

STRUCTURE-GUIDED INHIBITOR DESIGN OF
MYCOBACTERIUM TUBERCULOSIS DRUG TARGETS
FROM CENTRAL CARBON METABOLISM

A Dissertation

by

HAELEE KIM

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. Sacchettini
Committee Members,	David P. Barondeau
	Kevin Burgess
	Gregory D. Reinhart
Head of Department,	Simon North

August 2016

Major Subject: Chemistry

Copyright 2016 Haelee Kim

ABSTRACT

A key determinant of the pathogenicity of *M. tuberculosis* (*Mtb*) is its ability to conserve energy and resources by altering its intermediate metabolism response to host-derived hostile conditions. Growing evidence supports that the central carbon metabolism (CCM) of *M. tuberculosis* including glycolysis, gluconeogenesis, the pentose phosphate pathway, and the TCA cycle, is regulated to allow for simultaneous utilization of carbohydrates and fatty acids-derived carbon sources. The genes involved in CCM represent attractive anti TB drug-targets. Although the role of *Mtb* phosphoenolpyruvate carboxykinase (PEPCK) in gluconeogenesis pathway and pyruvate kinase (PykA) enzymes in glycolysis pathway are relevant for the *M. tuberculosis* pathogenesis, targeting these enzymes for the drug development is not straight forward because of the human orthologs existence. Here, X-ray crystallographic, biochemical, and inhibitory studies of *Mtb* PEPCK and PykA highlight distinct features of the pathogenic's enzyme that differ from those of the host orthologs' and provide opportunities to develop selective and potent inhibitors.

Structural data of *Mtb* PEPCK has revealed that the conformation changes of the flexible loops in response to substrates binding differ from the one reported for host versions of the enzyme. GTP-competitive inhibitors of human cytosolic PEPCK bind to *Mtb* PEPCK in similar fashion, but the discovery of two unique small pockets in *Mtb* has created an opportunity for the design of a selective *Mtb* PEPCK inhibitor. By using structure-guided medicinal chemistry for GTP-competitive inhibitor series with PEPCK,

we are able to improve the inhibitory effect of these inhibitors against the enzyme, creating inhibitors with a nano-molar range IC_{50} and develop 10 fold selective inhibitors against *Mtb* PEPCK over human PEPCK. Structural study of *Mtb* PykA also demonstrated that adenosine mono-phosphate (AMP) is an allosteric effector, and we identified the binding mode and interaction of AMP in the allosteric binding site for the first time. Screening diverse compound libraries gave us insight into the potential scaffolds for the development of more potent pathogen-specific inhibitors. The structural and inhibitor studies of *Mtb* PEPCK and PykA have allowed a better understanding of their role in gluconeogenesis and glycolysis, and provide a framework for the development of selective inhibitors.

ACKNOWLEDGEMENT

First, I wish to thank God, my Lord, who has provided me with spiritual guidance to the end of this long journey. Without the sentiment in this bible verse “I can do all this through him who strengthens me” (Philippians 4:13), this work would not have been possible. You are an awesome God!! Next, I would like to express the deepest appreciation to my advisor Dr. James D. Sacchetti, for engaging me in value-driven tuberculosis research. His supervision, support, guidance, and advice in this research have inspired me not only for my research but also for my life. I also give thanks to my committee members, Dr. Kevin Burgess, Dr. David P. Barondeau, and Dr. Gregory Reinhart. Their guidance and persistence helped me in the completion of this work. I am very thankful for the kindness, support, and encouragement of my lab members, particularly Dr. Inna Krieger, who provided me with valuable scientific discussion, and wise parenting advice. I am also thankful to Dr. Khisimuzi Mdluli and Dr. Nader Fotouhi in the TB Alliance, who provides valuable ideas and materials for the PEPCCK project. I am deeply obliged to Dr. Sabine Ehrt in Weill Cornell medical college, she has cheered me with her kindness, compliments, and advice through email.

This dissertation is sincerely dedicated to my family members. I wish to thank to my parents, Mr. Kiwon Kim and Mrs. Junghee Lee, who have aided and encouraged me throughout my life. Their love and dedication has always kept me going. I owe my deepest gratitude to my husband Mr. Bob (Gwanggyu Lee), who is my best friend, a great daddy, and family leader. Since 1999, he has provided with unconditional love to

fill my emotional tank and unlimited support to allow me achieve my dreams. He spent a year taking care of our second son so I could work, and without his dedication I could not have finished my Ph.D. There is no one like you in the world. Thank you so much, Bob. In particular, I am very thankful to my sons, Chris Lee (Jun-suh) and Daniel Lee (Jun-ho) for their solicitude and understanding. I felt very bad when you guys cried in front of the day-care center. The generosity and sacrifice shown by my family will remain a source of inspiration to me throughout my life.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
ACKNOWLEDGEMENT	iv
TABLE OF CONTENTS	vi
LIST OF FIGURES	viii
LIST OF TABLES.....	xiii
CHAPTER I INTRODUCTION AND LITERATURE REVIEW.....	1
Introduction	1
Mechanisms of currently available TB drugs	4
Problems of current TB treatment.....	27
New regimen(PaMZ) for efficient TB treatment	30
Central carbon metabolism in <i>M.tuberculosis</i> for targets for new TB drugs development	31
CHAPTER II STRUCTURAL STUDY OF PHOSPHOENOLPYRUVATE CARBOXYKINASE FROM <i>MYCOBACTERIUM TUBERCULOSIS</i>	41
Introduction	41
Result and discussion	48
Conclusion.....	78
Experimental procedures.....	79
Additional results	88
CHAPTER III STRUCTURAL ANALYSIS OF PEPCK INHIBITOR COMPLEX AND DEVELOPMENT OF SELECTIVE INHIBITORS.....	90
Introduction	90
Result and discussion	94
Conclusion.....	122
Experimental procedures.....	123
Additional results	130

CHAPTER IV STRUCTURAL STUDY OF PYRUVATE KINASE FROM <i>MYCOBACTERIUM TUBERCULOSIS</i>	136
Introduction	136
Result and discussion	142
Conclusion.....	153
Experimental procedures.....	154
Additional results	158
 CHAPTER V IDENTIFICATION OF POTENTIAL INHIBITORS TARGETING PYRUVATE KINASE FROM <i>MYCOBACTERIUM</i> <i>TUBERCULOSIS</i>	 164
Introduction	164
Result and discussion	166
Conclusion.....	174
Experimental procedures.....	174
 CHAPTER VI SUMMARY	 177
REFERENCES	179

LIST OF FIGURES

	Page
Figure 1-1. Chemical structure of isoniazid	5
Figure 1-2. Chemical structure of ethambutol	6
Figure 1-3. Chemical structure of ethioamide.....	8
Figure 1-4. Chemical structure of D-alanine and D-cycloserine	9
Figure 1-5. Chemical structure of aminoglycosides.....	11
Figure 1-6. Chemical structure of fluoroquinolones	14
Figure 1-7. Chemical structure of rifamycins	15
Figure 1-8. Chemical structure of bedaquiline.....	19
Figure 1-9. Chemical structure of para-aminosalicylic acid (PAS)	21
Figure 1-10. Chemical structure of PA-824 (Pretomanid) and OPC-7683 (Delamanid).....	22
Figure 1-11. Chemical structure of pyrazinamide.....	25
Figure 1-12. Tricarboxylic acid cycle and associated fatty acid degradation pathway (β -oxidation).....	35
Figure 2-1. The growth curve over time of three mutants (D75A, D78A, and E83A) on different carbon sources, glycerol or butyrate.....	47
Figure 2-2. Overall fold of <i>Mtb</i> PEPCK-GDP-Mn ²⁺ -Zn ²⁺ complex structure	49
Figure 2-3. The active site of <i>Mtb</i> PEPCK in complex with GDP-Zn ²⁺ -Mn ²⁺ -PO ⁴	50
Figure 2-4. Absorption K- and L-edges for metals in X-ray energy range from 6 to 20keV	55
Figure 2-5. Anomalous difference map.....	55
Figure 2-6. Zn ²⁺ vs Mn ²⁺ activity in <i>Mtb</i> PEPCK dependent upon pH.....	56
Figure 2-7. Active site in apo- <i>Mtb</i> PEPCK crystal structure, loops are colored by B-factor	57

Figure 2-8. Superimposed crystal structure of human cPEPCK GTP-bound (PDB : 1KHE) and <i>Mtb</i> PEPCK-GDP bound forms	58
Figure 2-9. Superimposition of the apo and GDP-bound <i>Mtb</i> PEPCK structures	59
Figure 2-10. The interaction between Ω -loop and P-loop.....	61
Figure 2-11. Multiple sequence alignment of GTP-dependent PEPCKs color coded according to sequence similarity to <i>Mtb</i> PEPCK.....	62
Figure 2-12. Different position of PEP in <i>Mtb</i> PEPCK structure	65
Figure 2-13. The interaction between the R-loop and Ω -loop in human cPEPCK.....	71
Figure 2-14. Activity of wild-type, E83A and D78A mutant of <i>Mtb</i> PEPCK	73
Figure 2-15. Interaction of D78 and E83 residues in R-loop with adjacent residues	73
Figure 2-16. The interaction between E83 and R81 side chain maintains catalytic efficient rotameric conformation of R81 side chain in the active site.....	75
Figure 2-17. Glycerol binding pocket in <i>Mtb</i> PEPCK structure	77
Figure 2-18. Comparison of a glycerol binding pocket of <i>Mtb</i> PEPCK and human cPEPCK	78
Figure 2-19. The coupled enzyme reactions (PK/LDH) used for the measurement of <i>Mtb</i> PEPCK enzyme activity.....	81
Figure 2-20. The enzyme reaction of coupled malate dehydrogenase (MDH) enzymes	83
Figure 2-21. Comparison pyruvate-kinase activity of <i>Mtb</i> PEPCK in absence and presence of GDP molecule	88
Figure 2-22. Crystal contacts in <i>Mtb</i> PEPCK crystal structure.....	89
Figure 3-1. The first scaffold from HTS hit is 1-allyl-3-butyl-8-methylxanthine ($IC_{50} = 225 \pm 25 \mu M$) against human cPEPCK	91
Figure 3-2. The modifications of Roche compounds targeting the human cPEPCK	92
Figure 3-3. The binding mode of a GTP-competitive PEPCK inhibitor.....	92
Figure 3-4. The binding site of Axon1165 in <i>Mtb</i> PEPCK.....	95

Figure 3-5. Comparison of binding interaction with Axon1165 in <i>Mtb</i> PEPCK and human cPEPCK (PDB ID: 1NHX).....	96
Figure 3-6. Comparison of the electrostatic surface of the small pocket near the inhibitor binding site.....	98
Figure 3-7. Superimposed inhibitor binding pocket of the human cPEPCK and <i>Mtb</i> PEPCK structures.....	98
Figure 3-8. Superimposition of PEPCK-inhibitor complex structures with RO4355622, RO4355668, RO4355703, and RO505244	102
Figure 3-9. Superimposed inhibitor complex structures of <i>Mtb</i> PEPCK with RO4355622, RO4355668, RO4355703, and RO505244 and human cPEPCK with N-(4- {[3-butyl-1-(2-fluorobenzyl)-2,6-dioxo-2,3,6,7-tetrahydro-1H-purin-8-yl]methyl}phenyl)-1-methyl-1H-imidazole-4-sulfonamide (orange color) (PDB code:2GMV)	103
Figure 3-10. The binding mode of RO4355707 in <i>Mtb</i> PEPCK-RO4355707 complex structure.....	109
Figure 3-11. The binding mode of RO0506473 in <i>Mtb</i> PEPCK and chemical structure of RO0506473	109
Figure 3-12. The binding mode of RO0506333 in <i>Mtb</i> PEPCK-RO0506333 complex structure and chemical structure of RO0506333	110
Figure 3-13. The binding mode of RO0505787 in <i>Mtb</i> PEPCK-RO0505787 complex structure and chemical structure of RO0505787	111
Figure 3-14. The binding mode of RO0504449 in <i>Mtb</i> PEPCK-RO504449 complex structure and chemical structure of RO504449.	111
Figure 3-15. <i>Mtb</i> PEPCK-RO507000 complex structure and chemical structure of RO507000.....	112
Figure 3-16. The chemical structures of PEPCK-0057 and PEPCK-0058	118
Figure 3-17. Superimposed structure of <i>Mtb</i> PEPCK (blue) and human PEPCK (pink) complex with Axon1165.....	118
Figure 3-18. The selectivity of PEPCK-0057 and PEPCK-0058 against <i>Mtb</i> PEPCK over human cPEPCK	119
Figure 3-19. The complex structure of <i>Mtb</i> PEPCK with PEPCK-0057	121

Figure 3-20. The superimposed structures of the PEPCK-0057- <i>Mtb</i> PEPCK complex (blue) and the Axon 1165-human PEPCK complex (pink)	121
Figure 3-21. The superimposed structures of the PEPCK-0058- <i>Mtb</i> PEPCK complex (pink) and the Axon 1165-human PEPCK complex (tan)	122
Figure 3-22. The chemical structure of the SRI hits from the HT screen against <i>Mtb</i> PEPCK by DSF and enzyme inhibition assay.....	131
Figure 3-23. The chemical structure of the phenoxyalkylimidazole compounds purchased from ChemBridge Corporation.....	132
Figure 3-24. The chemical structure of two hits (EN:T5251690, EN:T5292830) from Sac1 libraries, and IC ₅₀ value and dose-response curve	135
Figure 4-1. Chemical structure of common allosteric effectors for pyruvate kinase from diverse species.	138
Figure 4-2. Multiple sequence alignment of <i>Mtb</i> PykA with all four human isozymes (M1, M2, L, R).....	141
Figure 4-3. <i>Mtb</i> PykA structure.....	143
Figure 4-4. AMP molecule found in the allosteric effector binding site.....	146
Figure 4-5. The effect of AMP and FBP on <i>Mtb</i> PykA	148
Figure 4-6. Comparison of the allosteric binding site in the <i>Mtb</i> PykA (gray) structure with those of the human M2 pyruvate kinase structure (sky blue), (PDB: 4G1N).....	149
Figure 4-7. Phosphate binding pocket in <i>Mtb</i> PykA	151
Figure 4-8. The Fo-Fc density map (green) for the ADP bound in the active site of subunit G is shown contoured at 3.1 σ	152
Figure 4-9. ADP bound to the active site of the <i>Mtb</i> PykA enzyme.....	153
Figure 4-10. Schematic of the <i>Mtb</i> PykA enzyme assay coupled LDH enzyme reaction.	157
Figure 4-11. Trypsin digested peptide sequence identified from an LC/MS/MS comparison of unmodified and modified samples.....	159
Figure 4-12. The enzyme activity of wild-type (blue) and mutants of <i>Mtb</i> PykA, K110A (red), K110Q (green), and K110R (purple)	161

Figure 4-13. CD result of wild-type and mutants of <i>Mtb</i> PykA, K110Q, K110R, and K110A.....	162
Figure 4-14. Location of K110 in the B-domain.....	163
Figure 5-1. Sequence alignment of the C- C' interface composed of C α 1 helices (349-366, numbering in <i>Mtb</i> PykA (or PK) and C β 5 strands (458-466 residues) of each subunit in the <i>Mtb</i> and human pyruvate kinase isozymes	165
Figure 5-2. Predicted binding mode of F0808-1900 (N-(3-bromo-4-hydroxynaphthalen-1-yl)-4-ethylbenzenesulfonamide) with <i>Mtb</i> PykA discovered in docking.....	169
Figure 5-3. Four purchased compounds which are similar to “hit” structure (F0808-1900) with IC ₅₀ values	170
Figure 5-4. Comparison of the docked F0808-1900 compound in the <i>Mtb</i> PykA structure (pink) with the human M2 PK (tan; PDB 3GR4).....	171
Figure 5-5. The chemical structure of the four identified active compounds from the virtual screen done by AbbVie with IC ₅₀ and MIC ₉₀ values for each compound.	173

LIST OF TABLES

	Page
Table 1-1. Currently used first-line drugs and second-line drugs, target and effect	3
Table 2-1. Trace metal content of <i>Mtb</i> PEPCK by atomic absorption spectroscopy	54
Table 2-2. Kinetics data of the wild-type and the Q448A mutant for the gluconeogenic and anaplerotic reactions	67
Table 2-3. The K _d value (M) of GDP for WT (wild-type), Q448A mutant and human cPEPCK	68
Table 2-4. Secondary structure data of wild-type and mutant <i>M. smegmatis</i> PEPCK enzymes	72
Table 2-5. Data collection and refinement statistics for <i>Mtb</i> PEPCK in apo-form and complex with substrates	87
Table 3-1. 104 compound screen against the <i>Mtb</i> PEPCK enzyme	99
Table 3-2. Chemical structures, IC ₅₀ , and MIC value of selected five compounds	101
Table 3-3. Chemical structure of xanthine analogs, sulfonamide, and amide derivatives with MABA MIC and percent inhibition at 10 μM of compounds against the <i>Mtb</i> PEPCK enzyme	104
Table 3-4. The IC ₅₀ and MABA MIC values of the hit compounds with chemical structure	107
Table 3-5. The chemical structure of Roche compounds with the IC ₅₀ value against human cPEPCK and <i>Mtb</i> PEPCK	114
Table 3-6. Xanthine analogs – sulfonamides are tested on mc ² 7000 under different carbon source	116
Table 3-7. Data collection and refinement statistics for <i>Mtb</i> PEPCK in complex with inhibitors	127
Table 3-8. IC ₅₀ and ΔT _m value of seven SRI hits against <i>Mtb</i> PEPCK	130
Table 3-9. Melting temperature and thermal shift of samples, apo enzyme, apo enzyme in the presence of Mn ²⁺ , apo enzyme with inhibitors, and apo enzyme with inhibitors in presence of Mn ²⁺	133

Table 3-10. Chemical structure and thermal shift (ΔT_m with Mn) of fragment hits	134
Table 4-1. Data collection and refinement statistics for <i>Mtb</i> pyruvate kinase (PykA) in complex with AMP- PO_4^{2-} and AMP-ADP- PO_4^{2-}	145
Table 5-1. Distinct structural classes of four hits were identified that exhibit inhibition of both the <i>Mtb</i> PykA enzyme and the whole-cell (mc ² 7000)	167

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Introduction

Tuberculosis, caused by the bacterium *Mycobacterium tuberculosis* (*Mtb*), is one of the oldest described diseases and it is still stubbornly embedded in our population as a leading infectious disease. One-third of the world's population is latently infected with *Mtb* that can be activated at anytime. The 20th WHO global TB report of 2015 estimated that the incidence of tuberculosis was about 10 million cases, killing approximately 1.5 million people in 2014 ¹. Part of the reason for the global resurgence of TB is due to the emergence of resistant strains of TB; multi drug-resistant TB (MDR-TB), extensively drug-resistant TB (XDR-TB), and totally drug-resistant TB (TDR-TB). The AIDS epidemic has also contributed to the resurgence of TB as people infected with HIV/AIDS are more susceptible to TB infection due to immunodeficiency ¹. The currently accepted TB treatment regimen is a combination of first-line TB drugs, such as isoniazid, rifampicin, pyrazinamide, and ethambutol (Table 1-1) for 6-9 months ² (WHO) against a drug-sensitive TB strain. Second-line anti-TB drugs are used for drug-resistant TB treatment. The second-line drugs contain aminoglycosides (kanamycin, amikacin, streptomycin) classified as injectable anti-TB agents, para-aminosalicylic acid (PAS), ethionamide, fluoroquinolones (moxifloxacin, levofloxacin), cycloserine, and bedaquiline and delamanid.

