase the levels of Rel and (p)ppGpp. Reprinted with permission from Boutte, C., Crosson, S. (2013) Bacterial lifestyle shapes stringent response activation, *Trend in Microbiology* 21, 174-180. Copyright 2013 by Elsevier.

The Dos regulation model is another persistence related mechanism for *Mtb*. A different expression level between the virulent H37Rv and the avirulent H37Ra strains permitted the Dos regulation system to be first observed.<sup>80, 81</sup> Later, the Dos regulon was shown to facilitate the adaption for the hypoxia and the presence of oxidative species such as carbon monoxide and nitric oxide.<sup>82-84</sup> Dos regulation is a two-component regulatory system, composed of a regulation protein (DosR) and two sensor kinases (DosS and DosT) (Figure 3).<sup>85</sup> DosS and DosT activate DosR by conducting a phosphorylation reaction. DosS is a redox sensor and DosT is a hypoxia sensor. After receiving the signals, DosS and DosT first rapidly autophosphorylate at a histidine

residue and catalyze the phosphotransfer to an aspartate residue of DosR.<sup>86, 87</sup> The activated DosR influences the expression of over 50 genes. As a result, genes for triacylglycerol biosynthesis and electron acceptors for the anaerobic respiration, such as the nitrate/fumarate and succinate systems, are up-regulated. Alternatively, genes for the basal metabolisms are down-regulated.

DosR was found to be highly expressed in a hypoxia model and during human disease. DosR is dispensable for *in vitro* survival and replication in human monocytes, but mutations in DosR results in a virulence attenuation of *Mtb* in guinea pigs and rabbits.<sup>81</sup> However, a recent transcriptional profiling study demonstrated that the DosR regulon is not responsible for the long term survival under anaerobic conditions.<sup>88</sup> The DosR regulon was highly expressed at the outset of anaerobic conditions but returned to its basal level within 24 h. The expression level of another set of genes was altered in the enduring hypoxic situation. Notably, these enduring hypoxic response enzymes did not overlap with any enzymes belonging to the DosR regulon. These results underscore the role of DosR regulon in mediating the transition from aerobic to anaerobic conditions.

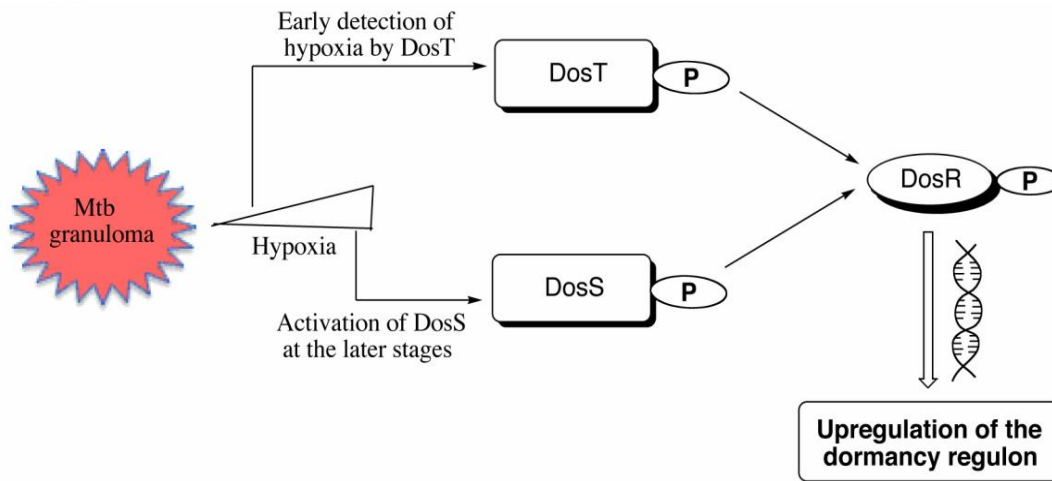


Figure 3. Dos regulation in *Mtb*.<sup>85</sup> Dos regulon is induced by DosT and DosS when *Mtb* is in hypoxia. After sensing the hypoxia, DosT is autophosphorylated in the early stage to mediate the metabolic switch from aerobic to hypoxia, and DosS gets autophosphorylated in the later stage to maintain the metabolic state in hypoxia. DosT and DosS transfer the phosphate group to DosR. The phosphorylated DosR binds DNA and causes an up-regulation of a series of enzymes called DosR regulon.

Although the exact mechanism of how NGMA cells are produced remains unclear, NGMA bacteria exhibit a unique phenotype and a differential metabolism. NGMA *Mtb* possesses a lipid body accumulation and is negative for acid-fast staining. To establish a low metabolic state, bacteria reduce nearly all basal metabolism such as amino acid, cofactor, ATP, and lipid syntheses, DNA replication and repair, translation, etc. Besides this reduction, some metabolism undergoes reprogramming including those of central carbon metabolism (CCM), respiratory chain, and redox state homeostasis ( $\text{NAD}^+/\text{NADH}$  ratio). For CCM, enzymes involved in glycolysis and the tricarboxylic acid cycle are down-regulated. Instead, a few enzymes involved in the fatty acid catabolism such as  $\beta$ -oxidation, glyoxylate shunt, and gluconeogenesis are up-regulated.

These gene expression changes highlight the fact that fatty acids are used as the major carbon source in the NGMA state.

To adapt to different oxygen tensions in its life cycle, *Mtb* has evolved three respiratory chains. A complete respiratory chain usually consists of three parts, an electron donor, an electron transport chain, and an electron acceptor. *Mtb* has redundant enzymes for each part (Table 2). Metabolic realignment is mainly observed within the electron transport chain. The cytochrome complexes and fumarate dehydrogenase (FRD) facilitate the oxidization of reduced menaquinol in mycobacteria individually. The cytochrome systems mediate the electron transfer in the presence of O<sub>2</sub>. In addition, FRD catalyzes the electron transfer to alternative electron acceptor, fumarate, under hypoxia. Two cytochrome complexes, cytochrome *bcc-aa<sub>3</sub>* and cytochrome *bd* oxidase, are present in *Mtb*. Transcriptional profiling analysis in the mouse infection model suggested that cytochrome *bcc-aa<sub>3</sub>* is used in the active mycobacteria, and nitrate reductase (NR) is predominately used in the NGMA mycobacteria. Cytochrome *bd* oxidase, however, is a transient enzyme to reshape electron flow and the redox balance during the environmental change,<sup>89</sup> and therefore is important for chronic infection.

Other than participating in energy production, some enzymes facilitate redox balance. Succinate dehydrogenase (SDH) and FRD were demonstrated to maintain the membrane potential and the redox balance in the hypoxia condition,<sup>90, 91</sup> and NR is used to maintain the intracellular redox balance when *Mtb* is in active macrophages.<sup>92</sup>

Notably, even though possessing an alternative ETC, and down-regulation of the F<sub>1</sub>F<sub>0</sub> ATP synthase, the remaining ATP synthesis activity of F<sub>1</sub>F<sub>0</sub> ATP synthase is

critical for cell survival in NGMA state.  $F_1F_0$  ATP synthase was found to be the molecular target of the recently approval anti-tubercular drug, TMC-207. TMC-207 shows potent inhibition against both multidrug resistant and NGMA *Mtb*.<sup>93, 94</sup> In fact, it is more effective to the NGMA *Mtb*. This result opens a door to target the NGMA *Mtb*. Enzymes involved in the energy metabolism and the alternative carbon metabolism therefore become an attractive drug targets of NGMA *Mtb*.

Table 2. Enzymes involved in the respiratory chains of *Mtb*.

Primary dehydrogenases and decarboxylases	Terminal oxidases and reductases	Redox carriers	ATP synthase
NADH dehydrogenase I	Cytochrome <i>bcc-aa<sub>3</sub></i>	Menaquinone	$H^+$ - $F_1F_0$
NADH dehydrogenase II (2)	complex		
Succinate dehydrogenase (2)	Cytochrom <i>bd</i> oxidase		
Formate dehydrogenase?	Fumarate reductase		
Carbon monoxide dehydrogenase	Nitrate reductase (2)		

### **Strategies to Identify Novel Chemotherapeutic Agents for Multidrug Resistant *Mtb* and Persistent *Mtb***

To efficiently combat TB, we need to first find drugs for the MDR- or XDR-TB to solve the current emergency. In the long term, finding drugs for persistent and dormant *Mtb* will decrease treatment duration and the occurrence of the drug resistance. A novel drug usually comes from (i) a structural modification of a currently used drug, (ii) a compound with a novel chemical scaffold but targeting the same molecular

mechanism as an existing pharmaceutical drug, or (iii) a compound with a novel chemical scaffold and a novel molecular mechanism.

The first antibiotic was discovered in the 1940s, and numerous antibiotics were discovered by screening soil-derived actinomycetes between 1940–1960, which was the so-called golden era of antibiotic discovery. In the golden era, the discovery of antibiotics was through natural product phenotypic screening in which molecules with therapeutic effects were identified and isolated. However, it became harder and harder to find a novel series of antibiotics using this approach in the 1990s. The increased knowledge of biochemistry and advances in technologies of molecular biology such as cloning and recombineering in the 1980s, moved antibiotic discovery into a rational design era, which is also known as the target-based drug discovery approach. This method was anticipated to fill the gap of finding novel antibiotics through the natural product screening approach. The target-based drug discovery approach starts from target selection. In the field of antibiotic discovery, a target is chosen based on two criteria, essentiality and selectivity. Essentiality means the target is critical for the pathogens' survival or virulence. Selectivity means the target is absent in humans or shows a large difference to its human counterpart. The breakthroughs of the genomic technology further laid the foundation of the target-based approach. After the mid 1990s, the improvement of sequencing skills allowed us to examine the sequence of the entire genome of an organism. In 1995, the first entire genome of *Haemophilus influenzae* was deciphered.<sup>95</sup> In addition, the first whole genomic sequence of a human was released in 2003.<sup>96</sup> The genomic study of pathogens illustrated the essential genes and their



corresponding gene products. The comparison with the human genome helps pre-eliminate those targets, whose inhibitors may lead to toxicity or side effect issues. Since then, the target-based approach has dominated the pharmaceutical industry including those working in antibiotic discovery.

However, the quantities of novel FDA approved drugs declined steadily after the 1990s.<sup>97</sup> A number of researchers believe that the decline is due to the overestimation of success in the target-based drug discovery approach.<sup>97</sup> Between 1999 and 2008, 50 out of 75 first-in-class (compounds with novel chemical scaffolds) FDA approved drugs were small-molecule drugs. Twenty-eight were derived from the phenotypic based drug discovery, and 17 derived from the target based drug discovery.<sup>98</sup> In addition, the approval of novel antibiotics dropped by 60%, from 33 between 1985 to 1999 to 13 between 2000 and 2014.<sup>99</sup> More seriously, of these only four are first-in-class.<sup>100</sup> The decline of the approval of novel antibiotics results from an overall decrease in research efforts for the antibiotic discovery due to the difficulty in licensing and poor return on investment. The productivity of the target-based approach is also questioned in the antibiotic discovery field. Multiple reviews have discussed the reasons for the failure of the target-based drug discovery.<sup>101</sup> The conclusions agree with the observations in a key report provided by the GSK for using the target-based approach to look for novel antibiotics.<sup>102</sup> Researchers in GSK summarized their findings after studying 300 genes and performing 70 high-throughput screens of individual targets from 1995–2001. They pointed out some explanations for why most hit molecules showed high potency against the chosen targets but without *in vivo* efficacy for the diseases. From the target

perspective, the target chosen based on the essentiality from the *in vitro* study often cannot pass the *in vivo* test, meaning inhibition of these selected targets would not affect the growth of bacteria *in vivo*. From the hit perspective, several hits suffer from bad pharmaceutical characteristics, such as low solubility or inability to penetrate bacterial cell walls, and also results in poor *in vivo* efficacy. In the end, the authors of this report<sup>102</sup> suggested that currently used pharmaceutical targets or targets that are obtained from the phenotypic screening are more valuable for a follow-up target-based drug discovery than those derived from genetic studies.

In response to the first suggestion, an examination of the molecular mechanism of PAS, a second-line anti-tubercular drug, is presented in Chapter II. PAS is one of the first discovered anti-tubercular drugs. It is currently used to treat the MDR- or XDR-TB, however, with an unclear molecular mechanism. Understanding the molecular mechanism is key to producing a new generation of PAS derivatives with improved efficacy and lower side effects. In addition, the definition of the molecular target also may provide a pharmacologically validated target for renewed target-based drug discovery.

### **Enzymes Involved in Central Carbon Metabolism (CCM) Are Promising Targets to Combat Persistent or Dormant *Mtb***

Lack of effective drugs against NGMA *Mtb* underlines the requirement to explore novel targets through phenotypic screening. However, it is challenging to

conduct a phenotypic screening for dormant *Mtb* since almost all models cannot reflect the true physiology. Hence, enzymes are considered to be potential targets when they are associated with the transition to the NGMA state and are essential for bacterial growth or virulence in the persistent models. However, a detailed biochemical and biophysical examination is required prior to further exploitation.

Enzymes involved in CCM are putative targets for NGMA *Mtb* since CCM is an important player for *Mtb* to transition between active and NGMA states. To adapt to the extracellular stresses in macrophages and granulomas, *Mtb* has evolved a flexible metabolic system to produce energy. A differential expression profile of CCM was observed when *Mtb* resided in macrophages. Several lipid catabolism enzymes are up-regulated; on the contrary, enzymes involved in glycolysis, the tricarboxylic acid (TCA) cycle, and aerobic respiration are down-regulated. These results suggest that *Mtb* relies on fatty acids as the predominant carbon source in macrophages. Noteworthy is that the replication capability of *Mtb* in immature macrophages implies that it evolves an optimal CCM to generate energy and required metabolites through fatty acids. In addition, the same alternations in the transcriptional levels of the CCM were observed while *Mtb* was cultured in the two well-known *in vitro* models for NGMA *Mtb*, the Loebel carbon starvation model and the Wayne hypoxia model.<sup>103</sup> These results reinforce the importance of CCM and therefore motivate researchers to develop inhibitors using these models.

The CCM plays important roles in all living organisms. It breaks down nutrients uptaken from the environment and converts them into energy and biomass for cell

survival. CCM is composed of multiple pathways including glycolysis, gluconeogenesis, the pentose phosphate shunt, the TCA cycle, and the methylcitrate cycle. Although it has been extensively studied in *E. coli* and also showed a high conservation with that in mammals, a relatively complicated and flexible CCM was observed in *Mtb* to adapt to the nutrient limiting environment during its life cycle. In fact, the rearrangement of CCM components in *Mtb* determines its fate in the host.

The advances in genomic sequencing, metabolomics and transcriptional profile analyses allow us to decipher the CCM in *Mtb*. Earlier studies were primarily conducted through the gene expression profile analysis via microarrays.<sup>83, 104-106</sup> In general, isocitrate lyase 1 (ICL-1) encoded by *icl* and *aceA*, phosphoenolpyruvate carboxykinase (PEPCK) encoded by *pcaA*, and enzymes involved in the  $\beta$ -oxidation are up-regulated in the intraphagosomal *Mtb* and become further up-regulated under the nutrient starvation or the hypoxia conditions. The genome of *Mtb* was deciphered in 1998,<sup>107</sup> and it provides an a glance of the overall picture of CCM in *Mtb*. Through homology-based sequence analysis, around 90 genes were annotated to be related to CCM and have been extensively detailed reviewed.<sup>108, 109</sup> Several proteins with functional redundancy are present in the CCM of *Mtb* that supports ideas about metabolism rerouting in response to different environments.

Another distinction of the CCM in *Mtb* is the ability to cocatabolize different carbon sources. Most bacteria use carbon sources under a carbon catabolite repression mechanism.<sup>110</sup> Bacteria utilize preferred carbon sources first, usually glucose, by repressing the enzymes involving the catabolism of other carbon sources when cultured

in a mixed carbon source. However, metabolites from different routes were observed simultaneously in  $^{13}\text{C}$  metabolomics experiments when *Mtb* was cultured *in vitro* with a mixed carbon source.<sup>111</sup> In addition, a recent  $^{13}\text{C}$  flux study demonstrated that *Mtb* used all available carbon sources such as C2 metabolites,  $\text{CO}_2$ , or amino acids at the same time in macrophages.<sup>112</sup>

The  $^{13}\text{C}$  flux study under hypoxia implied a general activity depletion of the TCA cycle enzymes and an activity increase of methylcitrate cycle enzymes that agrees with the transcriptional profiling. It also found a drastic increase of the intracellular levels of succinate.<sup>90</sup> The increased succinate served as an agent to maintain membrane potential, ATP synthesis, and anaplerosis. To summarize, *Mtb* uses fatty acid as the major carbon source during infection<sup>113</sup> and faces two challenges when using fatty acid as the carbon source: the generation of glucose through fatty acid and the detoxification of propionyl-CoA, a metabolites produced in the  $\beta$ -oxidation of odd-chain-length fatty acids and branched chain amino acids. To solve these challenges, ICL and PEPCK are in high demand. PEPCK is the first enzyme in gluconeogenesis, which mediates the generation of glucose from non-glucose carbon sources. In additional, *Mtb* ICL is a dual function enzyme, which functions as an isocitrate lyase in the glyoxylate shunt and a methylisocitrate lyase in the methylcitrate cycle.<sup>114</sup> The glyoxylate shunt serves as a bypass of the TCA cycle. The  $\text{CO}_2$  releasing steps are omitted through the glyoxylate shunt so it is a carbon conserved pathway compared with the TCA cycle. The methylisocitrate reduction is the critical step to detoxify the propionyl-CoA intermediate. Therefore, either or both functional roles can explain why ICL is up-regulated. In

addition, knocked-out *Mtb* ICLs<sup>115</sup> or PEPCK<sup>116</sup> impaired the ability to infect mice and demonstrated that *Mtb* ICLs and PEPCK are promising anti-tubercular drug targets. Compared with PEPCK, ICLs serve as better targets since there is no counterpart in humans.

Compelling research was performed on the *Mtb* ICLs, but little attention was placed on their substrate competing enzyme, IDHs. As mentioned before, the glyoxylate shunt is a bypass of the TCA cycle. ICL and IDH are in the niche of the branch point sharing the same substrate, isocitrate (ICT). Activities of ICL and IDH determine the carbon flux and therefore determine the fate of *Mtb* in different environments. *Mtb* possesses two IDHs, *Mtb* IDH1 and *Mtb* IDH2. *Mtb* IDH1 is annotated as a homodimeric enzyme encoded by the gene *Rv3339c*. To the contrary, *Mtb* IDH2 belongs to the monomeric IDH superfamily and is encoded by the gene *Rv0066c*.

Transposon mutagenesis experiments suggested that individual IDHs are non-essential for *Mtb* survival in 2003.<sup>117</sup> However, studies carried out in BCG underscored that IDH2 can be a potential synergetic drug target.<sup>118</sup> In this study, the individual IDH knock-out BCG strains were generated, which supported the idea that IDH1 and IDH2 are non-essential individually. Intriguingly, no IDH activity was detected for the IDH2 knock-out strain, suggesting that IDH2 was the only functional IDH *in vivo*. Moreover, the potency of the ICL inhibitor, 3-nitropropionate, increased in the IDH2 knock-out strain.

The co-catabolic characteristic of *Mtb* suggests that the co-inhibition of ICLs and IDHs is necessary to achieve bacterial death. The greater effect of ICL inhibitor in the IDH2 knock-out BCG strain further supports this hypothesis. A detailed introduction to IDH

followed by a thorough study of *Mtb* IDH2, including a complete biochemical and biophysical characterization and preliminary inhibitor screening, are presented in Chapter IV.

CHAPTER II  
MECHANISM ACTION STUDY OF A SECOND-LINE ANTI-TUBERCULAR  
DRUG *PARA*-AMINOSALICYLIC ACID (PAS)

**Introduction**

*Para*-aminosalicylic acid (PAS) (Figure 4A) was the first anti-tubercular drug<sup>119</sup> discovered, and it is still used as a second line drug for TB treatment, especially for MDR-TB. The anti-tubercular activity of PAS was first reported by Lehmann in 1946 when looking for bacteriostatic agents against TB within a collection of more than 50 derivatives of benzoic acid.<sup>120</sup> PAS was thereafter used as a first-line drug in combination with isoniazid and streptomycin until it was replaced by ethambutol in the 1960s as the combination with PAS showed higher frequency of vestibular and gastric side effects in patients.<sup>121</sup> PAS was reintroduced to treat MDR-TB and XDR-TB in 1992.<sup>122</sup>

PAS exhibited an EC<sub>90</sub> value of 1.25 µg/mL for *Mtb* H37Rv in an *in vitro* microplate alamar blue assay with a 7H9 buffer supplemented with glycerol, OADC (oleic acid, albumin, dextrose, and catalase), and tween80.<sup>123</sup> The recommended dosage of PAS granules is 4 g in a single dose and one dose twice per day. This dosage maintains PAS in an intracellular concentration higher than the reported MIC (> 1 µg/mL), for effectively controlling the MDR-TB.<sup>124</sup>



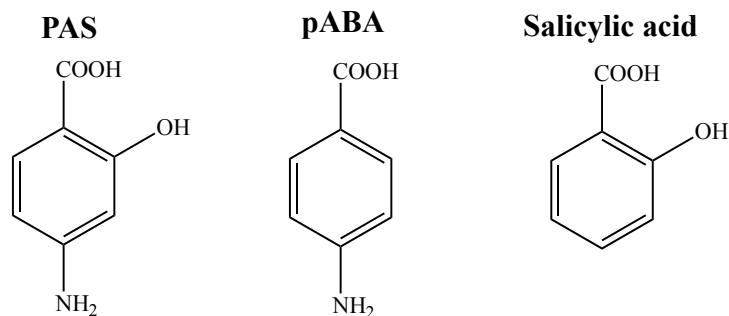


Figure 4. Chemical structures of PAS (left) and its structural related metabolites, *para*-aminobenzoic acid (pABA) (middle) and salicylic acid (right)

Intriguingly, PAS is an organism specific antibiotic. Diczfaluzi *et al.* found that it is not active against *E. coli*, *S. aureus*, or *S. haemolyticus*.<sup>125</sup> Moreover, it shows less potency against mycobacteria such as *M. avium* (EC<sub>90</sub>=80 µg/mL),<sup>123</sup> *M. smegmatis* (EC<sub>90</sub>=512 µg/mL),<sup>126</sup> *M. fortuitum* (EC<sub>90</sub>> 128 µg/mL),<sup>127</sup> and *M. chelonae* (EC<sub>90</sub>> 128 µg/mL).<sup>127</sup> Even though the anti-tubercular effect of PAS has been demonstrated over 70 years, its mode of action and molecular target were unclear. A lack of published data on these studies could be due to the discontinued use in the clinic during the 1960s to 1992.

One putative mechanism of action of PAS is to interfere with the folate biosynthetic pathway. Folates, tetrahydrofolate (THF) and its analogs, consist of three moieties, a pteridine, *p*-aminobenzoic acid (pABA) and L-glutamate. They function as carbon carriers in the one-carbon metabolism.<sup>128</sup> One-carbon groups from serine, glycine, histidine, and formate, are added to N<sup>5</sup>, N<sup>10</sup>, or both on the pyridine ring of folates. The common observed derivatives include 5-methyl-, 5,10-methylene-, 5,10-methenyl-, 5-formimino-, 5-formyl- and 10-formyl-THF. These derivatives are used as

carbon donors in several biosynthetic reactions, particularly the synthesis of purine, thymidylate, and methionine.<sup>128</sup>

Plants and most bacteria, including mycobacteria, generate THF by a *de novo* biosynthesis (Figure 5).<sup>129</sup> This pathway can be divided into two branches, the pteridine branch and the pABA branch. GTP is the precursor of the pteridine branch.<sup>130</sup> Following the consecutive action of FolE (EC 3.5.4.16), an unidentified Phosphatase, FolB (EC 4.1.2.25), and FolK (EC 2.7.6.3), GTP is converted into 7,8-dihydroneopterin 3'-triphosphate, 7,8-dihydroneopterin, 6-hydroxymethyl-7,8-dihydropterin, and then 6-hydroxymethyl-7,8-dihydropteridine pyrophosphate (HP).<sup>131</sup> Additionally, the pABA branch starts from chorismate, the product of the shikimate pathway.<sup>132</sup> First, chorismate is converted into 4-amino-4-deoxychorismate via a 4-amino-4-deoxychorismate synthase (EC 2.6.1.85) and further is turned into pABA by 4-amino-4-deoxychorimatalyase (EC 4.1.3.38).<sup>132</sup>

The two branches merge together when HP and pABA are condensed by dihydropteroate synthase (DHPS) (EC 2.5.1.15) forming 7,8-dihydropteroate (DHP). Dihydrofolate synthase (DHFS) (EC 6.3.2.12) then adds an L-glutamate to DHP to produce dihydrofolate (DHF). The produced DHF then reduces to the final product, THF, by dihydrofolate reductase (DHFR) (EC 1.5.1.3.). A polyglutamate tail (2–7 glutamyl residues<sup>133</sup>) is added to the  $\gamma$ -carboxy of the THF glutamate via folylpolyglutamate synthetase (FPGS) (EC 6.3.2.17). In *Mtb*, a bifunctional enzyme, FolC, catalyzes the reactions of DHFS and FPGS. In addition to the *de novo* synthesis, THF can be generated through recycling the reducing DHF in the thymidylate synthesis.

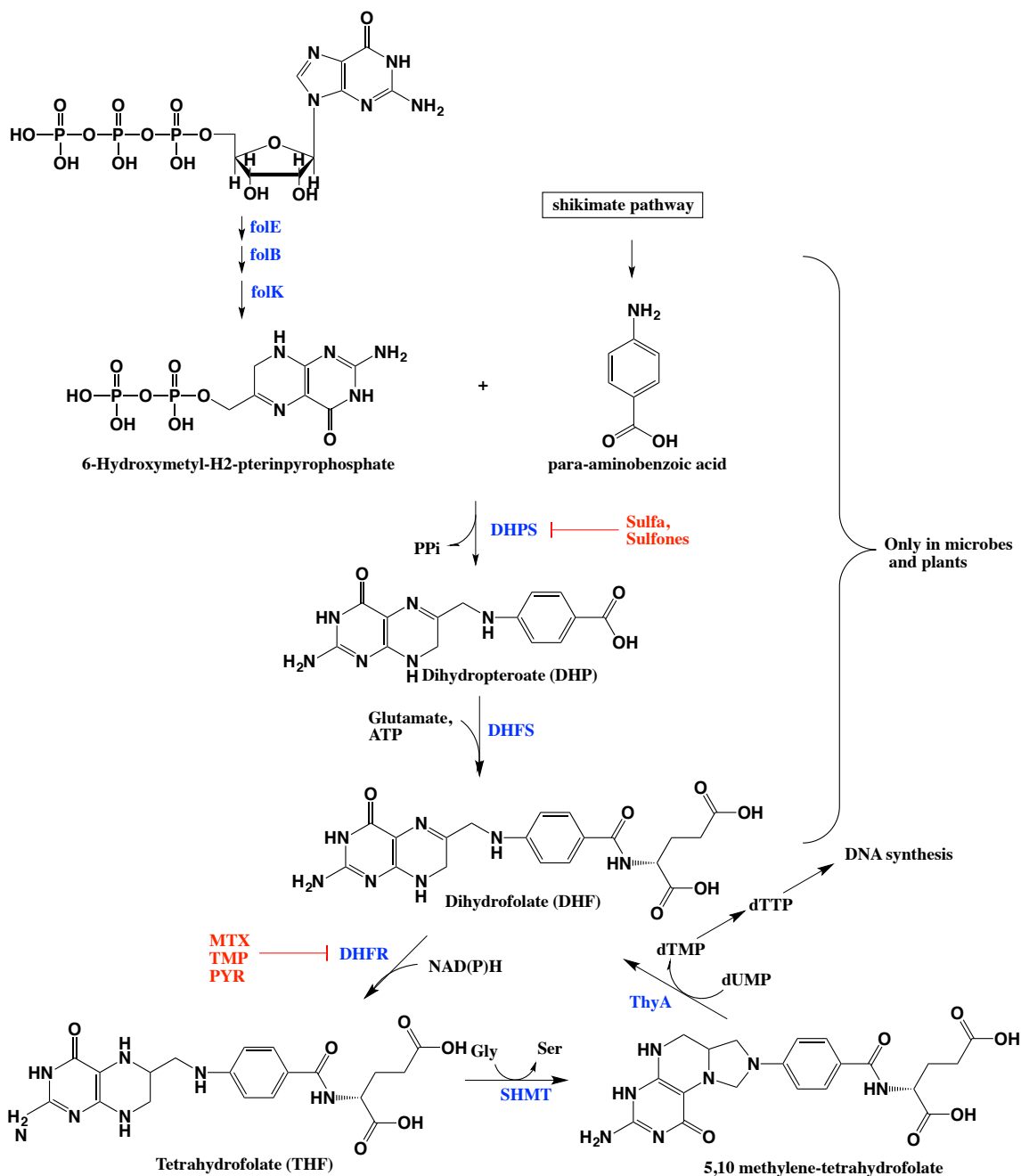


Figure 5. *De novo* folate metabolism and the sites at which antimetabolites act. Metabolites are shown in black, enzymes are labeled in blue, and inhibitors are colored in red.

PAS was first thought to be a competitive inhibitor of DHPS due to being a structural analogue of pABA (Figure 4B). However, PAS showed a different inhibition profile with sulfonamides and sulfones, the well studied inhibitors of DHPS (Table 3).<sup>134</sup> PAS exhibited a high inhibitory constant for *Mtb* DHPS but a better whole cell efficacy for *Mtb* when compared with that for the sulfonamides and sulfones. The unexpectedly low whole cell activities of sulfonamides and sulfones indicate the difficulty they face in accessing the intracellular target in *Mtb*. The different inhibition profiles therefore suggest that *Mtb* DHPS may not be the molecular target of PAS, unless there is a unique uptake system for PAS.

Table 3. Inhibition profile of PAS, sulfones and sulfonamides against the activity of *Mtb* DHPS and the growth of *Mtb*<sup>134</sup>

Compound	$K_i$ ( $\mu\text{g/mL}$ ) for <i>Mtb</i> DHPS	EC <sub>90</sub> ( $\mu\text{g/mL}$ )
PAS	0.18	0.3–1.0
Dapsone (sulfone)	0.002	>32
Sulfamethoxazole (sulfonamide)	0.004	8

Despite the fact that DHPS is not the target of PAS, the folate pathway is still associated with the mechanism of action of PAS due to the observed correlation between the PAS resistance and the mutation in thymidylate synthase (ThyA).<sup>135</sup> ThyA catalyzes the reductive methylation of 2'-deoxyuridine-5'-monophosphate (dUMP) to 2'-deoxythymidine-5'-monophosphate (dTMP).<sup>135, 136</sup> In this reaction, 5,10-methylene-tetrahydrofolate (5,10-methylene-THF) functions as a one-carbon donor and is converted

to a byproduct, dihydrofolate (DHF). 5,10-methylene-THF is a derivative of THF, the final product of the folate pathway. Mutations in gene *thyA* are one of the most common PAS resistant mechanisms in the clinical isolates. Around 37% of PAS resistant strains contain mutations in *thyA* in a study that took place in 2009.<sup>136</sup> A recombinant *Mycobacterium bovis* (*M. bovis*) BCG strain expressing *Mtb* ThyA showed a hypersensitivity to PAS, which indicates that ThyA can be an activator instead of a molecular target of PAS.<sup>135</sup> No previous study has investigated the biochemical relationship between ThyA and PAS. Therefore, the interconnection between PAS and the folate pathway remains unclear.

Alternatively, PAS is related to the iron acquisition mechanism. Microorganisms acquire irons from the environment via siderophores. Siderophores chelate  $\text{Fe}^{3+}$  and transfer it into cells. Two types of siderophores, carboxymycobactin and mycobactin, are found in *Mtb*.<sup>137</sup> Mycobactin (Figure 6) contains a long alkyl chain ( $n= 19$  in *Mtb*) that makes it lipophilic; in contrast, the alkyl chain of carboxymycobactin (Figure 6) is shorter ( $n= 2-9$ ) and with a carboxylate group at the end resulting in a hydrophilic characteristic.<sup>138</sup> Carboxymycobactin is secreted out from *Mtb* and collects the iron from the host.<sup>139</sup> The ferri-carboxymycobactin would either deliver the iron to the membrane associated mycobactin<sup>139</sup> or pass the cell membrane by an ABC transport system, IrtAB<sup>140</sup> (Figure 7). Salicylic acid plays a key role in the iron acquisition mechanism. It is not only a precursor of the siderophores but also a sensor to signal the intracellular iron status. As for pABA, PAS is a structural analogue of salicylic acid, and it shows potent inhibition against the mycobactin synthesis.<sup>141</sup>

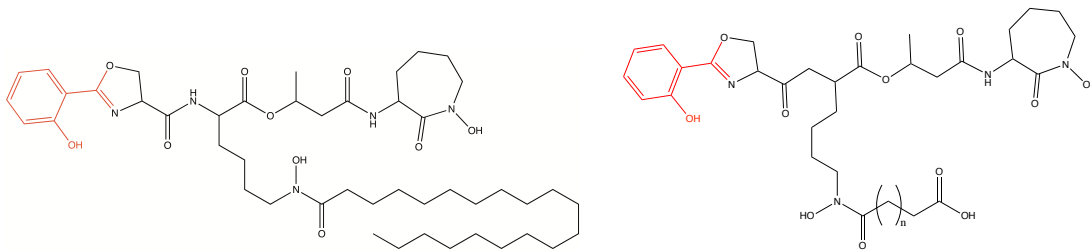


Figure 6. Chemical structures of siderophores produced by *Mtb*. Mycobactin (left) is membrane associated, and carboxymycobactin (right) is water soluble. The acyl chain on the hydroxylated lysine of the carboxymycobactin is 2–9 carbon long. Salicylic acid moiety is highlighted in red.

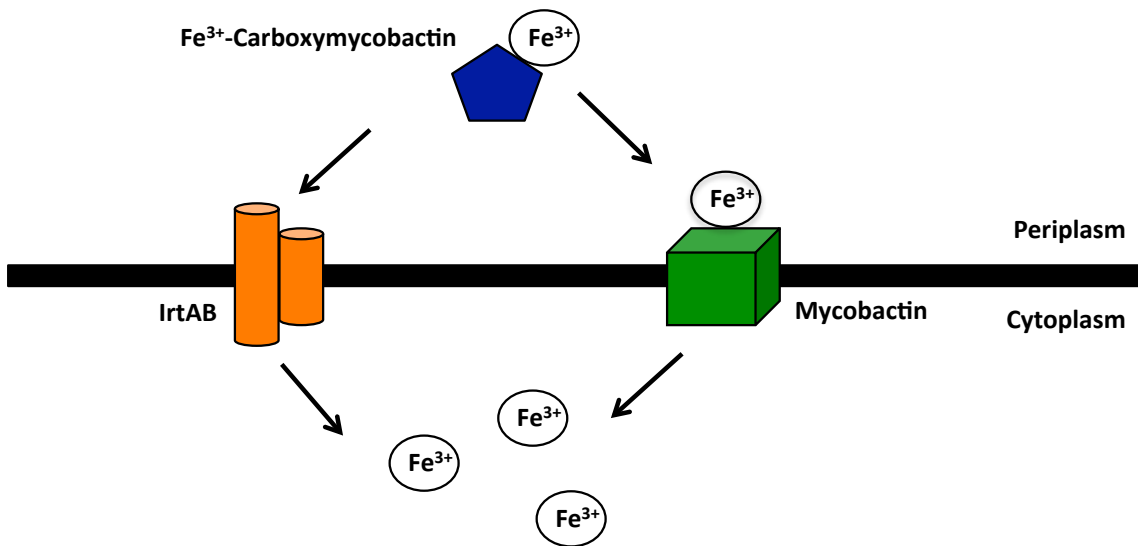


Figure 7. Iron acquisition system of *Mtb*. To obtain the iron, the water soluble carboxymycobactins are secreted by *Mtb* in an iron deficiency condition. The carboxymycobactins chelate  $\text{Fe}^{3+}$  from the host to form  $\text{Fe}^{3+}$ -carboxymycobactin complexes. The complexes are transported back through the mycobacterial outer membrane by the Msp family porins first. The complexes then either transfer the  $\text{Fe}^{3+}$  to the membrane associated mycobactins or are transported into the cell by the iron regulated ABC-transporter, IrtAB. Blue pentagon, orange rod, and green cube represent carboxymycobactin, IrtAB, and mycobactin, respectively.

Therefore, the mechanism of action of PAS is a controversial subject since no molecular based evidence supports the hypothesis that the folate pathway is the target, and no clinical evidence gets behind the idea that PAS targets the iron acquisition pathway. After the reintroduction of PAS, understanding the insight into the mechanism of action of PAS becomes important. The knowledge of the mechanism can provide a platform for designing PAS analogues with reduced toxicity and illustrate a molecular target with a validated clinical therapeutic ability for a target-based drug screening. In this chapter, a genetic study was implemented to characterize the putative molecular target of PAS. In addition, a structural investigation was carried out to justify the mechanism from a molecular basis.

## **Results and Discussion**

### *Mutations in the Folate Biosynthesis Genes Make *Mtb mc<sup>2</sup>7000* Resistant to PAS*

Advances in DNA sequencing technology allow a quick and affordable investigation of the entire genome of the drug resistant individual instead of examining the sequence of the pre-suggested target gene alone. Spontaneous laboratory PAS resistant mutants were isolated from the two *Mtb mc<sup>2</sup>7000* strains, wild-type (WT) and overexpressing exogenous ThyA strains. The additional copies of ThyA would compensate for resistance caused by the mutation of ThyA and presumably enhanced the chance to obtain the mutant strains with mutations other than ThyA.

Whole genome sequencing was conducted on five PAS resistant *Mtb mc*<sup>2</sup>7000 strains that were isolated from the WT *Mtb mc*<sup>2</sup>7000 experimental trial. Mutations in the *thyA* gene were detected in four of them, and a deletion across *dfrA* and *thyA* genes was observed in Strain 5 (Table 4). Strain 1 contained a deletion of 3 base pairs in *thyA*, which encoded Ala 40 in the WT ThyA. Strain 2 harbored a W60\* nonsense mutation in *thyA*, resulting in a premature ThyA protein. Strain 3 had a potential recombination event between *thyA* and its downstream gene, *PPE44*, which caused an inversion mutation. Similar to Strain 2, Strain 4 had a nonsense mutation of W83\* in *thyA*. A deletion of 824 base pairs, base pair numbers from 3073139 to 3073962, was observed on Strain 5 spanning genes *dfrA* (*Rv2763*) and *thyA* (*Rv2764c*).

As mentioned earlier, PAS resistance can emerge due to the point mutations in *thyA*, and a broad spectrum of mutation positions were reported. In the study conducted by Mathys *et al.*, 24 different mutation positions in *thyA* were observed within 118 clinical and laboratory PAS resistant isolates.<sup>136</sup> Many diverse mutation positions in *thyA* suggested that there is no mutation hot spot for this gene. Among the four *thyA* mutations in our study, only the W83\* mutation had been reported previously in five clinical isolates from the Beijing family<sup>142</sup> and one spontaneous mutant isolated from a 7H11 plate supplemented with 16 µg/mL PAS.<sup>136</sup> The low re-emergence of the reported mutants in our study could be due to the application of the mutagen, ethyl methanesulfonate (EMS), during the mutant isolation process for Strains 2–5. EMS mainly promotes the GC to AT substitution mutation and occasionally increases the frequency of the frameshift mutation event, which decreases the mutation diversity of



the selection pool. Moreover, the increased number of SNPs also corresponds to the EMS treatment.

To obtain PAS resistant mutants with mutations in genes other than *thyA*, PAS resistant strains were isolated in the presence of excessive ThyA by an overexpression plasmid. Two strains, Strain 6 and Strain 7, were obtained and sequenced (Table 4). Notably, mutations in another folate biosynthetic gene *folC* were observed in both strains, although the exact mutation positions were different. Gene *folC* encodes a bifunctional enzyme with DHFS and FPGS activities. DHFS catalyzes the conversion from DHP to DHF by adding an L-glutamate, and FPGS catalyzes the addition of poly-glutamate residues to DHF via the same reaction. Strain 6 and Strain 7 contained missense mutations of I43T and D135V, respectively. The mutations observed in *folC* are critical for the substrate binding of the encoded protein according to the crystal structure (2VOS<sup>143</sup>). I43 is located on the binding loop, and D135 is located in  $\alpha 4$ , which forms a salt bridge with the binding loop. These results indicate that FolC may be the molecular target of PAS or at least is involved in the targeting mechanism of PAS.

Table 4. The sequencing results of the PAS resistant strains

Strain	Strain manipulation	Conc. of PAS used for selection (µg/mL)	Gene modifications in folate biosynthetic genes			Total SNPs	MICs (µg/mL)
			<i>folC</i>	<i>thyA</i>	<i>dfrA</i>		
1	None	4	None	A40del	None	2	>8
2	Exposed with 0.1% EMS	4	None	W61*	None	15	>8
3	Exposed with 0.1% EMS	4	None	Potential recombination event	None	15	>8
4	Exposed with 0.1% EMS	4	None	W83*	None	9	>8
5	Exposed with 0.1% EMS	4	None	Deletion	Deletion	None	>8
6	Overexpression of ThyA	4	I43T	None	None	2	>8
7	Overexpression of ThyA	4	D135V	None	None	2	>8

#### *DHFR Is Essential Only When ThyA Is Active*

Before proceeding to examine the role of the mutation containing enzymes, it will be necessary to discuss why a strain with a deletion mutation in *dfrA* can survive. The existence of the *dfrA* deletion strain is surprising since it is thought to be essential for *Mtb* survival according to a transposon mutagenesis study.<sup>144</sup> It has been reported<sup>136</sup> that a PAS resistant strain contains point mutations in *dfrA* and *thyA* simultaneously; however, in our study, we observed a strain with a *thyA* and *dfrA* double deletion for the first time. Indeed, a double deletion of these two genes was recently observed in some clinical PAS-resistant isolates,<sup>145, 146</sup> demonstrating its clinical relevance.

A possible explanation would be the presence of ThyX, an alternative pathway to thymidine biosynthesis. A ThyA/DHFR coupled pathway was thought to be the only

pathway to synthesize dTMP. ThyA used 5,10-methylene-THF as the one-carbon unit donor and a reductant to catalyze the methylation of dUMP (Figure 8). 5,10-methylene-THF is therefore oxidized to DHF.<sup>147</sup> This resulting DHF is rapidly reduced to THF by DHFR. In 2002, the enzyme, ThyX, was first identified.<sup>148</sup> ThyX catalyzes the synthesis of dTMP in the presence of 5,10-methylene-THF, FAD and NADPH (Figure 8). 5,10-methylene-THF is only used as a carbon donor, and the FAD/NADPH couple provides the electrons for the methylation process of ThyX. Notably, instead of DHF, 5,10-methylene-THF is converted into THF in the reaction driven by ThyX. Therefore, ThyX is an alternative to the ThyA/DHFR coupled pathway not only in place of the ThyA.

dTMP biosynthesis is required for cell survival, and most bacteria have either a ThyA/DHFR or a ThyX pathway. However, corynebacteria, including *Mtb*, are found to possess both *thyA* and *thyX* genes.<sup>148</sup> Notably, it was demonstrated<sup>149</sup> that *Mtb* ThyA is non-essential for *Mtb* survival by a knock-out study. This result suggested that *Mtb* ThyX would compensate for the loss of *Mtb* ThyA. However, because *Mtb* ThyA shows a better catalytic activity and a higher expression level compared to *Mtb* ThyX,<sup>149, 150</sup> the ThyA/DHFR coupled pathway may dominate the thymidine biosynthesis in *Mtb*. *Mtb* DHFR is therefore required to recycle 5,10-methylene-THF and is essential for cell survival under normal conditions. For Strain 5, *Mtb* DHFR is dispensable since ThyA is dysfunctional, and ThyX is sufficient to replace the role of ThyA.

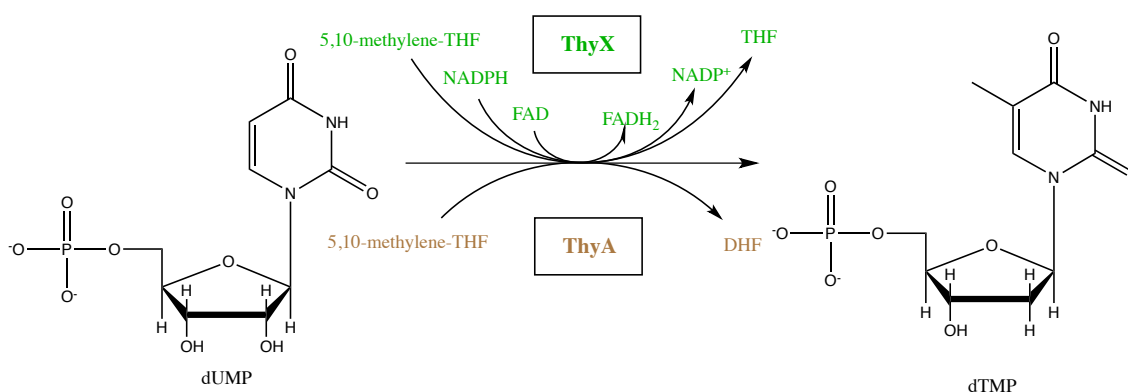


Figure 8. Thymidylate synthase reaction. There are two classes of thymidylate synthases, ThyA and ThyX. ThyA catalyzes the reductive methylation of dUMP using 5,10-methylene-THF as a one-carbon donor and a reductive power. Alternatively, ThyX catalyzes the reductive methylation of dUMP with 5,10-methylene-THF and an FAD/NADPH couple serving as the one-carbon donor and the reductant, respectively.

#### *MIC of mc<sup>2</sup>7000 against PAS Increased When DHFR Was Overexpressed*

Mechanisms of drug resistance resulting from mutations include (i) directly disrupting the binding between the drug and its molecular target; (ii) inducing the expression of the targeted enzyme or its functional counterpart; (iii) reducing the activity of the intracellular activator; and (iv) altering the drug transport system. Since mutations in *thyA*, *dfrA*, and *folC* are associated with the PAS resistance, their corresponding enzymes, ThyA, DHFR and FolC, are potential molecular targets or activators of PAS.

Gene knock-out and overexpression analyses are commonly used to investigate the influence of a specific enzyme on an inhibitor. In this study, PAS susceptibility tests were conducted on strains with the overexpression of ThyA, DHFR or FolC to evaluate if there is a true target or activator within these putative candidates. ThyA and FolC were overexpressed in *Mtb mc<sup>2</sup>7000* (ThyA<sup>ox</sup> and FolC<sup>ox</sup> strains) through a tetracycline

inducible expression vector<sup>151</sup> (pDT) with an N-terminal 6x His-tag. The overexpression of the proteins was confirmed by immunoblotting the 6x His-tag using an anti-His tag antibody. DHFR overexpression in mc<sup>2</sup>7000 (DHFR<sup>ox</sup> strain) was first constructed by the same method; however, neither the N-terminal nor the C-terminal 6x His tagged DHFR overexpression was observed. The effect of overexpressing DHFR on PAS resistance was therefore tested by cloning DHFR into a pMV261 plasmid, which expresses genes constitutively under the control of the *hsp60* gene promoter. MICs for PAS did not show significant differences when the FolC<sup>ox</sup> or ThyA<sup>ox</sup> strains were induced by 100 ng/mL of anhydrotetracycline when compared with that of both un-induced cells and the parental untransformed mc<sup>2</sup>7000 cells (MICs were 4 µg/mL for all experiments). To the contrary, MIC of PAS increased (>16 µg/mL) in the pMV261::*dhrA* transformed mc<sup>2</sup>7000 cells (Figure 9). This observation suggests that *Mtb* DHFR is the molecular target of PAS. This hypothesis explains the role of defective or dysfunctional ThyA in the PAS resistance. As discuss earlier, *Mtb* DHFR is a conditionally essential gene. It is essential when *Mtb* ThyA facilitates the dTMP synthesis. Therefore, a defective or dysfunctional ThyA makes *Mtb* DHFR dispensable and then causes resistance to a DHFR inhibitor, PAS.

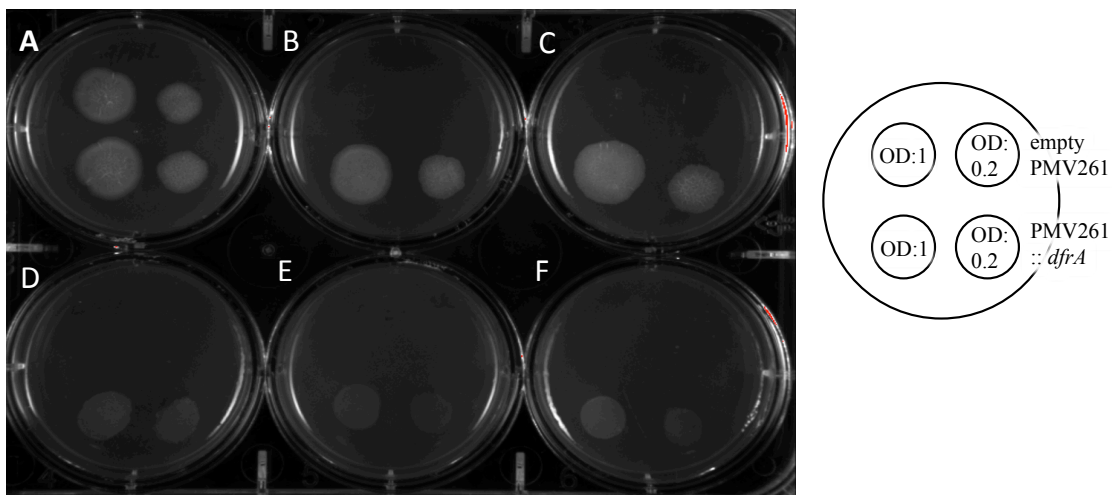


Figure 9. Whole cell assay of PAS to *Mtb mc<sup>2</sup>7000* strain with a DHFR overexpression vector. Each well (A-F) has different concentrations of PAS, wells A-C: 0, 1, 2 µg/mL, wells D-F: 4, 8, 16 µg/mL. For each well, the top left spot contains 10 µL of OD=1 cells with empty pMV261 vector; the top right spot contains 10 µL OD=0.2 cells with empty pMV261 vector; the bottom right spot contains 10 µL of OD=1 cells with pMV261::*dfrA* vector; and the bottom left spot contains 10 µL of OD=0.2 cells with pMV261::*dfrA* vector.

*PAS Is Incorporated into the Folate Pathway and DHPS or PAS Uptake Is the Rate-Limiting Step of the PAS Activation Process*

After our study was conducted, one important observation was reported by Chakraborty *et al.*<sup>152</sup> that PAS is incorporated into the folate pathway. An *Mtb* cell lysate after PAS treatment was analyzed by LC-MS. A series of hydroxylated folate metabolites were isolated and identified including the hydroxylated products of DHPS and FolC. Combining this result with our data, we concluded that PAS is a prodrug and targets DHFR. This hypothesis was soon confirmed by a follow-up paper published by Zheng *et al.* in 2014.<sup>153</sup> They further proposed that hydroxy-DHF is the toxic component inhibiting *Mtb* DHFR.<sup>153</sup>

An *in vitro* experiment was designed to investigate the effects of each hydroxylated metabolite against *Mtb* DHFR. Indeed, all hydroxylated metabolites are the putative active forms of PAS since the observation of the downstream metabolites does not rule out the possibility that one of the intermediate metabolites accumulates and results in toxicity. The hydroxylated metabolites are not commercially available. Thus, an enzymatically synthetic approach was used to obtain the putative compounds. The enzymes involved in the *Mtb* folate biosynthetic pathway were cloned and purified including FolK, FolP1, DHPS, and FolC. PAS was used as the substrate in the reaction of FolP1 to generate the hydroxy-substrate for the next step and so on. As a control, pABA was used as the substrate to evaluate the enzyme assay. After reaction, enzymes were removed via a 10kDa spin column, and the product of each reaction was monitored by LC-MS. The individual incorporated product was observed in the flow through by the LC-MS. The instability of the dihydropteridine group, however, makes it difficult to conduct a purification step. Nevertheless, the unpurified reaction flow through from each step was used as an inhibitor stock in the *Mtb* DHFR enzyme assay. The inhibitory activity of individual hydroxylated metabolites turned out to be inconclusive due to the low product yield after the multiple steps reaction and the unknown concentration in the unpurified metabolites solution. Therefore, the intracellular active chemical form awaits further investigation.

Alternatively, the whole cell efficacy of the FolC<sup>ox</sup> strain provides another insight into the activation procedure of PAS. Assuming that the hydroxy-DHF is the ultimate active chemical species, overexpression of FolC should induce the PAS

sensitivity of *Mtb* by increasing the production of the toxin. However, no MIC decrease was found when FolC was overexpressed. The non-correlation between the whole cell efficacy and the activity of FolC suggests that the determinant of the PAS activity occurs before this step, which are the uptake step or the conversion by DHPS.

*Molecular Modeling Revealing a Water-Mediated Hydrogen Bond Network Providing Selectivity between DHF and Hydroxy-DHF*

A structural study was performed to understand the binding mode of the hydroxy-DHF and the role of the hydroxyl group in terms of selectivity even though its biochemical inhibition awaits further investigation. An *in silico* docking experiment was conducted due to hydroxy-DHF being unavailable. Hydroxy-DHF was modeled by a Glide program in its flexible docking mode into a methotrexate (MTX) bound *Mtb* DHFR structure (Protein Databank (PDB) code: 1DF7<sup>154</sup>). In fact, a DHF bound *Mtb* DHFR would be the best model for this experiment since DHF is the closest structural analogue of the hydroxy-DHF. Whereas, the structure of the DHF bound *Mtb* DHFR has not been solved at this time, so the MTX bound *Mtb* DHFR structure was used instead. MTX is a tight binding DHF competitive inhibitor of DHFR. It exhibits an IC<sub>50</sub> value of 6.8 nM to *Mtb* DHFR.<sup>155</sup> A structural comparison was conducted to investigate whether the MTX bound *Mtb* DHFR structure is a good substitute. It has been demonstrated in *E. coli* DHFR that MTX and DHF are located at the same binding cavities. The difference is that the pteridine ring of MTX has a 180° rotation about the axis which is collinear with the line between C2 and N5 of DHF (Figure 10), and the remaining pABA and



glutamate moieties have the same binding mode comparing with DHF at the active site.<sup>156</sup> This binding correlation is conserved for the DHFRs with both MTX-bound and DHF-bound structures available.<sup>157-159</sup> These results suggests not only that MTX bound DHFR structure can be a good replacement but also that the docking results can be evaluated by the conserved binding mode alternation.

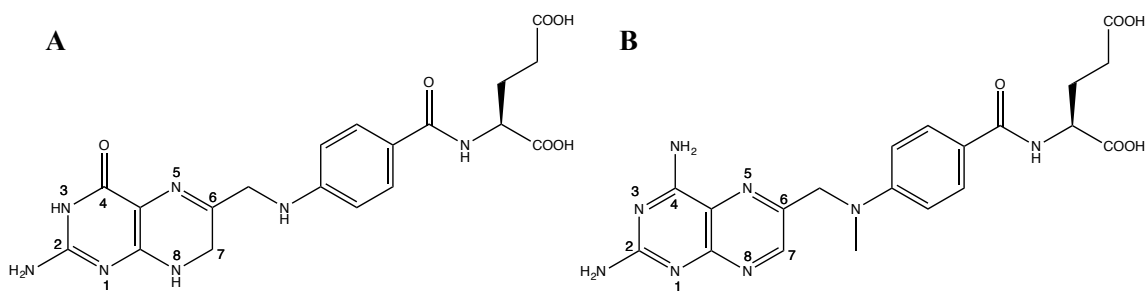


Figure 10. Chemical structure of (A) DHF and (B) MTX with numbering

Other than NADPH and MTX, a glycerol molecule is found nearby the N8 of MTX (Figure 10) in this MTX bound *Mtb* DHFR structure. To explore the most binding poses, NADPH and glycerol were removed from the structure before docking. The molecular modeling results revealed the interaction of the hydroxy-DHF with the MTX binding site in *Mtb* DHFR. Eight ligand poses were reported from the Glide docking, and the ligand strain energy of each ligand was calculated using ICM-browser pro (Table 5).

Table 5. Ligand conformations of the modelled hydroxy-DHF derived from the Glide docking and their corresponding strain energy calculated by ICM-browser pro

Ligand conformation	Ligand Strain Energy	Hydrogen bonding with H <sub>2</sub> O (644)
1	31	Yes
2	83	
3	96	
4	35	Yes
5	92	Yes
6	78	Yes
7	75	Yes
8	32	

In general, all conformations, except pose 8, are located at the MTX binding site. Specifically, the dihydropteridine moiety of the hydroxy-DHF appeared at the same position where the corresponding pteridine moiety of MTX is located, but the aforementioned 180° rotation was observed only in poses 1, 6 and 7. As shown in Figure 11, the additional hydroxyl groups pointed to the N-terminus of the helix B in pose 1 and poses 3–7. Alternatively, it pointed against the N-terminus in pose 2.

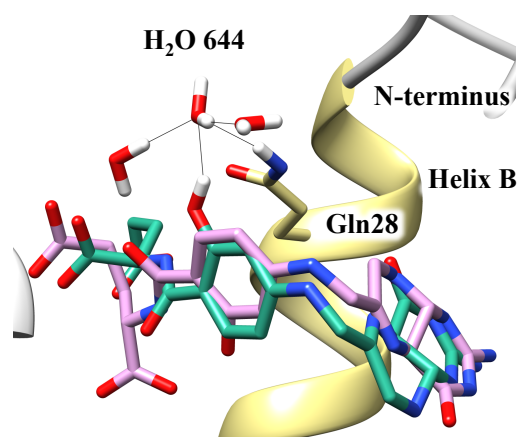


Figure 11. Two possible orientations of the hydroxyl groups of the hydroxy-DHFs as docking in the MTX binding site of *Mtb* DHFR. Helix B of *Mtb* DHFR is colored in yellow. The hydroxyl groups face the N-terminus of the helix B of *Mtb* DHFR in the majority docking poses except pose 2, which faces the C-terminus of the helix B. Pose 6 shown in green represents the major orientation, and the binding mode of pose 2 is shown in pink.

The hydroxyl group did not form a direct hydrogen bond with the surrounding residues of *Mtb* DHFR in any pose. However, a conserved water mediated hydrogen bond network was observed in poses 1, 4, 5, 6, and 7 (Figure 11). The hydroxyl group formed a hydrogen bond with an ordered water molecule (H<sub>2</sub>O 644). This water molecule also formed a hydrogen bond with the side chain of Gln28. This water mediated hydrogen bond network may lead to the selectivity between DHF and hydroxy DHF. Poses 1 and 4 were considered to be the better binding modes within eight poses. Hydroxy-DHFs in pose 1 and 4 had relatively low strain energy. In addition, the binding mode of pose 1 was similar to that of folate when compared with a reported folate bound *E. coli* DHFR structure.<sup>160</sup> The interaction of *Mtb* DHFR and the hydroxy-DHF in pose 1 and pose 4 are shown in Figure 12. In summary, the current mechanism of action of PAS

would be that PAS is incorporated into the folate pathway of *Mtb* and converted into hydroxy-DHF. Hydroxy-DHF then serves as an inhibitor to *Mtb* DHFR by competing with DHF through an extra water mediated hydrogen bond network.

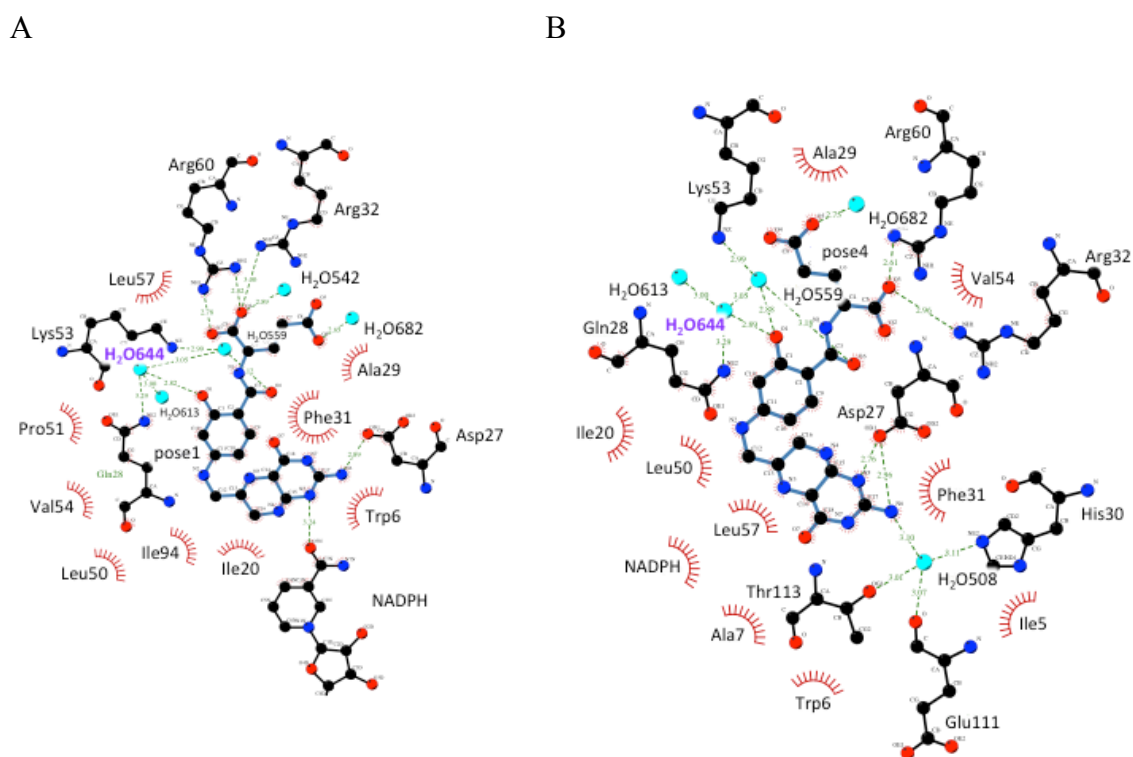


Figure 12. Detailed interactions between the top ranked poses and *Mtb* DHFR. Ligplot diagrams represent the interactions between *Mtb* DHFR and the top ranked poses (A) Pose 1 and (B) Pose 4. The hydrogen bonds are shown in green dashes line and the hydrophobic interactions in red arcs. The carbon atoms, oxygen atoms, and nitrogen atoms are colored by black, red, and blue, respectively. The ligand bonds are in light blue and protein bonds in black.

### PAS Targets More than One Pathway

Selectivity within different species of *Mycobacterium* is another question about the mechanism of action of PAS. Based on the proposed model, one possibility is that DHPS or FolC have critical differences in varied species of *Mycobacteria*. As a result, PAS cannot be converted into the active chemical form. After conducting a sequence analysis, we found that both DHPS and FolC showed high sequence identity (70%–99% with coverage more than 93%) among different species of *Mycobacterium*. These results rule out the first possibility. The second possibility would be a lack of residue, which can form the hydrogen bond with the conserved water as Gln28 in *Mtb*. A sequence alignment was performed on a collection of mycobacterial DHFRs composed of species of *Mycobacterium* for which PAS whole cell activity have been reported. As shown in Figure 13, three types of residues (glutamine, leucine, and methionine) were found at corresponding position of Gln28 in *Mtb* DHFR.

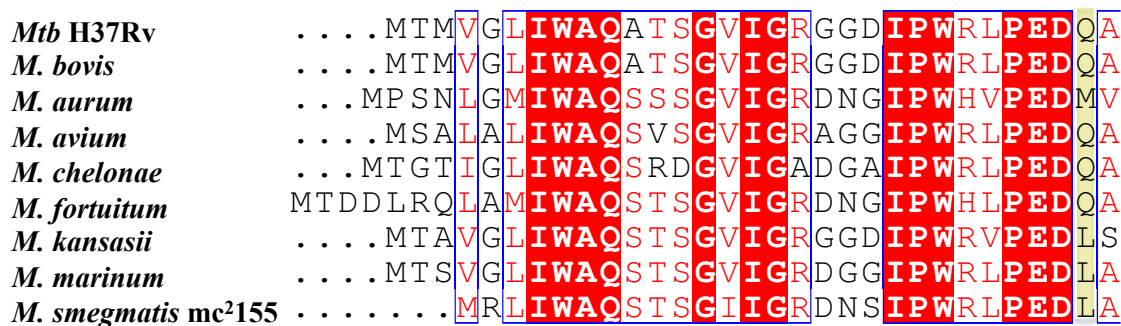


Figure 13. Multiple sequence alignment of a collection of DHFR from different species of *Mycobacterium*. Invariant residues (white on red background) and conservatively substituted residues (red) are indicated. Q28 and its corresponding residues are shaded in yellow. The figure is prepared using ESPript.<sup>161</sup>

PAS may not inhibit the group of DHFR with the leucine, which has neither a hydrogen bond donor nor an acceptor at its side chain to form the hydrogen bond with the ordered water. Therefore, the growth of this group of mycobacterium should be insensitive to PAS. A comparison of the PAS efficacy and the residues at Gln28 for different species is listed in Table 6. There is no significant correlation between the cell growth inhibition and the type of the residue. Particularly, *M. avium*, *M. chelonae*, and *M. fortuitum* have a conserved Gln but show high MIC values to PAS. This result indicates that the Gln28 hydrogen bond network is not the determinant for the selectivity.

The sequence analysis indicates that the current model is inadequate to explain the mechanism of action of PAS. Therefore, PAS may target another pathway such as the iron acquisition pathway, or the interaction between PAS and the folate pathway is more complicated than the current understandings.

Table 6. Whole cell activity of PAS against different species of *Mycobacterium* and the residue of their DHFR corresponding to Gln28 in *Mtb* DHFR

	EC <sub>90</sub> (µg/mL)	Putative water interacting residue
<i>Mtb</i> H37Rv	1.0 <sup>123</sup>	Q (Gln28)
<i>M. bovis</i>	1 <sup>162</sup>	Q
<i>M. aurum</i>	1 <sup>162</sup>	M
<i>M. avium</i>	>256 <sup>163</sup>	Q
<i>M. chelonae</i>	>128 <sup>127</sup>	Q
<i>M. fortuitum</i>	>128 <sup>127</sup>	Q
<i>M. kansasii</i>	>256 <sup>163</sup>	L
<i>M. marium</i>	>256 <sup>163</sup>	L
<i>M. smegmatis</i>	512 <sup>126</sup>	L

## Materials and Methods

### *Bacterial Strains, Media and Growth Condition*

*Mtb* mc<sup>2</sup>7000 strain was used in this study. *Mtb* mc<sup>2</sup>7000 is an *Mtb* strain without the primary attenuating mutation of BCG (RD1) and two genes for the pantothenate biosynthesis (*panC* and *panD*). The *Mtb* mc<sup>2</sup>7000 strain has been approved for use in Biosafety Level 2 laboratory since it fails to kill SCID mice following intravenous challenge and is ultimately cleared. *Mtb* mc<sup>2</sup>7000 was grown in Middlebrook 7H9 media (Difco, Sparks, MD) supplemented with 10% oleic acid-albumine-dextrose-catalase (OADC), 0.2% dextrose, 0.05% (vol/vol) tyloxapol (Sigma) and 25 µg/mL pantothenic acid (Sigma). The antibiotic hygromycin (50 µg/mL) was added as required for maintenance for the pDT plasmid. The culture was also grown in Middlebrook 7H10

media (Difco) supplemented with 0.5% dextrose and 25 µg/mL pantothenic acid when selections were conducted.

#### *Generation and Selection of PAS Resistant Mutants*

*Mtb mc*<sup>2</sup>7000 was first grown in supplemented 7H9 media to an optical density at 600 nm (OD<sub>600</sub>) of ~1.0. Then 1 mL of cells containing broth was plated onto a supplemented 7H10 media plates containing 10 µg/mL PAS (Sigma). The plate was incubated at 37°C and the colonies were picked after 1 month. Selected colonies were cultured in supplemented 7H9 media with 10 µg/mL PAS to confirm their PAS resistance.

#### *DNA Isolation*

To analyze the genomes of the PAS resistant strains, 10 mL of cultures were grown in the supplemented 7H9 media in the presence of 10 µg/mL PAS. Genomic DNAs of the PAS resistant strains were extracted by the cetyltrimethylammonium bromide (CTAB) method<sup>164</sup> when the cultures reached OD<sub>600</sub> of 1.0 and 1 mL of the cultures were flash frozen and stored for further use.

#### *Construction of an N-terminal 6x-His Tagged *ThyA* and *FolC* Overexpression Vector*

Genes of *thyA* and *folC* conjugated with 6x His-tag were first amplified from wild-type *Mtb* H37Rv genomic DNA using PCR with the following primers. Primer 5'-GGGGACAAGTTTGTACAAAAAAGCAG-3' was used for extending the additional



N-terminal segment. Primer pairs 5'-GGCTTCCACCATCATCACCACCATGT-3' and 5'-GGGGACCACTTTGTACAAGAAAGCTG-3' were used for the *thyA* gene. Primer pairs 5'-GCAGGCTTCCACCATCATCACCACCA-3' and 5'-GGGGACCACTTTGTACAAGAAAGCTG-3' were used for the *folC* gene. The genes were cloned into an anhydrotetracycline inducible pDTCF plasmid<sup>165, 166</sup> via a gateway cloning procedure.

#### *Construction of DHFR Overexpression Vector*

A *dfrA* containing pMV261 vector was obtained from Dr. Jacobs. An empty pMV261 vector was constructed to serve as a control. Gene *dfrA* was first removed from the vector by the enzymes, BamH1 and *Hind*III. The empty vector was then religated with a joining linker. The linker primers' sequences are as follows: forward primer: 5'-GGATCCAGCTGCAGAATTCGAAGCTT-3'; reverse primer: 5'-AGCTTCGAATTCTGCAGCTG-3'.

#### *MICs Determination of Mtb mc<sup>2</sup>7000 Overexpression Strains*

*Mtb mc<sup>2</sup>7000* cells were first transformed with overexpression vectors containing, *folC*, *thyA*, and *dfrA*, by electroporation. A single bacterial colony was picked and grown at 37°C to OD600 of 0.6–0.8 in supplemented 7H9 media with either kanamycin (20 µg/mL for pMV261) or hygromycin B (50 µg/mL for pDT system). *FolC*<sup>OX</sup> and *ThyA*<sup>OX</sup> strains were induced with 100 ng/mL anhydrotetracycline. MICs of

different overexpression strains were investigated on supplemented 7H10 plates containing different concentrations of PAS (0 µg/mL to 16 µg/mL).

### *Hydroxy-DHF Modeling Study*

The crystal structure of *Mtb* DHFR in complex with NADPH and methotrexate (1DF7) was used as the original model for the docking study. The downloaded structure was processed using the protein preparation wizard in the Schrodinger software. The protein preparation wizard<sup>167</sup> would add hydrogen atoms to the structure, optimize the hydrogen bond network, and minimize the restraint of the modified structure. Hydroxy-DHF was docked on the active site of *Mtb* DHFR with Glide.<sup>168</sup> A flexible ligand mode was selected, in which the ligand with different conformations were docking into a rigid protein structure. A systematic search was conducted by Glide for an optimal ligand conformation in response to the shape and properties of the binding cavity. The search relied on fitting the docking ligand with various conformations, orientations, and positions into the defined receptor. The information of the docked ligand, including its conformation, orientation and position, corresponding to the receptor is defined as “pose” by the Glide program. The searching can be divided into three steps. The initial step was a rough screening for all available space to the ligand. The pose of the ligand was then minimized within the receptor using the OPLS-AA force field function. Finally, 3 to 6 conformations with the lowest energy were further refined by a Monte Carlo procedure.

## CHAPTER III

# STRUCTURAL INSIGHTS INTO *MYCOBACTERIUM TUBERCULOSIS* RV2671 PROTEIN AS A DIHYDROFOLATE REDUCTASE FUNCTIONAL ANALOGUE CONTRIBUTING TO *PARA*-AMINOSALICYLIC ACID RESISTANCE\*

### Introduction

One of the first drugs developed to treat tuberculosis (TB) was *para*-aminosalicylic acid (PAS), and over 70 years later, it is still in use. While PAS has been replaced as a first line drug by isoniazid, pyrazinamide, and rifampin, it has found new usefulness as a second line drug to treat drug-resistant TB.<sup>122</sup>

PAS is a structural analogue of *p*-aminobenzoic acid (pABA), the substrate of dihydropteroate synthase (DHPS), and for years DHPS was thought to be the target. However, PAS was recently shown to be a prodrug,<sup>152, 153</sup> which is incorporated into the folate pathway through the conversion to hydroxy intermediates. *Mycobacterium tuberculosis* (*Mtb*) cytotoxicity is due to either a general poisoning of the folate-dependent pathway by hydroxy intermediates<sup>152</sup> or the inhibition of dihydrofolate reductase (DHFR) by 12-hydroxy-7,8- dihydrofolate (hydroxy dihydrofolate).<sup>153</sup> In the

---

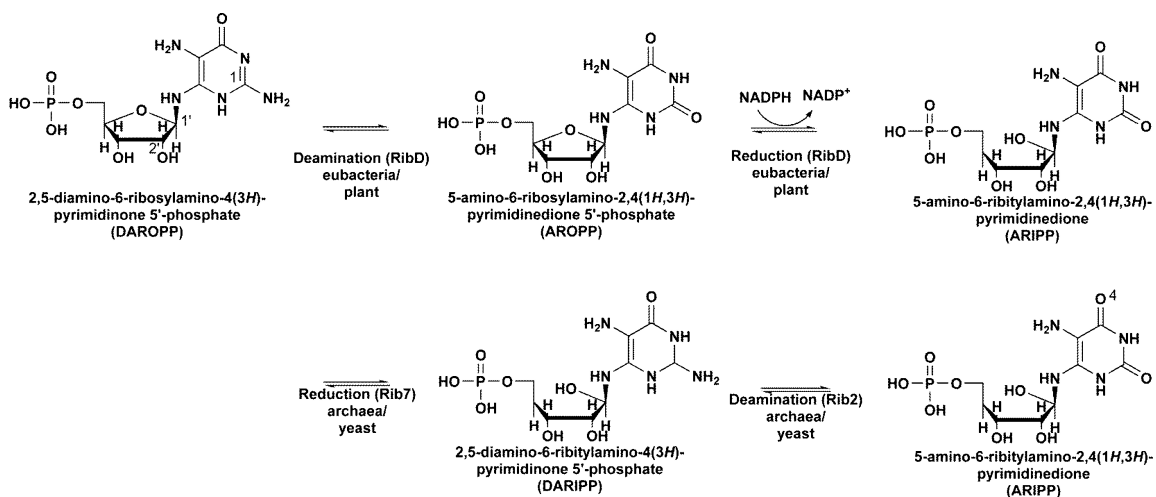
\* This work has been published in and is reprinted with permission from: Cheng, Y. S., and Sacchettini, J. C. (2016) Structural insights into *Mycobacterium tuberculosis* Rv2671 protein as a dihydrofolate reductase functional analogue contributing to *para*-aminosalicylic acid resistance, *Biochemistry* 55, 1107-1119. Copyright 2016 by American Chemical Society.

previously referenced study, a laboratory-derived *Mtb* mutant strain resistant to a DHFR inhibitor (NITD344,<sup>153</sup> a WR99210<sup>169</sup>-analogue ) was found to also be cross-resistant to PAS. Most notably, this same mutation has been reported in 10 out of 208 PAS resistant clinical isolates of *Mtb*, plainly demonstrating its association with PAS resistance *in vivo*.<sup>142</sup> This resistant strain (R7) contained a mutation in a noncoding region upstream of the *Mtb Rv2671*<sup>153</sup> and produced increased levels of *Mtb* RibD (Rv2671). *Mtb Rv2671* is currently annotated as RibD, an enzyme in the riboflavin biosynthetic pathway;<sup>107</sup> therefore, it was surprising to find that increased expression of an enzyme involved in riboflavin biosynthesis would confer resistance to PAS.

In eubacteria RibD, also referred to as RibG in some organisms, is typically a bifunctional enzyme which catalyzes the second and third reactions of the riboflavin biosynthetic pathway, that is, the conversion of 2,5-diamino-6-ribosylamino-4(3*H*)-pyrimidinone 5'-phosphate (DAROPP) to 5-amino-6-ribitylamino-2,4(1*H*,3*H*)-pyrimidinedione 5'-phosphate (ARIPP). The RibD catalyzes the reaction that involves the deamination of the 2-amino group of the pyrimidine ring followed by a reduction of the ribose ring using either NADPH or NADH.<sup>170</sup> In some eubacteria, the two activities are on separate enzymes. *Mtb Rv3752c* is annotated as the deaminase that would convert DAROPP to 5-amino-6-ribitylamino-2,4(1*H*,3*H*)-pyrimidinedione 5'-phosphate (AROPP), and *Mtb Rv2671* is annotated as an AROPP reductase converting AROPP to ARIPP. In fungi and some archaea, the reduction precedes the deamination, and the reactions are catalyzed by two monofunctional enzymes: the reductase Rib7 and the deaminase Rib2, respectively<sup>171, 172</sup> (Scheme 1). Other than *Rv2671* and *Rv3752c*, *Mtb*

*RibG* (*Rv1409*) encodes for a bifunctional RibD containing a DAROPP deaminase domain and an AROPP reductase domain. This is particularly unusual as the reactions, substrates, and products of *Rv1409* should be identical to that of *Rv2671* and *Rv3571c*.

Scheme 1



In order to gain insights into the molecular mechanisms of PAS resistance and define the catalytic activity of *Rv2671*, we have completed a biochemical and structural characterization of *Mtb Rv2671*. The biochemical assays indicated that *Mtb Rv2671* catalyzes the reduction of dihydrofolate (DHF) and the structural studies assisted in defining the molecular basis for the high  $K_m$  for DHF. Our results also explained the mechanism for the resistance to PAS in the R7 and clinical strains. Furthermore, activity assays using purified recombinant *Mtb Rv2671* revealed that it is neither an AROPP nor

a DAROPP reductase, suggesting that Rv2671 and other related orthologues are not involved in riboflavin biosynthesis.

## Results and Discussion

### *Determination of Dihydrofolate Reductase Activity of Rv2671*

The genome sequence of *Mtb* has Rv2671 annotated as an AROPP reductase (RibD) which is part of the riboflavin biosynthetic pathway that catalyzes the reduction of AROPP to ARIPP.<sup>107</sup> Its annotation was most likely due to the sequence similarity analysis showing that Rv2671 has a significantly better sequence similarity to the reductase domains of prokaryotic RibDs (AROPP reductases) (22% identity with *E. coli* RibD) and Rib7s (DAROPP reductases) (22% identity with *Methanocaldococcus jannaschii* Rib7), compared to that of any other enzyme family. The next closest was the dihydrofolate reductase family (14% identity with *E. coli* DHFR). Therefore, it was surprising that the overexpression of Rv2671 was able to replace *Mtb* DHFR function *in vitro*.<sup>153</sup> Specifically, a mutant strain of *Mtb* with increased levels of Rv2671 was shown to be resistant to the DHFR inhibitor NITD344,<sup>153</sup> an analogue of a well-known DHFR inhibitor WR99210.<sup>169</sup> In addition, a multicopy plasmid that expressed Rv2671 was sufficient to allow for the production of a knockout of the *dhfrA* gene,<sup>153</sup> which codes for DHFR and is essential in *Mtb* based on transposon hybridization.<sup>144</sup> Together, these studies suggested the possibility that Rv2671 was a bifunctional enzyme capable of catalyzing the reduction of DHF (Figure 14A) as well as AROPP.