Despite the global resurgence of TB, there has been a dearth of new TB drugs in the market over the past 40 years. Although bedaquiline (SIRTURO[®]) and delamanid (Delyba[®]) have been recently approved only for treatment of drug resistant TB, we still need more new TB drugs to overcome drug resistance, shorten the lengthy treatment times. For new TB drug discovery, understanding the molecular mechanisms of action and resistance of current TB drugs have provided valuable insights and accelerated the progress of the development of anti-TB drugs. For example, there are over 20 compounds in various stages of new drug development. These compounds include novel target drugs, re-purposed existing antibiotics, and modified existing anti-TB drugs. A review of the literature on current TB drugs may open an avenue for investigation of novel drug targets and development of new effective anti-TB chemotherapies. Knowledge of drug/target interaction-mediated cell death and resistant mutations in target-encoding or related genes may also help us avoid cross-resistance. This thesis starts with a review of targets of existing TB drugs, problems of current a standard regimen, and novel targets that promise to overcome the drawbacks of exiting TB drugs followed by a discussion of the use of structural studies and inhibitor identification in suggested novel targets.

Table 1-1. Currently used first-line drugs and second-line drugs, target and effect.

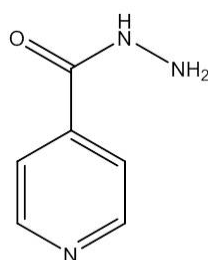
Class	Drug	Target and Effect
First-line drugs	Isoniazide	NADH-dependent enoyl-ACP reductase (InhA) Inhibition of mycolic acid biosynthesis
	Rifampicin	RNA polymerase, β subunit (Inhibition of transcription)
	Pyrazinamide	Multiple targets
	Ethambutol	Arabinosyl transferases (Inhibition of arabinogalactan synthesis)
Second-line drugs	Aminoglycoside (streptomycin, amikacin, kanamycin)	S12 and 16S rRNA components of 30S ribosomal subunit (Inhibition of protein synthesis)
	Para-amino salicylic acid	Dihydropteroate Synthase (Inhibition of folate synthesis)
	Ethionamide	NADH-dependent enoyl-ACP reductase (InhA) (Inhibition of mycolic acid synthesis)
	Levofloxacin, moxifloxacin (fluoroquinolone)	DNA gyrase (Inhibition of DNA synthesis)
	Cycloserine	D-alanine racemase and ligase (Inhibition of peptidoglycan synthesis)
	Bedaquiline	ATP synthase (Inhibition of ATP synthesis)
	Delamanid	Multiple targets

Mechanisms of currently available TB drugs

Current TB drugs can be classified based on the molecular mechanism or on the targets they affect. Most current drugs inhibit the cell wall, protein, DNA, RNA, or ATP, folate synthesis, while some inhibit multiple targets.

Inhibition of cell wall synthesis; isoniazide, ethambutol, ethionamide, D-cycloserine

The cell wall synthesis pathway is an ideal target for antimycobacterial drug development because those pathways are absent in the host. The unique cell wall architecture of *Mtb* is a major determinant of virulence for the mycobacterium. The cell is enclosed by layers of peptidoglycan, composed of extensively cross-linked peptide side chains consisting of L-Ala, D-Glu, meso-diaminopemelic acid, and D-Ala³. The high degree of peptidoglycan cross-linking contributes the structural integrity of the cell. The connection between peptidoglycan and mycolic acids is covalently linked with heteropolysaccharide, which is composed of arabinan and galactan (arabinogalactan)⁴. Mycolic acids are very long chain 2-alkyl, 3-hydroxy fatty acids whose length varies among species⁵. While *Corynebacterium* contains mycolic acid chains of up to 40 carbon atoms. *Mycobacterium* chains are much longer, 70-90 carbons, with the addition of different structures such as cyclopropanations (cis or trans), keto, or methoxy groups³. This lipid-rich mycobacterial cell wall contributes to the natural antibiotic resistance and impermeability of the cell envelope⁶. Isoniazid, ethambutol, ethionamide, and cycloserine are the classes of anti-TB drugs that inhibit cell wall biosynthesis. Moreover, the nitroimidazole antibiotics PA-824 (Pretomanid) and OPC-67683 (Delamanid) are known to target the cell wall.



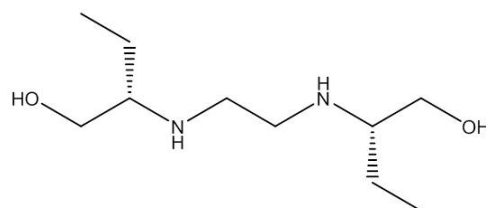
Isoniazid

Figure 1-1. Chemical structure of isoniazid.

Isoniazid

Isoniazid (isonicotinic acid hydrazide, INH) (Figure 1-1) is a pro-drug, activated by the catalase peroxidase, KatG, in *Mtb* to generate a hypothetical isonicotinoyl radical that binds to NAD(H). The active form, the INH-NAD adduct, inhibits the activity of NADH-dependent enoyl-ACP reductase (InhA) in the fatty acid synthesis type II system (FAS II), which is one of the mycolic acid biosynthetic pathways. Clinical isolates of isoniazid-resistant *Mtb* have been associated with a loss of catalase activity⁷ and transfer of the wild-type *M. tuberculosis katG* (catalase-peroxidase) into INH-resistant strains restored INH susceptibility⁸, suggesting that INH is activated by the *katG* gene. The crystal structure of the InhA complex with the INH-NAD adduct revealed that in clinical isolates isoniazid resistant mutations occur in the residues within or near the NADH binding site of InhA. Moreover, X-ray crystallography studies with mass spectrometry demonstrated that the KatG activated form of INH was covalently bonded to NADH in the active site of InhA, suggesting that the increased affinity of the INH-NAD adduct caused inhibition of isonicotinic acyl-NADH binding, thereby enabling the catalytic

action of InhA⁹. Since INH resistant mutations occur in KatG at a high rate, bypassing the activation step by creating a direct InhA inhibitor would be preferable. Therefore, interest is high in developing direct InhA inhibitors^{10, 11}. For example, GlaxoSmithKline performed an InhA target-based screen of a million compounds to identify inhibitors, and an SAR study was used to optimize the hit molecules. They found that tetrahydropyran analogs showed great InhA inhibitory potency in the nanomolar range with *in vitro* antibacterial activity¹² and thiadiazoles had low nanomolar InhA IC₅₀s with *in vivo* efficacy in a mouse model¹³. Those InhA inhibitors are currently under lead optimization in the TB drug pipeline¹⁴.



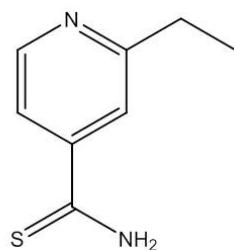
Ethambutol

Figure 1-2. Chemical structure of ethambutol.

Ethambutol

Ethambutol (2,2' ethylenediiminodi-1-butanol, EMB) (Figure 1-2) was discovered in 1961 by screening a collection of synthetic compounds and is a first-line TB drug. The primary target of EMB is a cluster of genes (*embCAB*) involved in arabinogalactan biosynthesis, it disrupts the polymerization of arabinan in the

mycobacterial cell wall, resulting in cell death. Metabolic labeling of cells with D-[¹⁴C] glucose exposed to EMB, revealed the accumulation of β-D-arabinofuranosyl-1-monophosphodecaprenol, which is the major intermediate of arabinogalactan biosynthesis⁴. Also, labeled [¹⁴C]-mycolic acid in the cell wall extracted from [¹⁴C] acetate grown cultures of *M. smegmatis* indicated that EMB inhibits the incorporation of mycolic acid into the cell wall¹⁵. Overexpression of the *embAB* gene leads to EMB-resistance due to a higher level of arabinosyl transferase activity¹⁶. Also, characterization of a clinical EMB-resistant *Mtb* isolates found mutations on the EmbB protein, and in EMB-resistant *M.smegmatis* there was a single point mutation on the *embB* gene¹⁷. Although, EMB has a modest effect against whole cells with an MIC of 10 μM, its low toxicity and synergistic effect with other drugs keep EMB in the first-line TB drugs. Therefore, significant efforts to develop EMB analogs have been made over the past 40 years to improve potency. SAR studies showed that the distance between the two nitrogens and the chirality of EMB is critical for activity. For example, the dextro (S,S) isomer shows 500 fold more efficacy than the (R,R) isomer¹⁸. Using combinatorial chemistry, diamine derivatives were synthesized, leading to identification of a novel EMB analog, SQ109, which exhibits a 35-fold greater potency than EMB¹⁹. SQ109 inhibits cell wall assembly using a different mechanism from EMB. SQ109 was discovered, it targets MmpL3, a transporter of trehalose monomycolate (TMM)²⁰ and clinical development is underway¹⁴.



Ethioamide

Figure 1-3. Chemical structure of ethioamide.

Ethionamide

Ethionamide (ETH) (Figure 1-3) is an important component of second-line therapy for the treatment of MDR and XDR-TB. There is structural similarity between ETH and INH, and over-expression of *inhA* confers resistance to both ETH and INH in *Mtb*²¹, indicating that InhA is ETH's target and is involved in mycolic acid biosynthesis, the same as INH. X-ray crystallographic data of InhA revealed that the ETH-NAD adduct binds in a similar position to INH-NAD²². Like INH, ETH is also a prodrug, but it is activated by EthA, a monooxygenase²³, and INH-resistant strains are susceptible when treated with ETH. This fact supported by overexpression *EthA* gene leads to ETH hypersensitivity and loss of function mutant strains were shown to be resistant to ETH²⁴. Baulard and his colleagues found that *ethA* gene expression is regulated by the *ethR* (Rv3855) gene, which is a transcriptional repressor for *ethA* gene²⁴. Low expression of *ethA* is responsible for the need for high dose treatment, leading to severe adverse effects. Therefore, Baulard and his colleagues attempted to identify EthR inhibitors to

increase ETH activation, and one of inhibitors resulted in reduced therapeutic dose of ETH²⁵.

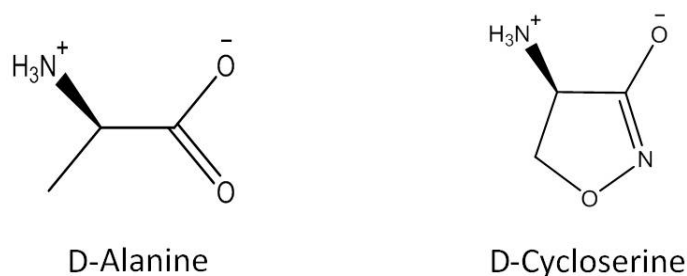


Figure 1-4. Chemical structure of D-alanine and D-cycloserine.

D-cycloserine

D-cycloserine (DCS) is a second-line antituberculosis drug for the treatment of MDR- and XDR-tuberculosis. DCS is a structural cyclic analogue of D-alanine, which is isolated from *Streptomyces orchidaceus* and *Streptomyces garyphalus* (Figure 1-4).

Shoolery and Rogers (1958) observed the accumulation of the uridine nucleotide, UDP-GNAc-lactyl-L-Ala-D-Glu-L-Lys, in growth inhibition of *S.aureus* by DCS and Strominger (1960) showed that D-alanine racemase (AlrA) and D-alanine: D-alanine ligase (Ddl) are competitively inhibited by DCS²⁶⁻²⁸. Also, Walter (1963) and Morrison (1962) reported that D-alanine rescued growth activity in presence of DCS growth inhibition in *Mycobacterium butyricum* and *Mtb*, respectively^{29 30}. Overexpression of the *AlrA* and *Ddl* genes in *M. smegmatis* supported that these overexpressed genes conferred resistance to D-cycloserine in *M. smegmatis*³¹. Therefore, it is thought that

DCS targets the AlrA and Ddl enzymes. AlrA catalyzes the conversion of L-Ala to D-Ala, and Ddl catalyzes the ATP-dependent ligation between two D-alanine to produce the D-alanyl-D-alanine dipeptide. Since D-alanine is one of the central components for the cross-linking, step of peptidoglycan assembly, inhibition of D-alanine racemase (AlrA) and D-alanine: D-alanine ligase (Ddl) blocks the peptidoglycan cross-linking and affects cell wall integrity. The severe adverse effects of DCS are attributed to non-specific inhibitions. It has been reported that DCS acts as a neuronal N-methyl-d-aspartate (NMDA)-receptor partial agonist³², and thus its side effects are associated with neural disorders including depression, memory loss, seizures, and suicidal tendencies, which makes it difficult for patients to complete the course of TB treatment³³. Therefore, finding the primary target of DCS is critical to develop less toxic and specific inhibitors and improve DCS's properties as an anti-TB drug. The arguments about this issue were found in several studies. Overexpressing the Ddl gene in *M.smegmatis* confers a lower level of resistance to DCS than overexpressing the alr gene, suggesting that Ddl is not the lethal target of DCS in *M.smegmatis*³⁴. In contrast, analysis of metabolomics, D-alanine:D-alanine, and Ala (D-, L-), from an H37Rv cell treated at the same DCS concentration, revealed that *Mtb* Ddl activity is strongly inhibited by DCS that is higher than Alr³⁵. Although the crystal structure of *Mtb* Ddl is available, the *Mtb* Ddl complex structure with DCS has not been solved, thus an SAR study has not yet been developed. The severity of the side effects limit its use, however, DCS remains a useful second-line TB drug because of its efficacy against MDR-,and XDR-TB and its lack of cross

resistance with existing anti-TB drugs. Improving the specificity of DCS could be an important toll in the treatment of TB.

Inhibition of protein synthesis: aminoglycosides

The ribosome is a complex molecular machine for protein synthesis. Significant structural differences between the bacterial and human counterparts makes the ribosome an attractive drug target for new antibiotics³⁶. The bacterial ribosome is composed of two subunits, the 50S and 30S, and each subunit is composed of a variety of ribosomal RNA and proteins. The 50S subunit catalyzes the formation of peptide bonds in polypeptide chains, and the 30S subunit is responsible for proofreading by discrimination of mismatched amino acyl transfer RNAs based on the mRNA codon³⁷.

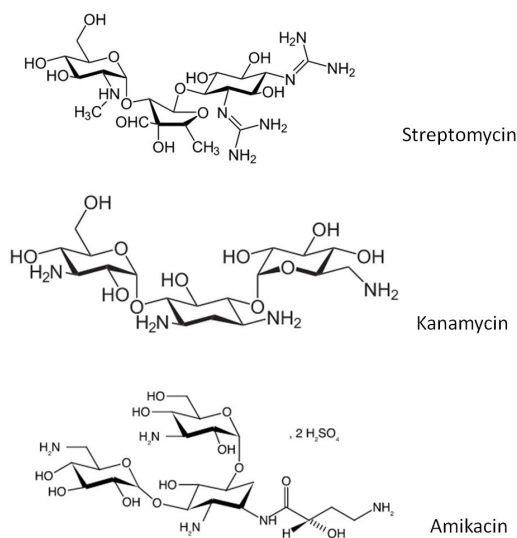


Figure 1-5. Chemical structure of aminoglycosides.

Aminoglycosides (Figure 1-5) are second-line antibiotics for MDR-TB treatment which inhibits bacterial protein synthesis by targeting the ribosome. They were isolated from soil microorganisms, and consist of an amino-modified sugar structure. The currently used aminoglycosides are streptomycin, kanamycin, and amikacin. Streptomycin was first introduced as a TB treatment in the early 1940s³⁸. Streptomycin-resistant clinical isolates harbored mutations in components of the 30S subunit of ribosome, 16S ribosomal RNA (or 16S rRNA, encoded by *rrs*) and ribosomal protein S12 (encoded *rpsL*)^{39 40}. Early experiment have demonstrated streptomycin interferes with the ribosome proofreading process⁴¹. A structure of the 30S subunit in complex with streptomycin showed that streptomycin binds to the phosphate backbone of 16S rRNA nucleotides, and also makes contact with protein S12. Binding to the ribosome causes non specific interaction with non-cognate tRNAs, resulting the protein mistranslation³⁷. The mechanism of action of amikacin and kanamycin action are same as that of streptomycin, but the efficacy of each aminoglycoside is reportedly different in practice. Due to the high frequency of streptomycin resistance in MDR-TB patients, streptomycin is not recommended to use for MDR-TB treatment. Kanamycin and amikacin are lower cost and less toxicity than streptomycin, thereby, both agents are the first choice of an injectable drugs⁴². The mutations in the *rrs* gene conferred a high-level of drug resistance to aminoglycosides. Recent studies on the mechanism of kanamycin resistance in *Mtb* identified a mutation in *whiB7*, a transcriptional activator, which increased the gene expression of both *eis* (Rv2416c) and *tap* (Rv1258c). Enhanced *eis* gene expression is associated with inactivation of kanamycin and increased *tap* gene

expression, a putative drug efflux encoding gene, also shows a low-level of streptomycin resistance⁴³. This study suggests that intrinsic mutations in *whiB7* can cause cross-resistance against clinically used aminoglycosides. However, not all resistant strains have mutations on *whiB7*, indicating that the mechanism of action of aminoglycoside antibiotic resistance is not clearly identified⁴⁴. Aminoglycosides have significant adverse effects, such as nephrotoxicity and ototoxicity in clinical use, resulting in patient non-compliance⁴⁵. The disadvantages of aminoglycosides make them a less than ideal MDR-TB treatment underscoring the requirement for more effective treatments.

Inhibition of DNA replication: fluoroquinolones (levofloxacin, moxifloxacin)

The process of DNA replication requires that the challenges arising from the topological issue of the circular double-stranded DNA helix structure be overcome. DNA topoisomerases overcome these challenges through breaking and rejoining DNA strands. There are two types of topoisomerases, types I and II. The type I topoisomerase breaks one DNA strand, while type II breaks double DNA strands at the same time. They are sub-divided into types IA, IB IIA, and IIB enzymes based on structure and mechanism. During DNA breakage, a covalent phosphotyrosyl bond between DNA and the enzyme forms to produce the covalent protein-DNA complex intermediates after nucleophilic attack by a tyrosine hydroxyl group in the active site toward the DNA phosphate^{46,47}.