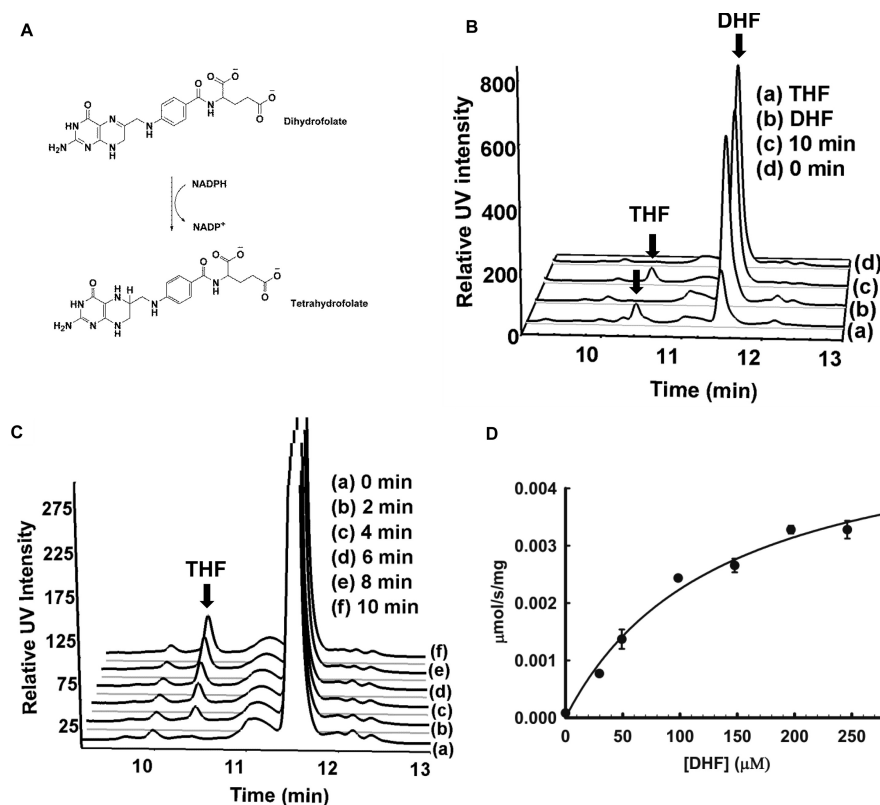


Figure 14. Characterization of Rv2671 as a DHFR. (A) Scheme of the DHF reduction reaction. (B) HPLC analysis of the *in vitro* reaction and the relevant control. (C) Time course for the reaction shows that the amount of THF increases with time. (D) Steady state kinetic reaction measured in the presence of Rv2671 and saturated NADPH. Reaction velocity plotted versus DHF concentration ( $K_m = 136 \pm 32 \mu\text{M}$ ;  $k_{cat} = 0.15 \pm 0.016 \text{ s}^{-1}$ ). Each DHF concentration was performed in triplicate, and the data represent mean values  $\pm$  SD.

We implemented a direct DHFR *in vitro* assay with pure recombinant *Mtb* Rv2671 to determine if Rv2671 could catalyze the reversible reduction of DHF to THF in the presence of NADPH. Tetrahydrofolate, the product of the reaction, was observed by HPLC spectrometry after the addition of DHF and NADPH to Rv2671 (Figure 14B and C). We detected an increase in a single species over time and the HPLC peak comigrated with a tetrahydrofolate standard ( $t_r = 10.4 \text{ min}$ ). The identity was further

confirmed to be tetrahydrofolate by MS. The  $K_m$  and  $k_{cat}$  of Rv2671 for DHF were determined to be  $136 \pm 32 \mu\text{M}$  and  $0.15 \pm 0.016 \text{ s}^{-1}$ , respectively (Figure 14D). A comparison of the kinetic parameters with DHFRs from different organisms is listed in Table 7. Briefly, the  $k_{cat}$  of the DHFRs ranged from 2 to  $30 \text{ s}^{-1}$ , at least 1 order of magnitude faster than the  $k_{cat}$  of Rv2671. Moreover, studied DHFRs have a lower  $K_m$  for DHF, in the range of 0.7 to  $4 \mu\text{M}$ , which is at least 25 times lower than that of Rv2671 ( $136 \pm 32 \mu\text{M}$ ). The  $K_m$  of DHF for *Mtb* DHFR is  $4.2 \pm 6 \mu\text{M}$  with a  $k_{cat}$  of  $2.3 \text{ s}^{-1}$ .<sup>173</sup> The  $K_m$  of DHF is about 30 times lower than that of Rv2671, and the reaction rate is 15 times faster. Therefore, Rv2671 can indeed catalyze the DHF reduction reaction but with less efficiency than the DHFRs of *Mtb* and other organisms.

Table 7. Kinetic parameters of Rv2671 and DHFRs

	$k_{cat} (\text{s}^{-1})$	$K_m (\mu\text{M})$	
		DHF	NADPH
<b>Rv2671</b>	$0.15 \pm 0.016$	$136 \pm 32$	-
<b>Organism</b>			
<i>Mycobacterium tuberculosis</i>	$2.3^{173}$	$1.6^{173}$	$< 1^{174}$
<i>Mycobacterium avium</i>	$20.5^{175}$	$0.7^{175}$	$1.55^{175}$
<i>Escherichia coli</i>	$12.6^{176}$	$0.7^{177}$	$4.8^{177}$
<i>Streptococcus pneumoniae</i>	$31.5^{178}$	$4.4^{178}$	$15.7^{178}$
<i>Mus musculus</i>	$7.2^{179}$	$0.9^{179}$	-
<i>Homo sapiens</i>	$11^{180}$	$0.12^{180}$	$4.2^{180}$



### *Overall Structures of Rv2671 with Co-factor*

The X-ray crystal structure of recombinant *E. coli*-derived Rv2671 was determined to 1.9 Å resolution. The crystal belongs to the orthorhombic space group  $C222_1$  with one molecule in the asymmetric unit. In the final refined structure (Table 8), a total of 246 out of 258 residues were built into the electron density map. Residues 1–9 and 91–95 were disordered in the structure. The structure of Rv2671 is a variant Rossmann fold, composed of eight parallel  $\beta$ -strands ( $\beta 8$ ,  $\beta 7$ ,  $\beta 1$ ,  $\beta 6$ ,  $\beta 2$ ,  $\beta 3$ ,  $\beta 4$ , and  $\beta 5$ ) and one antiparallel  $\beta$ -strand,  $\beta 9$ . This mixed  $\beta$ -sheet is surrounded by seven  $\alpha$ -helices and one  $3_{10}$ -helix ( $\eta 1$ ). Helices  $\alpha 1$  to  $\alpha 5$  are on one side of the central sheet and  $\alpha 6$ ,  $\alpha 7$ , and  $\eta 1$  are on the opposite side (Figure A-1A). Weak electron density for NADP(H) was observed in the active site suggesting NADP(H) copurified with the protein. The electron density of the NADP(H) indicates an occupancy of less than 50% in the crystal. Rv2671 is a homodimer in solution as determined by size-exclusion chromatography (Figure A-1B). In the crystal, the homodimer is generated by 2-fold crystallographic symmetry (Figure 15A). The free energy of assembly dissociation ( $\Delta G^{\text{diss}}$ ) is 44.3 kcal/mol; the positive value suggests that this dimer is thermodynamically stable. The buried surface area of the dimer interface accounts for 37% of the total surface area.<sup>181</sup> This large dimer interface contains residues from three loops (L1, L3, and L15) and four  $\beta$ -strands ( $\beta 1$ ,  $\beta 7$ ,  $\beta 8$ , and  $\beta 9$ ).

Table 8. Data collection and refinement statistics

	copurified NADP(H) bound Rv2671	NADP <sup>+</sup> bound Rv2671	THF bound Rv2671
<b>PDB ID</b>	4XRB	4XT5	4XT6
<b>Data collection</b>			
<b>Space group</b>	<i>C</i> 222 <sub>1</sub>	<i>C</i> 222 <sub>1</sub>	<i>C</i> 222 <sub>1</sub>
<b>Cell Dimensions</b>			
<b><i>a</i>, <i>b</i>, <i>c</i> (Å)</b>	71.4, 96.6, 75.7	69.7, 95.3, 73.9	70.5, 95.9, 76.1
<b><i>α</i>, <i>β</i>, <i>γ</i> (°)</b>	90, 90, 90	90, 90, 90	90, 90, 90
<b>Resolution (Å)<sup>a</sup></b>	27.4-1.9 (1.93-1.9)	44.75-2.13 (2.17-2.13)	36.61-1.85 (1.88-1.85)
<b><i>R</i><sub>sym</sub> or <i>R</i><sub>merge</sub></b>	0.045 (0.290)	0.069 (0.418)	0.053 (0.274)
<b><i>I</i>/σ(<i>I</i>)</b>	30.97 (8.33)	26.80 (7.02)	34.12(9.68)
<b>Completeness (%)</b>	99.36 (100)	99.96 (99.71)	99.34 (97.58)
<b>Redundancy</b>	7.1 (7.3)	9.0 (8.9)	8.7(9.0)
<b>Refinement</b>			
<b>No. reflections</b>	149934	127251	194557
<b><i>R</i><sub>work</sub>/ <i>R</i><sub>free</sub></b>	0.18/0.20	0.19/0.24	0.20/0.24
<b>No. atoms</b>	1974	1975	1919
<b>Protein</b>	1851	1863	1799
<b>Ligand/ion</b>	55	48	44
<b>Water</b>	68	64	73
<b>B-factors</b>	36.4	37.9	30.9
<b>Protein</b>	36.0	37.9	30.6
<b>Ligand/ion</b>	48.3	35.6	35.8
<b>Water</b>	38.9	38.9	33.7
<b>R.m.s deviations</b>			
<b>Bond lengths (Å)</b>	0.018	0.013	0.015
<b>Bond angles (°)</b>	1.96	1.36	1.24
<b>Ramachandran favored (%)</b>	99	97	97
<b>Ramachandran allowed (%)</b>	1	3	3
<b>Ramachandran outliers (%)</b>	0	0	0

Table 8. Data collection and refinement statistics (cont.)

	TMP bound Rv2671	TMQ bound Rv2671	DHP bound Rv2671
<b>PDB ID</b>	4XT7	4XT8	4XT4
<b>Data collection</b>			
<b>Space group</b>	<i>C222<sub>1</sub></i>	<i>C222<sub>1</sub></i>	<i>C222<sub>1</sub></i>
<b>Cell Dimensions</b>			
<i>a, b, c</i> (Å)	69.0, 93.2, 74.6	69.4, 93.6, 74.3	70.3, 94.4, 75.8
<i>α, β, γ</i> (°)	90, 90, 90	90, 90, 90	90, 90, 90
<b>Resolution</b> (Å) <sup>a</sup>	46.56-2.3 (2.34-2.3)	46.81-1.95 (1.98-1.95)	45.24-1.89 (1.92-1.89)
<i>R</i> <sub>sym</sub> or <i>R</i> <sub>merge</sub>	0.081 (0.331)	0.08 (0.288)	0.070 (0.585)
<i>I</i> / <i>σ</i> ( <i>I</i> )	21.31 (6.50)	34.82 (12.88)	21.72 (2.42)
<b>Completeness</b> (%)	99.94 (99.36)	99.89 (99.72)	99.43 (94.66)
<b>Redundancy</b>	7.3 (7.4)	8.6 (8.8)	9.3 (6.6)
<b>Refinement</b>			
<b>No. reflections</b>	81136	155567	189305
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.17/0.22	0.18/0.22	0.19/0.22
<b>No. atoms</b>	2006	2072	2002
<b>Protein</b>	1858	1885	1864
<b>Ligand/ion</b>	82	103	56
<b>Water</b>	66	84	82
<b>B-factors</b>	24.4	23.9	33.8
<b>Protein</b>	24.3	23.6	33.5
<b>Ligand/ion</b>	25.9	24.6	36.1
<b>Water</b>	25.5	28.2	38.2
<b>R.m.s deviations</b>			
<b>Bond lengths</b> (Å)	0.054	0.051	0.011
<b>Bond angles</b> (°)	1.43	1.43	1.36
<b>Ramachandran favored</b> (%)	98	98	97
<b>Ramachandran allowed</b> (%)	2	2	3
<b>Ramachandran outliers</b> (%)	0	0	0

<sup>a</sup> Resolution values for collection and refinement are the same.

The intersubunit interactions are dominated by hydrogen bonds and electrostatic interactions, with roughly 56 hydrogen bonds and 8 electrostatic interactions across the

interface. Each subunit of the homodimer contains one active site. The oligomeric state of Rv2671 is consistent with that of other RibD/Gs and Rib7s, except *Bacillus subtilis* RibG (*Bs* RibG), which is reported to be a tetramer in solution.<sup>182-185</sup> In contrast, most DHFRs are monomeric.<sup>186</sup>

The crystals of Rv2671 preincubated with NADP<sup>+</sup> were nearly isomorphous with the natively purified Rv2671 crystals and the structure was refined with diffraction data extending to 2.1 Å resolution (Table 8). The overall fold of the protein is highly conserved, with the partial NADP(H) bound structure having a root-mean-square deviation (rmsd) of 0.42 Å over 243 C $\alpha$  pairs. The disordered segment (residues 91–95) of the partial NADP(H) bound structure was clear in the electron density map and formed a loop in the NADP<sup>+</sup> bound structure. With the exception of the loop, NADP<sup>+</sup> bound Rv2671 exhibits no significant structural differences to the partial NADP(H) bound structure. This loop is near the NADP<sup>+</sup> binding site and interacts directly with the cofactor and is stabilized by both a hydrogen bond between the side chain oxygen of Glu91 and the 2' hydroxyl group of the nicotinamide ribose group (2.6 Å), and multiple van der Waals interactions (Figure 15B). The NADP<sup>+</sup> binds at the C-terminal edge of the central  $\beta$ -sheet ( $\beta$ 2,  $\beta$ 3, and  $\beta$ 6) and the pyrophosphate group locates between the  $\alpha$ 3 and  $\alpha$ 6 adjacent to the putative substrate binding cleft formed by  $\alpha$ 2,  $\alpha$ 3,  $\beta$ 2, and  $\beta$ 3 (Figure 15B). The conformation of NADP<sup>+</sup> and its binding site is very similar to that of RibD/G structures, as well as DHFR structures.<sup>184</sup>

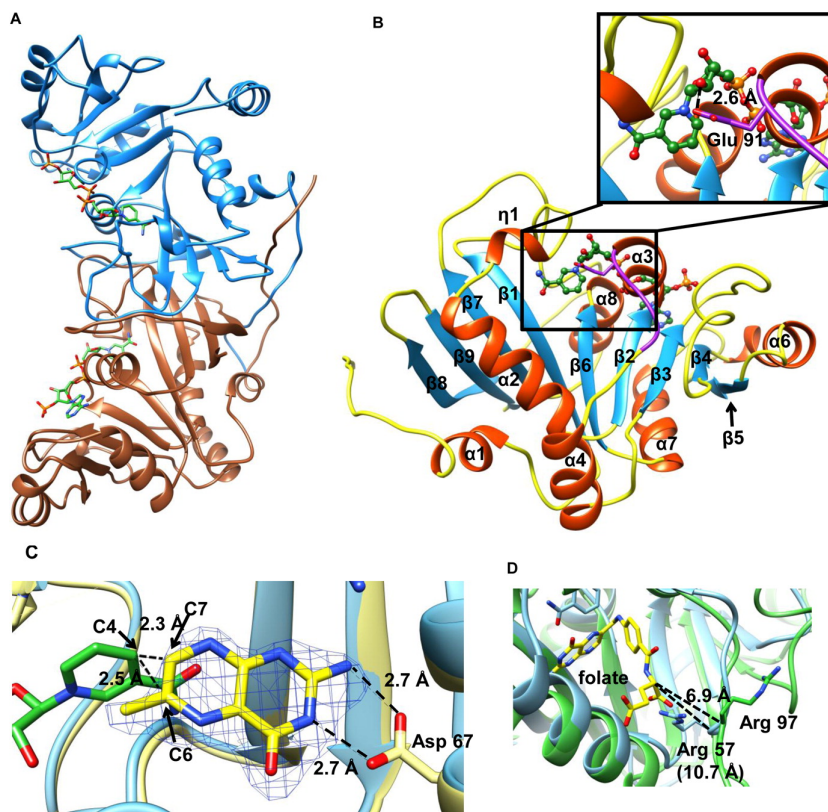


Figure 15. Crystal structure of Rv2671 and the THF binding mode. (A) Ribbon representation of the dimeric Rv2671. (B) Ribbon representation of the Rv2671 in complex with NADP<sup>+</sup>. The secondary structural elements are labeled sequentially. The loop (residue 91–95) disordered in the native purified structure is shown in purple, and NADP<sup>+</sup> is shown in the green ball and stick model. The hydrogen bond interaction between the ribose group of NADP<sup>+</sup> and Glu 91 in the loop (residue 91–95) is shown in the zoomed-in figure. (C) Superimposition of the THF bound Rv2671 structure (yellow) with the NADP<sup>+</sup> bound Rv2671 structure (blue). NADP<sup>+</sup> and the pteridine ring of THF are colored in green and yellow. The simulated annealing omit map contoured at 2.5 $\sigma$  around the pteridine ring is shown in blue. Distances between the C4 of nicotinamide and the C6 and C7 of the pteridine are labeled. (D) Superimposition of the NADP<sup>+</sup> bound Rv2671 structure (green) with the folate-NADPH bound *E. coli* RibD (4pdj, blue). Folate is colored in yellow. The side chain of the Arg 97 in the Rv2671 complex structure points to the other direction against the folate. The distances between  $\alpha$ -carbon of the glutamate group in folate and the main chain carbon of the arginines are 6.9 and 10.7 Å, respectively.

### *Complex Structure of Rv2671 with THF*

In order to understand the molecular basis for the relatively low DHFR activity of Rv2671, the structure of Rv2671 ternary complex with the products, NADP<sup>+</sup> and THF was solved. The crystals diffracted to about 2.1 Å resolution (Table 8). The backbone C $\alpha$  rmsd between 236 atom pairs is 0.49 Å with the NADP<sup>+</sup> bound structure. The electron density was not visible for residues 90–95, residues 56–57, or the nicotinamide ribose moiety of the NADP<sup>+</sup> in the THF bound structure. Notably, the disordering of the nicotinamide ribose moiety of the NADP<sup>+</sup> was also observed in a recent *Mtb* DHFR structural report<sup>187</sup> which showed that the density for the nicotinamide ribose moiety of NADPH was missing in the open conformation of the *Mtb* DHFR-NADPH complex structure.<sup>187</sup> We observed the same phenomenon in Rv2671. After soaking NADP<sup>+</sup> bound Rv2671 crystals with THF, the density for the nicotinamide ribose moiety disappeared along with the density for the adjacent loop residues 90–95. These residues correspond to loop2 (L2), a part of the ligand access control segment in *Mtb* DHFR,<sup>187</sup> suggesting that residues 90–95 serve the same function in Rv2671.

Unambiguous electron density for the tetrahydropteridine ring of THF was observed and built into the electron density (Figure 15C). However, electron density for the pABG group (pABA and the glutamate tail) of THF was not visible. NA2 and N3 of the pteridine ring interact with the side chain of Asp67 (2.7 and 2.7 Å), which is typically observed in DHF-complexed DHFR structures. Although the nicotinamide ribose moiety is disordered in the ternary structure, after we overlaid it with the NADP<sup>+</sup> Rv2671 structure the distances between the nicotinamide C4 and the pteridine C6 and

C7, where the hydride transfer would occur, were measured to be 2.5 and 2.3 Å (Figure 15C). Similar positions for these groups have been observed in the *E. coli* DHFR–DHF complex structure.<sup>188</sup> Upon the basis of the theoretical calculation, 2.6 Å is the optimal distance between the donor and acceptor atoms, while the hydride transfer occurs via a bent geometry.<sup>189, 190</sup>

The missing density for the pABG group in the ternary structure suggests that Rv2671 does not have the optimal amino acids to mediate its binding. In all DHF–DHFR complex structures, a conserved arginine forms an electrostatic interaction with the  $\alpha$ -carboxyl group of the glutamate tail. In the *E. coli* folate–DHFR structure (4pdj),<sup>191</sup> Arg57 corresponds to the conserved arginine in DHFRs. Superimposition of Rv2671 and *E. coli* folate-DHFR structures put *E. coli* folate–DHFR Arg57 at the cleft of loop 6 (L6) and loop 7 (L7) in Rv2671 (Figure 15D). The main chain of Arg97 at L6 in Rv2671 is in close proximity to that of *E. coli* DHFR Arg57. However, Arg97 is 3 Å closer to the  $\alpha$ -carboxyl group and points in the opposite direction against the DHF (Figure 15D). Therefore, it is unlikely that this Arg97 can fulfill the role of stabilizing the  $\alpha$ -carboxyl group of DHF. The position of Arg97 may be responsible for the 30 times higher  $K_m$  for Rv2671 compared to that of other DHFRs. The potential pABG binding cavity is also larger in Rv2671 (16 Å long, 12.5 Å wide) than it is in *Mtb* DHFR (17 Å long, 7.5 Å wide). This could reduce the affinity of pABG for Rv2671 but may also provide Rv2671 with the ability to bind another ligand.

Dihydropteroic acid (DHP) was tested as a potential substrate for Rv2671 since DHP contains a dihydropteridine ring and a benzoic acid group like THF, but does not

have the glutamate tail. The  $K_m$  of Rv2671 for DHP was  $180.2 \pm 25.5 \mu\text{M}$  with a  $k_{cat}$  of  $0.32 \pm 0.016 \text{ s}^{-1}$ , unexpectedly similar to the  $K_m$  and  $k_{cat}$  for DHF. Therefore, the crystal structure of DHP complexed to Rv2671 was solved and refined to 1.89 Å resolution (Table 8) to investigate the binding mode of DHP. The overall structure was nearly identical to the THF bound structure with an rmsd of 0.36 Å over 236 Cα pairs. The location of the dihydropteridine ring of DHP was also identical to the tetrahydropteridine ring of THF. The benzoic acid in the DHP and the nicotinamide ribose moiety of NADP<sup>+</sup> are disordered, as in the THF bound structure. The structural similarity and the missing benzoic acid density explain the similar steady state kinetic parameters for DHF and DHP. This result also suggests that Rv2671 can bind and reduce a broad spectrum of dihydropteridine containing analogues.

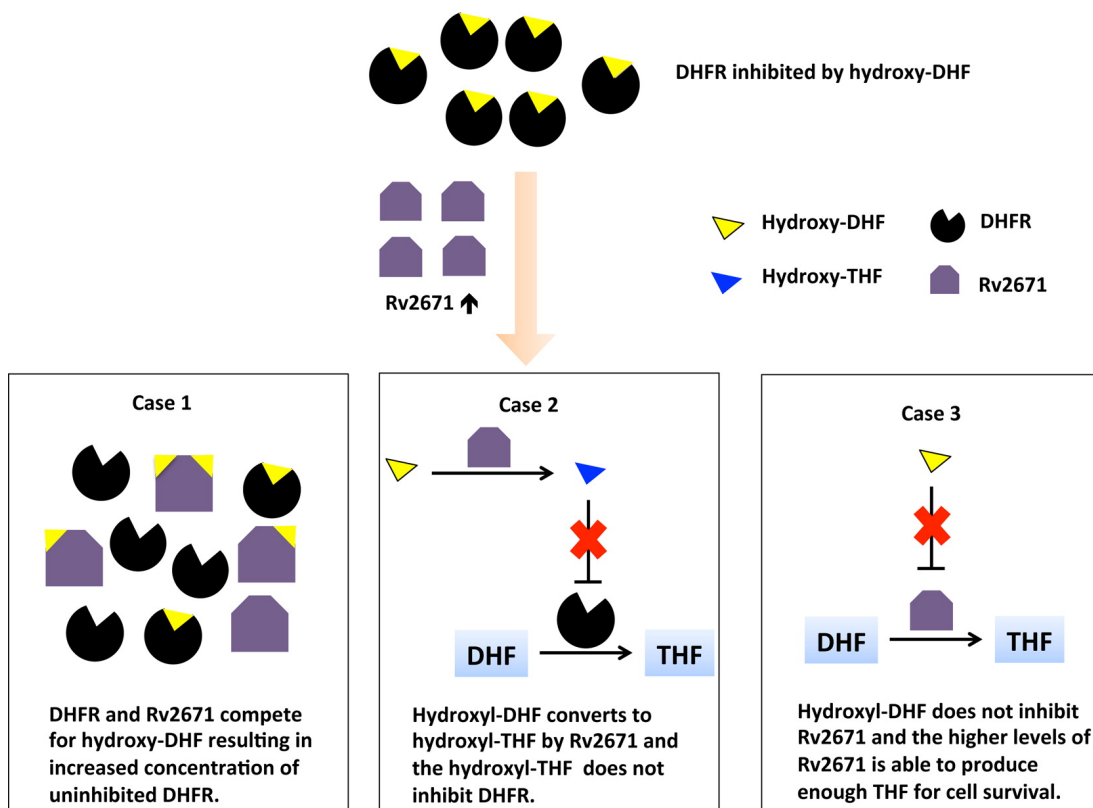
As previously stated, the drug PAS is converted in the cell to hydroxylated intermediates by enzymes in the folate biosynthetic pathway.<sup>152, 153</sup> The ability of Rv2671 to catalyze the reduction of DHF suggests several potential mechanisms for the observed resistance to PAS (Scheme 2). One scenario is that Rv2671, like DHFR, binds and is inhibited by the 12-hydroxy-7,8-dihydrofolate (hydroxy dihydrofolate) (Scheme 2, Case 1). In this case, the higher levels of Rv2671 would serve to sequester the inhibitor that would otherwise bind to DHFR within the cell. However, due to the high  $K_m$  and inability to see in the structure, our studies indicate that this scenario is unlikely to be the reason for resistance since the lack of affinity of Rv2671 for DHF suggests that hydroxy dihydrofolate would likely prefer binding with *Mtb* DHFR instead of Rv2671. A second and more likely scenario is that the hydroxy dihydrofolate is a substrate of



Rv2671 based on the less specific dihydropteridine reductase activity of Rv2671 (Scheme 2, Case 2). Thus, the higher levels of Rv2671 would effectively reduce the levels of DHFR inhibitor (hydroxy dihydrofolate) in the cell by converting it to the product hydroxy tetrahydrofolate. A third possibility is that Rv2671 does not bind to the hydroxy dihydrofolate (Scheme 2, Case 3). For this scenario, Rv2671 serves as the functional homologue of DHFR and overrides DHFR's inhibition by hydroxy dihydrofolate. It has been shown that *Mtb* can survive in a strain where > 97% of the DHFR enzyme had been depleted.<sup>192</sup> Therefore, the higher levels of Rv2671 would be able to sustain folate biosynthesis at a level sufficient to support *Mtb* even though Rv2671 has only about 10% of the wild-type DHFR activity.

Another outstanding question is whether PAS's cytotoxicity results from a general poisoning of folate-dependent pathways<sup>152</sup> or if it acts specifically on DHFR through the conversion of PAS to hydroxy dihydrofolate.<sup>153</sup> The nonspecific dihydropteridine reductase activity of Rv2671 indicates that the hydroxy dihydrofolate is likely a substrate of Rv2671 capable of converting a hydroxy dihydrofolate into a hydroxy tetrahydrofolate. This would mean that higher levels of Rv2671 in the cell would lead to the production of more hydroxy-folate intermediates which would in turn cause higher levels of sensitivity to PAS, if indeed PAS's cytotoxicity results from the assumption of general poison. Since this was not the case, and the cells were resistant to PAS when Rv2671 was at higher levels, the data suggest that hydroxy dihydrofolate inhibition of DHFR is the event that leads to cell death.

Scheme 2



### *Potential of Other Riboflavin Biosynthesis Reductases to Carry out DHFR's Function*

Our discovery that Rv2671 has DHFR activity led us to test to see if DHFR activity is a general feature of the AROPP reductase domains of RibD/Gs. DHFR activity was not detected for the recombinant *E. coli* RibD (the AROPP reductase domain shows 22% sequence identity to Rv2671). This is not surprising since overexpression of *Mtb* RibG with both DAROPP deaminase domain and AROPP reductase domain was unable to replace the function of *Mtb* DHFR.<sup>153</sup> In all known AROPP reductase sequences, there is a conserved lysine which interacts with the keto

group of the pyrimidine ring of the AROPP in the active site.<sup>182-185</sup> In the *Bs* RibG complex structure, the side chain of Lys151 forms a hydrogen bond with the keto group of the AROPP (3.4 Å).<sup>193</sup> Superimposing *E. coli* RibD (2OBC)<sup>184</sup> with Rv2671 shows that Asn44 in Rv2671 is in place of the aforementioned lysine (Lys152 in *E. coli* RibD) (Figure A-2). The side chain carbonyl group of Asn44 forms a hydrogen bond with the pteridine NA2 of THF (3.1 Å) in the Rv2671 structure. A clash of the lysine side chain with the THF pteridine ring was observed when superimposing the structures. Our results indicate that bacterial AROPP reductases are unable to catalyze the reduction of DHF. This is consistent with the inability of bifunctional *Mtb* RibG (Rv1409) to compensate for a loss in DHFR activity.<sup>153</sup>

*DHFR Inhibitors, Trimethoprim (TMP) and Trimetrexate (TMQ), Bind and Inhibit the DHFR Activity of Rv2671 in Vitro*

Two FDA approved DHFR inhibitors, TMP and TMQ, were tested for inhibitory activity against Rv2671. TMP is a selective DHFR inhibitor which typically shows about 2,500 times more potency against bacterial DHFR over human DHFR.<sup>194</sup> TMQ is a DHFR inhibitor used in cancer treatment<sup>195</sup> and does not show selectivity between bacterial and mammalian DHFRs.<sup>196</sup> A 350-fold difference was observed in the IC<sub>50</sub> values of TMP (Figure 16A) and TMQ (Figure 16B) to Rv2671 (93.5 ± 5.2 μM and 0.28 ± 0.02 μM, respectively). Although TMP is usually selective for bacterial DHFRs, it shows a low inhibition of not only Rv2671 but also *Mtb* DHFR (IC<sub>50</sub> for TMP is 16.5 μM<sup>197</sup>).

The structures of Rv2671 in complex with NADP<sup>+</sup> and either TMP or TMQ (Figure 16C) revealed the molecular basis for the differences in inhibitory activity. The overall structures of the protein with inhibitors are similar to the NADP<sup>+</sup> bound structure. For both inhibitors, hydrogen bonding interactions with Rv2671 appear to be the dominant binding interaction for the 2,4-diaminopyrimidine and the quinazoline group. The 2,4-diaminopyrimidine group of TMP forms six hydrogen bonds with residues of Rv2671 including Asn44, Asp67, Thr214, and a water molecule (Figure 16D). However, the quinazoline ring of TMQ forms seven H-bonds with residues of Rv2671 including Asn44, Phe45, Asp67, Glu193, Thr214, and Glu91 (Figure 16E).

When the two structures are superimposed, the 2,4-diaminopyrimidine group of TMP overlaps and is nearly coplanar (about 5 degrees out of plane) with the 2,4-diaminopyrimidine part of TMQ. The slight rotation results in one extra hydrogen bond between Rv2671 and TMQ, but the binding site of the two rings is identical. A more detailed analysis of the hydrogen bonding interactions between Rv2671 and the ligands is provided in Table A-1.

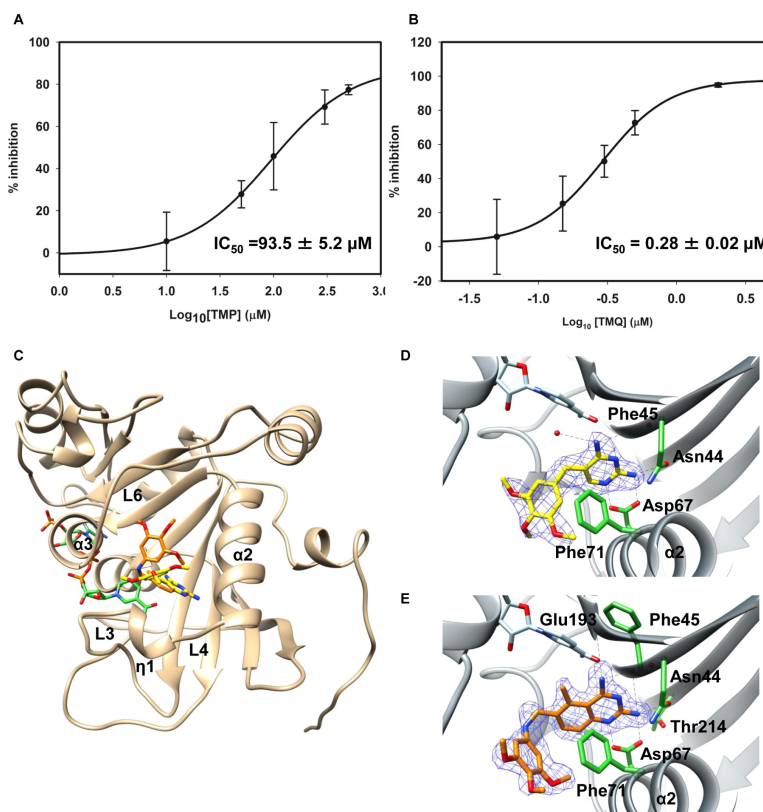


Figure 16. Crystal structure of Rv2671 in complex with TMP or TMQ. (A) The mean % inhibition values of Rv2671 are plotted against the  $\text{log}_{10}$  concentration of the DHFR inhibitor TMP. The calculated  $\text{IC}_{50}$  value is  $93.5 \pm 5.2 \mu\text{M}$ . (B) The mean % inhibition values of Rv2671 are plotted against the  $\text{log}_{10}$  concentration of the DHFR inhibitor TMQ. The calculated  $\text{IC}_{50}$  value is  $0.28 \pm 0.02 \mu\text{M}$ . (C) Ribbon representation of Rv2671 (tan) in complex with TMP (yellow) and TMQ (orange). (D) Binding of the TMP with Rv2671. The simulated annealing omit map contoured at  $2.5\sigma$  around the inhibitor TMP is shown in blue. The key residues interacting with TMP are labeled and shown as sticks in green. The hydrogen bonds are shown as dashed lines. (E) Binding of TMQ with Rv2671. The simulated annealing omit map contoured at  $2.5\sigma$  around the inhibitor TMQ is shown in blue. The main residues interacting with TMP are labeled and shown as green sticks. Different binding modes of the trimethoxybenzyl rings of TMP and TMQ in Rv2671 can be observed in D and E. The trimethoxybenzyl ring of TMP in the complex structure is perpendicular to Phe71. The trimethoxybenzyl ring of TMQ in the complex structure is parallel to Phe71.

The trimethoxybenzyl groups of TMP and TMQ were found to be in different binding orientations with Rv2671 (Figure 16C) due to the extra secondary amide in the linker region of TMQ which allows the trimethoxybenzyl ring to flip and move away from  $\alpha 2$ . The ring plane of the trimethoxybenzyl group of TMP is perpendicular to that of Phe71 in Rv2671 (Figure 16D), but the trimethoxybenzyl ring of TMQ in Rv2671 is parallel to the ring plane of Phe71 (Figure 16E). Different trimethoxybenzyl ring binding conformations result in different binding cavities, and the trimethoxybenzyl groups hold only via the van der Waals interactions with Rv2671, especially the hydrophobic effects. Since the hydrophobic interactions are the major contacts between the trimethoxybenzyl ring and the protein, the hydrophobicity of the ligand and its binding cavity was analyzed using the PLATINUM server, which calculates the hydrophobic properties of molecules and evaluates their match or mismatch in the protein–ligand complexes.<sup>198</sup> The calculated hydrophobic match surface ( $S_{LL}$ ) between TMP and the surrounding residues is  $0.06 \text{ \AA}^2$  compared with  $34.24 \text{ \AA}^2$  between TMQ and its environment, and the fraction of lipophilic match ( $\text{Match}^2$ ) of TMP is 0.0007 compared with 0.2175 of TMQ. The higher match numbers of TMQ indicate better hydrophobic contacts between TMQ and Rv2671. The additional hydrogen bond and better hydrophobic interactions with the trimethoxybenzyl ring results in significantly better inhibition by TMQ compared with that of TMP.

*Structural Comparison with Known RibDs/Rib7s from Other Organisms Reveals Differences in the Active Sites*

A structural similarity analysis using DALI<sup>199</sup> identified a number of closely related homologues of Rv2671, and over 700 entries were identified with a Z-score above 2. Out of the 700 entries, 7 pyrimidine reductases were found to be the closest structural homologues with Z-scores ranging from 16 to 27 and rmsd of C $\alpha$  ranging from 2.1 to 2.6 Å, 4 of which were RibD/Gs, 1 was a Rib7, and 2 were putative pyrimidine reductases from different species. DHFRs also exhibited a significant match with 34 dihydrofolate reductases or bifunctional dihydrofolate reductase–thymidylates with Z-scores ranging from 11.4 to 18.7 and rmsd ranging from 1.8 to 2.9 Å. A pyrimidine reductase-like protein from *C. diphtheria* (2p4g) belonging to the monofunctional RibD\_C family is the closest structural homologue of Rv2671 (Z-score of 27.1 and rmsd of 2.6 Å).

The AROPP reductase domain of RibD/Gs and Rib7s (DAROPP reductases) are often classified as pyrimidine reductases with identical reaction mechanism, overall structure, and substrate binding residues.<sup>200</sup> 2,5-Diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate (DAROPP) is the Rib7's substrate and 5-amino-6-ribosylamino-2,4(1H,3H)-pyrimidinedione 5'-phosphate (AROPP) is the substrate of the reductase domain of RibD/Gs (Scheme 1).<sup>171, 172</sup> DAROPP and AROPP differ at the C2 position of the pyrimidine ring; DAROPP has an amino group in place of AROPP's keto group. The distinct difference of RibD/Gs and Rib7s is a conserved lysine which interacts with the keto group of AROPP in the RibD/Gs. Pyrimidine reductases and

Rv2671 share a common core structure including the central  $\beta$ -sheet and the flanking  $\alpha$ -helices. The N-terminal region and  $\alpha 6$  (in Rv2671) show the most structural diversity. For RibD/Gs, the N-terminus of the reductase domain connects to the deaminase domain.<sup>184, 185</sup> For Rib7s and Rv2671, the N-terminal region participates in the dimerization of two subunits. Helix  $\alpha 6$  in Rv2671 is a loop in RibD/Gs, and in Rib7s the  $\alpha 6$  helices are located in different spatial positions depending on the length of the adjacent cofactor binding loop.

*Bs* RibG in complex with ([*(2R,3S,4R,5E)*-5-[(5-amino-2,6-dioxo-3*H*-pyrimidin-4-yl)imino]-2,3,4-trihydroxy-pentyl] dihydrogen phosphate (AIF) is the only structure available in the protein databank (PDB) of a RibG with a substrate (analogue in this case) of a pyrimidine reductase.<sup>193</sup> AIF is a proposed intermediate of the AROPP reduction reaction. The structures of Rv2671 and the *Bs* RibG-AIF complex were superimposed using the Matchmaker function of Chimera.<sup>201</sup> The two structures were very similar to an rmsd of 1.9 Å over 202 C $\alpha$  pairs. AIF was located at the cleft created by L3- $\eta 1$ -L4, L6,  $\alpha 2$ ,  $\alpha 3$ , and the central  $\beta$ -sheet;  $\alpha 2$ ,  $\alpha 3$  and the central  $\beta$ -sheets were in similar spatial positions between the two structures, but the L6 loop of Rv2671 was 2.8 Å away from the phosphate group of AIF and the region between L3- $\eta 1$ -L4 was 4.7 Å closer to  $\alpha 2$  compared to the analogous regions of *Bs* RibG.

The enzymatic reaction of pyrimidine reductases is initiated by a conserved ionizable residue to abstract the proton from the amine group next to C1 of DAROPP/AROPP to form a Schiff base intermediate at C1, followed by a direct hydride transfer to C1 from the nicotinamide group of NADPH.<sup>183, 184, 193</sup> The amino acids



proposed for the proton abstraction are Asp200 in *E. coli* RibD and Glu290 in *Bs* RibG. In the structure of Rv2671, Glu193 corresponds to *Bs* RibG Glu290, and Glu91 corresponds to Asp200 of *E. coli* RibD. Although the length of the side chains of Glu and Asp differ by one carbon, the distance between the side chain oxygen of Glu91 and the amine nitrogen beside the C1 of the modeled AIF is 2.8 Å away, within proton transfer distance. However, based upon the superimposed structures, two differences in the active site could interfere with the binding of DAROPP/AROPP to Rv2671. First, there could be a clash between the pyrimidine ring and the backbone of Rv2671. As stated earlier, the location of L3-η1-L4 is different between Rv2671 and *Bs* RibG, and the backbone of L3-η1 in Rv2671 showed the potential for overlap with the pyrimidine ring of AIF in the superimposed structure. Although the possibility of main chain conformational changes cannot be ruled out, the distance between the main chain carbonyl group of Ser59 of Rv2671 and the O4 of the AIF from the *Bs* RibG structure is 1.14 Å. Second, several conserved residues that are involved in substrate binding among functionally characterized pyrimidine reductases are not present in Rv2671 (Figure 17); for example, Thr171, Arg183, Asp199, Arg206, and Glu290 interact directly with AIF in *Bs* RibG and are not conserved in Rv2671. The backbone of Ala153 in Rv2671 overlaps with the side chain of Thr171 in *Bs* RibG. The pyrimidine ring of AIF forms hydrogen bonds with the hydroxyl group of the side chain (2.1 Å) and the main chain nitrogen (3.1 Å) of Thr171 in *Bs* RibG. The main chain of Glu91 in Rv2671 is in the position of the main chain of Asp199 in *Bs* RibG; the side chain oxygen of Asp199 forms hydrogen bonds with two ribosyl hydroxyl groups of AIF (2.8 Å and 3.1 Å). The side chains of

Arg183 and Arg206 in *Bs* RibG interact with the phosphate of AIF by hydrogen bonds and electrostatic interactions. Arg75 in Rv2671 was on top of Arg183 in *Bs* RibG. Met98 of Rv2671 corresponds to Arg206 in *Bs* RibG, but the distance between the two main chain carbons was 3.3 Å since Met98 is located on L6. The differences between the substrate binding residues, especially missing Thr171 and Arg206, may cause Rv2671 not to bind to DAROPP or AROPP.

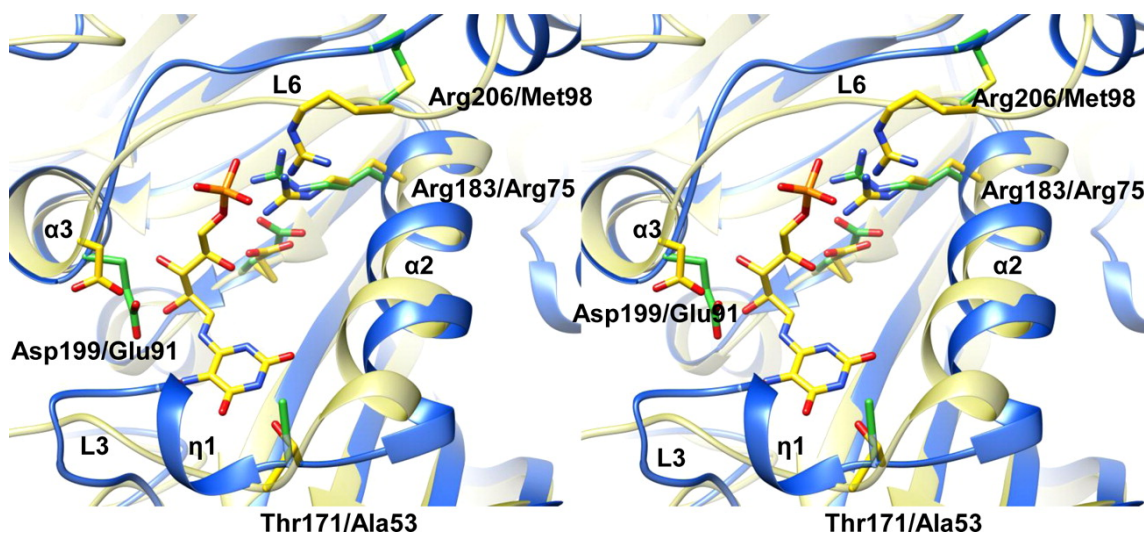


Figure 17. Residue differences between the substrate binding sites of Rv2671 and *Bs* RibG. Superposition of NADP<sup>+</sup> bound Rv2671 (blue) with the AIF bound *Bs* RibG (yellow) is shown in a stereo view. AIF and the residues directly interacting with AIF in *Bs* RibG are shown as yellow sticks, and the corresponding residues in Rv2671 are shown as green sticks.

#### *Rv2671 Is Not a DAROPP or AROPP Reductase*

As discussed above, the annotation and structural similarity suggest two possible biological substrates for Rv2671: DAROPP or AROPP. Due to the instability of

DAROPP and AROPP, these substrates were prepared enzymatically by *E. coli* RibA and *E. coli* RibD for the characterization of the pyrimidine reductase activity of Rv2671. Rv2671 showed no detectable reductase activity for either DAROPP or AROPP with different enzyme concentrations as well as substrate concentrations (Figure A-3). As a positive control, pure recombinant *E. coli* RibD showed both deaminase and reductase activity under the reaction condition described in the Experimental Procedure section. Rv2671's inability to catalyze the reduction of either DAROPP or AROPP under the tested conditions suggests that it is not a pyrimidine reductase and therefore not involved in the riboflavin biosynthetic pathway.

The evolutionary relationship between Rv2671, pyrimidine reductases, and DHFRs are apparent in the phylogenetic tree (Figure 18) constructed with 9 pyrimidine reductases from *Actinobacteria* which showed >50% sequence similarity to Rv2671, 15 DHFRs, 10 functionally identified RibDs or Rib7s, and 4 enoyl ACP reductases used as the out-group control. The full tree contained three well-supported monophyletic groups, enoyl ACP reductases, DHFRs, and RibDs/Rib7s. Although Rv2671 is part of the RibDs/Rib7s superfamily, the Rv2671-like pyrimidine reductases from *Actinobacteria* formed an independent clade within the RibD/Rib7 family. These enzymes have a conserved Asp (Asp67 in Rv2671) for pteridine recognition and a Glu (Glu91 in Rv2671) in place of the catalytically critical Asp of the pyrimidine reductases (Figure A-4). This indicates that this group of enzymes, with high sequence similarity to Rv2671 from *Actinobacteria*, is likely misannotated. The ability of Rv2671 to catalyze the reduction of dihydrofolate to tetrahydrofolate or dihydroptericoic acid to tetrahydroptericoic

acid further suggests that this particular group of proteins appears to have a fairly nonspecific dihydropteridine reductase activity. Moreover, the inability of *Mtb* Rv2671 to function as an AROPP reductase suggests that *Mtb* RibG (Rv1409) is likely to be the true AROPP reductase that generates ARIPP in the riboflavin biosynthetic pathway of *Mtb*.

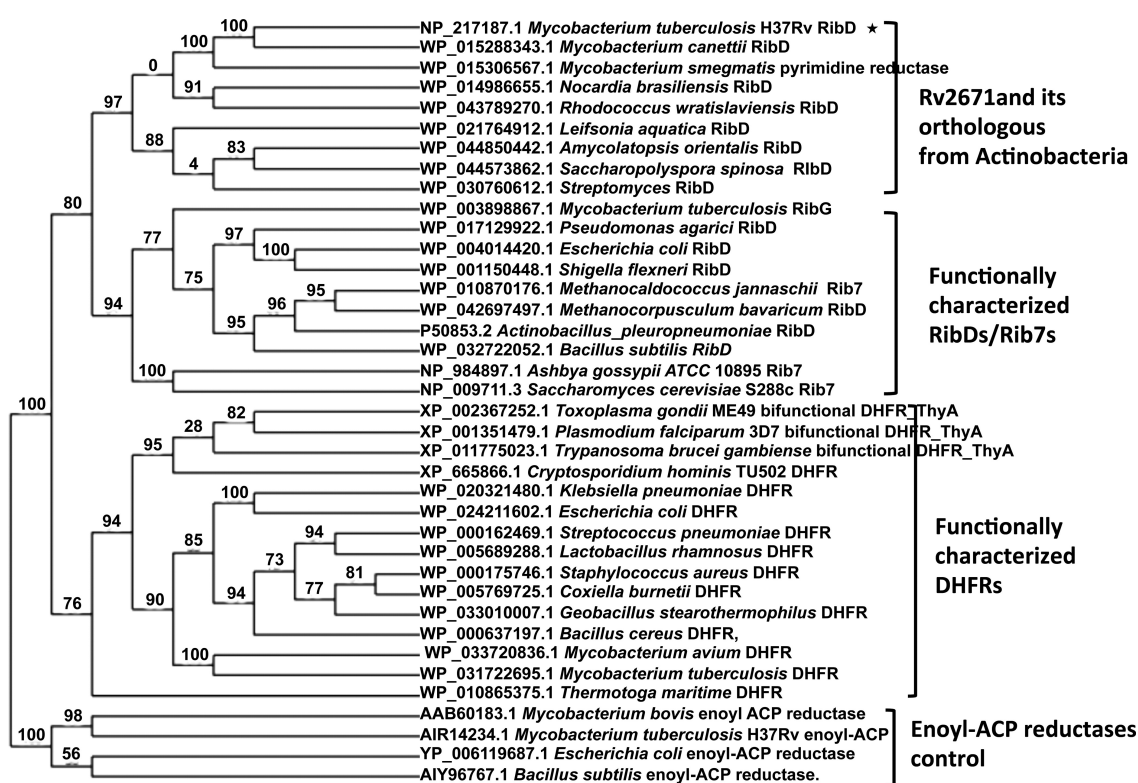


Figure 18. Phylogenetic tree of pyrimidine reductases and DHFRs. Sequences were aligned with MUSCLE; the tree was generated by the approximation of the standard Likelihood Ratio Test with PhyML. The numbers at the nodes indicate the level of confidence for the branches.

## Materials and Methods

### *Materials*

All chemical reagents used in buffers, protein purification, and enzymatic assays were purchased from Sigma-Aldrich (St. Louis, MO). Dihydropteroic acid was purchased from Schircks Laboratories (Switzerland).

### *Cloning, Expression, and Purification of Rv2671*

The sequence of full length *Rv2671* gene was amplified from the *Mtb* H37Rv genome by PCR. Primers were 5'-GGACCTCCATATGCCCGACTCTGGTCAGC-TCG-3' and 5'-CCCAAGCTTTCAGGTCTTGACGTAGCGGGTGTACAGG-3'. The amplified gene was inserted into a pET28b expression vector (Novagen) containing an N-terminal 6x-His tag with the TEV cleavage site using the NdeI and *Hind*III restriction sites. The plasmid was transformed into BL21(DE3) *Escherichia coli* (*E. coli*) cells (Novagen) for *Rv2671* expression. The cells with the plasmid were grown at 37 °C to OD<sub>600</sub> ~ 0.6–0.8 in LB medium with 50 µg/mL kanamycin followed by induction with 1 mM IPTG and grown overnight at 18 °C. The cells were lysed via a Microfluidizer M-100P (Microfluidics, Worcestershire, UK) in lysis buffer (50 mM Tris, pH 7.5, 500 mM NaCl, 5% glycerol, 1 mM PMSF, 10 µg/mL DNase, and 2 mM MgCl<sub>2</sub>) and centrifuged at 27,216g for 1 h. The supernatant was purified over a nickel column (GE Healthcare) with a 0 to 300 mM imidazole gradient and the 6x-His tag was cleaved using tobacco etch virus protease. The protein was >95% pure, as observed by SDS-PAGE, and was

concentrated to 7.5 mg/mL, flash frozen, and stored in dialysis buffer (50 mM Tris, pH 7, 50 mM NaCl, 5% glycerol, 1 mM DTT) at  $-80^{\circ}\text{C}$ . The concentration of the protein was determined by measuring  $A_{280}$  using NanoDrop 2000 UV-vis spectrophotometer with an extinction coefficient of the enzyme of  $17,545\text{ M}^{-1}\text{cm}^{-1}$

### *Crystallization*

In order to obtain cocrystal structures, Rv2671 was incubated with substrates or inhibitors and screened for crystallization conditions with a Mosquito liquid dispenser (TTP Labtech, Cambridge, MA, USA) using the sitting drop vapor diffusion technique in the CrystalMation Intelli-Plate 96-3 low-profile crystallization plate (Hampton Research). For each condition, 1  $\mu\text{L}$  of protein (7.5 mg/mL) and 1  $\mu\text{L}$  of crystallization formulation were mixed, and the mixture was equilibrated with 50  $\mu\text{L}$  of crystallization solution in the reservoir well. Full length Rv2671 protein crystals were further optimized via hanging drop vapor diffusion by incubating 1.5  $\mu\text{L}$  of purified protein solution with 1.5  $\mu\text{L}$  of crystallization solution (0.2 M NaCl, 0.1 M Bis-Tris, pH 6.5, and 25–27% PEG 3350) at  $18^{\circ}\text{C}$  for 2 d. Crystals were cryo-protected with 30% glycerol and flash frozen prior to data collection. For the  $\text{NADP}^{+}$  complex crystals, 2 mM  $\text{NADP}^{+}$  was cocrystallized with Rv2671 using the above procedure. The crystals of the Rv2671 with  $\text{NADP}^{+}$  and trimetrexate (TMQ) bound were also obtained from cocrystallization with 2 mM  $\text{NADP}^{+}$  and 2 mM TMQ. The Rv2671 crystal in complex with tetrahydrofolate (THF)- $\text{NADP}^{+}$ , trimethoprim (TMP)- $\text{NADP}^{+}$ , and dihydropteroic acid (DHP)- $\text{NADP}^{+}$  were obtained by soaking 5 mM of compounds into the  $\text{NADP}^{+}$  binary crystals.

### *Data Collection and Structure Determination*

Data were collected at Argonne National Laboratory using the Advanced Photon Source (APS) beamline 19ID or 23ID-B. All data were processed and reduced using HKL3000.<sup>202</sup> The structure of the copurified NADP(H) bound Rv2671 was solved by molecular replacement using MOLREP<sup>203</sup> in CCP4<sup>204</sup> with the protein atomic coordinates for the truncated pyrimidine reductase-like protein of *Corynebacterium diphtheria* from the Protein Data Bank (accession code 2P4G, deposited by Joint Center for Structural Genomics). The exact search model contains residues 36–86 and residues 174–258 of 2P4G. The complex structures were solved by molecular replacement using MOLREP<sup>203</sup> in CCP4<sup>204</sup> with the solved copurified NADP(H) bound Rv2671 structure. These crystals belong to the  $C222_1$  space group and contained one molecule in the asymmetric unit. Refinement and manual model building were performed with Phenix<sup>205</sup> and COOT,<sup>206</sup> respectively, and TLS was not used in the refinement.

### *Enzyme Kinetics*

The DHFR activity of Rv2671 was measured in a 1 mL solution containing 100 mM HEPES, 50 mM KCl, 200  $\mu$ M NADPH, 50 mM 2-mercaptoethanol, and 11.1  $\mu$ g Rv2671 at pH 7.5, which was adapted from a similar assay to what has been previously reported for *E. coli* DHFR.<sup>207</sup> Assays were started by adding DHF in concentration ranging from 0  $\mu$ M to 250  $\mu$ M. Enzyme activity was monitored using a Cary 100 spectrophotometer at 340 nm. Initial velocity data were fitted with the

Michaelis–Menten equation<sup>208</sup> using the SigmaPlot program (Systat Software, San Jose, CA).

The dihydropteroic acid reductase (DHPR) activity<sup>207</sup> of Rv2671 was measured in 200  $\mu\text{L}$  of enzyme solution containing 100 mM HEPES, 50 mM KCl, 200  $\mu\text{M}$  NADPH, 50 mM 2-mercaptoethanol, and 1.7  $\mu\text{g}$  of Rv2671 at pH 7.5. Assays were started by adding 10  $\mu\text{L}$  of DHP from 0  $\mu\text{M}$  to 530  $\mu\text{M}$ . Enzyme activity was monitored using a Cary 100 spectrophotometer at 340 nm. Initial velocity data were fitted with the Michaelis–Menten equation<sup>208</sup> using the SigmaPlot program (Systat Software, San Jose, CA).

The DHFR activity inhibition assay was performed under the same conditions described above but with the addition of 5  $\mu\text{L}$  of inhibitor in 100% DMSO. The inhibitor was preincubated with the protein for 10 min. After preincubation, 5  $\mu\text{L}$  of DHF was added to a final concentration of 200  $\mu\text{M}$  in order to start the reaction. The reaction activity with 5  $\mu\text{L}$  of DMSO alone served as a negative control. Activity data were plotted and fitted into the four parameter logistic equation (Eq. 1) by the SigmaPlot program (Systat Software, San Jose, CA).

$$\% \text{inhibition} = \min + \frac{\text{max} - \min}{1 + \left( \frac{\log_{10}[\text{inhibitor}]}{\text{IC}_{50}} \right)^{-\text{Hillslope}}} \quad \text{Eq. 1}$$

The pyrimidine reductase activity of Rv2671 was followed by monitoring NADPH oxidation at 340 nm in a solution containing 50 mM HEPES, 5 mM  $\text{MgCl}_2$ , 200



$\mu\text{M}$  NADPH, 30 mM DTT, and 1.32–26.50  $\mu\text{g}/\text{mL}$  Rv2671 at pH 8. The substrates, DAROPP and AROPP were synthesized enzymatically. DAROPP was generated by mixing 5 mM GTP, and 2 mg/mL *E. coli* GCH-II and AROPP was synthesized by incubating 1 mg/mL *E. coli* RibD with DAROPP in the absence of NADPH.<sup>209, 210</sup> The enzymes were removed by filtering through a 10 kDa cutoff concentrator (Vivaspin 500 centrifugal concentrator). DAROPP was analyzed as its diacetyl derivative and detected by fluorometric HPLC. AROPP was analyzed by LC-MS using an SHIMADZU LCMS-2010 spectrometer.

#### *HPLC Analysis of Tetrahydrofolate Formation*

Rv2671 (400 nM) was incubated with 500  $\mu\text{M}$  NADPH and 200  $\mu\text{M}$  DHF in the aforementioned buffer (100 mM HEPES, pH 7.5, 50 mM KCl, and 50 mM 2-mercaptoethanol). Aliquots were removed every 2 min over a 10 min period and were then added to a solution of 0.1 M TCA and 0.55 M ascorbic acid to stop the reaction.<sup>211</sup> To extend the stability of THF, the pH of the reaction mixture was immediately adjusted to pH 4 by adding NaOH.<sup>211</sup> The denatured protein was removed by centrifuging at 16,200g for 10 min at 4 °C. An Agilent technologies HPLC (1200 infinity) with a reversed-phase C18 Atlantis T3, 5  $\mu\text{m}$ , 4.6  $\times$  250 mm column (Waters) was used to separate the components in the reaction mixtures.<sup>211</sup> The column was eluted by the following optimized gradient at a 1 mL/min flow rate: solvent A is water, solvent B is 100 mM  $\text{KPi}$  pH 6.6, and solvent C is methanol; 0 min 100% B, 4 min 25% A 60% B 15% C, 8 min 20% A 40% B 40% C, 14 min 10% A 10% B 80% C, 18 min 10% A 10%

B 80% C, 20 min 25% A 50% B 25% C, 22 min 100% B, and 25 min 100% B. The absorbance at 290 nm was monitored. The fluorescent chromatograph was monitored with excitation at 290 nm and emission at 360 nm.

### *Evolutionary Tree Construction*

The amino acid sequences of thirty-eight representative homologues/orthologues were selected to generate the phylogenetic tree including 9 pyrimidine reductases from Actinobacteria, 15 DHFRs, 10 RibDs or Rib7s, and 4 enoyl ACP reductases used as the reference genes. Nine pyrimidine reductases from Actinobacteria were selected from the NCBI reference protein sequence database,<sup>212-214</sup> which showed >50% sequence similarity with *Mtb* Rv2671. The DHFRs and RibD/Gs were selected from functionally characterized proteins. The phylogenetic tree was generated by the Phylogeny.fr server<sup>215</sup> which used MUSCLE in order to carry out the multiple alignments of protein sequences,<sup>216</sup> PhyML to build the tree using the marginal likelihood method,<sup>217</sup> and TreeDyn in order to generate tree rendering.<sup>218</sup> The sequences used to generate the tree are listed in Supplementary information.

## CHAPTER IV

### KINETICS, STRUCTURE AND INHIBITORS OF *MTB* IDH2

#### Introduction

Enzymes in the central carbon metabolism (CCM) are newly identified anti-tubercular drug targets. CCM is an integration of pathways for the transport and oxidation of the carbon sources in all organisms. The tricarboxylic acid (TCA) cycle and the glyoxylate shunt are the pivots of the CCM in terms of deploying the carbon flux during catabolism of different carbon sources. These two pathways bifurcate at isocitrate (ICT) and determine its metabolic fate. ICT is the substrate of isocitrate dehydrogenase (IDH) in the TCA cycle and isocitrate lyase (ICL) in the glyoxylate shunt. The importance of ICL has been reviewed in Chapter I. It helps *Mtb* to sustain in macrophages and dormancy.<sup>219</sup> The ICL knock-out studies<sup>115</sup> have further demonstrated the potential of ICL to be a drug target for persistent *Mtb*. *Mtb* ICLs have been intensively studied; however, such studies on *Mtb* IDHs are lacking, although their activity may be the determinant of the carbon flux in *Mtb*.

IDH is a decarboxylating dehydrogenase (EC 1.1.1.42). It catalyzes a two-step reaction that converts ICT to  $\alpha$ -ketoglutarate (also called 2-oxoglutarate, 2OG), the latter being the precursor of glutamate (Figure 19). The first step of the reaction is initiated by a deprotonation at the 2-hydroxyl group of ICT, triggering a stereospecific hydride

transfer from C2 toward the cofactor NAD(P)<sup>+</sup> to generate NAD(P)H and an unstable oxalosuccinate intermediate. In the second step, this unstable intermediate undergoes a decarboxylation at the C3, and followed by a tautomerization of the enol intermediate to produce the keto product, 2OG.<sup>220-227</sup> NAD(P)H and CO<sub>2</sub> are therefore generated as by-products during the reaction. Divalent cations, typically Mg<sup>2+</sup> or Mn<sup>2+</sup>, are required to stabilize the oxalosuccinate and enolate intermediates during this catalytic reaction.<sup>228, 229</sup>

A lysine-tyrosine pair was proposed to be the acid-base pair involved in the catalytic mechanism. A lysine (Lys230' from the biologically neighboring subunit in *E. coli* IDH (*Ec* IDH)) is first deprotonated by the side chain of its neighboring Asp (Asp307 in *Ec* IDH) or a hydrogen bonded water. The deprotonated Lys230' then serves as a base to abstract the proton of the 2-hydroxyl group of ICT in the first step. In the decarboxylation step, the metal ion acts as a Lewis acid, and the protonated Lys230' provides a proton to the enol oxygen. In the tautomerization reaction, the deprotonated Lys230' again serves as a general base to accept the 2-hydroxyl group of the enol intermediate, and a tyrosine residue (Tyr160 in *Ec* IDH) functions as a general acid to protonate C3 to generate the ketone product (Figure 19).<sup>220, 230</sup>

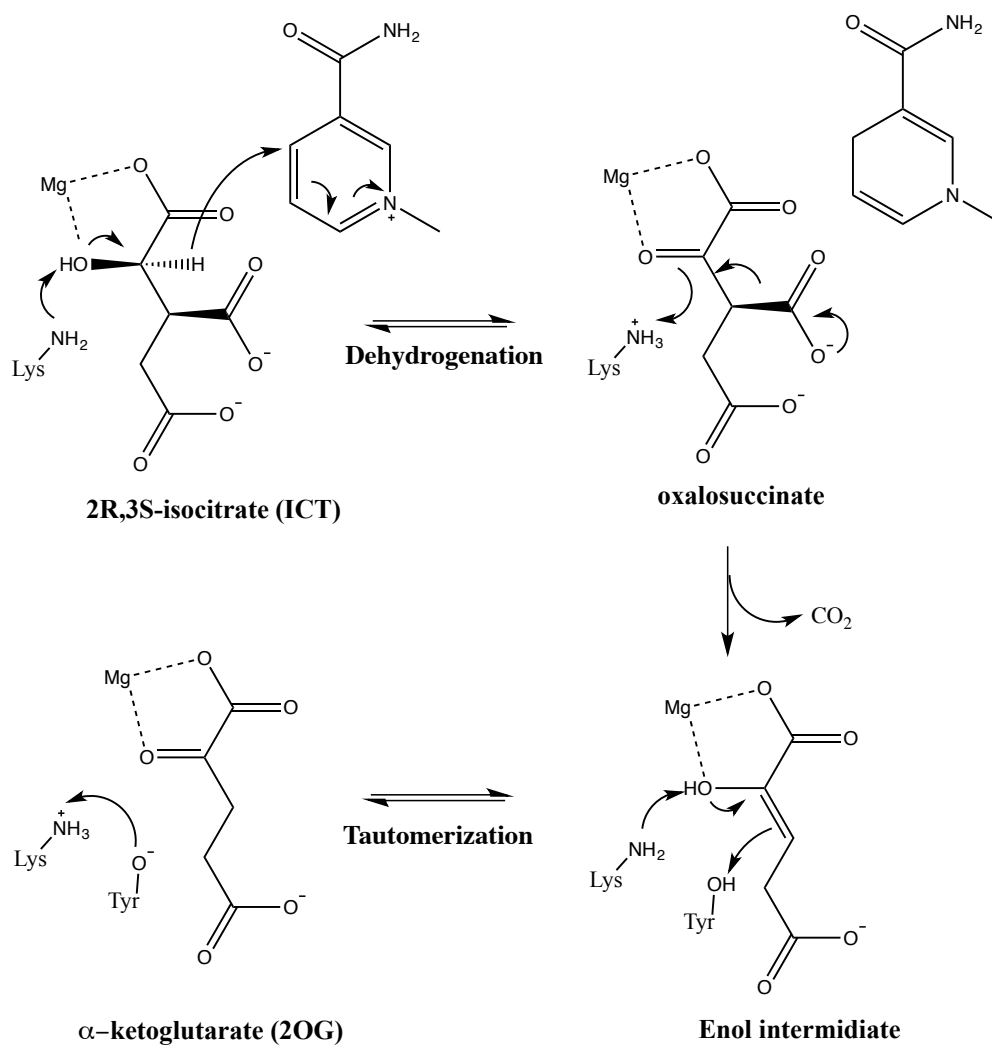


Figure 19. Proposed catalytic mechanism of IDH<sup>220, 230</sup>

IDHs can be classified into three main subgroups, type I, type II, and monomeric IDHs according to a sequence-based phylogenetic analysis.<sup>231</sup> The type I and type II subfamilies usually form homodimers. Two IDHs, *Mtb* IDH1 and *Mtb* IDH2, are presented in *Mtb*. *Mtb* IDH1 belongs to the type I subfamily, and *Mtb* IDH2 belongs to the monomeric IDH family. The structure, kinetics, and catalytic mechanisms of

homodimeric NADP-dependent IDHs, including *Mtb* IDH1, have been extensively studied. However, there has been little detailed investigation of monomeric IDHs. Although several monomeric NADP-dependent IDHs from prokaryotes have been isolated and functionally characterized, including *Acinetobacter lwoffii*,<sup>232</sup> *Azotobacter vinelandii* (*Av* IDH),<sup>233</sup> *Vibrio* sp. ABE-1,<sup>234</sup> *Vibrio parahaemolyticus* Y-4,<sup>235</sup> *Corynebacterium glutamicum* (*Cg* IDH),<sup>236</sup> *Colwellia maris*,<sup>237</sup> *Chlorobium limicola*,<sup>238</sup> *Streptomyces diastaticus* No. 7 M1033,<sup>239</sup> and *Streptomyces lividans* TK54,<sup>240</sup> they have not been studied as extensively as the homodimeric IDHs. To date, there has been only one kinetic study and two structures reported.<sup>241-245</sup> The kinetic study was established with *Av* IDH in 1972.<sup>244</sup> The reaction undergoes a rapid randomly sequential substrate binding and an ordered product releasing following the order of CO<sub>2</sub>, 2OG, and NADPH. In 2002, Yasutake *et al.* solved the first monomeric IDH's structure, the structure of *Av* IDH in complex with ICT and Mn<sup>2+</sup>.<sup>242</sup> The structures illustrated a highly structural conservation at the active sites between the monomeric and dimeric IDHs. However, the diversity was observed in NADP<sup>+</sup> binding cavity of monomeric IDHs, which explained the high NADP<sup>+</sup> specificity of monomeric IDHs.

*Escherichia coli* IDH (*Ec* IDH) studied as a prototype of the prokaryotic IDHs, is regulated by a unique bifunctional enzyme, isocitrate dehydrogenase kinase/phosphatase (*AceK*).<sup>246</sup> The gene *aceK* is located at the acetate operon, which also contains the genes encoding the glyoxylate shunt enzymes, ICL and malate synthase. *AceK* determines the activity of *Ec* IDH via a reversible phosphorylation (inactivate) and dephosphorylation (activate) of *Ec* IDH Ser113.<sup>247</sup> Ser113 mediates the binding of ICT during the reaction.

ICT binding was under the detection limit for the phosphorylated *Ec* IDH, and the activity dropped to 0.15% of that of the unphosphorylated *Ec* IDH.<sup>248</sup>

As mentioned above, *Mtb* has two IDHs. It is not clear why *Mtb* requires two IDHs, but the monomeric IDH2 is believed to participate in the TCA cycle since some mycobacterial species have only monomeric IDH. Moreover, no transposon insertion was observed in the gene encoding *Mtb* IDH2 in the transposon mutagenesis data, suggesting *Mtb* IDH2 is essential. However, an *Mtb* IDH2 knock-out strain can be established, suggesting that *Mtb* IDH2 is non-essential *in vivo*. The genetic observation can be a false positive in the transposon mutagenesis. The regulation of *Mtb* IDHs is also controversial. *Mtb* IDH1 has been proposed to be regulated by the phosphorylation and dephosphorylation based regulatory mechanism as *Ec* IDH as a phosphorylated, inactive *Mtb* IDH1 was observed after treated with mycobacterial eukaryotic-like Ser/Thr protein kinase G (pknG) *in vitro*.<sup>249</sup> However, some evidence did not support this hypothesis including that the IDH activity did not increase in the pknG knock-out strain, and no cognate dephosphorylating enzyme was identified. These results suggest a need to reinvestigate the regulation mechanism of *Mtb* IDH1. Dissimilar to the homodimeric IDHs, the regulatory mechanisms of monomeric IDHs are still unclear.

It has been shown that knock-out *Mtb* IDH2 would increase the efficacy of the ICL inhibitor, 3-nitropropionate.<sup>118</sup> Therefore, inhibitors of *Mtb* IDH2 may have synergistic effect with *Mtb* ICLs inhibitors. There are no reported small molecule inhibitors for monomeric IDHs. However, the research into the inhibitors of human IDH1 and IDH2 belonging to the type-II IDH subfamily has been intensively studied

recently.<sup>250</sup> More specifically, human mutated NADP-dependent IDHs (HIDH1 and HIDH2) are attractive targets for the cancer therapy since somatic missense mutations in HIDH1 or HIDH2 have been identified in several human cancers, especially glioblastoma,<sup>251, 252</sup> sarcoma,<sup>253, 254</sup> and acute myeloid leukemia (AML).<sup>255, 256</sup> The functional homology between *Mtb* IDH2 and HIDH1/HIDH2 provides the potential to repurpose the inhibitors of HIDH1/HIDH2 to prohibit the activity of *Mtb* IDH2.

In this chapter, studies on the kinetics, mechanism, and structure of *Mtb* IDH2 are discussed. The knowledge of *Mtb* IDH2 is not only the necessary building block for understanding the carbon utilization of *Mtb* during infection, but also an important information for the drug development targeting enzymes in the anaerobic glyoxylate shunt. In addition, the study of *Mtb* IDH2 enhances our understanding of monomeric IDHs. A structure comparison between *Mtb* IDH2 and human IDHs revealed the dissimilarity in the overall structure and the NADP<sup>+</sup> binding site as expected, suggesting the potential to find a selective inhibitor. The similarities in the ICT binding site and the catalytic residues, however, suggest that the inhibitors for HIDH1/2 mutants have the potential to inhibit *Mtb* IDH2. Here a preliminary inhibitor study for *Mtb* IDH2 was implemented. Four HIDH1/2 mutants' inhibitors were first tested for inhibition of *Mtb* IDH2. Next, a high-throughput screening (HTS) assay was established and optimized to explore structurally new *Mtb* IDH2 inhibitors from two anti-tubercular whole cell active compound collections.



## Results and Discussion

### *Enzyme Assay and Kinetic Mechanism of Mtb IDH2*

All IDHs, including *Mtb* IDH2, require metal cations, specifically  $Mg^{2+}$  or  $Mn^{2+}$ , to stabilize the reaction intermediate during the enzymatic reaction. The maximal activity of *Mtb* IDH2 was achieved when  $Mn^{2+}$  was used as the metal cofactor. Only 20% of the maximal enzyme activity was detected when  $Mn^{2+}$  was substituted by the same amount of  $Mg^{2+}$ . Additionally, *Mtb* IDH2 is specific for  $NADP^+$ , while does not exhibit activity with  $NAD^+$ .

A steady-state kinetics with varied concentrations of  $NADP^+$  and ICT were then conducted in the presence of  $Mn^{2+}$ . Initial velocity data of *Mtb* IDH2 presented in a double reciprocal plot with  $NADP^+$  as the variable and ICT as the fixed substrate yielded a series of parallel lines (Figure 20A). Likewise a plot of  $1/v$  versus  $1/[ICT]$  at different fixed levels of  $NADP^+$  also gave a parallel line pattern (Figure 20B). The data fit best to the equation for a Ping-Pong mechanism (Eq. 3). The kinetic parameters were obtained as  $k_{cat} = 116.6 \pm 2.05 \text{ s}^{-1}$ ,  $K_m$  for  $NADP^+ = 157.8 \pm 9.10 \text{ }\mu\text{M}$ , and  $K_m$  for ICT =  $36.7 \pm 2.55 \text{ }\mu\text{M}$ . It is unexpected to observe a Ping-Pong mechanism for an oxidative decarboxylation reaction, considering that a hydride transfer from one substrate intermediate (deprotonated ICT) to the cofactor ( $NADP^+$ ) is involved in the reaction that clearly suggests the existence of the transition complex.

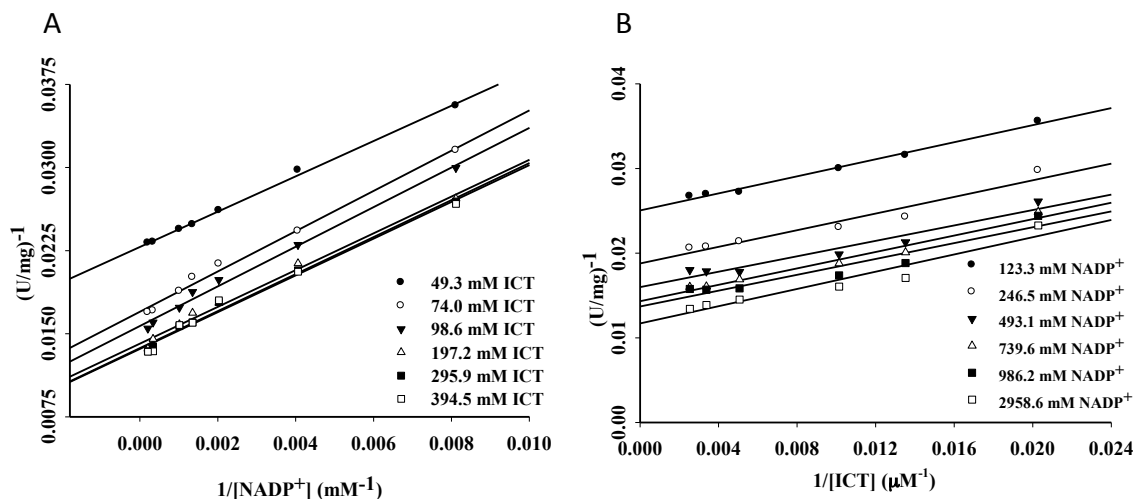


Figure 20. Double reciprocal plots of the initial velocity kinetics for *Mtb* IDH2. (A)  $1/v$  versus  $1/[NADP^+]$  at different fixed concentrations of ICT, (B)  $1/v$  versus  $1/[ICT]$  at different fixed concentrations of  $NADP^+$ .

The substrate inhibition and product inhibition experiments were used to further investigate the likelihood of a Ping-Pong mechanism. In the forward reaction, no substrate inhibition phenomenon was observed when the concentrations of either DL-ICT or  $NADP^+$  were up to 50-fold higher than the  $K_m$  values. Table 9 summarizes the inhibition parameters and the patterns of double-reciprocal plots while using the products, 2OG or NADPH, as the inhibitors. Due to the relatively high  $K_m$  for  $NADP^+$ , the product inhibition study was only conducted when each substrate was held at the sub-saturating level. NADPH was a competitive inhibitor against  $NADP^+$  and a mix-type inhibitor against ICT. 2OG was a mix-type inhibitor against either  $NADP^+$  or ICT. These data do not support a Ping-Pong mechanism but agree with an ordered sequential mechanism in which  $NADP^+$  binds first.

Table 9. Product inhibition patterns and kinetic parameters for *Mtb* IDH2

Inhibitor	Fixed substrate	Varied Substrate	Pattern	Fitted parameters (μM)
NADPH	ICT	NADP <sup>+</sup>	C	appK <sub>is</sub> = 22 ± 1
NADPH	NADP <sup>+</sup>	ICT	NC	appK <sub>is</sub> = 117 ± 28
				appK <sub>ii</sub> = 65 ± 2
2OG	NADP <sup>+</sup>	ICT	NC	appK <sub>is</sub> = 428 ± 84
				appK <sub>ii</sub> = 4,555 ± 1,001
2OG	ICT	NADP <sup>+</sup>	NC	appK <sub>is</sub> = 846 ± 168
				appK <sub>ii</sub> = 2,038 ± 149

The observation of parallel line pattern in the initial velocity study could be explicated by one of the special condition suggested by Segel that a  $K_{ia}$ , which is referred to a dissociation constant value of NADP<sup>+</sup>, is much smaller than its Michaelis-Menten constant ( $K_{mA}$ ).<sup>257</sup> Under this condition, the equation of the steady-state sequential mechanism Eq. 2 is reduced to Eq. 3 as the value of  $K_{ia} * K_b$  is hard to investigate and consequently resulting in a parallel line pattern for a sequential mechanism.

$$v = \frac{V_{max}[A][B]}{(K_a K_{mB} + K_{mB}[A] + K_{mA}[B] + [A][B])} \quad \text{Eq. 2}$$

The product releasing mechanism can also be determined via the product inhibition study. The competitive inhibition of NADPH with respect to NADP<sup>+</sup> and the mixed type inhibition with the respect to ICT suggest that NADPH is the last released product. The overall reaction scheme is shown in Scheme 3.

Scheme 3. Proposed kinetic mechanism of *Mtb* IDH2



### *Overall Structure of NADP<sup>+</sup> Bound Mtb IDH2*

The crystal structure of NADP<sup>+</sup> bound *Mtb* IDH2 was determined at a resolution of 2.67 Å by molecular replacement using the structure of *Av* IDH (1J1W<sup>241</sup>) as a search model and refined to a final  $R_{\text{work}}/R_{\text{free}}$  of 17/23%. The final model includes four protein chains, four NADP<sup>+</sup>, two malonate, one malate, and one succinate molecules in one asymmetric unit of the structure. The details of the refinement are shown in Table 10.