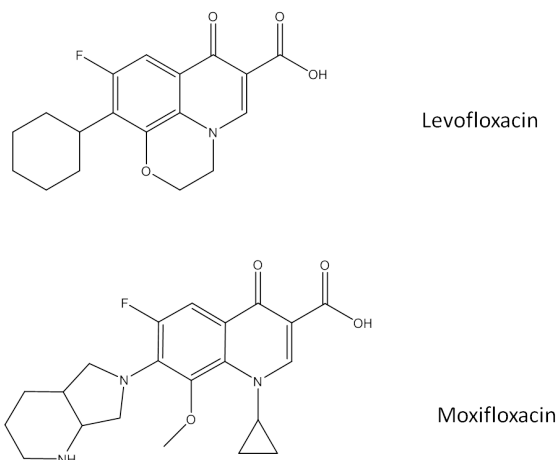


Figure 1-6. Chemical structure of fluoroquinolones.

The quinolones are synthetic molecules derived from nalidixic acid, a byproduct of the synthesis of chloroquine. It was discovered in 1962 and used to treat of Gram-negative urinary tract infections⁴⁸. Creating a fluoroquinolone by introducing of fluorine at the C6 position of the quinoline ring significantly improved bactericidal activity. The fluoroquinolones (levofloxacin and moxifloxacin) (Figure 1-6) are used for treatment of MDR-TB as second-line anti-TB drugs which target DNA-topoisomerase complexes formed by the type II DNA topoisomerase DNA gyrase⁴⁹. Stabilizing the covalent DNA-protein intermediate by fluoroquinolones arrests DNA replication by trapping topoisomerase at the DNA cleavage stage and preventing the resealing of the DNA breaks⁵⁰. Clinical isolates with mutations in *gyrA* or *gyrB* are associated with resistance to fluoroquinolone⁵¹. Due to its potential found in murin models of TB to shorten the length of treatment and active on both replicating and non-replicating tubercle bacilli, co-treatment of moxifloxacin with rifampin, pyrazinamide, was able to

eradicate *Mtb* from infected mice lung within 2 month⁵². Currently, novel TB regimen of PaMZ (PA-824, moxifloxacin pyrazinamide) for TB treatment is in clinical development^{14, 52}. However, the concern of fluoroquinolone-resistance leads to make efforts of finding new class of inhibitors targeting DNA gyrase are in progress, as DNA gyrase is a clinically validated therapeutic target.

Inhibition of RNA polymerase: rifamycins

Transcription by DNA-dependent RNA polymerase (RNAP), is a key step in gene expression and an essential step for bacterial growth and survival and the complexities of bacterial RNAPs differ from eukaryotic RNAPs⁵³, making bacterial RNA polymerase an attractive drug target. Most bacteria have a single RNAP with a core structure consisting of a multi-subunit, $\alpha_2\beta\beta'$ that forms a crab claw-like structure⁵⁴. The catalytic site of the enzyme is formed by the cleft between the β and β' subunits, which accommodates DNA templates to make phosphodiester bond during transcription elongation.

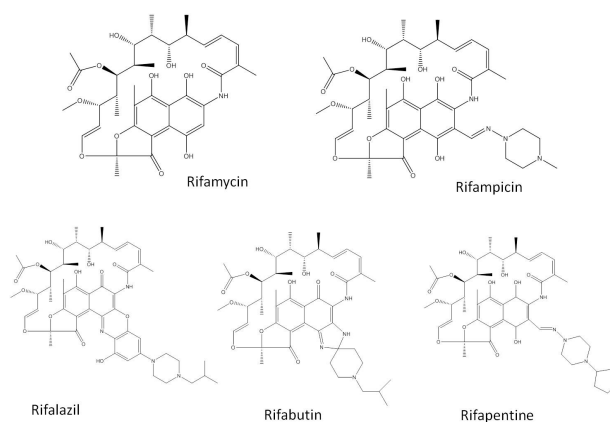


Figure 1-7. Chemical structure of rifamycins.

Rifamycins (Figure 1-7) are a part of the ansamycin family of antibiotics, whose backbone structure consists of a naphthoquinone core connected with an ansa polyhydroxylated bridge. It exhibits more potent activity on Gram-positive bacteria than on Gram-negative especially it is strongly affecting on mycobacteria. Rifamycins are an outstanding class of anti-tubercular drugs in that they shorten the treatment duration from 18 to 9 months, and have a sterilizing activity. Thereby, rifamycins are the mainstay of antituberculosis therapy, especially rifampin, one of rifamycins, which is currently used as a first-line of anti-TB drug. However, there are two major drawbacks to rifamycins (1) drug-resistance is rapidly acquired through mutations to the target enzyme, and (2) drug-drug interactions (with antiretroviral agents and bedaquiline) by inducing cytochromes P450 in the liver. For this reason, there have been numerous efforts to development analogs that minimize these drawbacks.

Understanding of the mechanism of action of rifamycins is critical for the development of novel agents or rifamycin analogs. Two different models from structure-function studies have been proposed to account for the mechanism of action of rifamycins: the steric occlusion model and the allosteric modulation model. Rifamycins target the β -subunit (encoded by *rpoB*) of DNA-dependent RNA polymerase (RNAP). The determination of the crystal structure of the *Thermus aquaticus* RNAP-rifampicin complex showed that rifampicin binds in a deep pocket 20Å away from the active center in the RNAP β -subunit. This study suggested the steric occlusion model for the mechanism, in which rifampicin sterically blocks the transcription elongation of RNA when the chain becomes 2 to 3 nucleotides long, but is not effective against longer

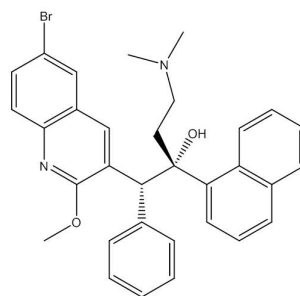
chains so it is only effective at a certain stage of transcription⁵⁵. However, the steric model could not explain all the mutations related to rifamycin-resistance. The mutations β -D516N and β -D516V on RNAP were found in the rifamycin-resistant strains, but do not decrease the affinity of rifamycins for RNAP. Moreover, rifabutin and rifalazil do not share cross-resistance with rifampin, indicating a different subset of resistant mutations⁵⁶. These observations prompted Artsimovitch et al. to propose the, allosteric modulation mechanism in which rifabutin and rifapentin bind allosterically to modulate a loss of affinity for catalytic Mg^{2+} binding in the active site resulting in transcription inhibition⁵⁷. However, Feklistov et al. re-tested the evidence in support of the allosteric modulation mechanism and found that the steric-interference model is the most likely real mechanism. Directly testing the affinity of rifamycins to Mg^{2+} in solution showed rifamycin had no effect on metal binding in the RNAP active site. Their study also showed that the rifamycin-resistant mutants, β -D516N and β -D516V, quantitatively affect on rifampicin-RNAP interaction, decreasing rifamycin affinity for RNAP. These results imply that tighter binding interactions between rifamycin analogs and the RNAP enzyme may have a chance to overcome rifamycin-resistant mutations. Of the numerous modifications and structure-activity relationships (SAR) studies on rifamycins, the most promising analogs are the piperidine ring-substituted rifabutin analogs and the rifalazil analogs because of their activities against a variety of rifampin-resistant clinical isolates of TB. Because these drugs target RNAP, it is being used as a target to search for additional novel compounds.

Another major problem of using rifamycins in TB treatment is its susceptibility to drug-drug interactions due to its ability to induce drug transporter proteins, P-glycoproteins and drug-metabolising enzymes, particularly cytochrome P450 (CYP) 3A4 in the liver and small intestine⁵⁸. These interactions cause rifampicin to decrease the plasma concentration of several drugs, especially HIV protease inhibitors, leading to treatment failure. A 600mg daily dose of rifampicin with a single protease inhibitor reduced the area under the concentration-time curve (AUC): indinavir was reduced by 92% , nelfinavir by 82% , and saquinavir by 70%⁵⁸. However, different rifamycins display different degrees of CYP induction; for example, co-administering with rifabutin (300 mg daily dose) improved the reduction of indinavir to only 34% below its original AUC⁵⁹. Co-administration with CYP3A inhibitors (HIV protease inhibitors) can increase the plasma concentration of rifabutin by inhibiting rifabutin metabolism by CYP3A4. Thereby, the different rifabutin profiles can alter the effectiveness of TB treatment for HIV co-infected patients. Though it may be possible to overcome two major drawbacks of rifamycins through structural modifications, rifamycin analogs still face the problem of drug-drug interaction with drugs other than antiretrovirals, such as other TB drugs (moxicloxacin⁶⁰ , bedaquiline⁶¹) creating more challenges to their use. Therefore, rifamycin-free regimes are being developed.

Inhibition of ATP synthase: bedaquiline

ATP synthase has a complex structure and reaction mechanism. It is composed of two domains, a membrane- bound F₀ domain, which is involved in proton translocation, and a cytoplasmic located hydrophilic F₁ domain that catalyzes ATP hydrolysis⁶².

Bacterial ATP production by F_1F_0 -ATP synthase is essential in a number of vital cellular processes. In aerobic bacteria, ATP synthase is involved in ATP generation through oxidative phosphorylation. In anaerobic bacteria, ATP synthase maintains pH homeostasis via generation of the proton motive force by pumping protons under non-respiratory growth conditions⁶³. In mycobacteria, ATP synthase is essential in both the replicating and non-replicating states. Hypoxic, non-replicating *Mtb* requires ATP synthesis activity to maintain ATP homeostasis⁶⁴. Deleting the *atpD* gene in *M.smegmatis* makes the bacterium incapable of growth on fermentable and non-fermentable carbon sources⁶⁵, and makes it a potential target for promising new antibacterial drugs.

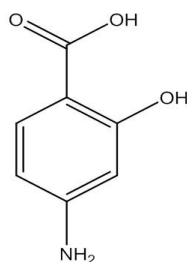


Bedaquiline

Figure 1-8. Chemical structure of bedaquiline.

Bedaquiline (BDQ) (Figure 1-8), the diarylquinoline TMC207 (also known as R207910), was approved by the FDA in 2012, specifically for MDR-TB treatment. It is effective not only on active, replicating *Mtb* but also against the dormant, non-

replicating state. The lead compound was discovered from a whole-cell screen, and then optimized into a series of diarylquinolines; the most potent compound, R207910, showed potent *in vitro* activity against *M. tuberculosis* H37Rv and clinical isolates, with MICs in the range of 0.030 to 0.120 µg/ml (rifampicin; 0.500 µg/ml and isoniazid 0.120 µg/ml). Its mechanism of action differs from pre-existing antitubercular drugs by in that it targets the energy metabolism of mycobacteria by inhibiting ATP synthase, leading to ATP depletion and causing an imbalance in pH homeostasis^{66 67}. Point mutations in the *atpE* gene, a part of the F₀ subunit of ATP synthase, conferred resistance to R207910⁶⁷. However, these target-based mutations were only observed in about 30% of the BDQ resistant isolates, indicating that multiple mechanisms of resistance existed⁶⁸. Rv0678, a transcriptional repressor of the gene encoding the MmpS5-MmpL5 efflux pump, was reported to be associated with non-target based resistance to BDQ. The BDQ resistant strains that lacked mutations on the *atpE* gene, showed cross-resistance to clofazimine (CFZ), and the sensitivity of CFZ and BDQ against the cell was increased in the presence of efflux inhibitors, verapamil and reserpine. Therefore, the MmpS5-MmpL5 efflux pump could be involved in non-target based acquired resistance of BDQ⁶⁹.



Para-aminosalicylic acid

Figure 1-9. Chemical structure of para-aminosalicylic acid (PAS).

Inhibition folate synthase: para-aminosalicylic acid (PAS)

para-aminosalicylic acid (PAS) (Figure 1-9) was introduced as an antituberculosis agent in 1940s⁷⁰, and is one of the second-line drugs used for MDR-TB treatment. PAS-resistant transposon mutants of *M. bovis* BCG harbor insertions in the thymidylate synthase (*thyA*) gene, which is involved in controlling intracellular folate levels and clinical PAS-resistant isolate strains contain mutations in *thyA*, showing that PAS action is associated with the folate pathway⁷¹. However, the detail mechanism of its action was not discovered over 60 years since it was introduced in TB treatment regimen. In 2013, Zheng, J. et al. has reported that PAS is a prodrug, its metabolite through the folate pathway working with dihydropteroate synthase (DHPS) and dihydrofolate synthase (DHFS) targets dihydrofolate reductase (DHFR, encoded by *dfrA*), inhibiting enzyme activity. They also showed that converted to the active form by enzymes in the folate pathway where it serves as an alternative substrate for dihydropteroate synthase (DHPS), resulting in an inhibition of folate metabolism^{72 73}. In addition, they found overexpressing RibD enzyme, known as a functional analogue of

DHFR, is resistant to PAS, suggesting that clinical isolates of PAS-resistance strains may contain mutations in *dhfrA* and *ribD* gene⁷². Due to PAS does not share cross-resistance with existing drugs; it can be still in second-line drugs for MDR- and XDR-TB treatment.

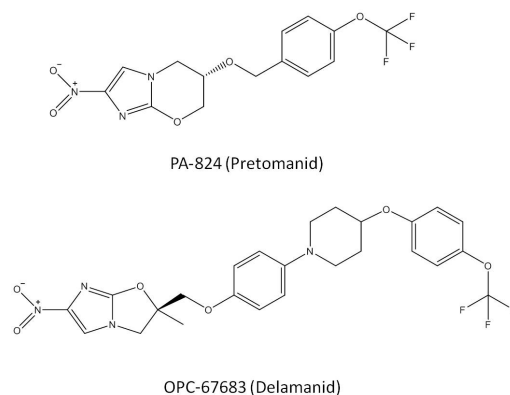


Figure 1-10. Chemical structure of PA-824 (Pretomanid) and OPC-67683 (Delamanid).

Inhibition of multiple targets: nitroimidazole, pyraziamide

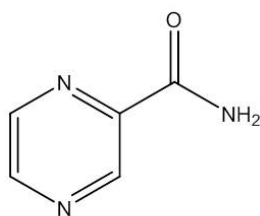
The novel antibacterial agents, PA-824 and OPC-67683 (Figure1-10), are a series of bicyclic nitroimidazoles, which were found to have anti-tubercular activity. These agents were developed using SAR studies by creating structural analogs of a radiosensitizer (CGI-17341) used in cancer chemotherapy. Due to the mutagenicity associated with bicyclic nitroimidazoles, its potential antibacterial properties were not investigated⁷⁴. However, Novartis found the lead compound PA-824 from screening 300 nitroimidazooxazines derivatives, which is active against *Mtb* without the mutagenicity. Otsuka Pharmaceuticals discovered and developed a series of nitroimidazooxazoles,

leading to the compound OPC-67683, Delamanid (Deltyba[®]), a potent compound against drug-susceptible TB and MDR-TB which first received approval for the treatment of MDR-TB in the European Union (EU) in April 2014.

PA-824 and OPC-67683 exhibit anti-tuberculosis activity against both non-replicating and actively replicating *Mtb*, which shortens the duration of TB treatment. Moreover, they do not share cross-resistance with current anti-TB drugs as seen by comparable susceptibility against MDR-TB strains⁷⁵. Both compounds are pro-drugs activated by a deazaflavin (cofactor F420) dependent nitroreductase (Ddn, Rv3547), and mutations in this gene confer resistance to both agents⁷⁶. Also, loss of a specific glucose-6 phosphate dehydrogenase (FGD1) or deazaflavin cofactor F420 is related to PA824 resistance, indicating they are necessary for sensitivity to nitroimidazole analogs⁷⁵. After intracellular bioreductive activation, the molecules inhibit mycolic acid synthesis under aerobic conditions, disrupting the formation of ketomycolates as seen by accumulation of hydroxymycolates. However, this mode of action does not explain the mechanism for killing non-replicating *Mtb*, since mycolic acid synthesis is not necessary for non-replicating *Mtb*⁷⁷. Singh et al. showed the PA-824 dependent production of HNO₂ by Ddn and demonstrated the des-nitro product from PA-824 activated by Ddn was correlated to the anaerobic killing activity of PA-824⁷⁸. The crystal structure of Ddn from *Mtb* and its homolog from *Nocardia farcinica* provided insights into the binding mode of PA-824 and allowed the development of novel analogues that bind to the mutant Ddn structure^{79, 80}. Analysis of ¹⁴C-labeled fatty acid and mycolic acid showed that OPC-67683 inhibits especially the synthesis of methoxy- and keto-mycolic

acid, with IC₅₀ values of 0.021-0.036 µg/ml, while INH inhibits non-specific mycolic acid subclasses (IC₅₀ = 0.630-1.851 µg/ml). Moreover, no effect of this compound on CYP enzymes was reported, indicating it can be used safely in HIV co-infected patients⁸¹. In clinical trials, the WHO reported that co-administration of delamanid with anti-retroviral drugs showed no clinically significant drug-drug interaction⁸².

Although the novel mechanism of bicyclic nitroimidazoles is an advantage against MDR- and XDR-TB treatment, the high rate of drug resistance is a major problem with these series of drugs. It has been reported that the mutation frequency to PA-824 resistance is comparable to that of INH⁷⁶, and two mutations in coenzyme F420 genes, *fbiA* and *fgd1*, conferred delamanid resistance as soon as delamanid was introduced to the regimen for MDR-TB treatment⁸³. However, the feature of nitroimidazoles activity on non-replicating strain is considered as important to be used in combination of other drugs for the TB treatment. Novel TB regimen including PA-824 is currently in clinical phase III, and results in shorter the length of treatment, safer to use in patients with HIV and TB than a standard regimen.



Pyrazinamide

Figure 1-11. Chemical structure of pyrazinamide.

Pyrazinamide (PZA) (Figure 1-11) is a first-line anti-TB drug discovered in 1952⁸⁴, and it is effective against semi-dormant TB within an acidic environment, contributing to shortening the required time for TB chemotherapy⁸⁵. Multiple targets of PZA in *Mtb* have been reported. PZA is a pro-drug, which is metabolized intercellularly to the active form pyrazinoic acid (POA) by pyrazinamidase/nicotinamidase (encoded by *pncA*). Mutations in the *pncA* gene result in PZA resistance in *Mtb* by preventing PZA conversion into POA^{86,87}. The ability of POA to disrupt the proton motive force and ATP synthesis has been observed^{88,89}, but the mechanism of action remains elusive. Ribosomal protein S1 (RpsA) was reported as a putative molecular target of POA in 2011 based on its binding activity to POA. RpsA is responsible for translation initiation of mRNAs and IS involved in *trans*-translation. *Trans*-translation in bacteria is a key process in high fidelity protein synthesis, arresting any ribosome with damaged mRNA free by removing the mRNA and the product of polypeptide chains⁹⁰. *Trans*-translation is performed by transfer-messenger RNA (tmRNA), which has the functions of both tRNA and mRNA. The RpsA C-terminus is involved in *trans*-translation by specifically

binding to the tmRNA and it has been proposed that POA binds to the C-terminal region of *Mtb* RpsA and disrupts the formation of the RpsA-tmRNA complex⁹¹. A structure of RpsA complexed with POA demonstrated that POA binds to the same RpsA residues involved in tmRNA binding⁹², thereby disrupting the interaction between RpsA and tmRNA, and causing inhibition of *trans*-translation. Recently the study to identify another target of PZA/POA reveals that the mutations in the *Mtb panD* gene (encoding aspartate decarboxylase) from *Mtb*, but not in the *pncA* and *rpsA* genes, conferred resistance to POA and PZA^{93,94}. Overexpression of the wild-type *Mtb* PanD resulted in POA resistance and the enzyme activity was inhibited by POA, but not by PZA. *Mtb* PanD catalyzes L-aspartate to β -alanine, which is required for pantothenate (vitamin B5) and CoA synthesis. The sterilizing effect of POA can be explained by the action of POA inhibiting this energy metabolism by targeting PanD⁹⁴. A complex structure of *Mtb* PanD with POA could elucidate the structural basis for inhibition of PanD by POA. Most frequent drug resistance mutant, M117I, in C-terminal of PanD is not involved in enzyme activity. From the structure of PanD from *E.coli* published in 2015, it reveals that C-terminus of PanD is critical for interacting with PanZ, which activates PanD⁹⁵. Therefore, it suggests that POA interrupts the important protein-protein interaction between PanD and other activators, leading to POA-resistance. However, the structure of PanD from *Mtb* has not been reported yet, thus it is unclear to demonstrate the molecular mechanism of POA/PZA action.