*Mtb* IDH2 is an  $\alpha/\beta$  protein composed of 27  $\alpha$ -helices, 6  $\eta$ -helices, and 24  $\beta$ -strands (Figure 21). It can be divided into a large domain (residues 136–567) and a small domain (residues 5–137 and 568–743). The large domain consists of 13  $\alpha$ -helices ( $\alpha 7$ – $\alpha 19$ ), 3  $\eta$ -helices ( $\eta 4$ – $\eta 6$ ), and 19  $\beta$ -strands ( $\beta 4$ – $\beta 22$ ). The small domain is composed of 14  $\alpha$ -helices ( $\alpha 1$ – $\alpha 6$  and  $\alpha 20$ – $\alpha 27$ ), 3  $\eta$ -helices ( $\eta 1$ – $\eta 3$ ), and 5  $\beta$ -strands ( $\beta 1$ – $\beta 3$  and  $\beta 23$ – $\beta 24$ ). The  $\beta$ -strands from the small domain and five  $\beta$ -strands ( $\beta 4$ ,  $\beta 15$ , and  $\beta 20$ – $\beta 22$ ) from the large domain form an interdomain  $\beta$ -sheet. This  $\beta$ -sheet consists of eight parallel  $\beta$ -strands and two anti-parallel  $\beta$ -strands. In addition, a pseudo two-fold axis was observed in the large domain (Figure 21). The catalytic site is located in the cleft between the two domains.

Table 10. Crystallographic data collection and refinement statistics of the NADP<sup>+</sup> bound *Mtb* IDH2 structure

NADP <sup>+</sup> bound <i>Mtb</i> IDH2	
<b>Data collection</b>	
Space group	<i>P1</i>
Cell dimensions	
<i>a, b, c</i> (Å)	74.8, 110.9, 114.9
$\alpha, \beta, \gamma$ (°)	86.0, 78.6, 76.8
Resolution (Å)	50- 2.67 (2.72-2.67) *
$R_{\text{sym}}$	0.115 (0.596)
$I/\sigma I$	16.72 (3.99)
Completeness (%)	98.35 (94.02)
Redundancy	3.9 (3.9)
<b>Refinement</b>	
No. reflections	390141
$R_{\text{work}}/ R_{\text{free}}$	0.17/0.23
No. atoms	
Protein	22932
Ligand/ion	263
Water	490
B-factors	
Protein	39.30
Ligand/ion	33.70
Water	30.90
R.m.s deviations	
Bond lengths (Å)	0.012
Bond angles (°)	1.29
Ramachandran plot statistics (%)	
Favored	96
Allowed	3.63
Outliers	0.37

\* Resolution values for collection and refinement are the same.

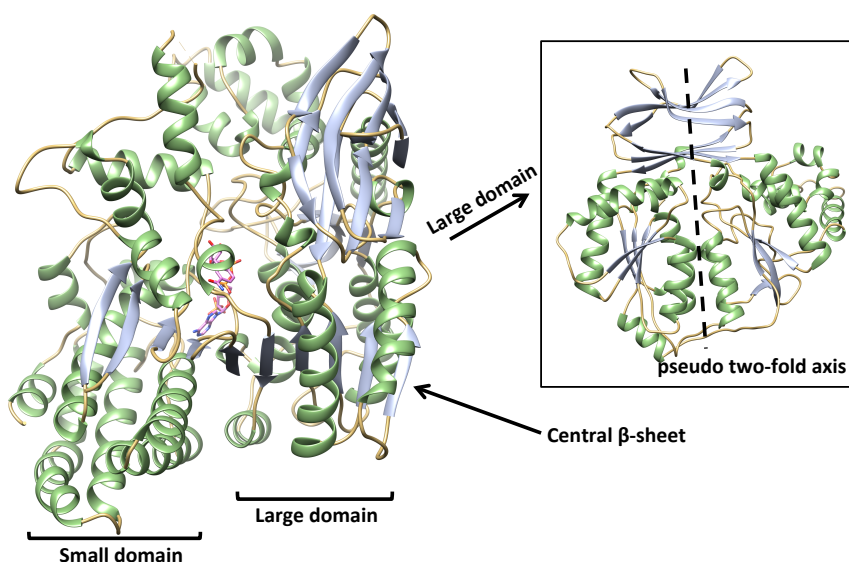


Figure 21. Overall structure of  $\text{NADP}^+$  bound *Mtb* IDH2 and the pseudo two-fold axis in the large domain.  $\text{NADP}^+$  bound *Mtb* IDH2 is shown in ribbon representation and colored by secondary structural elements; strands, grey ribbon; helices, green ribbon; loops, yellow tube.  $\text{NADP}^+$  is shown in pink and located in the cleft between the large domain and the small domain.

### *Structural Comparison with Other Monomeric IDHs*

In addition to the structures of *Av* IDH and *Cg* IDH, the structure of *Mtb* IDH2 was the third solved monomeric IDHs. A structural similarity search was performed by a pairwise comparison using the DALI structural similarity search server,<sup>f199</sup> which used Z-scores to estimate the results. A higher Z-score indicates a better similarity, and a structure with a Z-score greater than two suggests a significant similarity to the parental structure. A total of 116 structures were found with a Z-score greater than two, and as expected, *Av* IDH and *Cg* IDH were identified as the closest structural homologues of *Mtb* IDH2. *Av* IDH (1ITW<sup>242</sup>) exhibited a Z-score of 58.5 and a rmsd of 1.9 for 735 C $\alpha$  atoms; *Cg* IDH (3MBC<sup>243</sup>) exhibited a Z-score of 56.4 and a rmsd of 5.6 for 728 C $\alpha$

atoms. Overall, the structures are similar. The structural comparisons were conducted between individual domains as interdomain movements were observed in *Av* IDH and *Cg* IDH. For the small domain, these three structures contained the same secondary fold and topology (Figure 22A). For the large domain, spatial differences were observed in the region of  $\alpha 11$  to  $\alpha 14$  and the loop (residues 255–266) at the active site. The region of  $\alpha 11$  to  $\alpha 14$  in *Cg* IDH showed a  $22.7^\circ$  rotation compared with that of *Av* IDH and *Mtb* IDH (Figure 22B). The flexibility of the active site loop reinforces the substrate induced interdomain movement.

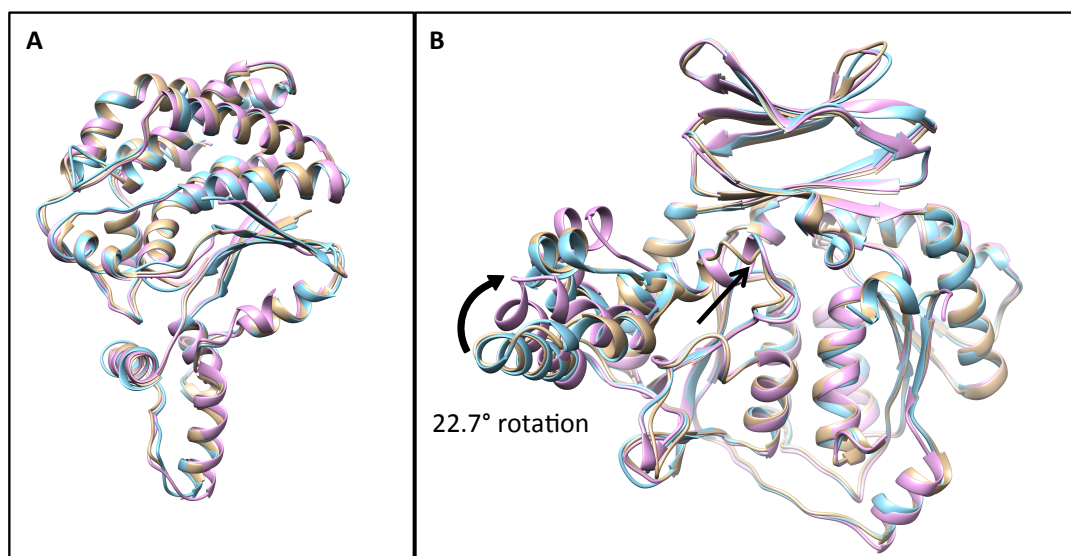


Figure 22. Structural comparisons of small and large domains. (A) Superimposed structures of the small domains of *Mtb* IDH2 (brown), *Av* IDH (blue), and *Cg* IDH (pink). (B) Superimposed structures of the large domains of *Mtb* IDH2 (brown), *Av* IDH (blue), and *Cg* IDH (pink).

Hinge-bending interdomain motions with two hinges were observed in *Av* IDH and *Cg* IDH (Figure 23). After compared with the structures of the reported *Av* IDH and *Cg* IDH, I concluded that this NADP<sup>+</sup> bound *Mtb* IDH2 is in a closed conformation. The  $\alpha 6$  (residues 137–143 in *Mtb* IDH2) and  $\beta 23$  (residues 568–571 in *Mtb* IDH2) were assigned as the corresponding hinges by the program *DYNDOM*<sup>258</sup> using the open conformational *Av* IDH structure (1ITW<sup>242</sup>) as the comparison model.

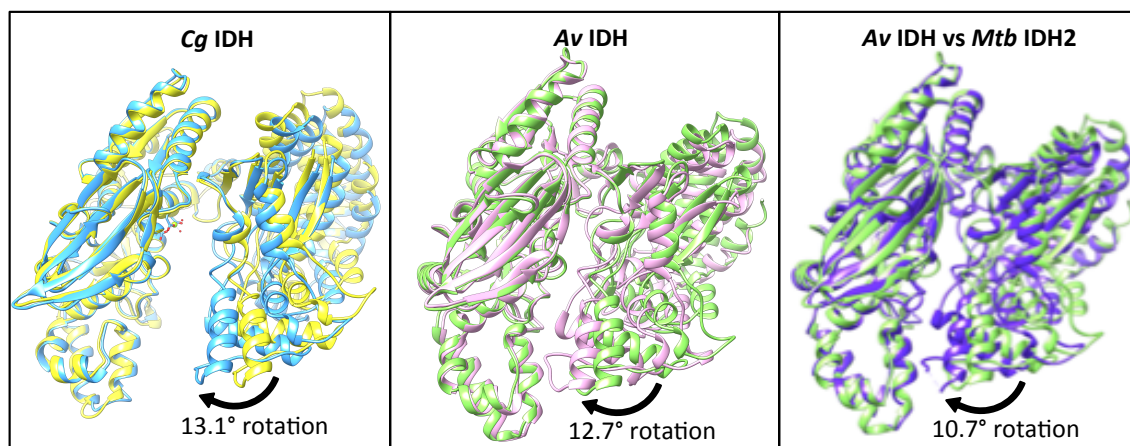


Figure 23. Interdomain motions of monomeric IDHs. The domain movement of *Cg* IDH is shown in left panel by superimposing the structures of *Cg* IDH with the open (in yellow) and closed (in blue) conformations. The middle panel presents the domain movement of *Av* IDH by superimposing the structures of *Av* IDH with the open (in green) and closed (in pink) conformations. In the right panel, the domain movement of *Mtb* IDH2 is predicted by overlaying the closed conformation *Mtb* IDH2 (in purple) with the open conformation *Av* IDH (in green).

#### *NADP*<sup>+</sup> Binding Site

Four NADP<sup>+</sup> molecules were modeled into unambiguous electron densities at the cleft between the two domains of all four molecules in the asymmetric unit (Figure 21).



Each NADP<sup>+</sup> forms 27 hydrogen bonds with the residues of its corresponding *Mtb* IDH2. Figure 24 illustrates the detailed hydrogen bonding and hydrophobic interactions between *Mtb* IDH2 and the NADP<sup>+</sup>. The nicotinamide group of the NADP<sup>+</sup> forms hydrogen bonds with the main chain amide nitrogen of Asn87 (3.1 Å) and a water molecule (3.3 Å). The adenine group of NADP<sup>+</sup> has hydrophobic interaction with the side chains of Thr564, Ala566, Trp604, and Asp605. Furthermore, the N1 and N6 atoms of the adenine group form hydrogen bonds with the main chain oxygen and nitrogen of Asp605 (3.0 Å and 3.0 Å). Five hydrogen bonds are observed between the residues His592, Arg603, and Arg652 and the 2'-phosphate group of the adenosine. More specifically, the phosphate group forms two hydrogen bonds with the guanidinium groups of Arg603 (2.5 Å and 3.0 Å) and Arg652 (2.9 Å and 3.3 Å) individually and a hydrogen bond with the side chain of His592 (2.8 Å). The hydrogen bonds between the 2'-phosphate group on the ribofuranose and the protein may lead to the selectivity of NADP<sup>+</sup> instead of NAD<sup>+</sup> in the *Mtb* IDH2.

### *Catalysis*

So far, the *Mtb* IDH2 structures in complex with ICT or 2OG have not been obtained. However, some ICT analogs including malonate, malate, and succinate, which existed as ingredients of the crystallization condition, were fortuitously crystallized in the active site of the NADP<sup>+</sup> bound *Mtb* IDH2 structure. A malate molecule was modeled into the electron density map of chain A. It forms six hydrogen bonds with its surrounding residues.

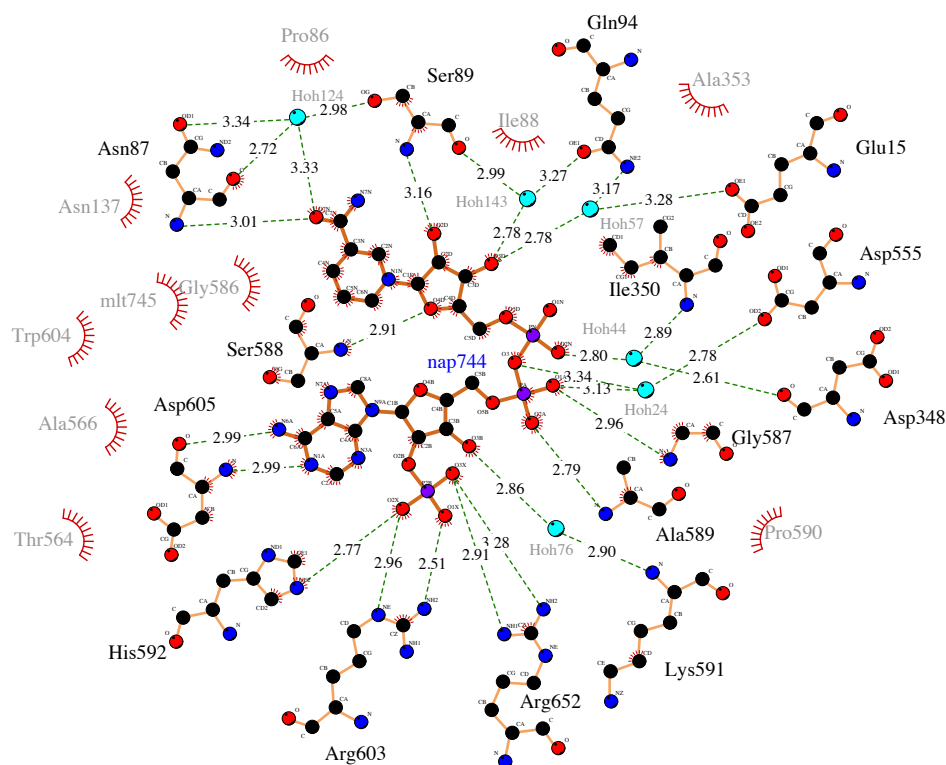


Figure 24. Detailed interactions between  $\text{NADP}^+$  and *Mtb* IDH2. A Ligplot<sup>259</sup> plot presents the interactions between  $\text{NADP}^+$  and *Mtb* IDH2. The bonds are shown in orange. The carbon atoms, oxygen atoms, nitrogen atoms, and phosphorus atoms are colored by black, red, blue, and purple. The hydrogen bonds are in green dashed lines.

The side chain of the following residues each forms one hydrogen bond with the modeled malate including Ser89 (3.3 Å), Ser134 (2.8 Å), Asn137 (3.1 Å), Arg141 (3.0 Å), Lys257 (2.7 Å), and Arg550 (2.2 Å). The hydrogen bond assignment was conducted by the LigPlot software<sup>259</sup> with its default parameter. The positioning of the malate suggests that Lys257, Tyr422 and Asp511 facilitate the catalytic reaction in the *Mtb* IDH2, after comparing with the spatial location of the proposed lysine-tyrosine pair and the Asp307 in the *Ec* IDH structure (4AJ3<sup>260</sup>) (Figure 25).

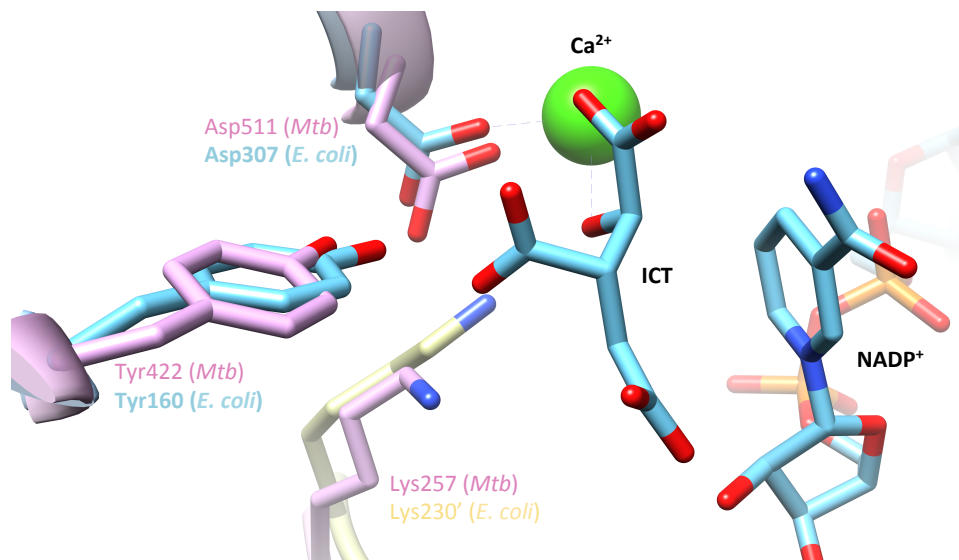


Figure 25. Active site comparison between *Mtb* IDH2 and *Ec* IDH illustrate the potential catalytic residues of *Mtb* IDH2. Superimposition of *Mtb* IDH2 (shown in pink) and *Ec* IDH (subunit A in blue and subunit B in yellow, PDB: 4AJ3) suggests that Lys257, Tyr422, and Asp511 in *Mtb* IDH2 are involved in the catalytic reaction.

### *Proposed Reaction Mechanism of Mtb IDH2*

The solved NADP<sup>+</sup> bound *Mtb* IDH2 structure is in a closed form with no accessible entrance for ICT, suggesting that ICT binds first as binding of NADP<sup>+</sup> would lead to a closed conformation. This structural observation is controversial to the kinetic mechanism in which NADP<sup>+</sup> binds first and followed by the ICT binding. To understand this binding order disagreement, we carefully re-investigated the structure. Notably, only the small domain contributes to the hydrogen bonding interactions between NADP<sup>+</sup> and *Mtb* IDH2. Combined with the knowledge that ICT binding induces the interdomain change, which result in the angle between the large domain and the small domain

decrease and the structure is converted from an open to a closed conformation, a model was proposed (Figure 26). In this model,  $\text{NADP}^+$  first binds to the small domain of open conformational *Mtb* IDH2 and the following binding of ICT then triggers the closure of *Mtb* IDH2. In the solved structure, the observed ICT analogs may serve as the trigger for the domain closure. This model will explain the observations from the structure and the kinetics.

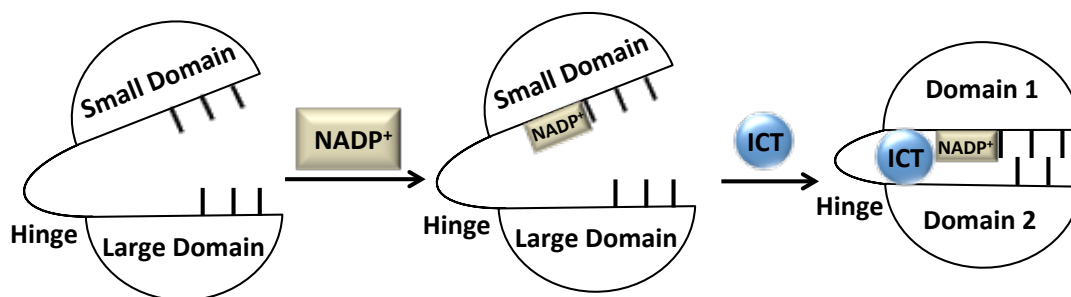


Figure 26. Models of the proposed intradomain closure mechanism.  $\text{NADP}^+$  binds first to the small domain of the free *Mtb* IDH2 producing the binary complex. ICT binds next and triggers an open-to-closed transition.

#### *Mtb* Monomeric IDH2 Is Not a Monomer

Size exclusion chromatography was applied to determine the oligomeric state of *Mtb* IDH2 in solution in a previous report.<sup>261</sup> *Mtb* IDH2 is the first homo-dimeric/homo-tetrameric member of the monomeric IDH family that has been reported to date. *Mtb* IDH2 existed as a tetramer in a buffer with 50 mM NaCl, a trimer in 500 mM NaCl buffer, and a tetramer in 1M NaCl buffer.<sup>261</sup> Trimer is not a constitutive oligomeric state of the enzyme; instead, it is a result of the dynamic equilibrium between the dimeric and

tetrameric state. The dissociation of the tetramer into dimers in the high salt condition indicates that the electrostatic interaction is the driving force of the dimerization between the two constituent dimers.

To examine the pH effect, the oligomeric state of *Mtb* IDH2 was investigated at pH 6.5, which is the lower bound of the intrabacterial pH of *Mtb*.<sup>262</sup> To obtain the desired pH, a phosphate buffer replaced the reported HEPES buffer. *Mtb* IDH2 existed as a dimeric form in the phosphate buffer at pH 6.5 and as a mixture of dimer and tetramer in the ratio 3:1 at pH 7.5. The ionic strength of the examining buffer (50 mM phosphate buffer, 50 mM NaCl, pH 7.5) is 179 mM, which is 321 mM less than 500 mM NaCl in the reported buffer. This result suggests that other than the ionic strength, the formation of the tetramer can be influenced by pH or buffers. To investigate these two factors, the oligomeric states were determined in the different buffers, which contain same ionic strength and pH (Table 11). Different oligomeric states were observed suggesting that the formation of the tetramer is sensitive to the buffer systems. The results of the oligomeric states investigation indicate that the tetramer is unstable and affects by environmental factors such as the ionic strength and the buffers, but the homodimer is stable.

Table 11. Oligomeric states of *Mtb* IDH2 in different buffer system

Buffer	NaCl (mM)	Ionic Strength (mM)	pH	Oligomeric state
50 mM phosphate	50	132	6.5	2
50 mM MOPS	50	132	6.5	2.5 (D:T= 3:1)
50 mM PIPES	50	132	6.5	3 (D:T=1:1)

By scrutinizing the crystal structure of *Mtb* IDH2, nine hydrogen bonds were observed between the assigned chains A and D (Figure 27). The side chains from Arg172, Arg179, Glu390, Thr397, and Arg529 of chain A and the side chain or main chain from Gln389, Glu390, Asn398 and Arg529 of chain D conducted the formation of hydrogen bonds in the dimer interface.

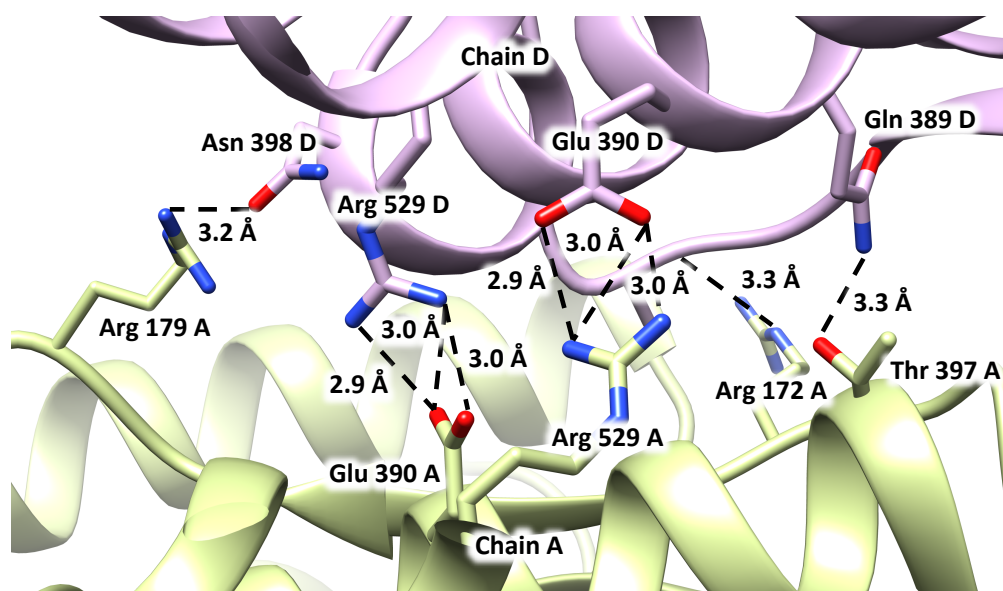


Figure 27. Interface interaction between chain A and chain D in the NADP<sup>+</sup> bound *Mtb* IDH2 structure. Hydrogen bonds between the residues of chain A (yellow) and chain D (pink) are shown in black dotted lines.

A site-directed mutagenesis study was implemented to validate if this dimer observed in the crystal structure is a result of the crystal packing or is physiologically relevant. Three *Mtb* IDH2 mutants were constructed and purified including *Mtb* IDH2 E390R, *Mtb* IDH2 R529E and *Mtb* IDH2 E390R/R529E double mutant. E390 and R529 were chosen as site-directed mutagenesis targets due to the symmetrical formation of an

electrostatic interaction in the structurally dimer interface. The oligomeric states of the purified *Mtb* IDH2 mutants were analyzed by the size exclusion chromatography (Figure 28). A phosphate buffer at pH 6.5 was used to conduct the experiment since the homotetramer is fully dissociated under this condition. E390R mutant exists as a monomer in solution, and the dimerization is restored when the R529E mutation was introduced. The monomeric form of the E390R mutant indicates that the dimer between chain A and D is biologically relevant, and E390 and R529 are crucial for the dimerization. Alternatively, the R529E mutant remained as a dimer in solution. This result implies that the residual interactions are sufficient to drive the dimerization. Therefore, the dimerization process disturbed by E390R mutation not only via the interruption of the electrostatic interaction between E390 and R529, but also the steric hindrance of two arginines between two subunits and consequently disturbing other interactions. To test if the biological dimer is activity relevant, the activity assay of the mutant *Mtb* IDHs were performed. The E390R and R529E mutants showed slightly higher activity but did not achieve a significant difference compared to the activity of the wild-type (WT) enzyme. This result indicates that the dimeric form is not required for the enzymatic activity.

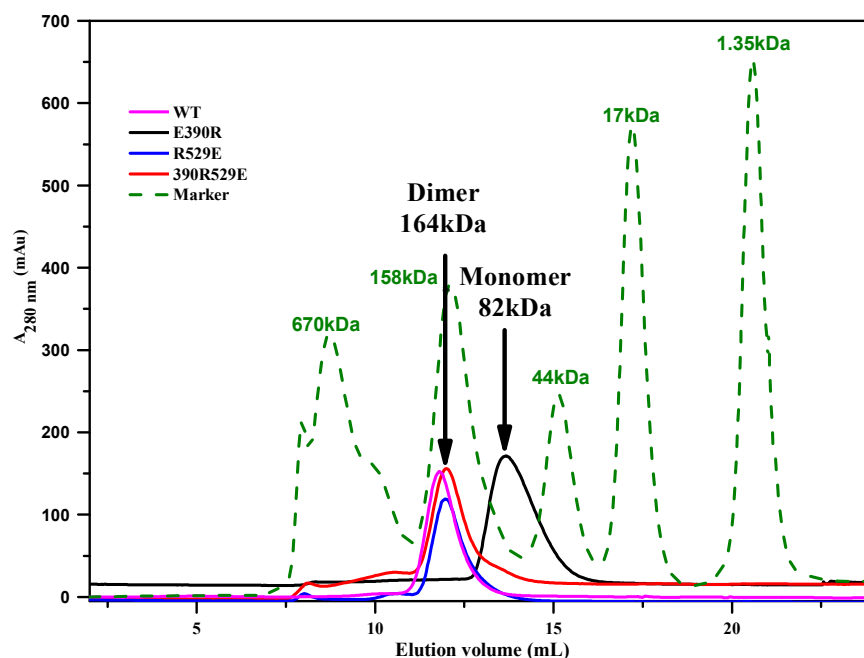


Figure 28. Size exclusion chromatography of wild-type and mutants of *Mtb* IDH2s. The purified proteins (0.1 mL, 10mg/mL) were loaded to a Superdex200 10/300 column equilibrated with binding buffer (50 mM sodium phosphate buffer pH 6.5 and 50 mM NaCl). WT (pink trace), R529E (blue), and double mutant (red) *Mtb* IDH2s were eluted from the column at molecular mass of ~ 164 kDa, which is corresponding to a dimer. Alternatively, E390R *Mtb* IDH2 (blue) were eluted from the column at molecular mass of ~ 82 kDa, suggesting that E390S *Mtb* IDH2 is a monomer under this buffer condition.

### *Structural Comparisons between Mtb IDH2 and Human IDHs*

Besides essentiality, the ability to provide the selectivity is another determinant for a good drug target. The “selectivity” can be exhibited when the enzyme of interest is (i) lack of human homologs, (ii) possessing a non-essential human homologue, or (iii) showing sufficient differences from its human counterpart. Sequence and structural comparisons were conducted to evaluate whether *Mtb* IDH2 is a good target from the selectivity perspective. Human and higher eukaryotes cells have three isoenzymes of



IDHs. Human IDH1 and IDH2 (HIDH1 and HIDH2) are NADP<sup>+</sup>-dependent homodimeric IDHs belonging to the type II superfamily. Human IDH3 (HIDH3) is an NAD<sup>+</sup>-dependent octamer and composed of three subunits in a ratio of  $\alpha 2\beta\gamma$ .<sup>263</sup> HIDH1 locates in the cytoplasm.<sup>264</sup> HIDH2 and HIDH3 are found in the mitochondrial.<sup>265</sup> HIDH3 participates in the TCA cycle for energy generation. The subunit  $\alpha$  is the catalytic subunit.<sup>266</sup> The subunits  $\beta$  and  $\gamma$  function as the regulatory units. The roles of HIDH1 and HIDH2 remain unclear, but it is believed that they serve to maintain the redox balance and protect the cell from the oxidative induced damage.<sup>131, 267</sup> Mtb IDH2 showed low sequence identity (5–12%) with human IDHs implying the difference for selection. Notably, it is usual that proteins show high structural similarity with low sequence similarity. Therefore, structural comparisons of the regions and the residues that are critical for the function of proteins are required to evaluate the similarity between two enzymes. The structural comparisons were performed between Mtb IDH2 and two structurally known human IDHs, HIDH1 and HIDH2. First, it is worth to mention that the active sites of HIDH1 and HIDH2 showed high similarity including the catalytic residues and the NADP<sup>+</sup> binding residues. As expected, the catalytic residues were conserved but were different in the NADP<sup>+</sup> binding pocket for Mtb IDH2 and HIDH1/HIDH2. Figure 29 shows the NADP<sup>+</sup> binding residues. The similarity of the catalytic residues implies the potential that mutant HIDHs' inhibitors can be inhibitors for Mtb IDH. The diversity of the NADP<sup>+</sup> binding site indicates the possibility to find a unique inhibitor for monomeric IDHs.

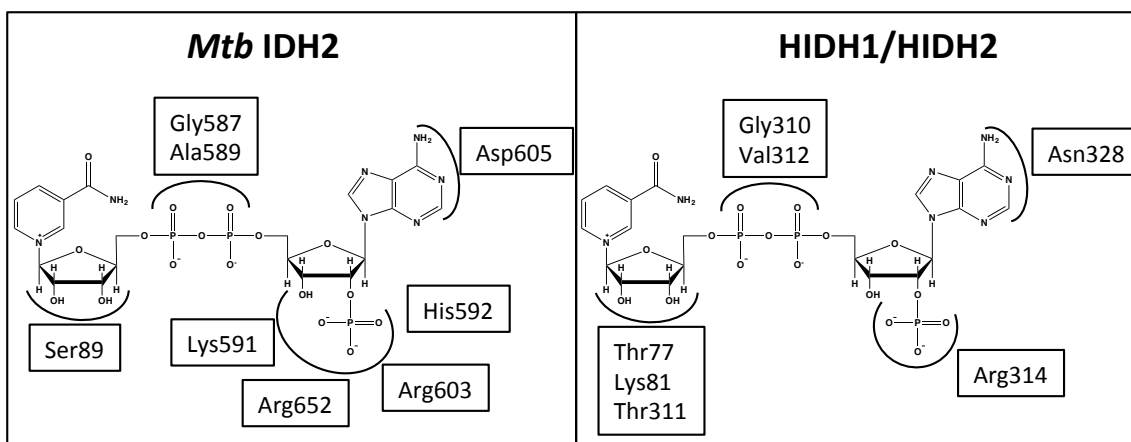


Figure 29. Schematic representations of the NADP<sup>+</sup> binding residues of *Mtb* IDH and HIDH1/HIDH2

#### *Inhibitory Tests for HIDHs' Inhibitors*

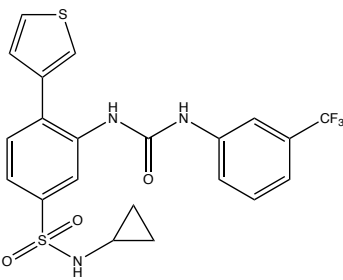
Although the sequence (5–12% sequence identity) and the overall structure of *Mtb* IDH2 are dissimilar to that of HIDHs, inhibitory test of the known HIDHs mutant inhibitors against *Mtb* IDH2 is still a good starting point since the active site residues of *Mtb* IDH2 and HIDHs are conserved. Four known HIDH1 inhibitors (Table 12) were assayed against the recombinant *Mtb* IDH2 at a single concentration point of 20  $\mu$ M in duplicates. AGI-5198,<sup>268</sup> AGI-6780,<sup>269</sup> and 1BZ<sup>270</sup> showed poor inhibitory activity against *Mtb* IDH2 (AGI-5198: 0%, 1BZ: 5.7% inhibition, and AGI 6780: 20% inhibition). Alternatively, SYC-3 demonstrated 100% inhibition against *Mtb* IDH2 at 20  $\mu$ M; hence, an IC<sub>50</sub> evaluation was conducted. SYC-3 exhibited an IC<sub>50</sub> of 5.2  $\pm$  3.1  $\mu$ M against *Mtb* IDH2 (Figure 30). A structure activity relationship (SAR) and protein-inhibitor complex structure study of SYC-3 are in progress.

Indeed, the *Mtb* IDH2 inhibition profiles of SYC-3 and AGI-6780 are expected based on their binding modes in HIDH1. SYC-3 binds at the relatively conserved active site,<sup>271</sup> so it is likely to be a shared inhibitor. To the contrary, AGI-6780 binds at the dimer interface,<sup>269</sup> which does not exist in *Mtb* IDH2. The structural difference consequently leads to an inhibition discrepancy. However, it is unclear why 1BZ is inactive against *Mtb* IDH2 since its binding site is similar to that of the SYC-3.<sup>270</sup> In fact, except for 1BZ, the overall inhibition pattern of *Mtb* IDH2 agrees the reported inhibition pattern against the WT HIDHs; compiled data are shown in Table 12.

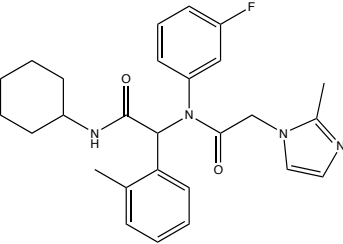
Table 12. Inhibitory activity of HIDHs' inhibitors against *Mtb* IDH2

Compd	% Inhibition at 20 $\mu$ M	IC <sub>50</sub> ( $\mu$ M)		K <sub>i</sub> ( $\mu$ M)
		HIDH1 (WT)	HIDH2 (WT)	HIDH1 (WT)
AGI-6780	21.2	>100 <sup>269</sup>	0.190 $\pm$ 0.0081 <sup>269</sup>	-
AGI-5198	-0.5	>100 <sup>268</sup>	>100 <sup>268</sup>	-
1BZ	5.7	-	-	16.8 <sup>270</sup>
SYC-3	100.0	-	-	8.8 $\pm$ 0.4 <sup>271</sup>

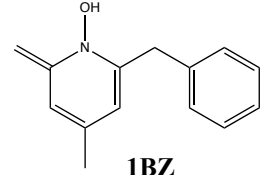
  



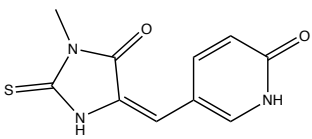
**AGI-6780**



**AGI-5198**



**1BZ**



**SYC-3**

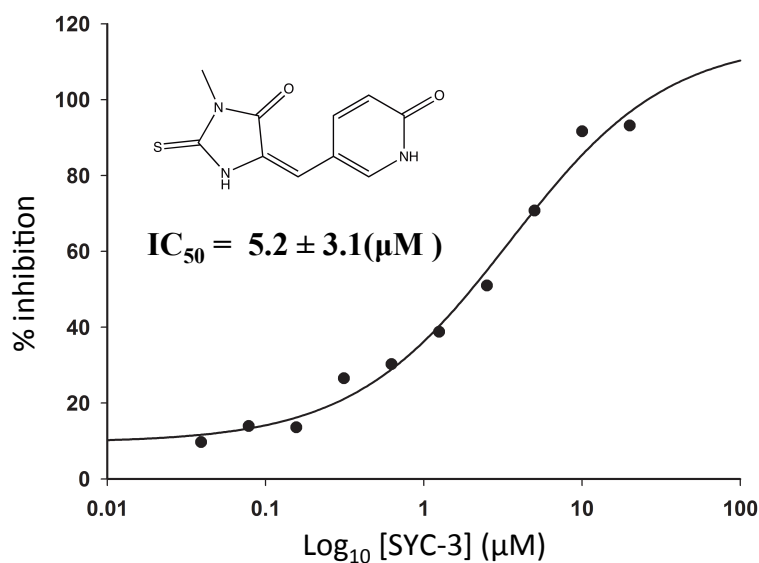


Figure 30. Inhibitory dose-response curve to determine the  $IC_{50}$  of SYC-3 for *Mtb* IDH2. Data represent mean values of duplicate experiments.

### *High-Throughput Screening*

To identify novel chemical inhibitors that are unique for *Mtb* IDH2 as well as monomeric IDHs, we further adapted the *Mtb* IDH2 enzyme assay to a high throughput format. The robustness and reproducibility of the assay for the HTS assay was evaluated by statistical analysis and yielded a  $Z'$  factor of 0.6. The reaction without enzyme was used as the positive control for the calculation since there was not a good *Mtb* IDH2 inhibitor when the experiment initially was performed. Using this assay, we screened 1,993 compounds including two collections of *Mtb* whole cell active compounds. The first collection is called the SRI library from Southern Research Institute. This library is composed of 1,113 *Mtb* H37Rv whole cell active compounds among 100,997 structurally different drug-like compounds from ChemBridge Corporation.<sup>272</sup> These

compounds established >90% inhibition against the H37Rv whole cell growth assay in 7H12 media at 10 µg/mL and did not show toxicity in the Vero cell-based cytotoxicity assay. The second collection is called the Sac\_1 Hit plates. The anti-tubercular whole cell activity screening of the Sac\_1 library was performed in house (Hughes, R. and Sacchettini, J., personal communication). A library containing 50,099 drug-like molecules was screened against *Mtb* mc<sup>2</sup>7000 strain in M9 media with sodium acetate as the carbon source. Compounds showing greater than 50% inhibition of the whole cell growth were selected as hits. A total of 880 compounds were found in the primary screen and 70% of them showed inhibition in the retest. The results of the HTS against *Mtb* IDH2 are summarized in Table 13. We obtained 14 candidates when a cutoff of >75% inhibition was applied. The 14 candidates were first retested at 50 µM for inhibition against *Mtb* IDH2 with material being drawn from the original stock solution plates. Two among 14 original hits were confirmed as true enzyme hits; shown in Figure 31A. Compound 7365674 was dropped due to low inhibitory activity (40% inhibition at 20 µM) and the higher potency against H37Rv than *Mtb* IDH2 implying that *Mtb* IDH2 is not the molecular target of this compound. This yielded one confirmed hit (6144788) was selected for resupply from ChemBridge.