Problems of current TB treatment

Despite the availability of effective anti-TB drugs that addressed above, the treatment that completely could lead to the elimination of the disease has proven to be very difficult for a number of complex factors.

Currently used TB treatment regimes require at least 6 months for success clinical cure using combination drugs. This lengthy multidrug treatment is attributed to patient difficult to complete the course and adherence to drugs, leading treatment failure, relapse of infection and development of drug resistance. The global incidence of MDR- and XDR-TB cases has contributed that TB treatment becomes more challenging. MDR-TB is defined as a TB strain which is not sensitive to isoniazid and rifampicin, the first-line anti-tubercular agents. XDR-TB is a multi drug-resistant strain that is resistant to first-line and the most effective second-line anti-TB drug, fluoroquinolone, together with one of injectables (i.e., amikacin, kanamycin, or capreomycin). TDR is a TB strain that is resistant to all of anti-TB drugs ⁹⁶. The occurrence of drug-resistant strains has gradually increased with estimated 10 million TB cases in 2014; of those an estimated 300,000 cases are multi-drug resistant (MDR), and that 9.7% individuals are extensively drug resistant (XDR-TB) ¹. The treatment for MDR-TB requires up to 18-24 months treatment duration with a combination of eight to ten pills daily, leading 30% of MDR-TB patients fails to complete this course ^{97,98}. Only 50% of successful results of MDR-TB treatment has been reported, and an estimated 190,000 people died from MDR-TB in 2014 ¹. XDR-TB is virtually untreatable and it is estimated that 70% of XDR-TB patients die within a month of diagnosis ¹. Most MDR- and XDR- TB drugs cause high

rates of adverse events including, ototoxicity, uncontrolled psychiatric disorders, and gastrointestinal effects^{99, 100}, leading of discontinuation of MDR- and XDR TB therapy.

People with untreated latent TB infection and HIV infection are much more likely to develop active TB disease during his or her lifetime than people without HIV infection. One-third of the more than 33 million people living with HIV/AIDS are also infected with tuberculosis, one-fourth of HIV positive people die due to TB infection⁹⁶. Co-infection of TB and HIV complicates the anti-TB treatment because of drug-drug interaction and overlapping drug toxicities. For example, the problem of drug-drug interaction between TB drugs (rifampicin) and anti-retroviral (ARV) drugs makes TB treatment for people with HIV/AIDS difficult. Rifampicin induces the expression of cytochrome P450 (CYP) 3A4 enzyme in the liver and intestines. This enzyme is drug-metabolising enzyme that can metabolize protease inhibitors, results in reduced of the therapeutic concentration of the protease inhibitors. Therefore, the use of rifampicin with these HIV protease inhibitors contributes HIV treatment failure and development of drug resistance¹⁰¹.

Heterogeneity of bacterial population and variety of niches they reside in the host is the reason that current therapy regimen does not completely eradicate all *Mtb* in the body. Although RIF and PZA have shorten the treatment duration from 12-18 months to 6 months, persisting after treatment subpopulation can be re-activated when host immune system is compromised. Those bacteria remain in non-replicating state, leading to latent TB infection. Mitchison has classified four *Mtb* subpopulations depending on

drug susceptibility, 1) actively growing bacilli which are most effectively killed by INH, 2) slow growing bacilli that have spurts of metabolism, which are killed by RIF, 3) less metabolically active bacilli present in the acid pH environment that are killed by PZA, and 4) dormant bacilli which are not killed effectively by current drugs¹⁰². Therefore, sterilizing effect of TB drugs is important to prevent re-activation of subpopulation, and activation of latent-TB among one third of global population.

Moreover, TB has complex pathology, and as a result bacilli reside in compartments with very different characteristics: cellular compartment enriched by macrophages and easily accessed by plasma and caseous necrotic core, rich in lipids, with limited diffusion from plasma. It is challenging to produce a drug with such physical-chemical properties that it will readily distribute in both very distinct compartments. Recent study showed that despite favorable pharmacokinetics of all current drugs in blood, only RIF and PZA were able to achieve sufficient concentrations across both compartments of lesions¹⁰³. It clearly suggests that use of any single anti-TB drugs is not sufficient to kill all *Mtb* bacilli, concomitant drug treatment is necessary to increase the success rate of TB therapy.

These challenges to control TB with current antituberculosis drugs highlight the urgent need for developing new anti-TB drugs that satisfy several important requirements for effective treatment. First, new TB drugs should have no cross-resistance with the first-line drugs to kill effectively MDR- and XDR-TB. For this, it is important to develop new drugs that inhibit novel targets that are different from those of currently used drugs. Second, new drug should be effective against dormant *Mtb* and

shorten the lengthy TB therapy to improve patient compliance. The drugs should have limited toxicity, suggesting the target of inhibition should be only present in bacteria or have sufficient structural difference between host and pathogen. Also, the new drugs should be affordable to ensure accessibility to people in the developing countries, where is the majority of the high TB burden resides¹⁰⁴. Lastly, drug-drug interactions between TB drugs and antiretroviral drugs should be eliminated or minimal to be useful for therapy for co-infected HIV and TB patients.

New regimen(PaMZ) for efficient TB treatment

The new regimen developed by TB Alliance which includes moxifloxacin, PA-824, and pyrazinamid, called PaMZ, is currently being evaluated in phase III clinical trials and expected to change a paradigm of standard TB treatment. In phase 2b trial done in 2015, this novel TB regimen showed improved treatment of both drug-susceptible and MDR-TB. In clinical study of PaMZ with over 200 patients, cultured sputum of 71% patients with 2 month course of PaMZ treatment shows no TB, while sputum of only 38% patients with standard treatment shows no TB¹⁰⁵. Moreover, this regimen treatment shortens the duration treatment time for both drug-susceptible and MDR-TB. A standard treatment currently takes 6 months for drug-susceptible TB and 18-24 months for MDR-TB. Due to powerful bactericidal activity, this regimen is shortened to 4 months for drug-susceptible TB and to 6 months for MDR-TB. It also proved to be efficient in patients co-infected with TB/HIV. In addition, this novel regimen is not required the injectable agents, so it is easy to store, distribute and administer which is crucial for the

third world countries where the burden of TB is highest. Therefore, the PaMZ regimen promises simpler, shorter, more affordable and safer therapy for both drug-susceptible and MDR-TB infected patients.

Central carbon metabolism in *M.tuberculosis* for targets for new TB drug development

As shown the new regimen, that includes new anti-TB drugs, dramatically improved TB treatment. Over past 50 years, only two drugs are approved recently. Therefore, we need more drugs to combat with TB. To date, there are over 20 potential compounds in the stage of lead optimization, and clinical development. They include re-purposed drugs, new classes drugs with unknown targets, and new drugs targeting novel pathways¹⁴. As we reviewed above, current targets of existing drugs are very limited. To pursuit finding promising new anti-TB drugs, especially in respect to overcome the problem of drug-resistance in TB treatment, it should be exploited the novel targets/pathways that are essential for TB survival and persistence not only *in vitro*, but also *in vivo*.

Increased research on pathogenicity of *Mtb* suggested that enzymes of *Mtb*'s CCM are essential for *Mtb* survival, and important to be validated as potential drug targets¹⁰⁶. *Mtb* central carbon metabolism (CCM) comprised glycolysis, gluconeogenesis, and the pentose phosphate pathways, as well as the glyoxylate shunt, the methylcitrate cycle, and the tricarboxylic acid (TCA) cycle has been shown to be regulated in a way distinct from other well studied bacteria, like *E.coli*¹⁰⁷. *Mtb* resides in

phagosomes of human macrophages, and its CCM has evolved to survive this unique environment, which made it unlike other bacterial pathogens¹⁰⁶. Most bacteria can selectively utilize the carbon source that supports the fastest growth through regulating the expression and/or function of enzymes involved in metabolism of secondary carbon sources, a regulatory mechanism known as carbon catabolite repression (CCR)¹⁰⁸. In contrast, *Mtb* can co-catabolize mixed carbon sources through the CCM to distinct metabolic fates. As demonstrated in ¹³C trace experiments, *Mtb*'s preferred carbon source is dextrose for the intermediates of glycolysis and the pentose phosphate pathway, while it prefers to use acetate for TCA cycle intermediates. *Mtb* also maintained a low level of acetate-derived ¹³C through gluconeogenesis¹⁰⁹.

Growing body of evidence support the conclusion that ability to reconfigure flux through CCM in response to environmental changes is closely linked to virulence and persistence of *Mtb*^{110, 106, 111}. *Mtb* strains lacking the gluconeogenic enzyme or components of the glyoxylate shunt were defective in replication and showed attenuated growth during the chronic phase of infection in a mouse model of pulmonary TB. During infection in a carbon limited conditions, *Mtb* upregulates ICL, first enzyme of the glyoxylate shunt of the tri-carboxylic acid (TCA) cycle to use fatty acids as a carbon source¹¹². The glyoxylate shunt allows the cycle to bypass the loss of two CO₂ molecules and allows for the condensation of the C2 intermediate (Acetyl Co-A) from β-oxidation and glyoxylate (C2 unit) to produce C4 intermediates, which incorporate through gluconeogenesis to reach glycolysis- or pentose-phosphate-derived metabolites¹¹³. The *Mtb* glyoxylate shunt pathway contains a malate synthase (GlcB) and two

isozymes of isocitrate lyase (Icl1/Icl2). ICL upregulation was demonstrated under low oxygen conditions, associated with non-replicating stage of *Mtb* as well as during infection¹¹⁴, and is an important enzyme not only for the utilization of fatty acid but also for propionyl-CoA metabolism because it also function as methyl-citrate lyase¹¹⁵. An *Mtb* ICL knock-out mutant failed to establish an infection in mice, supporting the requirement of ICL for *in vivo* growth and virulence¹¹². Together with the absence of this pathway in mammalian cells, the *Mtb* ICL and GlcB enzymes have been spotlighted as attractive drug targets for new anti-TB chemotherapies. ICL possesses a small charged active site which makes it a difficult target, thus not many efficacious ICL inhibitors have been found^{116,117}. Structure guided inhibitor studies of GlcB have shown that phenyl-diketoacids (PDKA) inhibit *Mtb* growth on acetate and in mice¹¹⁸. Similar to ICL, phosphoenolpyruvate carboxykinase (PEPCK), which is the first committed step enzyme in the gluconeogenesis pathway, is also up-regulated during *Mtb* growth on fatty acids¹¹⁹. A PEPCK-deficient *Mtb* strain was not able to grow on fatty acids as a sole carbon source *in vitro*, and failed to establish and maintain infection in mice, indicating that the gluconeogenesis pathway is required for *Mtb* virulence¹²⁰. Using *Mtb* PEPCK as a of target for development is hampered by the existence of a highly conserved human homolog. Thus, a detailed structural study of *Mtb* PEPCK is required to create inhibitors selective for the pathogen over human enzyme. The importance of gluconeogenesis in *Mtb* virulence was also confirmed by a fructose bisphosphatase (FBPase, rv1099c) and an alternative FBPase (rv3214) knock-out study. In the *Mtb* H37Rv genome, two genes, *glpX* (rv1099c) and *gpm2*, working together has FBPase activity that maintains

gluconeogenesis in *Mtb*. While GLPX is annotated as an FBPase, inhibiting it alone is not sufficient to block FBPase activity as GPM2 can maintain gluconeogenesis in the absence of GLPX activity; deletion of both *glpX* and *gpm2* disrupts gluconeogenesis and results in attenuation of *Mtb* growth in a mouse model of infection¹²¹. In addition, a glucokinase null mutant in *Mtb* revealed the role of glucose metabolism in *Mtb* pathogenesis. The glucokinase genes, *ppgK* (Rv2702) and *glkA* (Rv0650), are functionally related to glucokinases, that is, they are involved in the first reaction in glycolysis and allow *Mtb* to incorporate glucose into the glycolysis pathway. An *Mtb* with deletion of both genes failed to persist in mouse lungs, meaning that *Mtb* uses glucose during infection via glucose phosphorylation to persist during chronic infection. These results suggest that the *Mtb* CCM has made a distinct metabolic adaptation to the host and has shown to be flexible in its use of diverse host-derived carbon sources. Therefore, the metabolic plasticity of the *Mtb* CCM has emerged as one of key determinants to its pathogenicity.

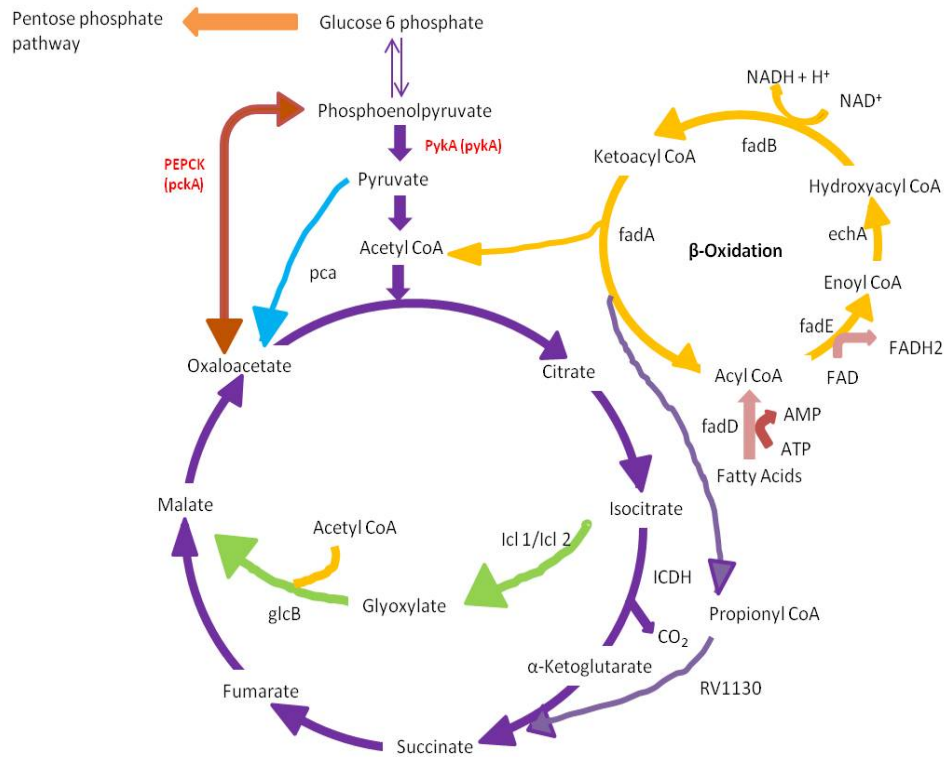


Figure 1-12. Tricarboxylic acid cycle and associated fatty acid degradation pathway (β -oxidation). Modified from Schnappinger ¹²².

PEP-oxaloacetate-pyruvate node in *Mtb* CCM

It has been reported by Sauer and Eikmanns that the junction between gluconeogenesis and glycolysis (interconversion of the intermediates PEP, pyruvate, and oxaloacetate) is important for carbon flux distribution in bacteria ¹²³. The intermediates PEP, pyruvate, and oxaloacetate are major metabolic precursors for the biosynthesis of several amino acids. As shown in Figure 1-12, oxaloacetate (OAA) can be used to synthesize aspartate and pyruvate can make alanine, valine, and leucine. Alanine is also crucial for peptidoglycan cross-linking, which is required for the rigidity of cell wall in *Mtb* that contributes its virulence ¹²⁴. PEP is used in the synthesis of the aromatic amino

acids phenylalanine and tryptophan ¹²⁵. In addition, they are also important energy precursors, entering the TCA cycle to produce cofactors NADH and FADH₂ that are used for respiratory ATP formation.

The PEP-oxaloacetate-pyruvate node is composed of several enzymes; Phosphoenolpyruvate carboxykinase (PEPCK), Pyruvate kinase (PykA), Pyruvate carboxylase (PCA), and Malic enzyme (MEZ) (Figure 1-12). PEPCK, preferentially catalyze PEP-production from oxaloacetate (OAA), and also can do the reversible reaction, OAA-formation from PEP by fixing CO₂. Pyruvate can be synthesized by the malic enzyme (MEZ) from malate, or by the pyruvate kinase (PykA) from PEP. In addition, pyruvate can be converted to oxaloacetate and malate by pyruvate carboxylase (PCA) and malic enzyme, respectively.

A carbon flux study of the *Mtb* CCM indicates that the major metabolic change in response to the macrophage environment occurs at the PEP-pyruvate-oxaloacetate node ^{109,110,120,121,126,127}, suggesting that the enzymes involved in this node represent good potential novel drug targets ¹⁰⁶. Hypoxia arrests the growth of *Mtb* and contributes to transition to the non-replicating state when granulomas develop. Low oxygen concentration induces up-regulation of PEPCK and malic enzymes, resulting in facilitating CO₂ fixation to replenish intermediates into TCA cycles ¹²⁸. Under slow growth rate conditions induced by limited nutrients, instead of fully turning TCA cycle, *Mtb* operates the GAS (Glyoxylate shunt and Anaplerotic reactions for oxidation of pyruvate, and Succinyl CoA synthetase) pathway for the generation of succinyl CoA. The pathway is represented by glyoxylate shunt, anaplerotic reaction for oxidation of

pyruvate, and succinyl CoA synthetase reaction, indicating that CO₂ fixation achieved by PEPCK, MEZ, and PCA is important for *Mtb*'s virulence¹²⁶.