Table 13. Summary of the HTS results

Total # of Compounds Screened	1,993
Hits from Primary Screen	14 (0.7%)
Secondary Validation	2- Confirmed hits (0.1%)
Individual Library Successful rate	
SRI Library Collection	
Total # of Compounds Screened	1,113
Hits from Primary Screen	13 (1.2%)
Secondary Validation	2- Confirmed hits (0.18%)
Sac_1 Hit Collection	
Total # of Compounds Screened	880
Hits from Primary Screen	1 (0.1%)
Secondary Validation	0- Confirmed hits (0%)

#### *Structure-Activity Relationship Study*

The sulfonamide containing molecule (6144788) is an inhibitor of *Mtb* IDH2 with an IC<sub>50</sub> value of 2.2 μM (Figure 31B). Compound 6144788 was then either soaked into the existence crystal or co-crystallized with *Mtb* IDH2. No intact drug density was observed in the calculated electron density map in all experimental trials. The low solubility of the compound 6144788 in the crystallization buffer may cause the failure to obtain the protein-inhibitor complex structure. The cLogP of 6144788 (Figure 31C) is higher than a lead like molecule, which is often between 1–3. Therefore, a preliminary structure-activity relationship (SAR) study was carried out with a series of commercially available sulfonamide structural derivatives (Table 14-17) to improve the potency of *Mtb* IDH2 inhibition and to address the solubility issue.

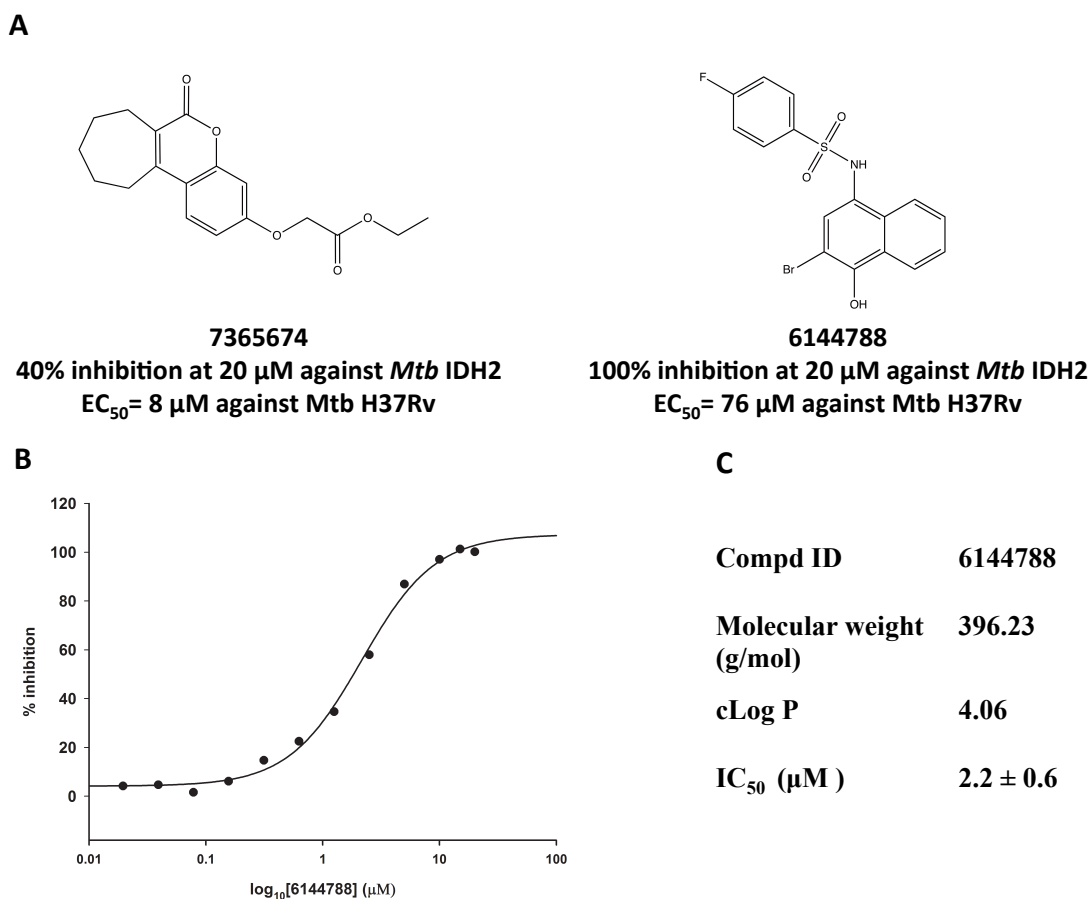


Figure 31. Detailed information of the hits from the HTS assay. (A) Chemical structures, enzyme inhibitory and the whole cell activity of the two obtained hits. (B)  $\text{IC}_{50}$  curve of 6144788 against *Mtb* IDH2 in the enzyme assay. (C) Molecular information of 6144788.

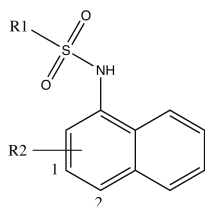
A total of 19 commercially available derivatives of the parental hit were collected and assayed against *Mtb* IDH2, but none of them showed significant improvement in inhibition in this preliminary SAR study, The tested compounds can be classified into five main structural sub-groups.

Molecules in sub-group 1 retained the N-(naphthalen-1-yl)benzenesulfonamide scaffold of the parental compound but incorporated multiple substitutions on the

aromatic rings. When 4-fluorophenyl is changed to 4-ethylphenyl and naphthyl in compounds 1-1 and 1-2, the inhibitory activity decreased slightly. In addition to the substitution of 4-fluorophenyl to 4-chlorophenyl and naphthyl, the bromo group was removed from compound 1-3 and 1-4. These modifications also reduced the inhibitory effects. Lower inhibitory activity of 1-3 comparing with 1-2 demonstrated the importance of the bromine substitution at the naphthalene group. Different functional groups were incorporated to investigate the role of the *para*-position of the phenyl ring including an alkyl, a different halogen or replacement with a fusing ring. Although the halogen atoms are commonly used to increase the hydrophobicity during the hit optimization, their abilities to form weak hydrogen bond (ex: CF...HN) or halogen bond interaction, orthogonal multipolar interactions and the cation-pi interaction are considered to play an important role to improve inhibitor's potency.<sup>273-275</sup> The inhibitory activity increases with the hydrophobicity increase in the comparison of 1-1 and 1-2 indicating the participation of the hydrophobic interaction. However, the fluoro-substituent parental hit and the chloro- substituted 1-4 showed better inhibition than the naphthyl modified 1-2 and 1-3 further suggesting that the electron withdrawing and weak hydrogen/halogen bond forming characteristic carried by the halogen atoms may be also involved in the binding mechanism.



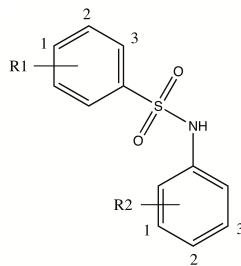
Table 14. Structures and inhibitory activity of sub-group1



Compd	R1	R2	% Inhibition at 20 $\mu$ M	IC <sub>50</sub> ( $\mu$ M)	clogP
1-1		1-Br, 2-OH	72.1		5.05
1-2		1-Br, 2-OH	124.5	3.9	5.20
1-3		2-OH	46.9		4.05
1-4		2-OH	76.9	15.8	4.10
1-5		1-H, 2-H	0.3		5.17
1-6		1-H, 2-H	-4.6		2.80

To investigate the importance of the naphthalene core connecting to the amino group of the sulfonamide, a series of benzenesulfonamide derivatives were collected and formed sub-group 2. None of these derivatives exhibited inhibition against *Mtb* IDH2. Notably, the replacement of the phenyl group is the only difference between the parental hit and compound 2-2. These results clearly highlight the importance of the naphthalene ring.

Table 15. Structures and inhibitory activity of sub-group2



Compd	R1	R2	% Inhibition at 20 $\mu$ M
2-1		1-Br, 2-OH	24.7
2-2	1-F	1-Br, 2-OH	3.9
2-3	1-OMe, 2-NH2	2-OEt	-18.9
2-4	1-COOH		-2.1
2-5			-12.1

One critical liability with this series of compounds is its poor solubility, which hinders the process of both testing in the binding assay and the structural study of the complex. Based on the results of sub-group 1 and 2, the naphthalene ring is essential for the activity but results in limited solubility of the compounds. Therefore, we attempted to overcome this issue by replacing the naphthalene with heterocyclic ring groups. As shown in Table 16 and Table 17, sub-groups 3 and 4 are the collections in which the naphthalene is replaced by a quinoline ring. Although all molecules in these two sub-groups exhibited lower potency compared to the parental hit, in general, better inhibitions were achieved by the sub-group 4 than the sub-group 3. This result again demonstrates the importance of substituent at the quinolone, which is in agreement with

the observation in sub-group1. These results demonstrated the possibility to obtain a potent compound with better solubility once a proper substituent is incorporated.

Table 16. Structures and inhibitory activity of sub-group3

Compd	R group	% Inhibition at 20 $\mu$ M	cLog P
3-1		25.7	3.39
3-2		23.4	4.43
3-3		10.4	3.13
3-4		29.4	3.28
3-5		24.7	2.50

Table 17. Structures and inhibitory activity of sub-group4

Compd	R1 group	R2 group	% Inhibition at 20 $\mu$ M	cLog P
4-1		1-Cl, 3-Cl	63.2	5.42
4-2		3-Br	63.3	3.83
4-3	CH <sub>3</sub>	2-OMe	12.2	1.62

## Conclusion

The biochemical and biophysical properties of *Mtb* IDH2 were thoroughly studied and characterized in this chapter. The kinetic data suggested that *Mtb* IDH2 underwent a sequential mechanism in which NADP<sup>+</sup> bound first and NADPH released last. The X-ray structure of *Mtb* IDH2 was solved for the first time, which is also the third reported structure of monomeric IDHs. The NADP<sup>+</sup> bound *Mtb* IDH2 was observed in a closed conformation. Intriguingly, instead of a monomer, *Mtb* IDH2 formed a dimer in solution. Through the structural investigation, E390 and R529 were

proposed to be the critical residues participating in the dimerization by providing the electrostatic interactions, which was further demonstrated by the site-directed mutagenesis study. The E390R *Mtb* IDH2 is a monomer in solution due to the interruption of the interactions and a steric hindrance. The mutation did not alter the activity of *Mtb* IDH2 indicating the dimer form was irrelevant in this activity.

The sequence and structural conservations of the catalytic residues between *Mtb* IDH2 and HIDHs suggest that the inhibitors of HIDHs can be inhibitors for *Mtb* IDH2 if they bind at the catalytic site. Alternatively, the diversity at the NADP<sup>+</sup> binding cavity indicates the potential to find inhibitors that are unique for monomeric IDHs. To identify inhibitors for *Mtb* IDH2, reported inhibitors of homologues (HIDHs) and two structurally diverse *Mtb* whole cell active libraries were assayed against *Mtb* IDH2. Two compounds, one from the HIDHs inhibitor collection and one from the whole cell active collections, showed low  $\mu\text{M}$  IC<sub>50</sub> values against *Mtb* IDH2. A preliminary SAR study was conducted to improve the potency and the solubility of the compound 6144788. However, the potency of the inhibitor did not improve through the optimization. Although none of the derivatives exhibited better potency in the preliminary SAR studies, these studies revealed the critical components of these inhibitors. Briefly, it is required for the naphthalene group to connect to the amino group of the sulfonamide linker, and the substituents at the naphthalene ring are also critical for the activity. Moreover, a *para*-halogen substituted phenyl ring attaching to the sulfur group of sulfonamide group is preferred. This information could serve as a platform for future inhibitor design. The failure of the structural and binding assay studies suggest that the

solubility issue needs to be addressed before the inhibition mechanism and binding mode of compound can be successfully illustrated.

## **Materials and Methods**

### *Protein Expression and Purification*

Full length *Mtb Rv0066c* gene was amplified from *Mtb* H37Rv genomic DNA by PCR with primer pairs, 5'-GGGAATTCCATATGATGAGCGCCGAACAGCCGACC-3' and 5'-CCCAAGCTTTCA GCCTTGGACAGCCTCCAGC-3'. The amplified gene was inserted into a pET28b (Novagen) expression vector fused with a TEV cleavage site and an N-terminal 6x-His tag using the NdeI and *Hind*III restriction sites. This engineered vector was transformed and expressed in the *E. coli* Rosseta (DE3) strain (EMD Millipore). *E. coli* cells with the vector grew at 37 °C in LB media with 50 µg/mL kanamycin (Gold Biotechnology, USA) and 34 µg/mL chloramphenicol (Gold Biotechnology, USA) to OD<sub>600</sub> of 0.6 and were then induced by 0.5 mM IPTG (Gold Biotechnology, USA) at 18 °C overnight. The induced cells were collected and centrifuged down at 4000 rpm for 20 min. The cell pellets were suspended in the lysis buffer (50 mM Tris, pH 8.0, 500 mM NaCl, 1 mM PMSF, 10 µg/mL DNase and 2 mM MgCl<sub>2</sub>), and lysed by a Microfluidizer M-100P machine (Microfluidics, Worcestershire, UK). The cell lysates were centrifuged at 15,000 rpm for 1 h. The supernatant was purified by affinity chromatography using a nickel column (GE Healthcare). The supernatant of the lysed cells was loaded on the column and washed by a wash buffer

(50 mM Tris, pH 8, 500 mM NaCl, and 30 mM imidazole). The target proteins were eluted by an elution buffer (50 mM Tris, pH 8, 500 mM NaCl, and 500 mM imidazole) with a gradient.

The TEV cleavage site and His-tag congregated *Mtb* IDH2 were incubated with the tobacco etch virus protease for 12 h at 4 °C and reloaded on the nickel column. The flowing through containing pure proteins was collected. The pure *Mtb* IDH2 was concentrated to 30–40 mg/mL, flash frozen, and stored in a storage buffer (50 mM Tris, pH 7.4, 50 mM NaCl) at –80 °C. The protein concentration was estimated using the  $A_{280}$  measured Nano Drop 2,000 with calculated extinction coefficient and molecular weight ( $67270 \text{ M}^{-1}\text{cm}^{-1}$  and 82.69 kDa).

#### *Enzyme Assays and Kinetic Studies*

The activity of *Mtb* IDH2 was determined by monitoring the reduction of  $\text{NADP}^+$  to NADPH. The absorbance changing of NADPH at 340 nm was measuring by a Cary 100 UV-Vis spectrophotometer (VARIAN, CA, USA) using a molar extinction coefficient of  $6,220 \text{ M}^{-1}\text{cm}^{-1}$ . Enzyme assays were conducted in 1 mL volume of 100 mM HEPES, pH 7.5, 20 mM  $\text{MnCl}_2$ , and 0.5–1 nM of *Mtb* IDH2. The final concentration of  $\text{NADP}^+$  varied from 125  $\mu\text{M}$  to 5 mM, and the final concentration of DL-ICT varied from 25  $\mu\text{M}$  to 600  $\mu\text{M}$ . The reaction was initiated by adding the ICT and was monitored at room temperature.

### *Product Inhibition Studies*

The inhibition effects of NADPH and 2OG were investigated in this work. Initial velocity patterns were obtained by measuring the initial rate at different concentrations of one reactant, with the concentration of other reactants fixed at a subsaturating concentration, and at different fixed concentrations of the inhibitor including zero.

### *Data Fitting*

The initial velocity data were fit to the equation of Ping-Pong mechanism (Eq. 3)

$$v = \frac{V_{max}[A][B]}{K_{mA}[B] + K_{mB}[A] + [A][B]} \quad \text{Eq. 3}$$

In Eq. 3,  $v$  is the reaction rate ( $\mu\text{mole/s/mg}$ );  $V_{max}$  is the maximum velocity of the reaction ( $\mu\text{mole/s/mg}$ );  $[A]$  and  $[B]$  are the concentration of substrate A and B ( $\mu\text{M}$ ), respectively;  $K_{mA}$  and  $K_{mB}$  are Michaelis-Menten constants for substrate A and B ( $\mu\text{M}$ ).

Product inhibition data were globally fit to the competitive inhibition equation (Eq. 4) or the mixed-type inhibition equation (Eq. 5).

$$v = \frac{V_{max}}{\left[1 + \left(\frac{K_m}{[S]}\right) \left(1 + \frac{[I]}{K_{is}}\right)\right]} \quad \text{Eq. 4}$$



$$v = \frac{V_{max}[S]}{\left[ K_m \left( 1 + \frac{[I]}{K_{is}} \right) + [S] \left( 1 + \frac{[I]}{K_{ii}} \right) \right]} \quad \text{Eq. 5}$$

In each equation,  $v$  is the reaction rate;  $V_{max}$  is the maximum velocity of the reaction,  $[S]$  is varied substrate concentration;  $[I]$  is varied inhibitor concentration.  $K_{is}$  is the inhibition constant for the slope.  $K_{ii}$  is the inhibition constant for the intercept and  $K_m$  is the Michaelis-Menten constant of substrate S.

### *Protein Crystallization*

Full length purified *Mtb* IDH2 pre-incubated with 2 mM NADP<sup>+</sup> were screened for the crystallization conditions with the Mosquito liquid dispenser (TTP Labtech, Cambridge, MA, USA) using the sitting drop vapor diffusion technique in the Crystallization Intelli-Plate 96-3 low-profile crystallization plate (Hampton Research). For each condition, 1  $\mu$ L of protein (10 mg/mL) and 1  $\mu$ L of crystallization formulation were mixed; the mixture was equilibrated with 50  $\mu$ L of the reservoir in the well. The NADP<sup>+</sup> bound *Mtb* IDH2 crystals were further reproduced and optimized via a hanging drop vapor diffusion method by incubating equal volume of 20 mg/mL *Mtb* IDH2 with 2 mM NADP<sup>+</sup> and the crystallization solution (8–10% Tacsimate pH 5.0, 18–20% PEG 3,350) at 18 °C for 1 to 2 weeks.

### *Data Collection, Structure Determination, and Refinement*

Crystals were cryo-protected with 30% glycerol and flash frozen prior to data collection. Diffraction data were collected at the Advanced Photon Source (APS) beamline 19ID, Argonne National Laboratory. Data Sets were indexed, integrated, and scaled by the HKL-3000<sup>276</sup> or HKL-2000 program package.<sup>202</sup> The space group of the NADP<sup>+</sup> complex crystal is *P1* and the unit cell parameters are  $a = 74.8 \text{ \AA}$ ,  $b = 110.9 \text{ \AA}$ ,  $c = 114.9 \text{ \AA}$ ,  $\alpha = 86.0^\circ$ ,  $\beta = 78.6^\circ$ , and  $\gamma = 76.8^\circ$ . There are four molecules in an asymmetric unit. Phase was obtained by molecular replacement using the known structure 1J1W<sup>241</sup> as the search model with the Molrep<sup>203</sup> in CCP4 packages.<sup>204</sup> The refinement was done by the Phenix<sup>205</sup> refinement packages, and the NCS function was applied during the refinement. The manual model building was done by the program Coot.<sup>206</sup>

### *Analytical Size Exclusion Chromatography*

The oligomeric state of the purified WT or mutated Mtb IDHs was estimated by the size exclusion chromatography. The samples (0.1 mL, 10mg/mL) were loaded to a Superdex200 10/300 column (GE Healthcare Bio-sciences) equilibrated with binding buffer (50 mM sodium phosphate buffer pH 6.5 and 50 mM NaCl). Protein standards (Bio-Rad) for calibrating gels were Thyroglobulin (bovine) (670 kDa),  $\gamma$ -globulin (bovine) (158 kDa), Ovalbumin (chicken) (44 kDa), Myoglobin (horse) (17 kDa), and Vitamin B<sub>12</sub> (1.35kDa).

### *Site-Directed Mutagenesis of Mtb IDH2*

*Mtb* IDH2 in pET28b harboring a TEV cleavage site and a 6× His tag at the N-terminus was used as template. Selected residues were altered to the desired residues using the QuikChange® site-directed mutagenesis kit (Stratagene). Primer sequences for the E390R are 5'-CGCGCATCTATCAGCGGATCATCAACTTCTG-3' and 5'-CAGAAGTTGATGATCCGCTGATAGATGCGCG-3'; for R529E are 5'-CAAGTGCGATCCATGGAATACACGTGCGAGCGC-3' and 5'-GCGCTCGCACGTGTATTCCATGGATCGCACTTG-3'. Correct insertion of all mutations was confirmed by DNA sequencing. The modified plasmids were transformed into the *E. coli* Rosseta (DE3) strain (EMD Millipore). The expression and purification of mutated enzymes were followed the aforementioned procedures as the WT enzyme.

### *Enzyme Assay in 96-Well Microplates*

*Mtb* IDH2 activity was evaluated by measuring the production of NADPH in a continuous spectrophotometric assay. The reaction was conducted in 96-well half area UV-Star plates (Greiner) in a final volume of 100 µL at 25 °C. The selected compounds were dissolved and diluted to a desired concentration with 100% DMSO. Briefly, 90 µL reaction buffer (100 mM HEPES, pH7.5, and 20 mM MnCl<sub>2</sub>), 2 µL inhibitor solution and 2 µL *Mtb* IDH2 with a final concentration 0.5 nM were added into each well of the assay plate and incubated at room temperature for 10 min. The reaction was initiated by adding 2 µL NADP<sup>+</sup> (500 µM) and 4 µL DL-ICT (200 µM) into each well. The reaction was monitored by measuring absorbance at 340 nm using a Multiskan GO microplate

reader (Thermo Fisher Scientific Inc.) under a kinetic mode over 10 min in 10 sec intervals. The linear region of the curve was used to determine the initial velocity parameters. The initial velocity of the no enzyme and 0.5 nM enzyme reactions were used as the 100% and 0% inhibition during calculation. The *Mtb* IDH2 inhibitory ability was determined by the following equation (Eq. 6):

$$\%inhibition = \left[ \frac{S - PC}{NC - PC} \right] \times 100 \quad \text{Eq. 6}$$

in which S represents the initial velocity of the reaction with the test compound, PC is the initial velocity of the positive control (reaction without enzyme), and NC denotes the initial velocity of the negative control (reaction with DMSO only).

### *High-Throughput Screening*

HTS assays were carried out in a 96-well format (Corning 3635) using a Cybio automation system. The 1st and 12th column were used for positive (DMSO, no protein, substrate) and negative controls (DMSO, protein, substrate) of the reaction. Screening compounds were filled in the inner 10 columns. Screening compounds (50  $\mu$ M) were preincubated with 50 nM purified *Mtb* IDH2 in the reaction mixture containing 200  $\mu$ M NADP<sup>+</sup>, 5 mM MgCl<sub>2</sub>, 25 mM Tris, pH 7.5, 25 mM NaCl at room temperature for 10 min. The reaction was initiated by the addition of 5  $\mu$ L of DL-ICT to a final concentration of 200  $\mu$ M. The final reaction condition contains 2% DMSO. The absorbance at 340 nm for NADPH was measured by a POLARstar OPTIMA plate reader

(BMG LABTECH) at two time points immediately and 20 min after the addition of the substrate. The difference between the two time points is the signal used in the calculation.  $Z'$  factor<sup>277</sup> was then calculated to determine the quality of the high-throughput screen using Eq. 7, in which  $\sigma$  represents the standard deviation,  $\mu$  is the average, + and – stand for the high and low controls.

$$Z' = 1 - \frac{3\sigma_+ + 3\sigma_-}{|\mu_+ - \mu_-|} \quad \text{Eq. 7}$$

A 75% inhibition was chosen as the initial hit cut-off threshold for this HTS assay. The hits were then picked from stock solutions and retested in duplicates. Validated hits were subsequently examined by a dose-response assay.

#### *Evaluation of $IC_{50}$ Values with Dose-Response Curves*

The setting of the dose-response assay is the same as the inhibition assay. To obtain a dose-response curve, compounds were tested at concentrations ranging from 30 nM to 50  $\mu$ M. Each reaction was performed in duplicates. The  $IC_{50}$  values were determined by the four parameter logistic equation (Eq. 1) using the SigmaPlot program (Systat Software, San Jose, CA).

## CHAPTER V

### CONCLUSIONS AND FUTURE DIRECTIONS

Tuberculosis (TB) is an infectious disease caused by *Mycobacterium tuberculosis* (*Mtb*). The emergence and prevalence of drug-resistant TB underscore an urgent need for novel anti-tubercular agents. Shortening the treatment duration is another important mission to control TB since the long therapeutic course is the major cause of the emergence of drug resistant strains. Antibiotic persisters, especially the environmental stress induced persisters, are thought to be responsible for the prolonged treatment. Therefore, the short term goal is to develop a drug for the drug resistant TB, and the ultimate goal is to find a way to effectively kill the persistent bacteria, which show high antibiotic tolerance.

Drug discovery is a long term project. The entire procedure takes on average about 7–10 years from early research to pre-clinical testing and then the clinical trials. The early research includes the target identification and the lead development. Two approaches are used for the early research: target-based and phenotypic approaches. The target-based approach starts from choosing a target and followed by a hit identification through an enzymatic based screening. The hit compound would be optimized in a structural-activity relationship (SAR) study and ultimately becomes a lead for the preclinical trials. Alternatively, the phenotypic approach initiates by screening for compounds with therapeutic effects in cells, tissues or whole organisms. Target

identification, mechanism of action determination, and lead optimization experiments will be performed on the compound of interest. The obtained lead then tests in the preclinical trials. In this dissertation, some basic science research involved in the target identification and evaluation were carried out to enhance the knowledge for developing a novel anti-tubercular drug.

After decades of study, the productivity of the target-based drug discovery is lower than expected. One common cause of failure is the overestimation of the genetic results. Frequently, the genetic essentiality does not correlate with the *in vivo* requirement of its gene product. Therefore, the targets of the currently used pharmaceutical drug are the best choice since they passed the clinic validation. The molecular targets of most clinically used anti-tubercular agents are well studied. However, the study about the mechanism of action of PAS was rare since it has been removed from the first line prescription. Understanding the mechanism of action of PAS provides the platform for developing the next generation drug by improving the potency and/or lowering the toxicity. Isolation and screening for the genomic changes associated with PAS resistance were applied to examine the mechanism of action of PAS. Three folate pathway related enzymes (ThyA, DHFR, and FolC) were considered to be potential candidates, and followed by the overexpression experiments. The results suggested that DHFR is most likely to be the intracellular target of PAS. Based on this assumption, it is surprised to find that a strain with a mutation in the non-coding region upstream of the gene *Rv2671*, which was annotated as a riboflavin reductase, would cause PAS resistance. A complete biochemical and structure study was performed to

elucidate the molecular based of this mutant. The purified Rv2671 turns out to possess a low DHFR activity. The DHFR activity of Rv2671 indeed reinforces the idea that *Mtb* DHFR is the target of PAS. The mechanism of the PAS resistant mutant, which mediates by the high levels of Rv2671, is through a metabolic bypass. This study suggested that *Mtb* DHFR is a good target for the target-based drug screening, and the Rv2671 inhibitor would resolve the PAS resistant strain caused by a high level of Rv2671.

As mentioned earlier, sometimes the gene products of the essential genes are not required for bacterial growth *in vivo*. Failure to adapt the essentiality results from the genetic study can be a consequence of various mechanisms. A metabolic bypass is one of them. *Mtb* ICLs are attractive targets for persistent *Mtb*. *Mtb* is unable to sustain in hosts and is avirulent when *Mtb* ICLs are knocked-out. *Mtb* ICLs are involved in the glyoxylate shunt and are located at the branch point of the glyoxylate shunt and the TCA cycle. Thus, the TCA cycle can serve as an alternative pathway to generate energy and biomass for cell survival when *Mtb* ICLs are inhibited. Hence, to effectively kill *Mtb*, it is necessarily to block both routes such as inhibiting both branch point enzymes, *Mtb* ICLs and *Mtb* IDHs. The biochemical and biophysical properties of *Mtb* IDH2 were detailed characterized to examine its druggability. The crystal structure of *Mtb* IDH2 in complex with NADP<sup>+</sup> was determined at 2.5 Å resolution. The differences of the NADP<sup>+</sup> binding site between *Mtb* IDH2 and human IDHs indicate the potential to find inhibitors that are specific for monomeric IDHs. Inhibitor studies were performed for *Mtb* IDH2 with a repurpose screen and an enzymatically high-throughput screening with two libraries composed of anti-tubercular agents. Two compounds, one from the HIDHs



inhibitor collection and one from the whole cell active collections (6144788), showed low  $\mu\text{M}$   $\text{IC}_{50}$  values against *Mtb* IDH2. The structural and inhibitor studies of *Mtb* IDH2 can be valuable starting points for other researchers.

In this dissertation, DHFR was identified to be a good target for the follow-up target-based screening. In addition, for *Mtb* IDH2, a high-throughput enzymatic assay was developed, and preliminary inhibitor studies were conducted. Despite the valuable findings, there are several unanswered questions. First, even though all the evidence suggests that DHFR is the molecular target of PAS, the intracellular active form of PAS is not yet identified. The interaction between *Mtb* DHFR and the intracellular active molecule also needs to be established. For *Mtb* IDH2, the yield compounds still awaits optimization. A synergistic effect test between the inhibitors for *Mtb* ICLs and *Mtb* IDHs would be the next step; once potent compounds are obtained.

## REFERENCES

- [1] Zumla, A., George, A., Sharma, V., Herbert, R. H., Oxley, A., and Oliver, M. (2015) The WHO 2014 global tuberculosis report--further to go, *Lancet Glob Health* 3, e10-12.
- [2] Crofton, J., and Mitchison, D. A. (1948) Streptomycin Resistance in Pulmonary Tuberculosis, *Br Med J* 2, 1009-1015.
- [3] Medical Research Council. (1948) Streptomycin Treatment of Pulmonary Tuberculosis: A Medical Research Council Investigation, *Br Med J* 2, 769-782.
- [4] Canetti, G. (1965) Present aspects of bacterial resistance in tuberculosis, *Am Rev Respir Dis* 92, 687-703.
- [5] David, H. L. (1970) Probability Distribution of Drug-Resistant Mutants in Unselected Populations of Mycobacterium tuberculosis, *Appl Microbiol* 20, 810-814.
- [6] Frieden, T. R., Sterling, T., Pablos-Mendez, A., Kilburn, J. O., Cauthen, G. M., and Dooley, S. W. (1993) The emergence of drug-resistant tuberculosis in New York City, *N Engl J Med* 328, 521-526.
- [7] World Health Organization (2006) Extensively drug-resistant tuberculosis (XDR-TB): recommendations for prevention and control, *Wkly Epidemiol Rec* 81, 430-432.
- [8] Parida, S. K., Axelsson-Robertson, R., Rao, M. V., Singh, N., Master, I., Lutckii, A., Keshavjee, S., Andersson, J., Zumla, A., and Maeurer, M. (2015) Totally drug-resistant tuberculosis and adjunct therapies, *J Intern Med* 277, 388-405.
- [9] Migliori, G. B., De Iaco, G., Besozzi, G., Centis, R., and Cirillo, D. M. (2007) First tuberculosis cases in Italy resistant to all tested drugs, *Euro Surveill* 12, E0705171.
- [10] Klopper, M., Warren, R. M., Hayes, C., Gey van Pittius, N. C., Streicher, E. M.,

- Muller, B., Sirgel, F. A., Chabula-Nxiweni, M., Hoosain, E., Coetzee, G., David van Helden, P., Victor, T. C., and Trollip, A. P. (2013) Emergence and spread of extensively and totally drug-resistant tuberculosis, South Africa, *Emerg Infect Dis* 19, 449-455.
- [11] Udhwadia, Z. F., Amale, R. A., Ajbani, K. K., and Rodrigues, C. (2012) Totally drug-resistant tuberculosis in India, *Clin Infect Dis* 54, 579-581.
- [12] Velayati, A. A., Masjedi, M. R., Farnia, P., Tabarsi, P., Ghanavi, J., Ziazarifi, A. H., and Hoffner, S. E. (2009) Emergence of new forms of totally drug-resistant tuberculosis bacilli: super extensively drug-resistant tuberculosis or totally drug-resistant strains in iran, *Chest* 136, 420-425.
- [13] Centers for Disease Control and Prevention. (2003) Treatment of tuberculosis, In *MMWR Recomm Rep* 2003/07/03 ed., pp 1-77.
- [14] Bigger, J. (1944) Treatment of Staphylosoccal infections with penicillin by intermittent sterilisation, *The Lancet* 244, 497-500.
- [15] Sharma, S. V., Lee, D. Y., Li, B., Quinlan, M. P., Takahashi, F., Maheswaran, S., McDermott, U., Azizian, N., Zou, L., Fischbach, M. A., Wong, K. K., Brandstetter, K., Wittner, B., Ramaswamy, S., Classon, M., and Settleman, J. (2010) A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations, *Cell* 141, 69-80.
- [16] Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L., and Leibler, S. (2004) Bacterial persistence as a phenotypic switch, *Science* 305, 1622-1625.
- [17] Luidalepp, H., Joers, A., Kaldalu, N., and Tenson, T. (2011) Age of inoculum strongly influences persister frequency and can mask effects of mutations implicated in altered persistence, *J Bacteriol* 193, 3598-3605.
- [18] Jain, S. K., Lamichhane, G., Nimmagadda, S., Pomper, M. G., and Bishai, W. R. (2008) Antibiotic treatment of tuberculosis: old problems, new solutions, *Microbe* 3, 285.
- [19] Dye, C., Scheele, S., Dolin, P., Pathania, V., and Ravigliione, M. C. (1999)

Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project, *Jama* 282, 677-686.

- [20] World Health Organization (2015) *Guidelines on the management of latent tuberculosis infection*, World Health Organization, Geneva, Switzerland.
- [21] Zhang, Y. (2004) Persistent and dormant tubercle bacilli and latent tuberculosis, *Front Biosci* 9, 1136-1156.
- [22] Manina, G., and McKinney, J. D. (2013) A single-cell perspective on non-growing but metabolically active (NGMA) bacteria, *Curr Top Microbiol Immunol* 374, 135-161.
- [23] Xu, H. S., Roberts, N., Singleton, F. L., Attwell, R. W., Grimes, D. J., and Colwell, R. R. (1982) Survival and viability of nonculturable *Escherichia coli* and *Vibrio cholerae* in the estuarine and marine environment, *Microb Ecol* 8, 313-323.
- [24] Kaprelyants, A. S., Gottschal, J. C., and Kell, D. B. (1993) Dormancy in non-sporulating bacteria, *FEMS Microbiol Lett* 104, 271-286.
- [25] Li, L., Mendis, N., Trigui, H., Oliver, J. D., and Faucher, S. P. (2014) The importance of the viable but non-culturable state in human bacterial pathogens, *Front Microbiol* 5, 258.
- [26] Hurst, A. (1977) Bacterial injury: a review, *Can J Microbiol* 23, 935-944.
- [27] Allegra, S., Berger, F., Berthelot, P., Grattard, F., Pozzetto, B., and Riffard, S. (2008) Use of flow cytometry to monitor *Legionella* viability, *Appl Environ Microbiol* 74, 7813-7816.
- [28] Alifano, P., Bruni, C. B., and Carlomagno, M. S. (1994) Control of mRNA processing and decay in prokaryotes, *Genetica* 94, 157-172.
- [29] J., B. (1993) mRNA degradation in prokaryotic cells: an overview., In *Control of messenger RNA stability*. (Belasco J, B. G., Ed.), pp 3-12, Academic Press, Inc., San

Diego, Calif.

- [30] Sheridan, G. E. C., Masters, C. I., Shallcross, J. A., and Mackey, B. M. (1998) Detection of mRNA by Reverse Transcription-PCR as an Indicator of Viability in *Escherichia coli* Cells, *Appl Environ Microbiol* 64, 1313-1318.
- [31] Biosca, E. G., Amaro, C., Marco-Noales, E., and Oliver, J. D. (1996) Effect of low temperature on starvation-survival of the eel pathogen *Vibrio vulnificus* biotype 2, *Appl Environ Microbiol* 62, 450-455.
- [32] Du, M., Chen, J., Zhang, X., Li, A., Li, Y., and Wang, Y. (2007) Retention of virulence in a viable but nonculturable *Edwardsiella tarda* isolate, *Appl Environ Microbiol* 73, 1349-1354.
- [33] Cunningham E., O. B. C., Oliver J. D., (2009) Effect of weak acids on *Listeria monocytogenes* survival: Evidence for a viable but nonculturable state in response to low pH, *Food Control* 20, 1141-1144.
- [34] Roszak, D. B., Grimes, D. J., and Colwell, R. R. (1984) Viable but nonrecoverable stage of *Salmonella enteritidis* in aquatic systems, *Can J Microbiol* 30, 334-338.
- [35] Bates, T. C., and Oliver, J. D. (2004) The viable but nonculturable state of Kanagawa positive and negative strains of *Vibrio parahaemolyticus*, *J Microbiol* 42, 74-79.
- [36] Pinto, D., Santos, M. A., and Chambel, L. (2015) Thirty years of viable but nonculturable state research: unsolved molecular mechanisms, *Crit Rev Microbiol* 41, 61-76.
- [37] Senoh, M., Ghosh-Banerjee, J., Ramamurthy, T., Hamabata, T., Kurakawa, T., Takeda, M., Colwell, R. R., Nair, G. B., and Takeda, Y. (2010) Conversion of viable but nonculturable *Vibrio cholerae* to the culturable state by co-culture with eukaryotic cells, *Microbiol Immunol* 54, 502-507.
- [38] Pinto, D., Almeida, V., Almeida Santos, M., and Chambel, L. (2011) Resuscitation of *Escherichia coli* VBNC cells depends on a variety of environmental or chemical stimuli, *J Appl Microbiol* 110, 1601-1611.

- [39] Mukamolova, G. V., Kaprelyants, A. S., Young, D. I., Young, M., and Kell, D. B. (1998) A bacterial cytokine, *Proc Natl Acad Sci U S A* 95, 8916-8921.
- [40] Sperandio, V., Torres, A. G., Jarvis, B., Nataro, J. P., and Kaper, J. B. (2003) Bacteria-host communication: the language of hormones, *Proc Natl Acad Sci U S A* 100, 8951-8956.
- [41] Schaible, U. E., Sturgill-Koszycki, S., Schlesinger, P. H., and Russell, D. G. (1998) Cytokine activation leads to acidification and increases maturation of Mycobacterium avium-containing phagosomes in murine macrophages, *J Immunol* 160, 1290-1296.
- [42] Fallahi-Sichani, M., Marino, S., Flynn, J. L., Linderman, J. J., and Kirschner, D. E. (2013) A systems biology approach for understanding granuloma formation and function in tuberculosis., In *Systems biology of Tuberculosis*, pp 127-155, Springer, New York.
- [43] Hunter, R. L. (2011) Pathology of post primary tuberculosis of the lung: An illustrated critical review, *Tuberculosis* 91, 497-509.
- [44] Tsai, M. C., Chakravarty, S., Zhu, G., Xu, J., Tanaka, K., Koch, C., Tufariello, J., Flynn, J., and Chan, J. (2006) Characterization of the tuberculous granuloma in murine and human lungs: cellular composition and relative tissue oxygen tension, *Cell Microbiol* 8, 218-232.
- [45] Wayne, L. G. (1977) Synchronized replication of Mycobacterium tuberculosis, *Infect Immun* 17, 528-530.
- [46] Schumacher, M. A., Piro, K. M., Xu, W., Hansen, S., Lewis, K., and Brennan, R. G. (2009) Molecular mechanisms of HipA-mediated multidrug tolerance and its neutralization by HipB, *Science* 323, 396-401.
- [47] Yamaguchi, Y., Park, J. H., and Inouye, M. (2011) Toxin-antitoxin systems in bacteria and archaea, *Annu Rev Genet* 45, 61-79.
- [48] Wang, X., Lord, D. M., Cheng, H. Y., Osbourne, D. O., Hong, S. H., Sanchez-Torres, V., Quiroga, C., Zheng, K., Herrmann, T., Peti, W., Benedik, M. J., Page, R.,

- and Wood, T. K. (2012) A new type V toxin-antitoxin system where mRNA for toxin GhoT is cleaved by antitoxin GhoS, *Nat Chem Biol* 8, 855-861.
- [49] Masuda, H., Tan, Q., Awano, N., Wu, K. P., and Inouye, M. (2012) YeeU enhances the bundling of cytoskeletal polymers of MreB and FtsZ, antagonizing the CbtA (YeeV) toxicity in *Escherichia coli*, *Mol Microbiol* 84, 979-989.
- [50] Sala, A., Bordes, P., and Genevoux, P. (2014) Multiple toxin-antitoxin systems in *Mycobacterium tuberculosis*, *Toxins (Basel)* 6, 1002-1020.
- [51] Sala, A., Bordes, P., Fichant, G., and Genevoux, P. (2012) Toxin-Antitoxin Loci in *Mycobacterium tuberculosis*, In *Prokaryotic Toxin-Antitoxins*, pp 295-314, Springer Berlin Heidelberg.
- [52] Ramage, H. R., Connolly, L. E., and Cox, J. S. (2009) Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution, *PLoS Genet* 5, e1000767.
- [53] Albrethsen, J., Agner, J., Piersma, S. R., Hojrup, P., Pham, T. V., Weldingh, K., Jimenez, C. R., Andersen, P., and Rosenkrands, I. (2013) Proteomic profiling of *Mycobacterium tuberculosis* identifies nutrient-starvation-responsive toxin-antitoxin systems, *Mol Cell Proteomics* 12, 1180-1191.
- [54] 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) The role of RelMtb-mediated adaptation to stationary phase in long-term persistence of *Mycobacterium tuberculosis* in mice, *Proc Natl Acad Sci U S A* 100, 10026-10031.
- [55] Keren, I., Minami, S., Rubin, E., and Lewis, K. (2011) Characterization and transcriptome analysis of *Mycobacterium tuberculosis* persisters, *mBio* 2, e00100-00111.
- [56] Sharp, J. D., Cruz, J. W., Raman, S., Inouye, M., Husson, R. N., and Woychik, N. A. (2012) Growth and translation inhibition through sequence-specific RNA binding by *Mycobacterium tuberculosis* VapC toxin, *J Biol Chem* 287, 12835-12847.
- [57] Zhu, L., Zhang, Y., Teh, J. S., Zhang, J., Connell, N., Rubin, H., and Inouye, M.