In this dissertation, we chosen PEPCK and PykA enzymes as potential drug targets for the development of efficient anti-TB chemotherapies. The PEPCK enzyme exists across all species and has a highly conserved active site, and *Mtb* PEPCK is highly homologous to the human enzymes (cytosolic and mitochondrial versions). Thus, a detailed structural study of *Mtb* PEPCK is required to support a selective inhibitor development, which would be active against *Mtb* PEPCK and not to be toxic to the host. Unlike PEPCK, the essentiality of *Mtb* PykA in infection has not been thoroughly studied. PykA is located at the junction of glycolysis and gluconeogenesis. This point is a pivot to switch between both pathways, and the enzyme is likely to be tightly regulated by the relevant intermediates. Transcriptional regulation for *Mtb* PykA has not been observed in response to different carbon sources, which can be interpreted that either *Mtb* PykA function change is not required for adaptation of *Mtb* to the host environment or that the activity can be regulated on a different level, such as phosphorylation, acetylation, or allosteric regulation by effectors. Transposon mutagenesis study in H37Rv grown on glycerol and fatty acids resulted in no mutations in the *Mtb* PykA gene¹²⁹, suggesting that *Mtb* PykA is required for *in vitro* growth on glycerol and fatty acids and it has good potential as a drug target for TB chemotherapy. Interestingly, *Mycobacterium bovis* (*M. bovis*) harbors an inactivate version of pyruvate kinase despite sharing over 99.95% of its genome with *Mtb*¹³⁰. As a result, glycerol and glucose are not sufficient to support the growth of *M. bovis* in the absence of a pyruvate. An *Mtb* mutant

with disrupted PykA gene (*pykA*) also failed to grow on glycerol, but it grows on either pyruvate or oleic acid as a carbon source. Increased expression of proteins associated with fatty acids β -oxidation was observed in *Mtb* lacking PykA, suggesting that it shifts to fatty acid degradation upon the loss of the *pykA* gene¹³⁰. In addition, there was an increase in the expression of isocitrate dehydrogenase (Icd2: Rv0066c) and a decrease in expression of isocitrate lyase (Icl: Rv0467) upon the loss of the *pykA* gene. Flux balance analysis also indicates increased flux through the TCA cycle in the *pykA*-deficient *Mtb*¹³⁰. The considerable changes of the proteomic profile in response to the loss of *pykA* suggest that a better understanding of the role of *pykA* in the metabolic flux of the CCM is required.

Mtb PykA catalyzes the production of ATP from ADP during the conversion of PEP to pyruvate. Interestingly, PykA in *Trypanosomatidae* has been identified as a potential drug target because *Trypanosoma brucei* heavily relies on the glycolysis pathway for ATP production¹³¹. In *Mtb*, functional *icl* is required to reduce the intracellular ATP concentration which is associated with successfully entering the non-replicating state, suggesting a link between the CCM and the regulation of the intracellular ATP level¹³². An *Mtb* PEPCK knock-out mutant is still capable of producing pyruvate and alanine via the malic enzyme. Thus, *Mtb* PykA is not required to maintain intracellular level of pyruvate, but under certain conditions, ATP production via *Mtb* PykA may be important for maintenance of its intracellular ATP level. As Ehrt and co-workers described, pyruvate cannot be a precursor for PEP synthesis in *Mtb*, as PEPCK is the sole enzyme capable of PEP biosynthesis¹²⁷. It is possible that one more

important role of *Mtb* PEPCK is to provide the PEP intermediate which enables *Mtb* PykA to contribute to the maintenance of intracellular ATP levels. Despite *Mtb* PEPCK and PykA are critical for pathogenicity roles in *Mtb*'s metabolic adaptation to the host environment are critical to pathogenicity, these enzymes were not considered as drug targets due to the existence of their highly conserved host orthologs. Finding the differences in the atomic structures and regulatory mechanisms of these enzymes compared to those of the host orthologs would provide insights into the development of potent and specific inhibitors.

For the rest of chapter II and IV, we focused on the X-ray crystal structural studies for *Mtb* PEPCK (Rv0211) and *Mtb* PykA (Rv1617) to examine the structural differences between human and *Mtb* versions of these enzymes. We discovered regions that are distinct from human orthologs and can be exploited for the design of selective inhibitors. Moreover, biochemical studies for two enzymes provided information on regulatory events mechanism for better understanding of their role in CCM. Chapter III and V present the results and discussion of inhibitor study on two enzymes; for example, high-throughput screening was employed to identify their inhibitors followed by SAR study based on co-crystal structure with inhibitors were performed, especially for *Mtb* PEPCK. We developed selective *Mtb* PEPCK inhibitors over human enzyme and described the basis for selectivity in chapter III. For the inhibitor study on *Mtb* PykA, the interactions between a hit from HT screening and the enzyme were analyzed to propose a design of selective and potent inhibitors. This result may initiate the SAR study to develop more potent inhibitors. Our works represented in this dissertation contribute to

TB drug discovery and development, positioning PEPCK and PykA as potential drug targets and laying out an initial ground for development of selective inhibitors.

CHAPTER II
STRUCTURAL STUDY OF PHOSPHOENOLPYRUVATE CARBOXYKINASE
FROM *MYCOBACTERIUM TUBERCULOSIS*

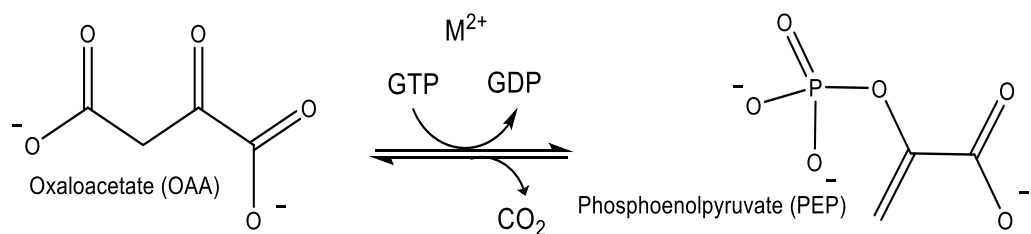
Introduction

Phosphoenolpyruvate carboxykinase (PEPCK) is the first and rate limiting enzyme in the *M.tuberculosis* gluconeogenic pathway, a part of the central carbon metabolism, that incorporates intermediates from the TCA cycle into glucose, amino acids, DNA, and RNA¹³³. The *Mtb* PEPCK enzyme is essential for TB growth and persistence during infection and is an attractive anti-TB target¹²⁰. Since the PEPCK enzyme analogs are conserved in the host, detailed structural study of *Mtb* PEPCK is needed for selective inhibitor design to specifically target the pathogenic enzyme.

PEPCK reaction

PEPCK (EC 4.1.1.32) reversibly catalyzes both the gluconeogenic (PEP-forming) and anaplerotic (OAA-forming) reactions. It preferentially operates the gluconeogenic PEP-forming reaction, the conversion of oxaloacetate (OAA) to phosphoenolpyruvate (PEP) in the presence of divalent metals and GTP. There are two steps in synthesis of PEP by PEPCK, first, the OAA is decarboxylated to become enolate anion intermediate, and then the γ -phosphoryl group of GTP transferred to make PEP. The metals play important roles to stabilize anion intermediate and GTP conformation¹³⁴. This reaction is only pathway that the intermediate (oxaloacetate) from the TCA cycle can be incorporated into the intermediate (PEP) in gluconeogenesis pathway. C¹³

tracing analysis shows that the knock-out PEPCK *Mtb* strain does not make the PEP, serine, and hexose-P intermediates when grown on an acetate carbon source, indicating gluconeogenesis is the only pathway to produce those intermediates¹²⁰. In the reverse, anaplerotic, reaction PEP is converted to OAA in the presence of divalent metals and GDP (Scheme 2-1). This reaction is required for *M.tuberculosis* survival under *in vitro* glycerol limited growth conditions, similar to the *in vivo* conditions of the macrophages, to maintain the availability of substrates for central carbon metabolism^{126,135}. During cell growth, intermediates from the TCA cycle are constantly withdrawn to supply necessary anabolic pathways; depleting the supply of C4 intermediates which must be restored. Replenishment of OAA into TCA cycle in most organisms can be accomplished, irreversibly, by phosphoenolpyruvate carboxylase (PEPC) (EC 4.1.1.31) or by the reversible PEPCK (EC 4.1.1.32). Interestingly, PEPC is found in most plants and bacteria, and even in other mycobacteria such as *M. smegmatis*, *M. avium* and *M. leprae*, but not in *M. bovis* and *M. tuberculosis*. Furthermore, it has been reported that PEPCK alone is sufficient for anaplerosis of OAA in *Saccharomyces cerevisiae* and *Corynebacterium glutamicum*^{136,137}. It is possible that *M.tuberculosis* relies on PEPCK for OAA anaplerosis due to the absence of PEPC, which makes this reaction important for *M.tuberculosis* survival during infection.



Scheme 2-1. *Mtb* PEPCK's reactions: PEP-forming reaction (gluconeogenic reaction) and OAA-forming reaction (anaplerotic reaction).

PEPCK isoforms

There are two types of PEPCKs, GTP-dependent and ATP-dependent, named for nucleotide specificity¹³⁸. *E. coli*, yeast, and plants have the ATP-dependent PEPCK isozymes, while eukaryotes, such as human, rats, and chickens, have the GTP-dependent PEPCK isoform¹³⁹. Despite low sequence homology between the primary sequences of ATP- and GTP-dependent PEPCK, their overall folded structure is highly similar and all active site residues are conserved¹⁴⁰. *Mtb* PEPCK is an exception, it is classified as a GTP-dependent PEPCK [EC 4.1.1.32] with an approximately 70% sequence similarity with eukaryotic PEPCKs and optimal activity in the presence of GTP¹⁴¹. Human GTP-dependent PEPCK has two isoforms; cytosolic PEPCK (cPEPCK) and mitochondrial PEPCK (mPEPCK). In the human liver, 30-50% of the PEPCK is the cytosolic isoform and 50-70% is the mitochondrial isoform.

PEPCK regulation

Human c PEPCK (GTP-dependent PEPCK) gene transcription is induced by starvation and its synthesis can be regulated by a variety of hormonal stimuli such as

glucagon, glucocorticoids, adrenaline, retinoic acid and thyroid hormone but reduced by insulin¹⁴². There is little evidence to suggest that enzyme can be regulated by protein modification¹⁴³. Recent study revealed human cPEPCK has a post-translational mechanism of regulation by lysine acetylation¹⁴⁴. In mycobacteria, the transcription of PEPCK is up-regulated when *Mtb* growth on fatty acid as a carbon source¹¹⁹. Moreover, it has been reported that hypoxic conditions, protein-protein interactions with thioredoxin, and Fe²⁺ ions induce *Mtb* PEPCK activity in the direction of the anaplerotic reaction (OAA-formation)^{145,146}. PEPCK catalyze the anaplerotic reaction in pyruvate kinase-deficient mutant of *B.subtilis*, and also functions CO₂ fixing enzyme in anaerobic bacteria. On the other hand, the preferred reaction in mammalian PEPCK is thought to be gluconeogenesis. This evidence suggests that PEPCK can play both gluconeogenesis and anaplerotic reaction in highly varied species. It implies structural difference between organisms may provide the key point to understand those differences.

Metals in PEPCK activity

Two divalent metal ions are required for optimal PEPCK activity. Kinetic studies have been carried out with various metals in different organisms and indicate the effect of divalent cations differs widely. Mn²⁺ is the preferred activator in mitochondrial PEPCK in many species^{147,148, 149,150}. In *C. glutamicum* PEPCK, Mn²⁺ and Co²⁺ both highly activated the enzyme reaction, while Cr³⁺ and Zn²⁺ showed an inhibitory effect¹⁵¹. Incubation of a low concentration of Fe²⁺, Co²⁺, Mn²⁺, or Cd²⁺ with rat liver PEPCK in the presence of Mg²⁺ showed improved enzyme activity in the PEP-forming reaction but maximal activity was obtained combining with Mg²⁺ and Mn²⁺ in the PEP-forming

reaction of PEPCK from cytosolic rat liver. The synergetic effect of Mg^{2+} and Mn^{2+} has been reported for GTP-dependent PEPCKs¹⁵². In a crystal study of *E. coli* PEPCK, Mn^{2+} placed in the catalytic active site to stabilize the transition state and Mg^{2+} lay in the nucleotide binding site to create the Mg^{2+} -ATP complex¹⁵³. A crystal study of human cPEPCK also defined these two metals as Mn^{2+} in the catalytic site and Mg^{2+} in the GTP-binding site based on their maximal activity¹⁴⁰.

Dynamics in PEPCK

Crystallographic studies of GTP-dependent PEPCK have revealed the presence of a three flexible loop architecture; the P-loop, the Ω -loop (the active site lid), and the R-loop whose movements are important to function¹⁵⁴. The P-loop is the phosphate binding region, which is a common structure found in nucleotide binding proteins, such as protein kinase and glutathione synthetase which have an ATP-binding domain¹⁵⁵. The R-loop is composed of residues 84-92 in rat cPEPCK, containing an arginine residue (R87 in human cPEPCK) that binds directly to the ligand (OAA/PEP) through electrostatic interaction. Significant movement of this loop has not been observed in the wild-type PEPCK enzyme. However, the absence of the Ω -loop in a mutant PEPCK structure shows that the conformational changes are preserved while the enzyme activity is abolished¹⁵⁶. The P- and R-loop participate directly in substrate binding: the phosphate group of the GTP or GDP molecule binds to the P-loop, and the OAA or PEP molecules bind to the R-loop. On the other hand, the flexible, Ω -loop on the protein surface, rather than directly interacting with the substrate, changes from an open to a closed conformation upon substrate binding, acting like an active site lid¹⁵⁴.

The Ω -loop consists of 11 residues (A464-V474 residue in rat cPEPCK) whose conformational changes between an open and closed form are depending upon ligand binding in the active site. In the rat cPEPCK, Ω -loop closure occurs when all substrates form a Michaelis complex in the active site. Closing of the Ω -loop helps to stabilize and protect of the reactive intermediate resulting in an increase in enzyme accuracy (specificity).

The R-loop in *Mycobacterium* consists of residues 79-86. In *Mycobacterium smegmatis*, PEPCK mutations to D75, D78 and E83 (D75A, D78A, and E83A) were generated and the D78A and E83A mutants decreased enzyme activity in both directions¹⁵⁷. In Dr. Sabine Ehrt's group in Cornell University, one of our collaborators, especially study on *Mtb* PEPCK enzyme. They generated these mutants in order to see whether these mutants can support *Mtb* growth on different carbon source. As shown Figure 2-1, the growth of the wild-type PEPCK shows normal growth on glycerol and butyrate, but the E83A and D78A mutants cannot support bacterial growth on butyrate, but D75A supports growth as with the wild-type bacteria. Therefore, the E83A and D78A mutants in *Mtb* are expected to display impaired enzyme activity. Further studies are needed to completely understand the role of these residues, and of the R-loop, in *Mtb* PEPCK.

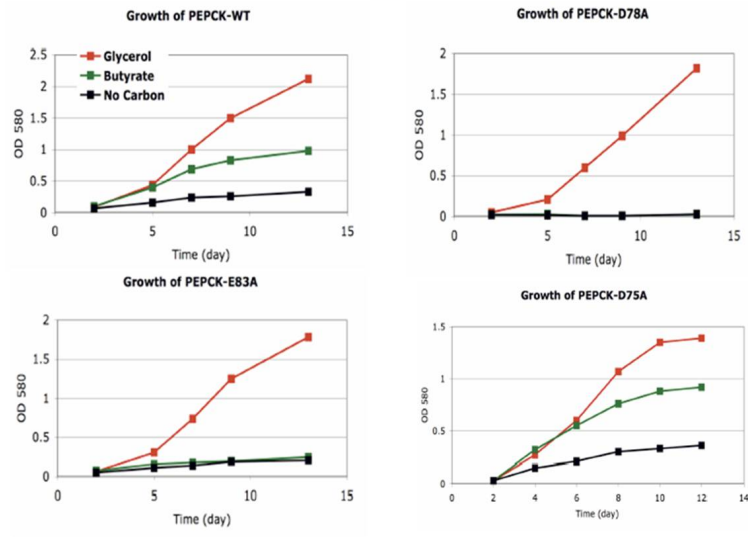


Figure 2-1. The growth curve over time of three mutants (D75A, D78A, and E83A) on different carbon sources, glycerol or butyrate. Data was provided from Dr. Sabine's group. D78A and E83A mutants are not able to support TB growth on butyrate as a sole carbon source, while D75A mutant shows similar growth rate to wild-type PEPCK.

Sequence similarity between the Mtb and human PEPCK isozymes (cytosolic PEPCK and mitochondrial PEPCK)

The gluconeogenesis pathway is shared between host and pathogen, and all PEPCKs have a similar overall structure with a conserved active site; both *Mtb* PEPCK and the human cPEPCK/mPEPCK are also GTP-dependent, indicating a high conservation. *Mtb* PEPCK contains 606 amino acid residues while the human protein has 622 residues for the cytosolic isoform, and 640 residues for the mitochondrial isoform. The sequence identity of *Mtb* PEPCK is 52% and 51% with the human cPEPCK and mPEPCK, respectively, and the similarity is approximately 67% and 65% for cPEPCK and mPEPCK. Therefore, a detailed comparison of *Mtb* and human PEPCK structures is

required to develop selective molecules to inhibit only *Mtb* PEPCK without affecting the host enzymes.

In this study, the crystal structures of *Mtb* PEPCK in the apo form and in complex with substrates were determined at high-resolution. The structural data of *Mtb* PEPCK reveal that the behavior of the flexible loops upon substrate binding differs from the reported rat cPEPCK structure. A unique interaction between the P-loop and the Ω -loop in *Mtb* PEPCK increases the efficiency of catalysis in the direction of the anaplerotic reaction. Since *Mtb* PEPCK has disparate flexible loop properties in the active site, they provide an opportunity to design ligands that can selectively inhibit *Mtb* PEPCK over the human enzyme.

Result and discussion

Overall structure of Mtb PEPCK and substrate binding site

The *Mtb* PEPCK structure displays high conservation with known PEPCK structures. It consists of an N-terminal and a C-terminal domain with an α/β topology that brings the termini into close proximity. The co-purified GDP-metal complex is shown in Figure 2-2, the active site of *Mtb* PEPCK is located in the cleft between the two domains as reported in other GTP-dependent PEPCK structures¹⁴⁰. The key residues in the active site involved in substrate binding and metal ion coordination are highly conserved in the *Mtb* PEPCK enzyme.