- (2006) Characterization of mRNA interferases from *Mycobacterium tuberculosis*, *J Biol Chem* 281, 18638-18643.
- [58] Schifano, J. M., Edifor, R., Sharp, J. D., Ouyang, M., Konkimalla, A., Husson, R. N., and Woychik, N. A. (2013) Mycobacterial toxin MazF-mt6 inhibits translation through cleavage of 23S rRNA at the ribosomal A site, *Proc Natl Acad Sci U S A* 110, 8501-8506.
- [59] Zhu, L., Phadtare, S., Nariya, H., Ouyang, M., Husson, R. N., and Inouye, M. (2008) The mRNA interferases, MazF-mt3 and MazF-mt7 from *Mycobacterium tuberculosis* target unique pentad sequences in single-stranded RNA, *Mol Microbiol* 69, 559-569.
- [60] Huang, F., and He, Z. G. (2010) Characterization of an interplay between a *Mycobacterium tuberculosis* MazF homolog, Rv1495 and its sole DNA topoisomerase I, *Nucleic Acids Res* 38, 8219-8230.
- [61] Neubauer, C., Gao, Y. G., Andersen, K. R., Dunham, C. M., Kelley, A. C., Hentschel, J., Gerdes, K., Ramakrishnan, V., and Brodersen, D. E. (2009) The structural basis for mRNA recognition and cleavage by the ribosome-dependent endonuclease RelE, *Cell* 139, 1084-1095.
- [62] Zhang, Y., and Inouye, M. (2009) The inhibitory mechanism of protein synthesis by YoeB, an *Escherichia coli* toxin, *J Biol Chem* 284, 6627-6638.
- [63] Hurley, J. M., and Woychik, N. A. (2009) Bacterial toxin HigB associates with ribosomes and mediates translation-dependent mRNA cleavage at A-rich sites, *J Biol Chem* 284, 18605-18613.
- [64] Jiang, Y., Pogliano, J., Helinski, D. R., and Konieczny, I. (2002) ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of *Escherichia coli* gyrase, *Mol Microbiol* 44, 971-979.
- [65] Avarbock, D., Salem, J., Li, L. S., Wang, Z. M., and Rubin, H. (1999) Cloning and characterization of a bifunctional RelA/SpoT homologue from *Mycobacterium tuberculosis*, *Gene* 233, 261-269.



- [66] Boutte, C. C., and Crosson, S. (2013) Bacterial lifestyle shapes stringent response activation, *Trends in Microbiology* 21, 174-180.
- [67] Sureka, K., Dey, S., Datta, P., Singh, A. K., Dasgupta, A., Rodrigue, S., Basu, J., and Kundu, M. (2007) Polyphosphate kinase is involved in stress-induced mprAB-sigE-rel signalling in mycobacteria, *Mol Microbiol* 65, 261-276.
- [68] Chuang, Y. M., Bandyopadhyay, N., Rifat, D., Rubin, H., Bader, J. S., and Karakousis, P. C. (2015) Deficiency of the novel exopolyphosphatase Rv1026/PPX2 leads to metabolic downshift and altered cell wall permeability in Mycobacterium tuberculosis, *MBio* 6, e02428.
- [69] Thayil, S. M., Morrison, N., Schechter, N., Rubin, H., and Karakousis, P. C. (2011) The role of the novel exopolyphosphatase MT0516 in Mycobacterium tuberculosis drug tolerance and persistence, *PLoS One* 6, e28076.
- [70] Chuang, Y. M., Belchis, D. A., and Karakousis, P. C. (2013) The polyphosphate kinase gene ppk2 is required for Mycobacterium tuberculosis inorganic polyphosphate regulation and virulence, *MBio* 4, e00039-00013.
- [71] Choi, M. Y., Wang, Y., Wong, L. L., Lu, B. T., Chen, W. Y., Huang, J. D., Tanner, J. A., and Watt, R. M. (2012) The two PPX-GppA homologues from Mycobacterium tuberculosis have distinct biochemical activities, *PLoS One* 7, e42561.
- [72] Avarbock, D., Avarbock, A., and Rubin, H. (2000) Differential regulation of opposing RelMtb activities by the aminoacylation state of a tRNA.ribosome.mRNA.RelMtb complex, *Biochemistry* 39, 11640-11648.
- [73] Jain, V., Saleem-Batcha, R., China, A., and Chatterji, D. (2006) Molecular dissection of the mycobacterial stringent response protein Rel, *Protein Sci* 15, 1449-1464.
- [74] Potrykus, K., and Cashel, M. (2008) (p)ppGpp: still magical?, *Annu Rev Microbiol* 62, 35-51.
- [75] Paul, B. J., Berkmen, M. B., and Gourse, R. L. (2005) DksA potentiates direct activation of amino acid promoters by ppGpp, *Proc Natl Acad Sci U S A* 102, 7823-

7828.

- [76] Kriel, A., Bittner, A. N., Kim, S. H., Liu, K., Tehranchi, A. K., Zou, W. Y., Rendon, S., Chen, R., Tu, B. P., and Wang, J. D. (2012) Direct regulation of GTP homeostasis by (p)ppGpp: a critical component of viability and stress resistance, *Mol Cell* 48, 231-241.
- [77] Tare, P., Mallick, B., and Nagaraja, V. (2013) Co-evolution of specific amino acid in sigma 1.2 region and nucleotide base in the discriminator to act as sensors of small molecule effectors of transcription initiation in mycobacteria, *Mol Microbiol* 90, 569-583.
- [78] Primm, T. P., Andersen, S. J., Mizrahi, V., Avarbock, D., Rubin, H., and Barry, C. E., 3rd. (2000) The stringent response of *Mycobacterium tuberculosis* is required for long-term survival, *J Bacteriol* 182, 4889-4898.
- [79] Klinkenberg, L. G., Lee, J. H., Bishai, W. R., and Karakousis, P. C. (2010) The stringent response is required for full virulence of *Mycobacterium tuberculosis* in guinea pigs, *J Infect Dis* 202, 1397-1404.
- [80] Dasgupta, N., Kapur, V., Singh, K. K., Das, T. K., Sachdeva, S., Jyothisri, K., and Tyagi, J. S. (2000) Characterization of a two-component system, devR-devS, of *Mycobacterium tuberculosis*, *Tuber Lung Dis* 80, 141-159.
- [81] Malhotra, V., Sharma, D., Ramanathan, V. D., Shakila, H., Saini, D. K., Chakravorty, S., Das, T. K., Li, Q., Silver, R. F., Narayanan, P. R., and Tyagi, J. S. (2004) Disruption of response regulator gene, devR, leads to attenuation in virulence of *Mycobacterium tuberculosis*, *FEMS Microbiol Lett* 231, 237-245.
- [82] Voskuil, M. I., Schnappinger, D., Visconti, K. C., Harrell, M. I., Dolganov, G. M., Sherman, D. R., and Schoolnik, G. K. (2003) Inhibition of respiration by nitric oxide induces a *Mycobacterium tuberculosis* dormancy program, *J Exp Med* 198, 705-713.
- [83] Muttucumaru, D. G., Roberts, G., Hinds, J., Stabler, R. A., and Parish, T. (2004) Gene expression profile of *Mycobacterium tuberculosis* in a non-replicating state, *Tuberculosis (Edinb)* 84, 239-246.

- [84] Kumar, A., Deshane, J. S., Crossman, D. K., Bolisetty, S., Yan, B. S., Kramnik, I., Agarwal, A., and Steyn, A. J. (2008) Heme oxygenase-1-derived carbon monoxide induces the Mycobacterium tuberculosis dormancy regulon, *J Biol Chem* 283, 18032-18039.
- [85] Sivaramakrishnan, S., and Ortiz de Montellano, P. R. (2013) The DosS-DosT/DosR Mycobacterial Sensor System, *Biosensors (Basel)* 3, 259-282.
- [86] Saini, D. K., Malhotra, V., Dey, D., Pant, N., Das, T. K., and Tyagi, J. S. (2004) DevR-DevS is a bona fide two-component system of Mycobacterium tuberculosis that is hypoxia-responsive in the absence of the DNA-binding domain of DevR, *Microbiology* 150, 865-875.
- [87] Roberts, D. M., Liao, R. P., Wisedchaisri, G., Hol, W. G., and Sherman, D. R. (2004) Two sensor kinases contribute to the hypoxic response of Mycobacterium tuberculosis, *J Biol Chem* 279, 23082-23087.
- [88] Rustad, T. R., Harrell, M. I., Liao, R., and Sherman, D. R. (2008) The enduring hypoxic response of Mycobacterium tuberculosis, *PLoS One* 3, e1502.
- [89] Shi, L., Sohaskey, C. D., Kana, B. D., Dawes, S., North, R. J., Mizrahi, V., and Gennaro, M. L. (2005) Changes in energy metabolism of Mycobacterium tuberculosis in mouse lung and under in vitro conditions affecting aerobic respiration, *Proc Natl Acad Sci U S A* 102, 15629-15634.
- [90] Eoh, H., and Rhee, K. Y. (2013) Multifunctional essentiality of succinate metabolism in adaptation to hypoxia in Mycobacterium tuberculosis, *Proc Natl Acad Sci U S A* 110, 6554-6559.
- [91] Watanabe, S., Zimmermann, M., Goodwin, M. B., Sauer, U., Barry, C. E., 3rd, and Boshoff, H. I. (2011) Fumarate reductase activity maintains an energized membrane in anaerobic Mycobacterium tuberculosis, *PLoS Pathog* 7, e1002287.
- [92] Sohaskey, C. D., and Wayne, L. G. (2003) Role of narK2X and narGHJI in hypoxic upregulation of nitrate reduction by Mycobacterium tuberculosis, *J Bacteriol* 185, 7247-7256.

- [93] Andries, K., Verhasselt, P., Guillemont, J., Göhlmann, H. W. H., Neefs, J.-M., Winkler, H., Gestel, J. V., Timmerman, P., Zhu, M., Lee, E., Williams, P., Chaffoy, D. d., Huitric, E., Hoffner, S., Cambau, E., Truffot-Pernot, C., Lounis, N., and Jarlier, V. (2005) A Diarylquinoline Drug Active on the ATP Synthase of Mycobacterium tuberculosis, *Science* 307, 223-227.
- [94] Koul, A., Dendouga, N., Vergauwen, K., Molenberghs, B., Vranckx, L., Willebrords, R., Ristic, Z., Lill, H., Dorange, I., Guillemont, J., Bald, D., and Andries, K. (2007) Diarylquinolines target subunit c of mycobacterial ATP synthase, *Nat Chem Biol* 3, 323-324.
- [95] Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., and et al. (1995) Whole-genome random sequencing and assembly of Haemophilus influenzae Rd, *Science* 269, 496-512.
- [96] Collins, F. S., Green, E. D., Guttmacher, A. E., and Guyer, M. S. (2003) A vision for the future of genomics research, *Nature* 422, 835-847.
- [97] Sams-Dodd, F. (2005) Target-based drug discovery: is something wrong?, *Drug Discov Today* 10, 139-147.
- [98] Swinney, D. C., and Anthony, J. (2011) How were new medicines discovered?, *Nature Reviews Drug Discovery* 10, 507-519.
- [99] Ventola, C. L. (2015) The Antibiotic Resistance Crisis: Part 1: Causes and Threats, *P T* 40, 277-283.
- [100] Centers for Disease Control and Prevention (2013) Antibiotic resistance threats in the United States, 2013 | Antibiotic/Antimicrobial Resistance | CDC, CDC, Atlanta.
- [101] Brown, D. (2007) Unfinished business: target-based drug discovery, *Drug Discov Today* 12, 1007-1012.
- [102] Payne, D. J., Gwynn, M. N., Holmes, D. J., and Pompliano, D. L. (2006) Drugs for bad bugs: confronting the challenges of antibacterial discovery, *Nature Reviews Drug Discovery* 6, 29-40.

- [103] Wayne, L. G., and Hayes, L. G. (1996) An in vitro model for sequential study of shutdown of *Mycobacterium tuberculosis* through two stages of nonreplicating persistence, *Infect Immun* 64, 2062-2069.
- [104] Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A., and Duncan, K. (2002) Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling, *Mol Microbiol* 43, 717-731.
- [105] Hampshire, T., Soneji, S., Bacon, J., James, B. W., Hinds, J., Laing, K., Stabler, R. A., Marsh, P. D., and Butcher, P. D. (2004) Stationary phase gene expression of *Mycobacterium tuberculosis* following a progressive nutrient depletion: a model for persistent organisms?, *Tuberculosis (Edinb)* 84, 228-238.
- [106] Karakousis, P. C., Yoshimatsu, T., Lamichhane, G., Woolwine, S. C., Nuermberger, E. L., Grosset, J., and Bishai, W. R. (2004) Dormancy phenotype displayed by extracellular *Mycobacterium tuberculosis* within artificial granulomas in mice, *J Exp Med* 200, 647-657.
- [107] Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence, *Nature* 393, 537-544.
- [108] Baughn, A. D., and Rhee, K. Y. (2014) Metabolomics of Central Carbon Metabolism in *Mycobacterium tuberculosis*, *Microbiol Spectr* 2, 1-16.
- [109] Rhee, K. Y., de Carvalho, L. P., Bryk, R., Ehrt, S., Marrero, J., Park, S. W., Schnappinger, D., Venugopal, A., and Nathan, C. (2011) Central carbon metabolism in *Mycobacterium tuberculosis*: an unexpected frontier, *Trends Microbiol* 19, 307-314.
- [110] Gorke, B., and Stulke, J. (2008) Carbon catabolite repression in bacteria: many ways to make the most out of nutrients, *Nat Rev Microbiol* 6, 613-624.

- [111] de Carvalho, L. P. S., Fischer, S. M., Marrero, J., Nathan, C., Ehrt, S., and Rhee, K. Y. (2010) Metabolomics of *Mycobacterium tuberculosis* Reveals Compartmentalized Co-Catabolism of Carbon Substrates, *Chemistry & Biology* 17, 1122-1131.
- [112] Beste, D. J., Noh, K., Niedenfuhr, S., Mendum, T. A., Hawkins, N. D., Ward, J. L., Beale, M. H., Wiechert, W., and McFadden, J. (2013) <sup>13</sup>C-flux spectral analysis of host-pathogen metabolism reveals a mixed diet for intracellular *Mycobacterium tuberculosis*, *Chem Biol* 20, 1012-1021.
- [113] Ehrt, S., and Rhee, K. (2013) *Mycobacterium tuberculosis* metabolism and host interaction: mysteries and paradoxes, *Curr Top Microbiol Immunol* 374, 163-188.
- [114] Gould, T. A., van de Langemheen, H., Munoz-Elias, E. J., McKinney, J. D., and Sacchettini, J. C. (2006) Dual role of isocitrate lyase 1 in the glyoxylate and methylcitrate cycles in *Mycobacterium tuberculosis*, *Mol Microbiol* 61, 940-947.
- [115] McKinney, J. D., Bentrup, K. H. z., Muñoz-Eliás, E. J., Miczak, A., Chen, B., Chan, W.-T., Swenson, D., Sacchettini, J. C., Jacobs, W. R., and Russell, D. G. (2000) Persistence of *Mycobacterium tuberculosis* in macrophages and micerequires the glyoxylate shunt enzyme isocitrate lyase, *Nature* 406, 735-738.
- [116] Marrero, J., Rhee, K. Y., Schnappinger, D., Pethe, K., and Ehrt, S. (2010) Gluconeogenic carbon flow of tricarboxylic acid cycle intermediates is critical for *Mycobacterium tuberculosis* to establish and maintain infection, *Proc Natl Acad Sci U S A* 107, 9819-9824.
- [117] Sassetti, C. M., and Rubin, E. J. (2003) Genetic requirements for mycobacterial survival during infection, *Proc Natl Acad Sci U S A* 100, 12989-12994.
- [118] Mueller, C. (2009) Regulation of the branch point between the glyoxylate shunt and the TCA cycle in mycobacteria, In *Department of Microbiology*, pp 1-139, National University of Singapore, Singapore.
- [119] Murray, J. F., Schraufnagel, D. E., and Hopewell, P. C. (2015) Treatment of Tuberculosis. A Historical Perspective, *Ann Am Thorac Soc* 12, 1749-1759.
- [120] Lehmann, J. (1946) Para-aminosalicylic acid in the treatment of tuberculosis,

*Lancet* 250, 15-16.

- [121] British Medical Research Council. (1973) Co-operative controlled trial of a standard regimen of streptomycin, PAS and isoniazid and three alternative regimens of chemotherapy in Britain. A report from the British Medical Research Council, *Tubercle* 54, 99-129.
- [122] Centers for Disease Control and Prevention (1992) Update: availability of streptomycin and para-aminosalicylic acid—United States., MMWR ed., p 482, Center for Disease Control and Prevention, U.S.
- [123] Collins, L., and Franzblau, S. G. (1997) Microplate alamar blue assay versus BACTEC 460 system for high-throughput screening of compounds against *Mycobacterium tuberculosis* and *Mycobacterium avium*, *Antimicrob Agents Chemother* 41, 1004-1009.
- [124] Peloquin, C. A., Berning, S. E., Huitt, G. A., Childs, J. M., Singleton, M. D., and James, G. T. (1999) Once-daily and twice-daily dosing of p-aminosalicylic acid granules, *Am J Respir Crit Care Med* 159, 932-934.
- [125] Diczfaluzny E. (1947) Action of salicylic acid derivatives. I. Investigations on the action of PAS, *Ark Kemi* 24B, 1-8.
- [126] Taneja, N. K., and Tyagi, J. S. (2007) Resazurin reduction assays for screening of anti-tubercular compounds against dormant and actively growing *Mycobacterium tuberculosis*, *Mycobacterium bovis* BCG and *Mycobacterium smegmatis*, *J Antimicrob Chemother* 60, 288-293.
- [127] Swenson, J. M., Thornsberry, C., and Silcox, V. A. (1982) Rapidly growing mycobacteria: testing of susceptibility to 34 antimicrobial agents by broth microdilution, *Antimicrob Agents Chemother* 22, 186-192.
- [128] Henderson, J. F. (1979) Teaching one-carbon metabolism, *Biochemical Education* 7, 51-52.
- [129] de Crecy-Lagard, V., El Yacoubi, B., de la Garza, R. D., Noiriel, A., and Hanson, A. D. (2007) Comparative genomics of bacterial and plant folate synthesis and

salvage: predictions and validations, *BMC Genomics* 8, 245.

- [130] Brown, G. M., and Williamson, J. M. (1982) Biosynthesis of riboflavin, folic acid, thiamine, and pantothenic acid, *Adv Enzymol Relat Areas Mol Biol* 53, 345-381.
- [131] Bodaness, R. S. (1982) The potential role of NADPH and cytoplasmic NADP-linked dehydrogenases in protection against singlet oxygen mediated cellular toxicity, *Biochem Biophys Res Commun* 108, 1709-1715.
- [132] Dosselaere, F., and Vanderleyden, J. (2001) A metabolic node in action: chorismate-utilizing enzymes in microorganisms, *Crit Rev Microbiol* 27, 75-131.
- [133] Lu, Y. Z., Aiello, P. D., and Matthews, R. G. (1984) Studies on the polyglutamate specificity of thymidylate synthase from fetal pig liver, *Biochemistry* 23, 6870-6876.
- [134] Nopponpunth, V., Sirawaraporn, W., Greene, P. J., and Santi, D. V. (1999) Cloning and expression of Mycobacterium tuberculosis and Mycobacterium leprae dihydropteroate synthase in Escherichia coli, *J Bacteriol* 181, 6814-6821.
- [135] Rengarajan, J., Sasseti, C. M., Naroditskaya, V., Sloutsky, A., Bloom, B. R., and Rubin, E. J. (2004) The folate pathway is a target for resistance to the drug para-aminosalicylic acid (PAS) in mycobacteria, *Mol Microbiol* 53, 275-282.
- [136] Mathys, V., Wintjens, R., Lefevre, P., Bertout, J., Singhal, A., Kiass, M., Kurepina, N., Wang, X. M., Mathema, B., Baulard, A., Kreiswirth, B. N., and Bifani, P. (2009) Molecular genetics of para-aminosalicylic acid resistance in clinical isolates and spontaneous mutants of Mycobacterium tuberculosis, *Antimicrob Agents Chemother* 53, 2100-2109.
- [137] De Voss, J. J., Rutter, K., Schroeder, B. G., and Barry, C. E., 3rd. (1999) Iron acquisition and metabolism by mycobacteria, *J Bacteriol* 181, 4443-4451.
- [138] Ratledge, C., and Dover, L. G. (2000) Iron metabolism in pathogenic bacteria, *Annu Rev Microbiol* 54, 881-941.
- [139] Gobin, J., and Horwitz, M. A. (1996) Exochelins of Mycobacterium tuberculosis



remove iron from human iron-binding proteins and donate iron to mycobactins in the *M. tuberculosis* cell wall, *J Exp Med* 183, 1527-1532.

- [140] Braibant, M., Gilot, P., and Content, J. (2000) The ATP binding cassette (ABC) transport systems of *Mycobacterium tuberculosis*, *FEMS Microbiol Rev* 24, 449-467.
- [141] Ratledge, C., and Brown, K. A. (1972) Inhibition of mycobactin formation in *Mycobacterium smegmatis* by p-aminosalicylate. A new proposal for the mode of action of p-aminosalicylate, *Am Rev Respir Dis* 106, 774-776.
- [142] Zhang, X., Liu, L., Zhang, Y., Dai, G., Huang, H., and Jin, Q. (2015) Genetic determinants involved in p-aminosalicylic acid resistance in clinical isolates from tuberculosis patients in northern China from 2006 to 2012, *Antimicrob Agents Chemother* 59, 1320-1324.
- [143] Young, P. G., Smith, C. A., Metcalf, P., and Baker, E. N. (2008) Structures of *Mycobacterium tuberculosis* folylpolyglutamate synthase complexed with ADP and AMPPCP, *Acta Crystallogr D Biol Crystallogr* D64, 745-753.
- [144] Sassetti, C. M., Boyd, D. H., and Rubin, E. J. (2003) Genes required for mycobacterial growth defined by high density mutagenesis, *Mol Microbiol* 48, 77-84.
- [145] Moradigaravand, D., Grandjean, L., Martinez, E., Li, H., Zheng, J., Coronel, J., Moore, D., Torok, M. E., Sintchenko, V., Huang, H., Javid, B., Parkhill, J., Peacock, S. J., and Koser, C. U. (2016) DfrA-thyA double deletion in para-aminosalicylic acid resistant *Mycobacterium tuberculosis* Beijing strains, *Antimicrob Agents Chemother* 60, 3864-3867.
- [146] Martinez, E., Holmes, N., Jelfs, P., and Sintchenko, V. (2015) Genome sequencing reveals novel deletions associated with secondary resistance to pyrazinamide in MDR *Mycobacterium tuberculosis*, *J Antimicrob Chemother* 70, 2511-2514.
- [147] Carreras, C. W., and Santi, D. V. (1995) The catalytic mechanism and structure of thymidylate synthase, *Annu Rev Biochem* 64, 721-762.
- [148] Myllykallio, H., Lipowski, G., Leduc, D., Filee, J., Forterre, P., and Liebl, U.

(2002) An alternative flavin-dependent mechanism for thymidylate synthesis, *Science* 297, 105-107.

- [149] Fivian-Hughes, A. S., Houghton, J., and Davis, E. O. (2012) Mycobacterium tuberculosis thymidylate synthase gene thyX is essential and potentially bifunctional, while thyA deletion confers resistance to p-aminosalicylic acid, *Microbiology* 158, 308-318.
- [150] Hunter, J. H., Gujjar, R., Pang, C. K., and Rathod, P. K. (2008) Kinetics and ligand-binding preferences of Mycobacterium tuberculosis thymidylate synthases, ThyA and ThyX, *PLoS One* 3, e2237.
- [151] Ehrt, S., Guo, X. V., Hickey, C. M., Ryou, M., Monteleone, M., Riley, L. W., and Schnappinger, D. (2005) Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor, *Nucleic Acids Res* 33, e21.
- [152] Chakraborty, S., Gruber, T., Barry, C. E., 3rd, Boshoff, H. I., and Rhee, K. Y. (2013) Para-aminosalicylic acid acts as an alternative substrate of folate metabolism in Mycobacterium tuberculosis, *Science* 339, 88-91.
- [153] Zheng, J., Rubin, E. J., Bifani, P., Mathys, V., Lim, V., Au, M., Jang, J., Nam, J., Dick, T., Walker, J. R., Pethe, K., and Camacho, L. R. (2013) para-Aminosalicylic acid is a prodrug targeting dihydrofolate reductase in Mycobacterium tuberculosis, *J Biol Chem* 288, 23447-23456.
- [154] Li, R., Sirawaraporn, R., Chitnumsub, P., Sirawaraporn, W., Wooden, J., Athappilly, F., Turley, S., and Hol, W. G. (2000) Three-dimensional structure of M. tuberculosis dihydrofolate reductase reveals opportunities for the design of novel tuberculosis drugs, *J Mol Biol* 295, 307-323.
- [155] Nixon, M. R., Saionz, K. W., Koo, M. S., Szymonifka, M. J., Jung, H., Roberts, J. P., Nandakumar, M., Kumar, A., Liao, R., Rustad, T., Sacchetti, J. C., Rhee, K. Y., Freundlich, J. S., and Sherman, D. R. (2014) Folate pathway disruption leads to critical disruption of methionine derivatives in Mycobacterium tuberculosis, *Chem Biol* 21, 819-830.
- [156] Bystroff, C., Oatley, S. J., and Kraut, J. (1990) Crystal structures of Escherichia coli dihydrofolate reductase: the NADP<sup>+</sup> holoenzyme and the folate.NADP<sup>+</sup> ternary

complex. Substrate binding and a model for the transition state, *Biochemistry* 29, 3263-3277.

- [157] Volpato, J. P., Yachnin, B. J., Blanchet, J., Guerrero, V., Poulin, L., Fossati, E., Berghuis, A. M., and Pelletier, J. N. (2009) Multiple conformers in active site of human dihydrofolate reductase F31R/Q35E double mutant suggest structural basis for methotrexate resistance, *J Biol Chem* 284, 20079-20089.
- [158] Franklin, M. C., Cheung, J., Rudolph, M. J., Burshteyn, F., Cassidy, M., Gary, E., Hillerich, B., Yao, Z. K., Carlier, P. R., Totrov, M., and Love, J. D. (2015) Structural genomics for drug design against the pathogen *Coxiella burnetii*, *Proteins* 83, 2124-2136.
- [159] Hay, S., Evans, R. M., Levy, C., Loveridge, E. J., Wang, X., Leys, D., Allemann, R. K., and Scrutton, N. S. (2009) Are the catalytic properties of enzymes from piezophilic organisms pressure adapted?, *ChemBiochem* 10, 2348-2353.
- [160] Sawaya, M. R., and Kraut, J. (1997) Loop and subdomain movements in the mechanism of *Escherichia coli* dihydrofolate reductase: crystallographic evidence, *Biochemistry* 36, 586-603.
- [161] Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server, *Nucleic Acids Res* 42, W320-324.
- [162] Gupta, A., and Bhakta, S. (2012) An integrated surrogate model for screening of drugs against *Mycobacterium tuberculosis*, *J Antimicrob Chemother* 67, 1380-1391.
- [163] Li, G., Lian, L. L., Wan, L., Zhang, J., Zhao, X., Jiang, Y., Zhao, L. L., Liu, H., and Wan, K. (2013) Antimicrobial susceptibility of standard strains of nontuberculous mycobacteria by microplate Alamar Blue assay, *PLoS One* 8, e84065.
- [164] Larsen, M. H., Biermann, K., Tandberg, S., Hsu, T., and Jacobs, W. R., Jr. (2007) Genetic manipulation of *Mycobacterium tuberculosis*, *Curr Protoc Microbiol*, 10A12.11–10A12.21.
- [165] Galagan, J. E., Minch, K., Peterson, M., Lyubetskaya, A., Azizi, E., Sweet, L.,

Gomes, A., Rustad, T., Dolganov, G., Glotova, I., Abeel, T., Mahwinney, C., Kennedy, A. D., Allard, R., Brabant, W., Krueger, A., Jaini, S., Honda, B., Yu, W. H., Hickey, M. J., Zucker, J., Garay, C., Weiner, B., Sisk, P., Stolte, C., Winkler, J. K., Van de Peer, Y., Iazzetti, P., Camacho, D., Dreyfuss, J., Liu, Y., Dorhoi, A., Mollenkopf, H. J., Drogaris, P., Lamontagne, J., Zhou, Y., Piquenot, J., Park, S. T., Raman, S., Kaufmann, S. H., Mohny, R. P., Chelsky, D., Moody, D. B., Sherman, D. R., and Schoolnik, G. K. (2013) The Mycobacterium tuberculosis regulatory network and hypoxia, *Nature* 499, 178-183.

[166] Korkegian, A., Roberts, D. M., Blair, R., and Parish, T. (2014) Mutations in the essential arabinosyltransferase EmbC lead to alterations in Mycobacterium tuberculosis lipoarabinomannan, *J Biol Chem* 289, 35172-35181.

[167] Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R., and Sherman, W. (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments, *J Comput Aided Mol Des* 27, 221-234.

[168] Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., and Shenkin, P. S. (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, *J Med Chem* 47, 1739-1749.

[169] Rieckmann, K. H. (1973) The in vitro activity of experimental antimalarial compounds against strains of Plasmodium falciparum with varying degrees of sensitivity to pyrimethamine and chloroquine., In *Chemotherapy of malaria and resistance to antimalarials.*, p 58, World Health Organization, Geneva, Switzerland., WHO Tech. Rep.

[170] Bacher, A., Eberhardt, S., Fischer, M., Kis, K., and Richter, G. (2000) Biosynthesis of vitamin b2 (riboflavin), *Annu Rev Nutr* 20, 153-167.

[171] Graupner, M., Xu, H., and White, R. H. (2002) The pyrimidine nucleotide reductase step in riboflavin and F(420) biosynthesis in archaea proceeds by the eukaryotic route to riboflavin, *J Bacteriol* 184, 1952-1957.

[172] Behm-Ansmant, I., Grosjean, H., Massenet, S., Motorin, Y., and Branlant, C. (2004) Pseudouridylation at position 32 of mitochondrial and cytoplasmic tRNAs requires two distinct enzymes in Saccharomyces cerevisiae, *J Biol Chem* 279, 52998-53006.

- [173] Czekster, C. M., Vandemeulebroucke, A., and Blanchard, J. S. (2011) Kinetic and chemical mechanism of the dihydrofolate reductase from *Mycobacterium tuberculosis*, *Biochemistry* 50, 367-375.
- [174] Argyrou, A., Vetting, M. W., Aladegbami, B., and Blanchard, J. S. (2006) *Mycobacterium tuberculosis* dihydrofolate reductase is a target for isoniazid, *Nat Struct Mol Biol* 13, 408-413.
- [175] Bock, R. A., Soulages, J. L., and Barrow, W. W. (2007) Substrate and inhibitor specificity of *Mycobacterium avium* dihydrofolate reductase, *Febs J* 274, 3286-3298.
- [176] Maglia, G., Javed, M. H., and Allemann, R. K. (2003) Hydride transfer during catalysis by dihydrofolate reductase from *Thermotoga maritima*, *Biochem J* 374, 529-535.
- [177] Fierke, C. A., Johnson, K. A., and Benkovic, S. J. (1987) Construction and evaluation of the kinetic scheme associated with dihydrofolate reductase from *Escherichia coli*, *Biochemistry* 26, 4085-4092.
- [178] Lee, J., Yennawar, N. H., Gam, J., and Benkovic, S. J. (2010) Kinetic and structural characterization of dihydrofolate reductase from *Streptococcus pneumoniae*, *Biochemistry* 49, 195-206.
- [179] Thillet, J., Adams, J. A., and Benkovic, S. J. (1990) The kinetic mechanism of wild-type and mutant mouse dihydrofolate reductases, *Biochemistry* 29, 5195-5202.
- [180] Appleman, J. R., Beard, W. A., Delcamp, T. J., Prendergast, N. J., Freisheim, J. H., and Blakley, R. L. (1990) Unusual transient- and steady-state kinetic behavior is predicted by the kinetic scheme operational for recombinant human dihydrofolate reductase, *J Biol Chem* 265, 2740-2748.
- [181] Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state, *J Mol Biol* 372, 774-797.
- [182] Lv, Z., Sun, J., and Liu, Y. (2013) Structural and functional insights into *Saccharomyces cerevisiae* riboflavin biosynthesis reductase RIB7, *PLoS One* 8,

e61249.

- [183] Chatwell, L., Krojer, T., Fidler, A., Romisch, W., Eisenreich, W., Bacher, A., Huber, R., and Fischer, M. (2006) Biosynthesis of riboflavin: structure and properties of 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate reductase of *Methanocaldococcus jannaschii*, *J Mol Biol* 359, 1334-1351.
- [184] Stenmark, P., Moche, M., Gurmu, D., and Nordlund, P. (2007) The crystal structure of the bifunctional deaminase/reductase RibD of the riboflavin biosynthetic pathway in *Escherichia coli*: implications for the reductive mechanism, *J Mol Biol* 373, 48-64.
- [185] Chen, S. C., Chang, Y. C., Lin, C. H., and Liaw, S. H. (2006) Crystal structure of a bifunctional deaminase and reductase from *Bacillus subtilis* involved in riboflavin biosynthesis, *J Biol Chem* 281, 7605-7613.
- [186] Wilquet, V., Gaspar, J. A., van de Lande, M., Van de Castele, M., Legrain, C., Meiering, E. M., and Glansdorff, N. (1998) Purification and characterization of recombinant *Thermotoga maritima* dihydrofolate reductase, *Eur J Biochem* 255, 628-637.
- [187] Dias, M. V., Tyrakis, P., Domingues, R. R., Paes Leme, A. F., and Blundell, T. L. (2014) *Mycobacterium tuberculosis* dihydrofolate reductase reveals two conformational states and a possible low affinity mechanism to antifolate drugs, *Structure* 22, 94-103.
- [188] Reyes, V. M., Sawaya, M. R., Brown, K. A., and Kraut, J. (1995) Isomorphous crystal structures of *Escherichia coli* dihydrofolate reductase complexed with folate, 5-deazafoolate, and 5,10-dideazatetrahydrofolate: mechanistic implications, *Biochemistry* 34, 2710-2723.
- [189] Wu, Y., and Houk, K. (1987) Transition structures for hydride transfers, *J Am Chem Soc* 109, 906-908.
- [190] Wu, Y., and Houk, K. (1987) Theoretical transition structures for hydride transfer to methyleniminium ion from methylamine and dihydropyridine - on the nonlinearity of hydride transfers, *J Am Chem Soc* 109, 2226-2227.

- [191] Wan, Q., Bennett, B. C., Wilson, M. A., Kovalevsky, A., Langan, P., Howell, E. E., and Dealwis, C. (2014) Toward resolving the catalytic mechanism of dihydrofolate reductase using neutron and ultrahigh-resolution X-ray crystallography, *Proc Natl Acad Sci U S A* *111*, 18225-18230.
- [192] Wei, J. R., Krishnamoorthy, V., Murphy, K., Kim, J. H., Schnappinger, D., Alber, T., Sasseti, C. M., Rhee, K. Y., and Rubin, E. J. (2011) Depletion of antibiotic targets has widely varying effects on growth, *Proc Natl Acad Sci U S A* *108*, 4176-4181.
- [193] Chen, S. C., Lin, Y. H., Yu, H. C., and Liaw, S. H. (2009) Complex structure of *Bacillus subtilis* RibG: the reduction mechanism during riboflavin biosynthesis, *J Biol Chem* *284*, 1725-1731.
- [194] Margosiak, S. A., Appleman, J. R., Santi, D. V., and Blakley, R. L. (1993) Dihydrofolate reductase from the pathogenic fungus *Pneumocystis carinii*: catalytic properties and interaction with antifolates, *Arch Biochem Biophys* *305*, 499-508.
- [195] Jackson, R. C., Fry, D. W., Boritzki, T. J., Besserer, J. A., Leopold, W. R., Sloan, B. J., and Elslager, E. F. (1984) Biochemical pharmacology of the lipophilic antifolate, trimetrexate, *Adv Enzyme Regul* *22*, 187-206.
- [196] Gangjee, A., Jain, H. D., Phan, J., Guo, X., Queener, S. F., and Kisliuk, R. L. (2010) 2,4-Diamino-5-methyl-6-substituted arylthio-furo[2,3-d]pyrimidines as novel classical and nonclassical antifolates as potential dual thymidylate synthase and dihydrofolate reductase inhibitors, *Bioorg Med Chem* *18*, 953-961.
- [197] White, E. L., Ross, L. J., Cunningham, A., and Escuyer, V. (2004) Cloning, expression, and characterization of *Mycobacterium tuberculosis* dihydrofolate reductase, *FEMS Microbiol Lett* *232*, 101-105.
- [198] Pyrkov, T. V., Chugunov, A. O., Krylov, N. A., Nolde, D. E., and Efremov, R. G. (2009) PLATINUM: a web tool for analysis of hydrophobic/hydrophilic organization of biomolecular complexes, *Bioinformatics* *25*, 1201-1202.
- [199] Holm, L., and Rosenström, P. (2010) Dali server: conservation mapping in 3D, *Nucleic Acids Res* *38*, W545-549.

- [200] Romisch-Margl, W., Eisenreich, W., Haase, I., Bacher, A., and Fischer, M. (2008) 2,5-diamino-6-ribitylamino-4(3H)-pyrimidinone 5'-phosphate synthases of fungi and archaea, *Febs J* 275, 4403-4414.
- [201] Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis, *J Comput Chem* 25, 1605-1612.
- [202] Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol* 276, 307-326.
- [203] Vagin, A., and Teplyakov, A. (1997) MOLREP: an Automated Program for Molecular Replacement, *Journal of Applied Crystallography* 30, 1022-1025.
- [204] Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments, *Acta Crystallogr D Biol Crystallogr* 67, 235-242.
- [205] Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr D Biol Crystallogr* 66, 213-221.
- [206] Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr D Biol Crystallogr* 60, 2126-2132.
- [207] Stone, S. R., and Morrison, J. F. (1982) Kinetic mechanism of the reaction catalyzed by dihydrofolate reductase from *Escherichia coli*, *Biochemistry* 21, 3757-3765.
- [208] Michaelis, L., and Menten, M. L. (1913) Die kinetik der invertinwirkung, *Biochem. z* 49, 352.