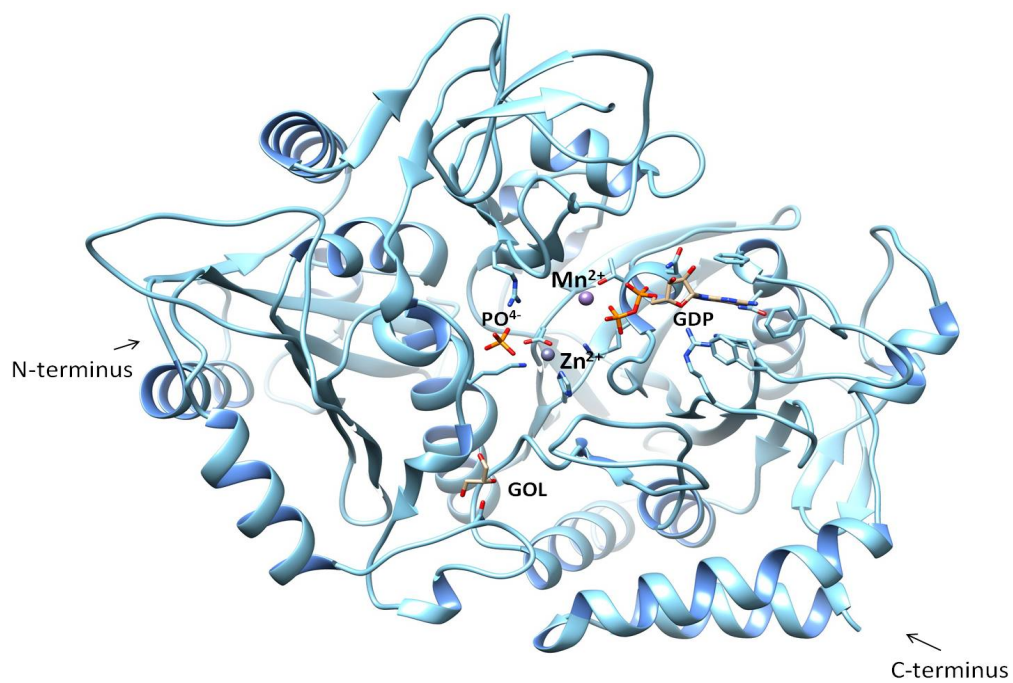


Figure 2-2. Overall fold of *Mtb* PEPCK-GDP-Mn²⁺-Zn²⁺ complex structure. GDP is bound to the C-terminal domain, the phosphate ion is bound to the N-terminal domain, and there are two divalent metals between the two domains.

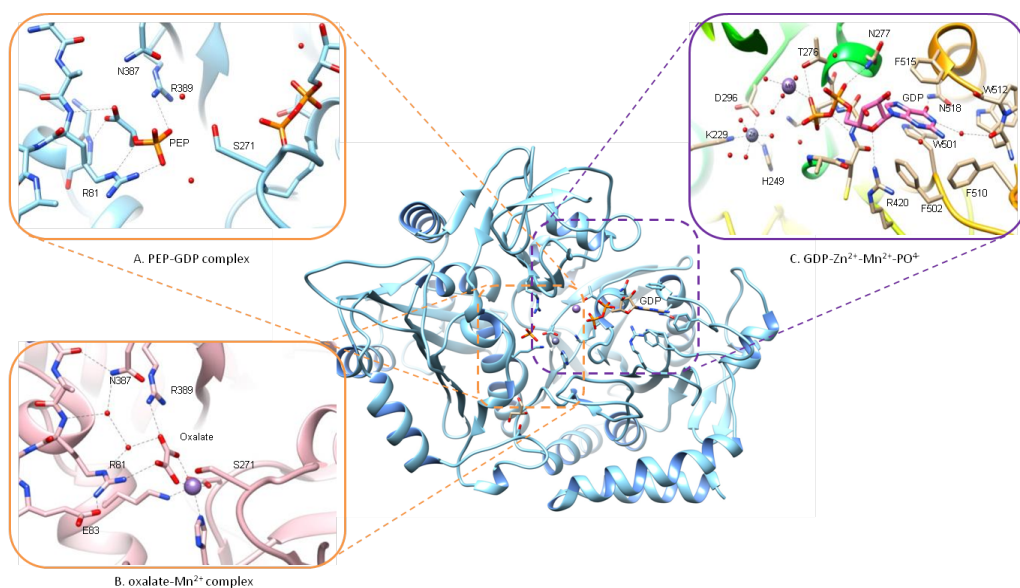


Figure 2-3. The active site of *Mtb* PEPCK in complex with GDP-Zn²⁺-Mn²⁺-PO⁴⁻. The anionic substrate, PEP or OAA, binds respectively to the N-terminal domain in the (A) PEP-GDP or (B) OAA-Mn²⁺ complex structure. The nucleotide substrate, (C) GTP or GDP, binds to the C-terminal domain. Two divalent metals are located between the N- and C-terminal domains.

Mtb PEPCK has two separate substrate binding pockets, PEP or OAA binds in the N-terminal domain, while the nucleotide (GTP, GDP) binds in the C-terminal domain, there are also two metal binding sites located between the N- and C-terminal (Figure 2-3). In the oxalate-Mn²⁺ *Mtb* PEPCK complex structure, the two oxalate oxygens interact via hydrogen bonding to the H-N^o of the side chains of R81 (2.9 Å) and R389 (2.7 Å). The rest of the oxalate oxygens chelate a catalytic metal at distances of 2.6 Å and 2.4 Å. In the PEP-GDP complex structure, the two PEP phosphate oxygens interact with the H-N^o of the R81 and R389 side chains at 3.1 Å and 2.8 Å, respectively. The carboxyl oxygen of PEP hydrogen bonds to the backbone amide hydrogen atom of the G222 and R81 residues, which is consistent with the human

cPEPCK-PEP complex. However, the interaction with N387 and the movement of Y220 observed in human cPEPCK were not observed in *Mtb* PEPCK structures.

In the *Mtb* PEPCK-GDP complex structure, the guanine base of the nucleotide is surrounded by the three benzene rings of the phenylalanine residues, F515, F510, and F502, in the C-terminal domain. The carbonyl oxygen atom from the guanine base ring hydrogen bonds to the side chain -NH₂ group of N518 (3.0 Å) and the backbone amide group of F515 (2.90 Å). The oxygen atom of the ribose ring hydrogen bonds to the ω – NH₂ group of the R420 side chain (3.1 Å). The GDP phosphate group is stabilized by the backbone amides of residues from the P-loop (S271-N277) including A272 (2.9 Å), G274 (3.1 Å), T276 (3.0 Å), and N277 (2.9 Å). This nucleotide interaction is also observed in the recently reported *Mtb* PEPCK-GDP-bound structure ¹⁴⁶.

Identification of co-purified metals with Mtb PEPCK suggests Zn²⁺ bound in the catalytic site and Mn²⁺ bound in the nucleotide binding site

The PEPCK enzyme is dependent upon divalent metal ions for activity. Although the active site is very highly conserved, including the two metal sites found in PEPCKs across species from bacteria to humans, studies of metal usage have not provided a cohesive picture of PEPCK's metal requirements. Kinetic data for divalent metals in different PEPCKs show a highly divergent effect from different divalent cations. Mn²⁺ has been used as an efficient activator in mitochondrial PEPCKs of many species ^{139,147,158}, while maximal PEPCK activity in cytosolic rat liver was obtained using a combination of Mg²⁺ and Mn²⁺ for the PEP-forming reaction ¹⁵². Two metals in reported crystal structures were not previously identified but have now been denoted Mn²⁺ and

Mg²⁺ based on activity observations, metal coordination geometry, or crystallization conditions^{140,153}. Recently, the structure of *Mtb* PEPCK in complex with GDP -Mn²⁺ was released. That paper showed that the Electron Paramagnetic Resonance (EPR) from Mn²⁺ could tightly bind to the nucleotide metal binding site, but not the catalytic metal binding site¹⁴⁶. The factors governing PEPCK metal selection requires more study for insight into its activity. As we did not include any metals during purification or crystallization, we made an effort to identify the metals we observed bound to *Mtb* PEPCK.

The density for both metal sites shows similar coordination. The metal density in the catalytic site is octahedrally coordinated with the N^{ε2} atom of H249 at a distance of 2.3 Å, the Nζ atom of K229 at 2.4 Å, one oxygen from the side chain of D296 at 2.2 Å, and with three water molecules at an average distance of 2.3 Å. The nucleotide binding site metal is also octahedrally coordinated with the hydroxyl oxygen from the side chain of T276 at 2.2 Å, one oxygen atom from the β-phosphate of GDP at a distance of 2.3 Å, and four water molecules at an average distance of 2.3 Å.

To identify the two co-purified metals in the structure, we used inductively coupled plasma mass spectrometry (ICP-MS) in combination with anomalous scattering. The ICP-MS analysis of the trace metals in the protein sample, purified directly from *E. coli* without added metals, showed a mixture of divalent metals bound: Ni, Zn, Mg, Fe, Mn, and Cu, each with a less than 1:1 molar ratio with the enzyme. They were found at relatively high concentrations in the order, Ni, Zn, Fe, Mn, Mg, and Cu, as shown in Table 2-1.

Anomalous difference scattering generated at a wavelength of 1.54 Å showed a strong anomalous signal for the metal in the nucleotide binding site. Mn²⁺, Fe²⁺, and Co²⁺ were the only metals from our analysis that had a significant anomalous signal at 1.54 Å, as shown in Figure 2-4. At 0.97 Å, a strong anomalous scattering signal was seen in the catalytic metal site, with no anomalous signal observable in the nucleotide binding metal site. The only transition metals corresponding with anomalous scattering at this wavelength are: Cu²⁺, Ni²⁺, and Zn²⁺ (Figure 2-5). We designated the catalytic metal Zn²⁺ due to the relatively high concentration seen in ICP-MS, Ni²⁺ was disregarded as being most likely an artifact of the His tag purification step. The probability of binding of Mn²⁺ and Fe²⁺ in the nucleotide binding site is similar and it is difficult distinguish them due to their similar properties as transitional metals, but we designated it Mn²⁺ based on the ICP-MS result, which showed a slightly higher molar ratio than Fe²⁺.

Identifying zinc as the catalytic binding site metal binding site was surprising as it is not the preferred metal for activity in other PEPCK enzymes. In the literature, the PEPCK enzyme from *Thermococcus kodakaraensis* showed no activity in the presence of Zn²⁺ ¹⁵² and Zn²⁺ has an inhibitory effect on *C. glutamicum* PEPCK ¹⁵¹. Since we found that Zn²⁺ binds to the catalytic metal binding site, we attempted to test *Mtb* PEPCK activity using Zn²⁺ or Mn²⁺ as the solo divalent cation across a range of pHs from 5.8-8.5 (Figure 2-6). We noticed that the HEPES molecule in the buffer showed 50% inhibition at 50 mM HEPES at pH 7.0, thus we used Tris buffer, which has no inhibition (data not shown). Mn²⁺ and Zn²⁺ show similar activity, about 0.5

$\mu\text{mole/mg}\cdot\text{min}$ at pH 7.0, but at pHs below 7.0, Zn^{2+} shows higher activity than Mn^{2+} . Above pH 7.0, the reverse is true and Mn^{2+} is preferred to Zn^{2+} . This result suggests that Mn^{2+} and Zn^{2+} activity for *Mtb* PEPCK is comparable at pH 7.0, indicating a similar probability of binding to the catalytic site. Since we purified *Mtb* PEPCK in a pH 7.0 buffer condition, it is likely that the mixture of metals in the catalytic metal site contained Zn^{2+} .

Table 2-1. Trace metal content of *Mtb* PEPCK by atomic absorption spectroscopy.

Metal	ppb($\mu\text{g/l}$)	μM	Molar ratio, metal/enzyme
Ni	64	1.090	0.153
Zn	52	0.795	0.111
Fe	12	0.215	0.030
Mn	12	0.218	0.031
Mg	12	0.494	0.069
Cu	6	0.094	0.013

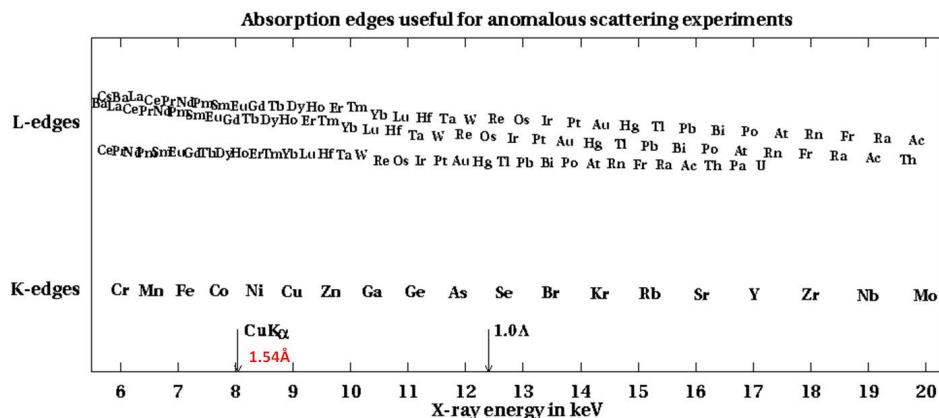


Figure 2-4. Absorption K- and L-edges for metals in X-ray energy range from 6 to 20keV. This chart is made by Ethan Merritt¹⁵⁹.

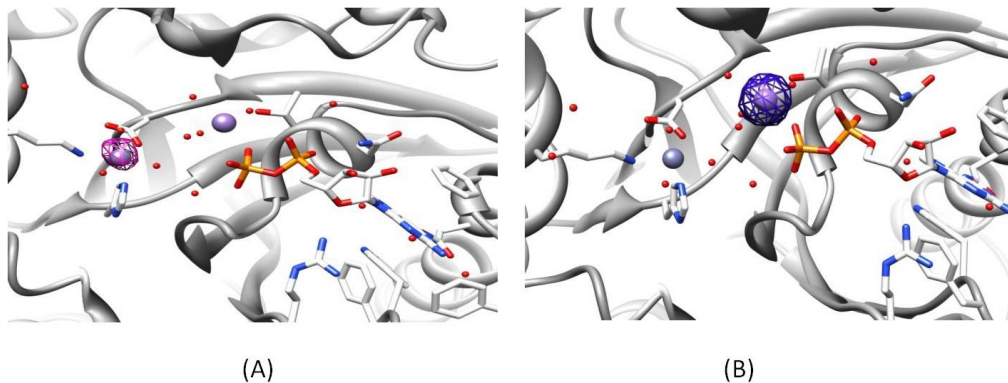


Figure 2-5. Anomalous difference map. (A) The anomalous signal appeared in the catalytic metal site, which coordinated with D296, K229, and H249, and water molecules (data collected at 0.97 Å), and (B) The anomalous signal appeared in the nucleotide binding metal site with GDP, T276, and water molecules (data collected at 1.54 Å).

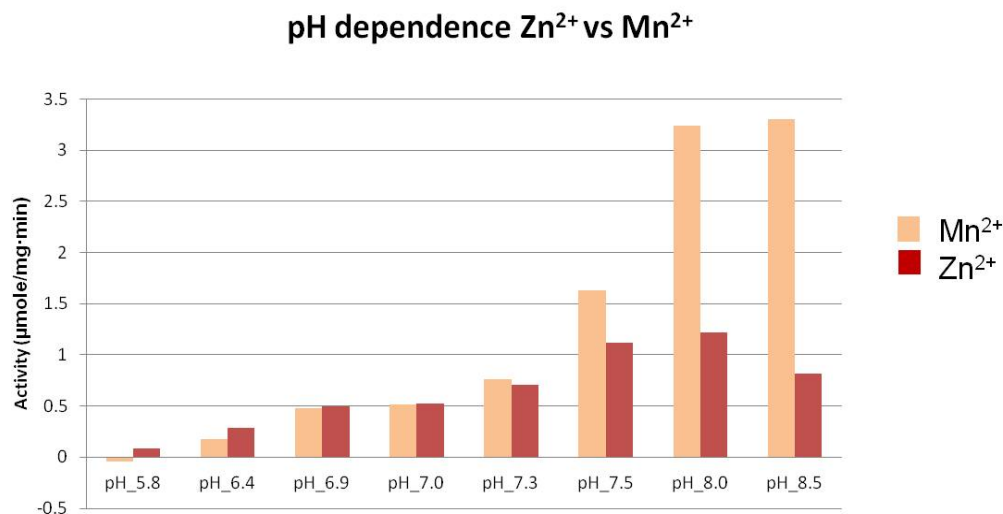


Figure 2-6. Zn²⁺ vs Mn²⁺ activity in *Mtb* PEPCK dependent upon pH.

Overall structural comparison of Mtb PEPCK and human cPEPCK reveals intriguing differences in flexible loops

Structural differences in the active site of PEPCK between the host and pathogen are of particular importance for the design of selective inhibitors that specifically target *Mtb* without affecting the host. The sequence identities in the active and substrate binding site are significantly conserved between the pathogen and host enzyme. *Mtb* PEPCK contains 606 amino acid residues while the human protein has 622 for the cytosolic isoform, and 640 for the mitochondrial isoform. The sequence identity of *Mtb* PEPCK is 52% and 51% with the human cytosolic PEPCK(cPEPCK) and mitochondrial PEPCK(mPEPCK), respectively, and the similarity is approximately 67% and 65% for cPEPCK and mPEPCK. The structure of human cPEPCK is only available in PDB website. A superposition of the PEPCK structures of the pathogen (GDP-bound *Mtb*

PEPCK) and host (1KHB, human cPEPCK) shows that the overall fold of *Mtb* PEPCK is almost identical to that of the human enzyme. The root mean square displacement (r.m.s.d.) of the *Mtb* PEPCK is about 1.0 Å over 491 C^α atoms. The reported rat cytosolic PEPCK crystal structure (91% sequence identity with human cPEPCK) contains three flexible elements in the active site, the P-loop, Ω-loop, and R-loop, which are important components for enzyme catalysis¹⁵⁶. In our study, the apo-*Mtb* PEPCK structure with colored B-factors exhibits these three flexible loops in the active site; the P-loop (residues 269-277), Ω-loop (active site lid, residues 448-458), and R-loop (residues 79-86) (Figure 2-7).

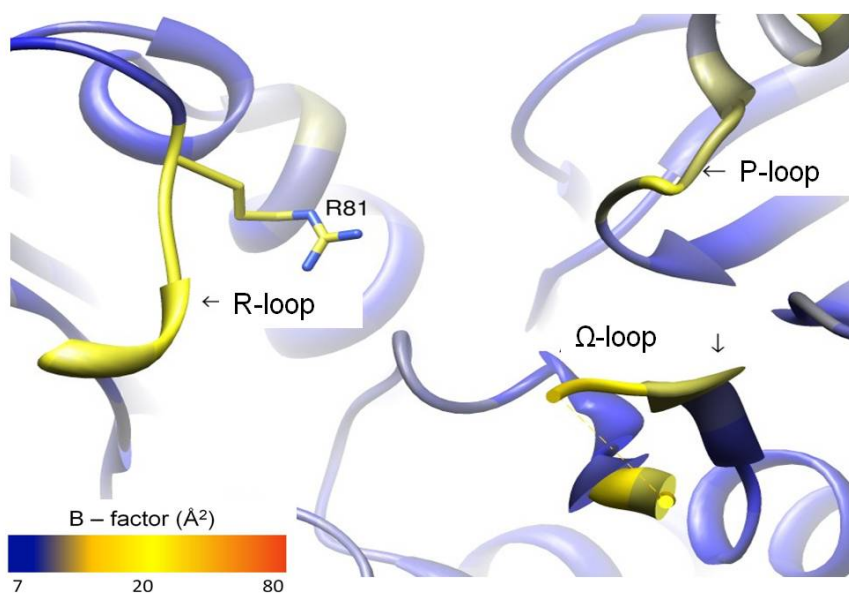


Figure 2-7. Active site in apo-*Mtb* PEPCK crystal structure, loops are colored by B-factor. Three flexible loops, P-loop, R-loop, and Ω-loop have relatively high level of B-factor.