- [209] Magalhães, M. L., Argyrou, A., Cahill, S. M., and Blanchard, J. S. (2008) Kinetic and mechanistic analysis of the *Escherichia coli* ribD-encoded bifunctional deaminase-reductase involved in riboflavin biosynthesis, *Biochemistry* 47, 6499-6507.
- [210] Hasnain, G., Frelin, O., Roje, S., Ellens, K. W., Ali, K., Guan, J. C., Garrett, T. J., de Crécy-Lagard, V., Gregory, J. F., McCarty, D. R., and Hanson, A. D. (2013) Identification and characterization of the missing pyrimidine reductase in the plant riboflavin biosynthesis pathway, *Plant Physiol* 161, 48-56.
- [211] Bailey, S. W., and Ayling, J. E. (2009) The extremely slow and variable activity of dihydrofolate reductase in human liver and its implications for high folic acid intake, *Proc Natl Acad Sci U S A* 106, 15424-15429.
- [212] Pruitt, K. D., Tatusova, T., Brown, G. R., and Maglott, D. R. (2012) NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy, *Nucleic Acids Res* 40, D130-135.
- [213] Pruitt, K. D., Brown, G. R., Hiatt, S. M., Thibaud-Nissen, F., Astashyn, A., Ermolaeva, O., Farrell, C. M., Hart, J., Landrum, M. J., McGarvey, K. M., Murphy, M. R., O'Leary, N. A., Pujar, S., Rajput, B., Rangwala, S. H., Riddick, L. D., Shkeda, A., Sun, H., Tamez, P., Tully, R. E., Wallin, C., Webb, D., Weber, J., Wu, W., DiCuccio, M., Kitts, P., Maglott, D. R., Murphy, T. D., and Ostell, J. M. (2014) RefSeq: an update on mammalian reference sequences, *Nucleic Acids Res* 42, D756-763.
- [214] Tatusova, T., Ciuffo, S., Fedorov, B., O'Neill, K., and Tolstoy, I. (2014) RefSeq microbial genomes database: new representation and annotation strategy, *Nucleic Acids Res* 42, D553-559.
- [215] Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J. F., Guindon, S., Lefort, V., Lescot, M., Claverie, J. M., and Gascuel, O. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist, *Nucleic Acids Res* 36, W465-469.
- [216] Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Res* 32, 1792-1797.

- [217] Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0, *Syst Biol* 59, 307-321.
- [218] Chevenet, F., Brun, C., Banuls, A. L., Jacq, B., and Christen, R. (2006) TreeDyn: towards dynamic graphics and annotations for analyses of trees, *BMC Bioinformatics* 7, 439.
- [219] Gengenbacher, M., Rao, S. P., Pethe, K., and Dick, T. (2010) Nutrient-starved, non-replicating *Mycobacterium tuberculosis* requires respiration, ATP synthase and isocitrate lyase for maintenance of ATP homeostasis and viability, *Microbiology* 156, 81-87.
- [220] Aktas, D. F., and Cook, P. F. (2009) A lysine-tyrosine pair carries out acid-base chemistry in the metal ion-dependent pyridine dinucleotide-linked beta-hydroxyacid oxidative decarboxylases, *Biochemistry* 48, 3565-3577.
- [221] Rudolph Steinberger, F. H. W. (1951) Metal ion-catalyzed decarboxylation: A model for an enzyme system, *J Am Chem Soc* 73, 429-435.
- [222] Lienhard, G. E., and Rose, I. A. (1964) The mechanism of action of 6-phosphogluconate dehydrogenase, *Biochemistry* 3, 190-195.
- [223] Dalziel, K., and Londesborough, J. C. (1968) The mechanisms of reductive carboxylation reactions. Carbon dioxide or bicarbonate as substrate of nicotinamide-adenine dinucleotide phosphate-linked isocitrate dehydrogenase and malic enzyme, *Biochem J* 110, 223-230.
- [224] Siebert, G., Carsiotis, M., and Plaut, G. W. (1957) The enzymatic properties of isocitric dehydrogenase, *J Biol Chem* 226, 977-991.
- [225] Cook, P. F., and Cleland, W. W. (1981) pH variation of isotope effects in enzyme-catalyzed reactions. 2. Isotope-dependent step not pH dependent. Kinetic mechanism of alcohol dehydrogenase, *Biochemistry* 20, 1805-1816.
- [226] Ehrlich, R. S., and Colman, R. F. (1987) Ionization of isocitrate bound to pig heart NADP<sup>+</sup>-dependent isocitrate dehydrogenase: <sup>13</sup>C NMR study of substrate binding,

*Biochemistry* 26, 3461-3466.

- [227] Grissom, C. B., and Cleland, W. W. (1988) Isotope effect studies of chicken liver NADP malic enzyme: role of the metal ion and viscosity dependence, *Biochemistry* 27, 2927-2934.
- [228] Rudolph Steinberger, F. H. W. (1951) Metal ion-catalyzed decarboxylation: A model for an enzyme system, *J Am Chem Soc* 73, 429-435.
- [229] Hurley, J. H., Dean, A. M., Koshland, D. E., Jr., and Stroud, R. M. (1991) Catalytic mechanism of NADP(+)-dependent isocitrate dehydrogenase: implications from the structures of magnesium-isocitrate and NADP+ complexes, *Biochemistry* 30, 8671-8678.
- [230] Miller, S. P., Goncalves, S., Matias, P. M., and Dean, A. M. (2014) Evolution of a transition state: role of Lys100 in the active site of isocitrate dehydrogenase, *Chembiochem* 15, 1145-1153.
- [231] Zhu, G., Golding, G. B., and Dean, A. M. (2005) The selective cause of an ancient adaptation, *Science* 307, 1279-1282.
- [232] Self, C. H., and Weitzman, P. D. (1972) The isocitrate dehydrogenases of *Acinetobacter lwoffii*. Separation and properties of two nicotinamide-adenine dinucleotide phosphate-linked isoenzymes, *Biochem J* 130, 211-219.
- [233] Chung, A. E., and Franzen, J. S. (1969) Oxidized triphosphopyridine nucleotide specific isocitrate dehydrogenase from *Azotobacter vinelandii*. Isolation and characterization, *Biochemistry* 8, 3175-3184.
- [234] Ochiai, T., Fukunaga, N., and Sasaki, S. (1979) Purification and some properties of two NADP+-specific isocitrate dehydrogenases from an obligately psychrophilic marine bacterium, *Vibrio* sp., strain ABE-1, *J Biochem* 86, 377-384.
- [235] Fukunaga, N., Imagawa, S., Sahara, T., Ishii, A., and Suzuki, M. (1992) Purification and characterization of monomeric isocitrate dehydrogenase with NADP(+)-specificity from *Vibrio parahaemolyticus* Y-4, *J Biochem* 112, 849-855.

- [236] Bai, C., Fernandez, E., Yang, H., and Chen, R. (1999) Purification and stabilization of a monomeric isocitrate dehydrogenase from *Corynebacterium glutamicum*, *Protein Expr Purif* 15, 344-348.
- [237] Watanabe, S., Yasutake, Y., Tanaka, I., and Takada, Y. (2005) Elucidation of stability determinants of cold-adapted monomeric isocitrate dehydrogenase from a psychrophilic bacterium, *Colwellia maris*, by construction of chimeric enzymes, *Microbiology* 151, 1083-1094.
- [238] Kanao, T., Kawamura, M., Fukui, T., Atomi, H., and Imanaka, T. (2012) Characterization of isocitrate dehydrogenase from the green sulfur bacterium *Chlorobium limicola*, *European Journal of Biochemistry* 269, 1926-1931.
- [239] Zhang, B. B., Wang, P., Wang, A., Wang, W. C., Tang, W. G., and Zhu, G. P. (2013) Expression and characterization of a novel isocitrate dehydrogenase from *Streptomyces diastaticus* No. 7 strain M1033, *Mol Biol Rep* 40, 1615-1623.
- [240] Zhang, B., Wang, B., Wang, P., Cao, Z., Huang, E., Hao, J., Dean, A. M., and Zhu, G. (2009) Enzymatic characterization of a monomeric isocitrate dehydrogenase from *Streptomyces lividans* TK54, *Biochimie* 91, 1405-1410.
- [241] Yasutake, Y., Watanabe, S., Yao, M., Takada, Y., Fukunaga, N., and Tanaka, I. (2003) Crystal structure of the monomeric isocitrate dehydrogenase in the presence of NADP<sup>+</sup>: insight into the cofactor recognition, catalysis, and evolution, *J Biol Chem* 278, 36897-36904.
- [242] Yasutake, Y., Watanabe, S., Yao, M., Takada, Y., Fukunaga, N., and Tanaka, I. (2002) Structure of the monomeric isocitrate dehydrogenase: evidence of a protein monomerization by a domain duplication, *Structure* 10, 1637-1648.
- [243] Sidhu, N. S., Delbaere, L. T., and Sheldrick, G. M. (2011) Structure of a highly NADP<sup>+</sup>-specific isocitrate dehydrogenase, *Acta Crystallogr D Biol Crystallogr* 67, 856-869.
- [244] Wicken, J. S., Chung, A. E., and Franzen, J. S. (1972) Isocitrate dehydrogenase from *Azotobacter vinelandii*. Order of substrate addition and product release, *Biochemistry* 11, 4766-4778.

- [245] Imabayashi, F., Aich, S., Prasad, L., and Delbaere, L. T. (2006) Substrate-free structure of a monomeric NADP isocitrate dehydrogenase: an open conformation phylogenetic relationship of isocitrate dehydrogenase, *Proteins* 63, 100-112.
- [246] LaPorte, D. C., and Koshland, D. E., Jr. (1982) A protein with kinase and phosphatase activities involved in regulation of tricarboxylic acid cycle, *Nature* 300, 458-460.
- [247] Hurley, J. H., Dean, A. M., Sohl, J. L., Koshland, D. E., Jr., and Stroud, R. M. (1990) Regulation of an enzyme by phosphorylation at the active site, *Science* 249, 1012-1016.
- [248] Dean, A. M., Lee, M. H., and Koshland, D. E., Jr. (1989) Phosphorylation inactivates Escherichia coli isocitrate dehydrogenase by preventing isocitrate binding, *J Biol Chem* 264, 20482-20486.
- [249] Balganes, T., Datta, S., and Ghosh, I. (2006) Method.
- [250] Wang, P., Lv, C., and Zhu, G. (2015) Novel type II and monomeric NAD<sup>+</sup> specific isocitrate dehydrogenases: Phylogenetic affinity, enzymatic characterization, and evolutionary implication, *Sci Rep* 5, 9150-9161.
- [251] Parsons, D. W., Jones, S., Zhang, X., Lin, J. C., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Siu, I. M., Gallia, G. L., Olivi, A., McLendon, R., Rasheed, B. A., Keir, S., Nikolskaya, T., Nikolsky, Y., Busam, D. A., Tekleab, H., Diaz, L. A., Jr., Hartigan, J., Smith, D. R., Strausberg, R. L., Marie, S. K., Shinjo, S. M., Yan, H., Riggins, G. J., Bigner, D. D., Karchin, R., Papadopoulos, N., Parmigiani, G., Vogelstein, B., Velculescu, V. E., and Kinzler, K. W. (2008) An integrated genomic analysis of human glioblastoma multiforme, *Science* 321, 1807-1812.
- [252] Yan, H., Parsons, D. W., Jin, G., McLendon, R., Rasheed, B. A., Yuan, W., Kos, I., Batonic-Haberle, I., Jones, S., Riggins, G. J., Friedman, H., Friedman, A., Reardon, D., Herndon, J., Kinzler, K. W., Velculescu, V. E., Vogelstein, B., and Bigner, D. D. (2009) IDH1 and IDH2 mutations in gliomas, *N Engl J Med* 360, 765-773.
- [253] Amary, M. F., Bacsi, K., Maggiani, F., Damato, S., Halai, D., Berisha, F., Pollock, R., O'Donnell, P., Grigoriadis, A., Diss, T., Eskandarpour, M., Presneau, N., Hogendoorn, P. C. W., Futreal, A., Tirabosco, R., and Flanagan, A. M. (2011) IDH1

and IDH2 mutations are frequent events in central chondrosarcoma and central and periosteal chondromas but not in other mesenchymal tumours, *J Pathol* 224, 334-343.

- [254] Borger, D. R., Tanabe, K. K., Fan, K. C., Lopez, H. U., Fantin, V. R., Straley, K. S., Schenkein, D. P., Hezel, A. F., Ancukiewicz, M., Liebman, H. M., Kwak, E. L., Clark, J. W., Ryan, D. P., Deshpande, V., Dias-Santagata, D., Ellisen, L. W., Zhu, A. X., and Iafrate, A. J. (2012) Frequent mutation of isocitrate dehydrogenase (IDH)1 and IDH2 in cholangiocarcinoma identified through broad-based tumor genotyping, *Oncologist* 17, 72-79.
- [255] Paschka, P., Schlenk, R. F., Gaidzik, V. I., Habdank, M., Kronke, J., Bullinger, L., Spath, D., Kayser, S., Zucknick, M., Gotze, K., Horst, H. A., Germing, U., Dohner, H., and Dohner, K. (2010) IDH1 and IDH2 mutations are frequent genetic alterations in acute myeloid leukemia and confer adverse prognosis in cytogenetically normal acute myeloid leukemia with NPM1 mutation without FLT3 internal tandem duplication, *J Clin Oncol* 28, 3636-3643.
- [256] Mardis, E. R., Ding, L., Dooling, D. J., Larson, D. E., McLellan, M. D., Chen, K., Koboldt, D. C., Fulton, R. S., Delehaunty, K. D., McGrath, S. D., Fulton, L. A., Locke, D. P., Magrini, V. J., Abbott, R. M., Vickery, T. L., Reed, J. S., Robinson, J. S., Wylie, T., Smith, S. M., Carmichael, L., Eldred, J. M., Harris, C. C., Walker, J., Peck, J. B., Du, F., Dukes, A. F., Sanderson, G. E., Brummett, A. M., Clark, E., McMichael, J. F., Meyer, R. J., Schindler, J. K., Pohl, C. S., Wallis, J. W., Shi, X., Lin, L., Schmidt, H., Tang, Y., Haipek, C., Wiechert, M. E., Ivy, J. V., Kalicki, J., Elliott, G., Ries, R. E., Payton, J. E., Westervelt, P., Tomasson, M. H., Watson, M. A., Baty, J., Heath, S., Shannon, W. D., Nagarajan, R., Link, D. C., Walter, M. J., Graubert, T. A., DiPersio, J. F., Wilson, R. K., and Ley, T. J. (2009) Recurring mutations found by sequencing an acute myeloid leukemia genome, *N Engl J Med* 361, 1058-1066.
- [257] Segel, I. H. (1975) *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*, Wiley-Interscience, New York.
- [258] Poornam, G. P., Matsumoto, A., Ishida, H., and Hayward, S. (2009) A method for the analysis of domain movements in large biomolecular complexes, *Proteins* 76, 201-212.
- [259] Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a

program to generate schematic diagrams of protein-ligand interactions, *Protein Eng* 8, 127-134.

- [260] Gonçalves, S., Miller, S. P., Carrondo, M. A., Dean, A. M., and Matias, P. M. (2012) Induced fit and the catalytic mechanism of isocitrate dehydrogenase, *Biochemistry* 51, 7098-7115.
- [261] Banerjee, S., Nandyala, A., Podili, R., Katoch, V. M., Murthy, K. J., and Hasnain, S. E. (2004) Mycobacterium tuberculosis (Mtb) isocitrate dehydrogenases show strong B cell response and distinguish vaccinated controls from TB patients, *Proc Natl Acad Sci U S A* 101, 12652-12657.
- [262] Vandal, O. H., Pierini, L. M., Schnappinger, D., Nathan, C. F., and Ehrt, S. (2008) A membrane protein preserves intrabacterial pH in intraphagosomal Mycobacterium tuberculosis, *Nat Med* 14, 849-854.
- [263] Ramachandran, N., and Colman, R. F. (1980) Chemical characterization of distinct subunits of pig heart DPN-specific isocitrate dehydrogenase, *J Biol Chem* 255, 8859-8864.
- [264] Jennings, G. T., Sechi, S., Stevenson, P. M., Tuckey, R. C., Parmelee, D., and McAlister-Henn, L. (1994) Cytosolic NADP(+)-dependent isocitrate dehydrogenase. Isolation of rat cDNA and study of tissue-specific and developmental expression of mRNA, *J Biol Chem* 269, 23128-23134.
- [265] Reitman, Z. J., and Yan, H. (2010) Isocitrate dehydrogenase 1 and 2 mutations in cancer: alterations at a crossroads of cellular metabolism, *J Natl Cancer Inst* 102, 932-941.
- [266] Cupp, J. R., and McAlister-Henn, L. (1993) Kinetic analysis of NAD(+)-isocitrate dehydrogenase with altered isocitrate binding sites: contribution of IDH1 and IDH2 subunits to regulation and catalysis, *Biochemistry* 32, 9323-9328.
- [267] Jo, S. H., Son, M. K., Koh, H. J., Lee, S. M., Song, I. H., Kim, Y. O., Lee, Y. S., Jeong, K. S., Kim, W. B., Park, J. W., Song, B. J., and Huh, T. L. (2001) Control of mitochondrial redox balance and cellular defense against oxidative damage by mitochondrial NADP<sup>+</sup>-dependent isocitrate dehydrogenase, *J Biol Chem* 276, 16168-16176.

- [268] Rohle, D., Popovici-Muller, J., Palaskas, N., Turcan, S., Grommes, C., Campos, C., Tsoi, J., Clark, O., Oldrini, B., Komisopoulou, E., Kunii, K., Pedraza, A., Schalm, S., Silverman, L., Miller, A., Wang, F., Yang, H., Chen, Y., Kernytsky, A., Rosenblum, M. K., Liu, W., Biller, S. A., Su, S. M., Brennan, C. W., Chan, T. A., Graeber, T. G., Yen, K. E., and Mellinghoff, I. K. (2013) An inhibitor of mutant IDH1 delays growth and promotes differentiation of glioma cells, *Science* 340, 626-630.
- [269] Wang, F., Travins, J., DeLaBarre, B., Penard-Lacronique, V., Schalm, S., Hansen, E., Straley, K., Kernytsky, A., Liu, W., Gliser, C., Yang, H., Gross, S., Artin, E., Saada, V., Mylonas, E., Quivoron, C., Popovici-Muller, J., Saunders, J. O., Salituro, F. G., Yan, S., Murray, S., Wei, W., Gao, Y., Dang, L., Dorsch, M., Agresta, S., Schenkein, D. P., Biller, S. A., Su, S. M., de Botton, S., and Yen, K. E. (2013) Targeted inhibition of mutant IDH2 in leukemia cells induces cellular differentiation, *Science* 340, 622-626.
- [270] Zheng, B., Yao, Y., Liu, Z., Deng, L., Anglin, J. L., Jiang, H., Prasad, B. V., and Song, Y. (2013) Crystallographic Investigation and Selective Inhibition of Mutant Isocitrate Dehydrogenase, *ACS Med Chem Lett* 4, 542-546.
- [271] Wu, F., Jiang, H., Zheng, B., Kogiso, M., Yao, Y., Zhou, C., Li, X. N., and Song, Y. (2015) Inhibition of Cancer-Associated Mutant Isocitrate Dehydrogenases by 2-Thiohydantoin Compounds, *J Med Chem* 58, 6899-6908.
- [272] Ananthan, S., Faaleolea, E. R., Goldman, R. C., Hobrath, J. V., Kwong, C. D., Laughon, B. E., Maddry, J. A., Mehta, A., Rasmussen, L., Reynolds, R. C., Secrist, J. A., 3rd, Shindo, N., Showe, D. N., Sosa, M. I., Suling, W. J., and White, E. L. (2009) High-throughput screening for inhibitors of Mycobacterium tuberculosis H37Rv, *Tuberculosis (Edinb)* 89, 334-353.
- [273] Dalvi, V. H., and Rossky, P. J. (2010) Molecular origins of fluorocarbon hydrophobicity, *Proc Natl Acad Sci U S A* 107, 13603-13607.
- [274] Bissantz, C., Kuhn, B., and Stahl, M. (2010) A Medicinal Chemist's Guide to Molecular Interactions, *J Med Chem* 53, 5061-5084.
- [275] Biffinger, J. C., Kim, H. W., and DiMagno, S. G. (2004) The polar hydrophobicity of fluorinated compounds, *Chembiochem* 5, 622-627.



- [276] Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006) HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes, *Acta Crystallogr D Biol Crystallogr* 62, 859-866.
- [277] Zhang, J. H., Chung, T. D., and Oldenburg, K. R. (1999) A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays, *J Biomol Screen* 4, 67-73.

## APPENDIX A

### SUPPORTING INFORMATION FOR CHAPTER III<sup>†</sup>

#### Materials and Methods

##### *Sequences Used in the PhylogeneticTree<sup>1-3</sup>*

NP\_217187.1 *Mycobacterium tuberculosis* RibD (strain H37Rv),  
WP\_015306567.1 *Mycobacterium smegmatis* pyrimidine reductase, WP\_015288343.1  
*Mycobacterium canettii* RibD, WP\_043789270.1 *Rhodococcus wratislaviensis* 5-amino-  
6-(5-phosphoribosylamino)uracil reductase, WP\_014986655.1 *Nocardia brasiliensis* 5-  
amino-6-(5-phosphoribosylamino)uracil reductase, WP\_030760612.1 *Streptomyces*  
diaminohydroxyphosphoribosylaminopyrimidine reductase, WP\_044573862.1  
*Saccharopolyspora spinosa* 5-amino-6-(5-phosphoribosylamino)uracil reductase,  
WP\_021764912.1 *Leifsonia aquatica* RibD, WP\_044850442.1 *Amycolatopsis orientalis*  
5-amino-6-(5-phosphoribosylamino)uracil reductase, WP\_017129922.1 *Pseudomonas*  
*agarici* diaminohydroxyphosphoribosylaminopyrimidine deaminase, WP\_042697497.1  
*Methanocorpusculum bavaricum* 5-amino-6-(5-phosphoribosylamino)uracil reductase,  
WP\_004014420.1 *Escherichia coli* RibD, WP\_001150448.1 *Shigella flexneri*

---

<sup>†</sup> This work has been published in and is reprinted with permission from: Cheng, Y. S., and Sacchettini, J. C. (2016) Structural insights into *Mycobacterium tuberculosis* Rv2671 protein as a dihydrofolate reductase functional analogue contributing to *para*-aminosalicylic acid resistance, *Biochemistry* 55, 1107-1119. Copyright 2016 by American Chemical Society.

diaminohydroxyphosphoribosylaminopyrimidine deaminase, NP\_984897.1 *Ashbya gossypii* (strain ATCC 10895) AER037Cp, WP\_010870176.1 *Methanocaldococcus jannaschii* 5-amino-6-(5-phosphoribosylamino)uracil reductase, NP\_009711.3 *Saccharomyces cerevisiae* (strain S288c) Rib7p, WP\_032722052.1 *Bacillus subtilis* RibD, WP\_003898867.1 *Mycobacterium tuberculosis* RibG, WP\_000162469.1 *Streptococcus pneumoniae* DHFR, WP\_010865375.1 *Thermotoga maritima* DHFR, WP\_011499411.1 *Methanococcoides burtonii* deaminase, XP\_665866.1 *Cryptosporidium hominis* (strain TU502) Dhfr, WP\_000637197.1 *Bacillus cereus* DHFR, WP\_033720836.1 *Mycobacterium avium* DHFR, WP\_005769725.1 *Coxiella burnetii* DHFR, WP\_033010007.1 *Geobacillus stearothermophilus* DHFR, WP\_020321480.1 *Klebsiella pneumoniae* DHFR, WP\_000175746.1 *Staphylococcus aureus* DHFR, WP\_005689288.1 *Lactobacillus rhamnosus* DHFR, WP\_024211602.1 *Escherichia coli* DHFR, XP\_002367252.1 *Toxoplasma gondii* ME49 bifunctional dihydrofolate reductase thymidylate synthase, XP\_011775023.1 *Trypanosoma brucei gambiense* (strain DAL972) dihydrofolate reductase thymidylate synthase, XP\_001351479.1 *Plasmodium falciparum* (strain 3D7) bifunctional dihydrofolate reductase thymidylate synthase, WP\_031722695.1 *Mycobacterium tuberculosis* DHFR, gi|1155270|gb|AAB60183.1| *Mycobacterium bovis* enoyl ACP reductase, gi|690312203|gb|AIR14234.1| *Mycobacterium tuberculosis* (strain H37Rv) enoyl-ACP reductase, gi|387616665|ref|YP\_006119687.1| (*Escherichia coli* strain O83:H1 str. NRG 857C) enoyl-ACP reductase, gi|728888373|gb|AIY96767.1| *Bacillus subtilis* enoyl-ACP reductase.

## Tables and Figures

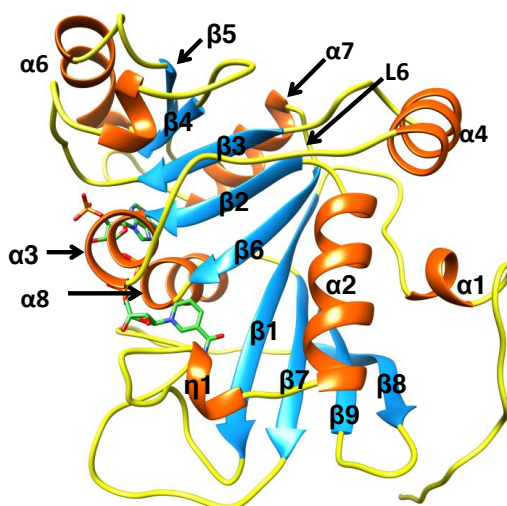
Table A-1. Hydrogen bond interactions between Rv2671 and TMP/TMQ

Rv2671	TMP	Distance (Å)
Asp67 OD2	N2	2.53
Asp67 OD1	N4	2.55
Asn44 OD1	N4	3.24
Thr214 OG1	N4	3.32
Asn44 O	N7	2.70
Water 16	N7	3.15

Rv2671	TMP	Distance (Å)
Asp67 OD2	N1	2.54
Phe45 O	N24	3.08
Asn44 OD1	N24	2.98
Thr214 OG1	N24	3.30
Asp67 OD1	N24	2.75
Asn44 O	N25	2.67
Glu193 O	N25	3.14
Water 52	O20	3.11

A



B

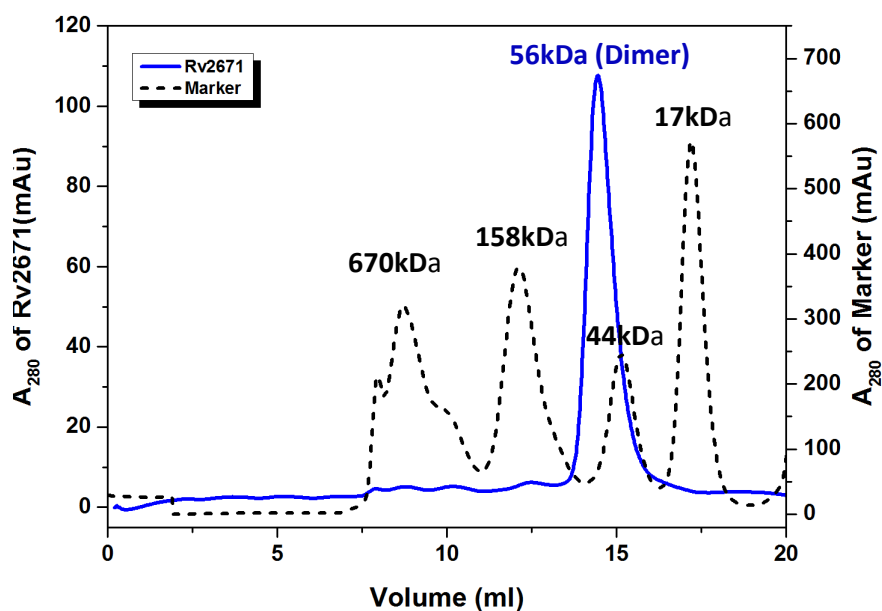


Figure A-1 Ribbon representation of Rv2671 and the gel filtration chromatography of Rv2671. (A) Ribbon representation of Rv2671 in complex with NADP<sup>+</sup>. The secondary structural elements are labeled sequentially. (B) Purified Rv2671 were used to investigate its oligomerization state in solution. The samples (0.1 mL, 14.5 mg/mL) were loading to Superdex200 10/300 GL column (GE Healthcare Bio-sciences) with 50 mM Tris pH7.5, 50 mM NaCl at a flow rate of 0.3 mL/min. For calibration of the column the following size markers (Bio-Rad) were used: Thyroglobulin (bovine) (670 kDa),  $\gamma$ -globulin (bovine) (158 kDa), Ovalbumin (chicken) (44 kDa) and Myoglobin (horse) (17 kDa). The figure depicts the elution profiles at 280 nm for Rv2671 (line) and standards (dot). Rv2671 were eluted from the column at molecular mass of  $\sim$  56 kDa, suggesting that Rv2671 exists as a dimer in solution.

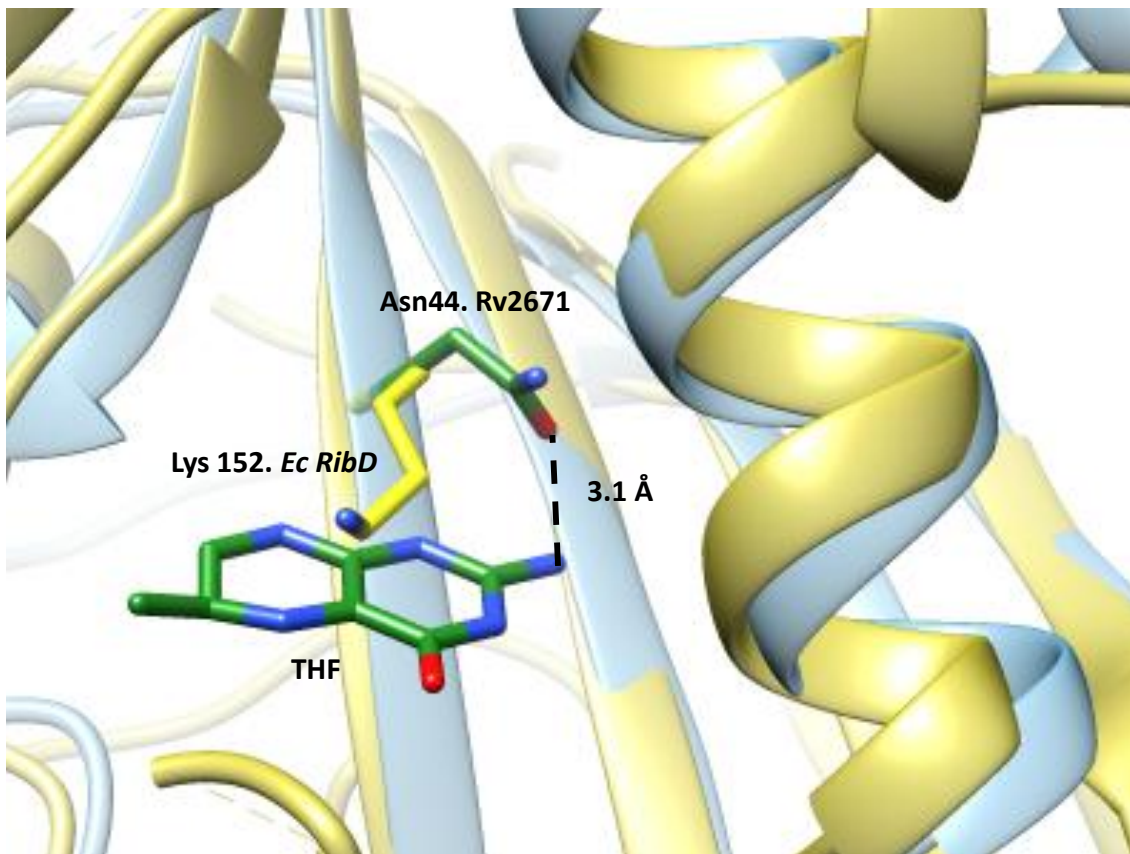


Figure A-2 Superimposition of the THF bound Rv2671 structure (blue) with the *E. coli* RibD structure (blue, 2OBC<sup>4</sup>). Pteridine ring of THF is colored in green. The conserved lysine of RibD (lys152 in *E. coli* RibD) clashed with the THF in the superimposition structures. Asn44 however is in the Rv2671 at the same position and forms a hydrogen bond with THF. Distance between the side chain carbonyl group of Asn44 from Rv2671 and the NA2 atom of the pteridine is 3.1 Å.

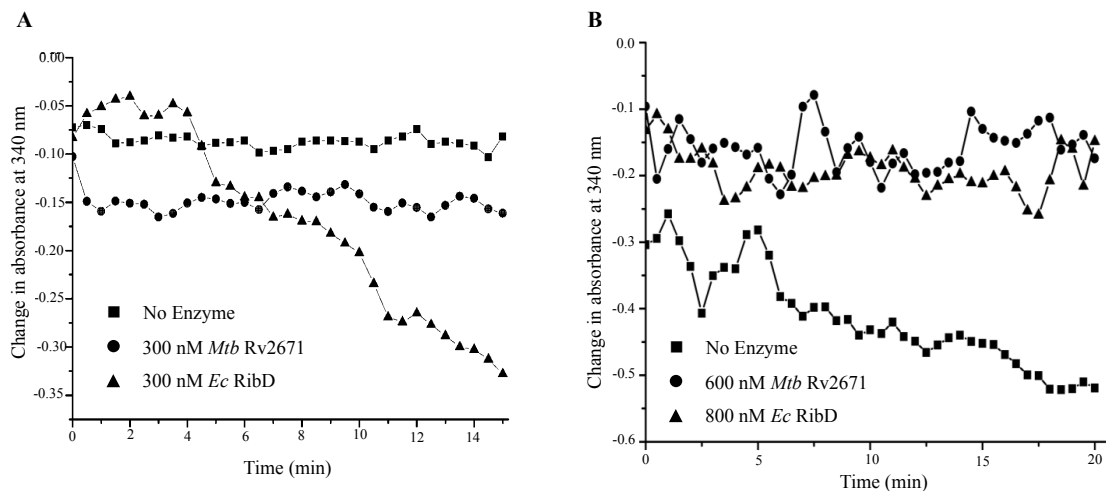


Figure A-3 Pyrimidine reductase activity assays. The reaction was followed by the consumption of NADPH by pyrimidine reductase activity. The absorbance at 340 nm was monitored for 20 min using a Cary100. (A) Rv2671 at 300 nM was tested for DAROPP reductase activity using DAROPP as substrate (circles). The *in situ* reaction of *E. coli* RibD was used as a positive control (triangles); DAROPP was added into the reaction mixture in the presence of full length *E. coli* and NADPH. The reaction without enzymes was conducted as a negative control (squares). (B) Rv2671 at 600 nM was tested for AROPP reductase activity using AROPP as substrate (triangles). The reaction with *E. coli* RibD was used as a positive control (squares) and the reaction without enzymes was conducted as a negative control (circles).

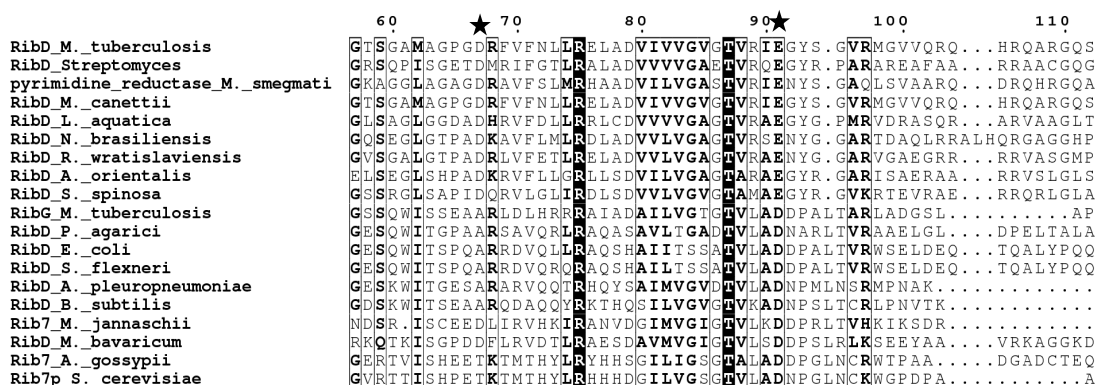


Figure A-4 Multiple sequence alignment of Rv2671 with its close orthologs and functionally characterized pyrimidine reductases. The alignment was generated by Muscule<sup>5</sup> and visualized by ESript 3.x.<sup>6</sup> The sequences were numbered by the residue number of Rv2671 and Asp67 and Glu91 are marked with star, respectively. The sequences used in the alignment are: *Mycobacterium tuberculosis* RibD (strain H37Rv), *Mycobacterium smegmatis* pyrimidine reductase, *Mycobacterium canettii* RibD, *Rhodococcus wratislaviensis* 5-amino-6-(5-phosphoribosylamino)uracil reductase, *Nocardia brasiliensis* 5-amino-6-(5-phosphoribosylamino)uracil reductase, *Streptomyces* diaminohydroxyphosphoribosylaminopyrimidine reductase, *Saccharopolyspora spinosa* 5-amino-6-(5-phosphoribosylamino)uracil reductase, *Leifsonia aquatica* RibD, *Amycolatopsis orientalis* 5-amino-6-(5-phosphoribosylamino)uracil reductase, *Pseudomonas agarici* diaminohydroxyphosphoribosylaminopyrimidine deaminase, *Methanococcus bavaricum* 5-amino-6-(5-phosphoribosylamino)uracil reductase, *Escherichia coli* RibD, *Shigella flexneri* diaminohydroxyphosphoribosylaminopyrimidine deaminase, *Ashbya gossypii* (strain ATCC 10895) AER037Cp, *Methanocaldococcus jannaschii* 5-amino-6-(5-phosphoribosylamino)uracil reductase, *Saccharomyces cerevisiae* (strain S288c) Rib7p, *Bacillus subtilis* RibD. The detail sequence accession numbers were listed in the supplementary materials and methods section.

## References for Appendix Materials

- [1] Pruitt, K. D., Tatusova, T., Brown, G. R., and Maglott, D. R. (2012) NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy, *Nucleic Acids Res* 40, D130-135.
- [2] Pruitt, K. D., Brown, G. R., Hiatt, S. M., Thibaud-Nissen, F., Astashyn, A.,



Ermolaeva, O., Farrell, C. M., Hart, J., Landrum, M. J., McGarvey, K. M., Murphy, M. R., O'Leary, N. A., Pujar, S., Rajput, B., Rangwala, S. H., Riddick, L. D., Shkeda, A., Sun, H., Tamez, P., Tully, R. E., Wallin, C., Webb, D., Weber, J., Wu, W., DiCuccio, M., Kitts, P., Maglott, D. R., Murphy, T. D., and Ostell, J. M. (2014) RefSeq: an update on mammalian reference sequences, *Nucleic Acids Res* 42, D756-763.

[3] Tatusova, T., Ciufo, S., Fedorov, B., O'Neill, K., and Tolstoy, I. (2014) RefSeq microbial genomes database: new representation and annotation strategy, *Nucleic Acids Res* 42, D553-559.

[4] Stenmark, P., Moche, M., Gurmu, D., and Nordlund, P. (2007) The crystal structure of the bifunctional deaminase/reductase RibD of the riboflavin biosynthetic pathway in *Escherichia coli*: implications for the reductive mechanism, *J Mol Biol* 373, 48-64.

[5] Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Res* 32, 1792-1797.

[6] Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server, *Nucleic Acids Res* 42, W320-324.