The major structural differences between human and the *Mtb* enzyme occur in the two active site loops, the Ω -loop and R-loop. The Ω -loop in human cPEPCK was not visible in the electron density, indicating it was disordered in GTP analog complex form (PDB: 1KHE), whereas it forms ordered/closed conformation with full electron density in *Mtb* PEPCK-GDP bound structure (Figure 2-8). The R-loop flips over toward solvent area in *Mtb* PEPCK-GDP bound structure, however this loop was buried in human cPEPCK¹⁴⁰. The intriguing differences in these flexible loops may provide an opportunity for drug development targeting *Mtb* PEPCK, therefore we look into in depth of “cause of effect” on these loops movements in this chapter.

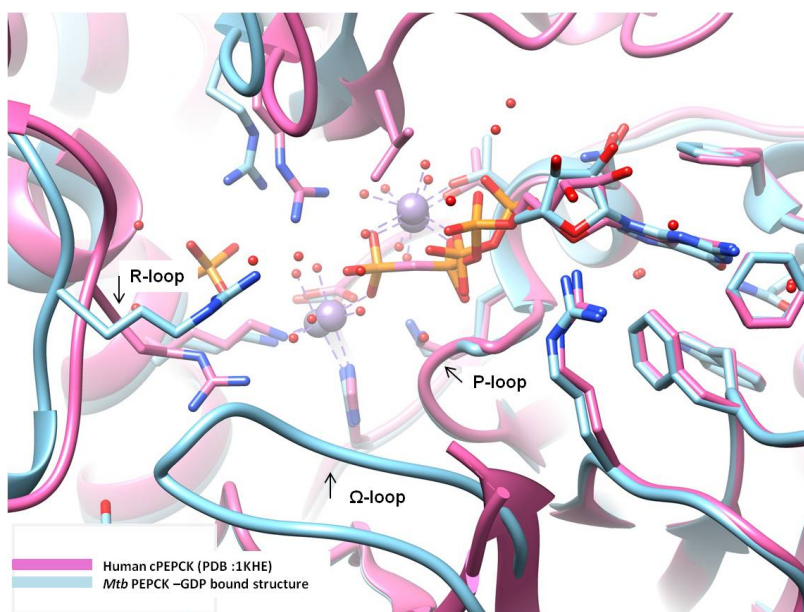


Figure 2-8. Superimposed crystal structure of human cPEPCK GTP-bound (PDB : 1KHE) and *Mtb* PEPCK-GDP bound forms. The R-loop and Ω -loop show different conformations. The Ω -loop in *Mtb* PEPCK adopts closed conformation, while its conformation in human cPEPCK is disordered. The R-loop in *Mtb* PEPCK stays way from the active site while it is packed inside of the active site in human cPEPCK.

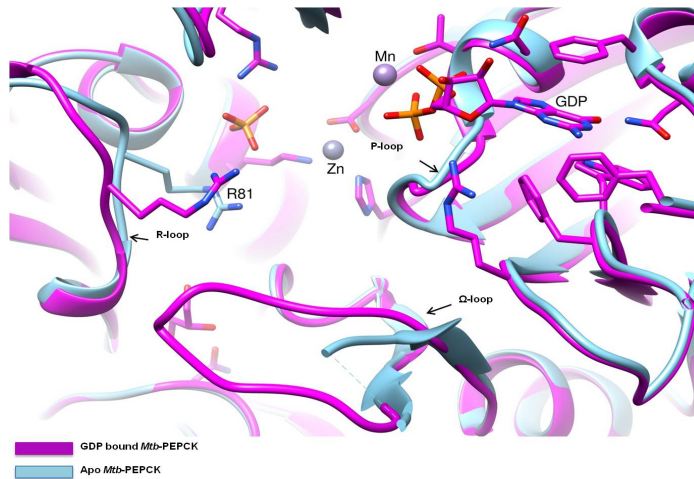


Figure 2-9. Superimposition of the apo and GDP-bound *Mtb* PEPCK structures. It represents the P-, Ω - (active site lid), and R-loops adopt different conformation upon GDP binding.

The Mtb PEPCK Ω -loop undergoes a conformational change upon GDP binding

Superimposition of the *Mtb* PEPCK apo and GDP-bound crystal structures reveals different conformational changes in the Ω -loop (residues 448-458) as shown Figure 2-9. Density for the Ω -loop was different between the apo and GDP-bound *Mtb* PEPCK structures. In the apo *Mtb* PEPCK structure, there was no density for residues 450-457 (part of the Ω -loop) meaning it was disordered, while in the GDP-complex *Mtb* PEPCK structure, density was fully observable for the entire loop and is in a closed position over the active site. In the GDP-bound structure, the GDP phosphate group binds to the P-loop (residues 269-276) and the GDP ribose binds to the guanidinium group of R420, which also links to the P-loop. Specifically, the amino -NH₂ group of the R420 side chain hydrogen bonds to the ribose ring oxygen (3.1 Å), and the H-N^ε of R420 hydrogen bonds to the backbone carbonyl oxygen of A272 (3.0 Å) from the P-

loop. Consequently, the P-loop moves 1.9 Å (measured from the C α of A272) to the other substrate (OAA, PEP) binding sites. This movement allows two additional residues (E447 and Q448) on the Ω -loop to form hydrogen bonds: the carbonyl oxygen atom of E447 side chain interacts with the backbone amide hydrogen from K422 (2.9 Å), the hydrogen atom in the –NH₂ group of the Q448 side chain forms a hydrogen bond with the backbone carbonyl oxygen atom of S271 (2.9 Å), and the carbonyl oxygen of the Q448 side chain hydrogen bonds with the P-loop via a hydrogen atom on the sulfhydryl group of the C273 side chain (3.1 Å) (Figure 2-10). In the apo structure, the side chains of E447 and Q448 instead protrude into the solvent.

Comparison of the apo and GDP-bound structures reveals that the P-loop movement upon GDP binding facilitates closure of the Ω -loop through hydrogen bonds between the P-loop and Ω -loop mediated by Q448 of the Ω -loop. The C273 residue is highly conserved in all PEPCKs and is known as a “hyper-reactive” residue, but the Q448 residue is not conserved in eukaryotic GTP-dependent PEPCKs including human cPEPCK, rat cPEPCK, and chicken mPEPCK (mitochondria), which instead contain an alanine in this position as shown in sequence alignments (Figure 2-11). Thus, the interaction between the P-loop and the Ω -loop cannot occur in other species, making it a unique feature of *Mtb* PEPCK.

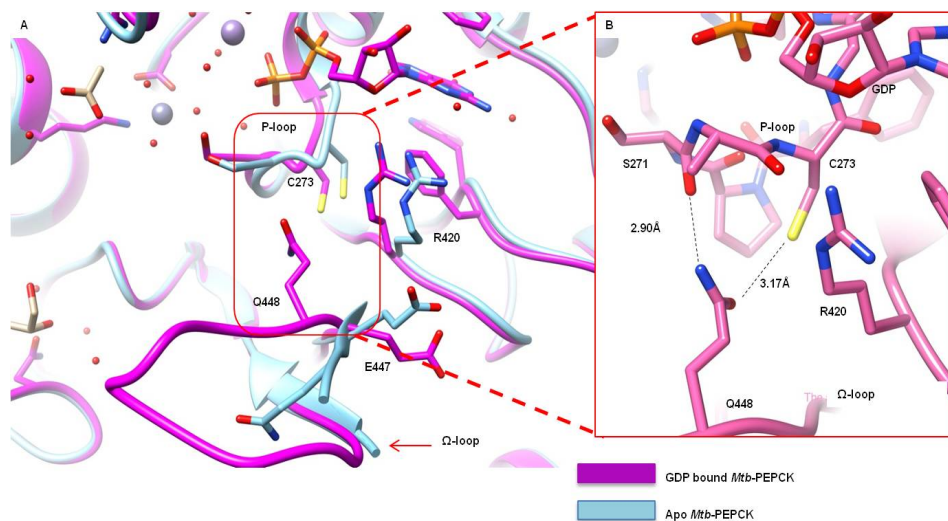


Figure 2-10. The interaction between Ω -loop and P-loop. (A) GDP-bound structure (magenta), apo structure (blue); (B) The interaction between Q448 in the closed Ω -loop and the backbone carbonyl of S271 (2.90 Å), and the thiol group of C273 (3.17 Å) from the P-loop.

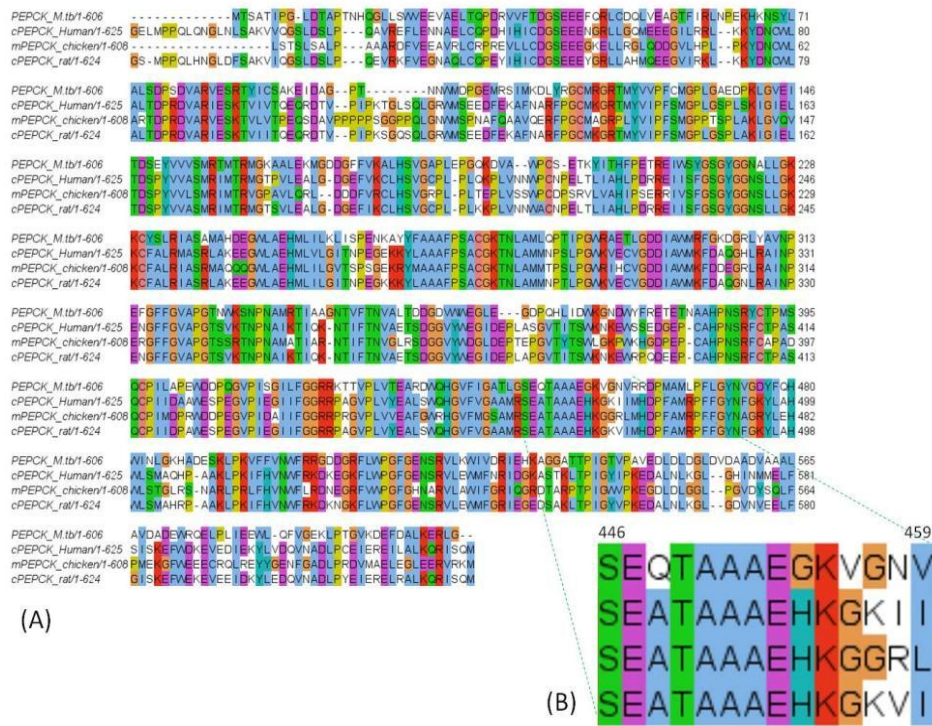


Figure 2-11. Multiple sequence alignment of GTP-dependent PEPCs color coded according to sequence similarity to *Mtb* PEPC. (A) Human cytosolic type (cPEPCK_human), chicken mitochondria (mPEPCK_chicken), and rat cytosolic type PEPC (cPEPCK_rat). (B) Sequence alignment of residues 446 to 459 in the Ω -loop (numbering in *Mtb* PEPC), shows that the Q448 residue in *Mtb* PEPC is not conserved in other GTP-dependent PEPCs.

Stability of the Ω -loop is achieved differently between the rat and *Mtb* PEPCs.

The *Mtb* PEPC stabilizes the closed Ω -loop conformation through the interactions between the P-loop and Ω -loop. However the rat cPEPCK crystal structure seems to pull the Ω -loop closer through interactions between the Ω -loop and R-loop. In the rat cPEPCK structure, there are hydrogen bonds between the two oxygens of the E469 side chain in the Ω -loop and two hydrogens from the backbone amides of E89 and S90 in the R-loop at distances of 2.7 Å and 3.1 Å, respectively. Another hydrogen bond occurs

between the N^{ε2} atom of H249 and a carbonyl oxygen of the E89 side chain (3.0 Å)¹⁵⁶ which stabilizes Ω-loop closure. This interaction between the Ω-loop and R-loop is absent in the *Mtb* PEPCK structure.

Both the R-loop and P-loop are directly involved in substrate binding in *Mtb* PEPCK, as the P-loop interacts with GTP or GDP while the R-loop binds OAA or PEP in the active site. The Ω-loop does not make any direct substrate contacts, but it has significant roles during catalysis through 1) placing substrates in the active site at a reasonable distance for catalysis to occur, 2) protecting intermediates from solvent¹⁵⁶. Therefore, the interaction between the Ω-loop and R-loop in rat cPEPCK likely makes Ω-loop closure dependent on substrate binding (PEP or OAA) in the active site, while the interaction between the Ω-loop and P-loop in *Mtb* PEPCK induces Ω-loop closure upon nucleotide binding alone. This implies that eukaryotic PEPCK activity relies on PEP or OAA concentration for catalysis whereas *Mtb* PEPCK activity may depend on nucleotide concentration.

Ω-loop closure supports the precise positioning of PEP in the Mtb PEPCK active site

In order to understand the importance of Ω-loop closure in *Mtb* PEPCK, we examined the substrate-complex structures of *Mtb* PEPCK. In the structure of *Mtb* PEPCK with PEP alone, the electron density for PEP was not in a position consistent with catalysis, instead, the PEP phosphate group overlapped with the catalytic metal binding site position despite the high amount of metal in the soaking condition. In figure 2-12, the PEP phosphate oxygen interacts with three metal binding residues, the N^{ε2} atom from H249 (2.9 Å), the N^ζ atom from K229 (3.0 Å), and an oxygen from the D296

side chain (2.7 Å). The PEP carboxyl group binds with the side chains of Asp462 and Arg460 through water mediated hydrogen bonds at distances of 2.4-2.8 Å (Figure 2-12B). In addition, the density for part of the Ω -loop, residues 450 to 457, was not visible in the electron density map and the Q448 residue pointed to the solvent area, suggesting that the Ω -loop was open.

Unlike the structure with PEP alone, the *Mtb* PEPCK PEP-GDP complex structure clearly shows PEP in the catalytic position, consistent with reported PEPCK structures (Figure 2-12C). The crystal structure of the PEP-GDP complex was obtained by soaking a mixture of PEP and GDP into the E83A mutant crystal. We used the E83A mutant crystal to prevent catalysis during soaking. In the PEP-GDP-bound structure, the PEP carboxylate group hydrogen bonds to the -NH on the backbone of G222 (3.0 Å) and R81 (3.0 Å), and the PEP phosphate interacts with the side chain of R81 and R389 (2.7-3.1 Å) with no density observed for the metal ions, indicating that metal coordination is not critical for PEP positioning in the catalytic site, and the binding of GDP was solely responsible for PEP repositioning. The GDP induced Ω -loop closure resulted in a 2.3 Å shift of the P-loop (measured from the C α of the A272) toward the catalytic metal binding site, making contacts between the sulfhydryl group of C273 on the P-loop and the carbonyl group from the side chain of Q448 on the Ω -loop (3.2 Å). This P-loop position overlaps the PEP binding position in the absence of GDP, consequently, the non-catalytic PEP binding site observed in the *Mtb* PEPCK PEP complex structure becomes unavailable after nucleotide binding. In the human cPEPCK PEP complex structure, PEP takes position in the catalytic site without requiring conformational

changes to the active site.

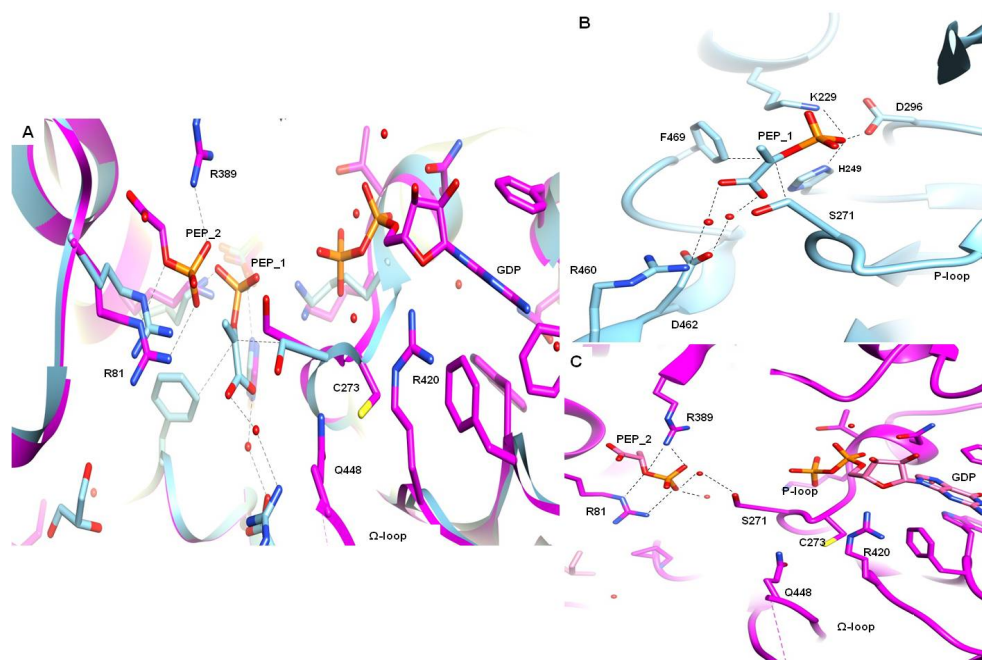


Figure 2-12. Different position of PEP in *Mtb* PEPCK structure. (A) Superimposed PEP complex structures, PEP_1 is misplaced vs PEP_2 correctly positioned. (B) The PEP complex crystal structure shows PEP misplaced in the absence of GDP [C] PEP binds in the correct catalytic position in the presence of GDP.

The structure of *Mtb* PEPCK in complex with oxalate (OAA analog) and Mn^{2+} shows that oxalate can be positioned in the catalytic site by directly coordinating to a metal (Figure 2-3) even in the absence of GTP and when the Ω -loop is disordered. The Mn^{2+} is coordinated by K229, H249, and D296. The two oxalate ion oxygens bind Mn^{2+} at distances of 2.4 Å and 2.6 Å. The negative charge of the oxalate oxygen atoms is stabilized by the side chains of R389 and R81. The binding mode of oxalate in *Mtb* PEPCK is identical to that reported in other GTP-PEPCK structures. OAA directly

coordinates with the divalent metal in the active site as three residues support metal interactions, indicating that the metal facilitates OAA positioning for catalysis. However, PEP does not directly coordinate with the catalytic metal, instead using metal outer sphere coordination to bind in the active site¹⁶⁰. In the human cPEPCK PEP complex structure, two of the water molecules mediate the interaction between the Mn²⁺ ion and the oxygen atoms of PEP's phosphate group¹⁴⁰. Therefore, structural information point out that the factors influencing substrate position depend on the direction of the reaction; we hypothesize that the divalent metal coordinates OAA positioning for the gluconeogenic reaction, while Ω -loop closure of in the active site are critical to correctly orient PEP for the anaplerotic reaction.

Loss of interaction between the P-loop and Ω -loop affects Ω -loop closure, decreasing the anaplerotic activity of *Mtb* PEPCK

To test our hypothesis that the position of the anaplerotic substrate (PEP) is significantly controlled by the conformation of the Ω -loop in *Mtb* PEPCK, we disrupted Ω -loop closure by removing the hydrogen bonding interaction between the P-loop and Ω -loop. The mutation of Q448 residue to an alanine causes loss of interaction between the P-loop and Ω -loop. Kinetic experiments were performed to determine the effect of the mutation in both directions of the gluconeogenic and anaplerotic reaction (Table 2-2).

Table 2-2. Kinetics data of the wild-type and the Q448A mutant for the gluconeogenic and anaplerotic reactions.

	V _{max} (μmol min ⁻¹ mg ⁻¹)	K _{cat} (s ⁻¹)
Gluconeogenic reaction (PEP-formation)		
Wild-type	9.2 ± 0.6	10.6 ± 0.7
Q448A	14.1 ± 0.5	16.2 ± 0.6
Anaplerotic reaction (OAA-formation)		
Wild-type	9.8 ± 0.8	11.3 ± 0.9
Q448A	2.8 ± 0.3	3.2 ± 0.3

By abolishing the hydrogen bond between the P-loop and the Ω-loop, the Q448A mutant demonstrated that Ω-loop closure has a distinct effect on the kinetics of the anaplerotic reaction. For the gluconeogenic reaction, the Q448A mutant results in an approximately 1.3-1.6 fold increase in the K_{cat} compared to wild-type PEPCK.

In the direction of the anaplerotic reaction, the Q448A mutant in *Mtb* PEPCK showed a significant three fold reduced K_{cat}, meaning that the wild-type enzyme is three times more efficient at catalyzing the anaplerotic reaction than the mutant. The kinetic data also demonstrates that Ω-loop closure is more important for PEP positioning than for OAA positioning as the mutant showed only a decrease in the rate of the anaplerotic reaction. The reduced turnover rate of the mutant shows the importance of the hydrogen bond between the P-loop and Ω-loop for Ω-loop closure and it suggests that the catalytic step of the Ω-loop closing is the rate-limiting step of the anaplerotic reaction.

There is a possibility that the Q448A mutant affects GDP binding affinity in direction of the OAA forming reaction. However, there is no difference between the K_d

value of the wild-type and the Q448A, indicating that the loss of GDP affinity (Table 2-3) is not the reason for the loss of enzyme activity in direction OAA formation.

Table 2-3. The Kd value (M) of GDP for WT (wild-type), Q448A mutant and human cPEPCK.

	WT	Q448A	Human cPEPCK
Kd	0.0015	0.0015	0.0028

Moreover our results are in agreement with a recent study that found that a C273S mutant displayed greater efficiency in anaplerotic activity¹⁴⁵. Our understanding is that the interaction between the Q448 residue and S273 in the C273S mutant is stronger than in the wild-type because serine makes a tighter hydrogen bond than cysteine resulting in a more efficient Ω -loop closure. Therefore, the interaction between the P-loop and the Ω -loop, unique to *Mtb* PEPCK, accelerates Ω -loop closure and is important for the anaplerotic reaction because it plays a role in effectively positioning PEP for catalysis.

The flexible loops play a key role in the anaplerotic reaction of Mtb PEPCK

The eukaryotic GTP-dependent PEPCK preferentially catalyzes the reaction only in the forward, PEP-forming reaction, due to the Km values of the forward reaction substrates, GTP and OAA, is in the range of the physiologic intracellular concentration. Also the presence of pyruvate carboxylase, which is better suited to OAA formation from pyruvate than PEPCK, so the role of PEPCK in reverse-reaction, OAA formation

has not been focused among eukaryotic PEPCKs¹⁶¹. Most bacteria and yeast are ATP-dependent enzymes, but the crystal structure of *Mtb* PEPCK reveals that it is a GTP-dependent PEPCK. The PEPCK gluconeogenesis reaction in *Mtb* is critical to utilize the macrophage's fatty acid carbon source for maintenance and persistence during infection¹²⁰. However, under slow growth conditions *Mtb* PEPCK has been observed catalyzing the anaplerotic reaction, converting PEP to OAA¹²⁶. Slow growth *in vitro* mimics the conditions of macrophage during the course of infection¹¹⁰. Therefore, the anaplerotic reaction of *Mtb* PEPCK might also play a key role in *Mtb* pathogenicity.

The conformational changes of the Ω -loop closure upon GDP binding differs from the host enzyme due to the unique interaction between the Ω -loop and P-loop in *Mtb* PEPCK. Disruption of that interaction resulted in a three-fold decrease rate of the anaplerotic reaction compared to the wild-type. This interaction is critical for the PEPCK anaplerotic reaction and is not present in other GTP-dependent PEPCKs, including the host enzyme, representing a unique feature of *Mtb* PEPCK. *Mtb* PEPCK improved anaplerotic reaction is the result of a one amino acid difference from the host, an important metabolic adaption in response to the host-driven environment. This structural and biological evidence provides better insight into the anaplerotic role of *Mtb* PEPCK in *Mtb* pathogenicity.

The flexibility of the R-loop (residues 79-86) is unique to Mtb PEPCK among wild-type PEPCK enzymes

The R-loop in *Mtb* PEPCK is more flexible than that in the human PEPCK due to differences in key interactions. In the apo *Mtb* PEPCK structure, the R-loop packs into

the active site by adopting a closed conformation, while in the GDP-bound structure it forms an open conformation that is exposed to solvent (Figure 2-13). The superimposed apo and GDP-bound structures reveal that the displacement of the R-loop is about 4.3 Å (measured from the C α of R81). This R-loop contains catalytic residue R81, which binds directly to the substrate (OAA or PEP), and the flexibility of this loop has never been observed indicating that the R-loop was not flexible in the reported structures of rat wild-type cPEPCK and human cPEPCK. However, in a rat cPEPCK mutant missing the Ω -loop, the R-loop moved away from the active site and adopted an open conformation. In the wild-type rat cPEPCK structure, H470 from the Ω -loop interacts via a hydrogen bond with E89 from the R-loop (3.0 Å), but this interaction becomes unavailable in the mutant rat cPEPCK structure. The absence of the interaction between the R-loop and Ω -loop makes the R-loop in the rat cPEPCK adopt an open conformation. Superimposing the human cPEPCK and *Mtb* PEPCK structures (Figure 2-13) reveals that G454 in *Mtb* PEPCK is in the equivalent position to H470 in the rat protein and is unable to sustain the same R-loop/ Ω -loop interaction. Furthermore, the salt bridge interaction between E89 and R483 (3.1 Å) in the human cPEPCK seems to determine the rigidity of the R-loop regardless of substrate presence. *Mtb* PEPCK has L467 instead of the human cPEPCK's Arg in the same position, preventing the formation of a salt bridge in *Mtb* PEPCK. Thus, structural evidence demonstrates that the R-loop in *Mtb* PEPCK is more flexible than its human PEPCK counterpart due to the absence of both the salt bridge and the R-loop/ Ω -loop interaction. Further detailed study is required to address about the importance of the R-loop's flexibility on catalysis.

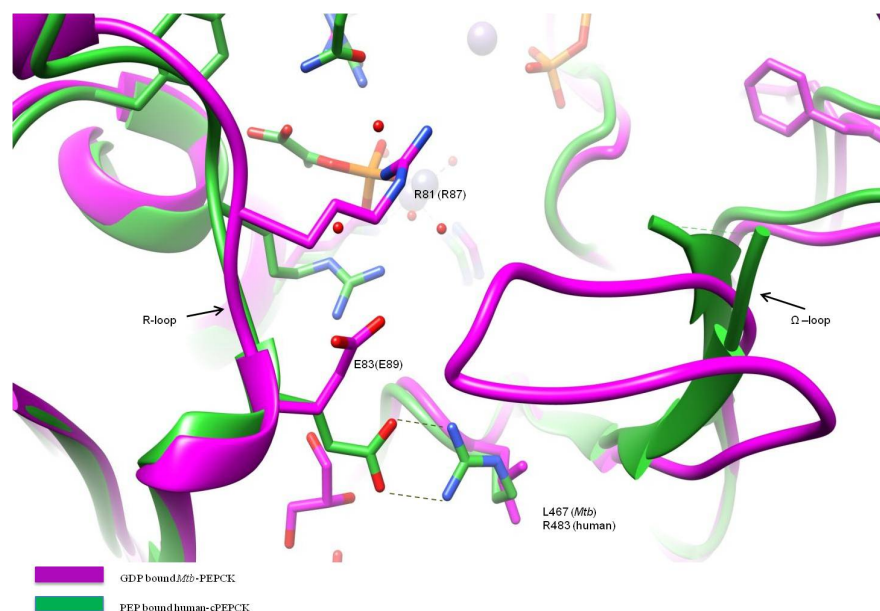


Figure 2-13. The interaction between the R-loop and Ω -loop in human cPEPCK. Superimposed GDP complex and *Mtb* PEPCK structure (violet) with human cPEPCK structure (green). The salt bridge interaction between E89 and R438 in human cPEPCK is missing in the *Mtb* PEPCK structure due to the equivalent L467. *Mtb* PEPCK has G454 instead of histidine, removing the interaction between the R-loop and Ω -loop.

Differences in the conformational changes of the R-loop in *Mtb* PEPCK compared to other GTP-dependent PEPCKs led us to investigate the catalytic role of the R-loop in *Mtb* PEPCK. Mutations of the *M. smegmatis* PEPCK, E83A, and D78A, reduced enzyme activity in both the OAA and PEP forming reactions¹⁵⁷. Dr. Sabine's group demonstrated the E83A and D78A mutants do not support *Mtb* growth on butyrate as a carbon source (Figure 2-1). The *Mtb* PEPCK E83A and D78A mutants showed almost no enzyme activity compared with the WT enzyme (Figure 2-14), while the D75A mutant retains the full activity of the wild-type PEPCK (data not shown). It is in agreement with previously published *M.smeg* mutants and Dr. Sabine's data

(unpublished). Intriguingly, the reported CD data from *M.smeg* mutants¹⁵⁷ indicates the protein structure of the D75A, D78A, and E83A mutants show no major differences compared to wild-type enzyme (Table 2-4). These results suggest that those residues are critical for *Mtb* PEPCK activity but not for the structure. However, those residues do not directly interact with substrates in the OAA/PEP binding site, begging the question of how they contribute to enzyme activity.

Table 2-4. Secondary structure data of wild-type and mutant *M.smegmatis* PEPCK enzymes¹⁵⁷.

Enzyme	α -Helix	β -Sheet	β -Turn	Other	NRMSD
WT	30	34	16	20	0.001
E83A	31	34	15	20	0.001
D78A	32	29	18	21	0.001
D75A	29	32	17	22	0.001

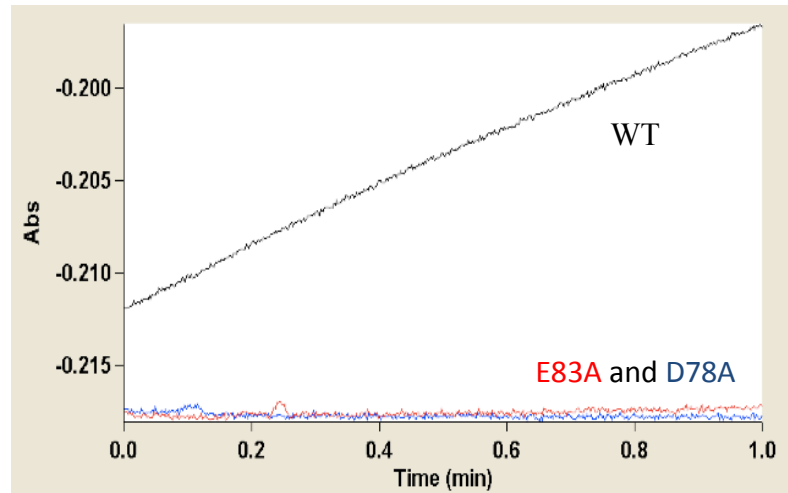


Figure 2-14. Activity of wild-type, E83A and D78A mutant of *Mtb* PEPCK. The increased absorbance (340 nm) over time represents the enzyme activity, E83A and D78A mutant shows no activity.

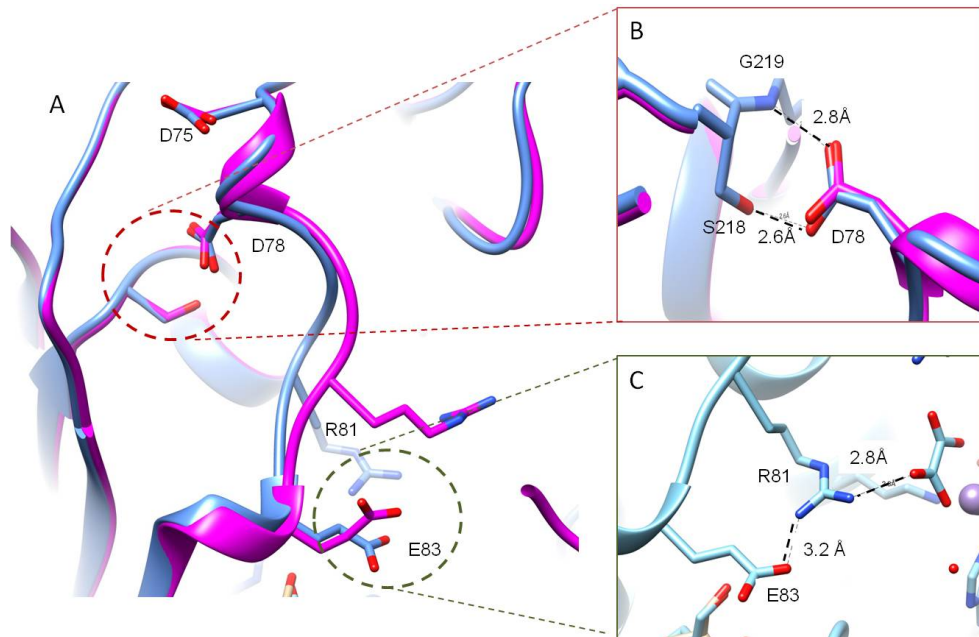


Figure 2-15. Interaction of D78 and E83 residues in R-loop with adjacent residues. Superimposed structures of apo- (blue) and GDP bound *Mtb* PEPCK (pink) shows different conformation of R-loop. The D78 and E83 residues contribute to the conformational maintenance.

The crystallographic evidence of the loop's flexibility made us postulate that stabilizing the mobile R81 loop helps the R81 residue bind substrates during the reaction. In Figure 2-15, the D75 residue is located in the surface area, pointing out toward solvent without interacting with the adjacent residues. However, the D78 residue has an intermolecular interaction with the S218 and G219 residues. The carbonyl oxygen of the D78 side chain functions as hydrogen bond acceptor interacting with the backbone amide of the G219 residue (distance 2.8 Å) and it also hydrogen bonds to the oxygen atom of the D78 side chain (distance 2.6 Å). Considering no significant structural changes on the D75A, D78A, and E83A mutants¹⁵⁷, the D78 residue may play a key role in maintaining R-loop conformation during catalysis, while the D75 residue plays no significant role.

We aimed to determine the effect of mutations on R81 loop conformation with crystal structures of mutants, E83A and D78A. Unfortunately, only E83A mutant was crystallized, while D78A mutant was aggregated during purification. The purified E83A *Mtb* PEPCCK was crystallized with co-purified GDP molecule as same as wild-type *Mtb* PEPCCK enzyme. This fact indicates that the loss of activity in the E83A mutant is not due to the loss of nucleotide binding and GDP bound E83A PEPCCK shows only minor differences from the WT structure. In GDP bound E83A *Mtb* PEPCCK structure, the electron density for the R81 side chain was not clearly visible. In contrast, the structure of the E83A mutant with the PO₄ ion bound to the hydrogen of HN^ε- R81 side chain (distance 2.9Å) shows full density for the R81 residue with an additional coordinated water molecule to -NH of R81 side chain (distance 2.9Å), which binds in place of the

oxygen of the E83 carboxylate (Figure 2-16). Comparison of known PEPCK structures with our ligand complex structures (PO₄ and oxalate) shows that the R81 side chain (R87 in human cPEPCK and rat cPEPCK) always adopts an identical rotameric conformation in the active site.

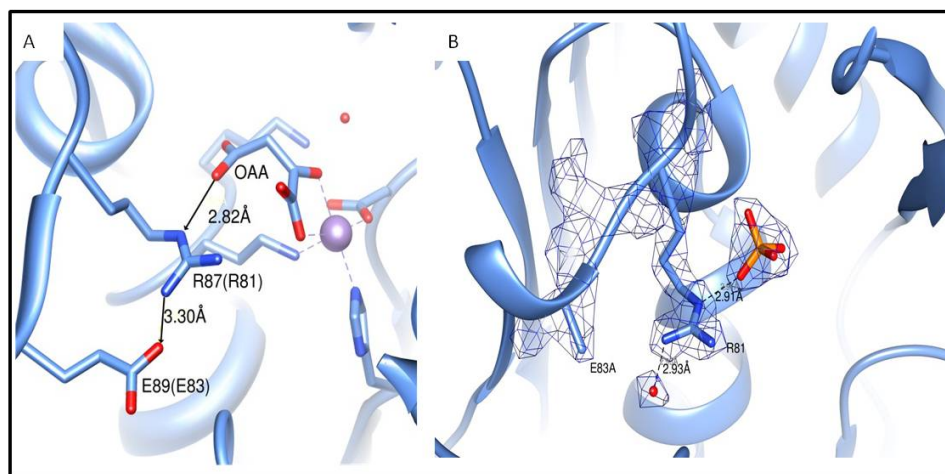


Figure 2-16. The interaction between E83 and R81 side chain maintains catalytic efficient rotameric conformation of R81 side chain in the active site. (A) Rat cytosolic PEPCK in complex with OAA (PDB entry 2QF1), (B) E83A mutant structure with full density for R81 in the presence of ligand and water.

In Figure 2-15C, the carboxyl oxygen of the E83 side chain accepts hydrogen of H-N1 of the R81 side chain to form a hydrogen bond (distance 3.2 Å), and N^ε-R81 is positioned for substrate binding, oxygen of oxalate ion hydrogen bonds to –NH of R81 side chain (distance 2.8 Å). However, the R81 residue forms differently when there is no interaction with E83 as shown in Figure 2-15A. In addition, the R87 residue (numbered 81 in *Mtb* PEPCK) in rat cPEPCK structure forms two main interactions with substrate

and E89 residue (numbered 83 in *Mtb* PEPCK). The oxygen atom of carboxyl group on oxaloacetate (OAA) accepts H-N^c of R87 side chain to form H-bond interaction (distance 2.8Å) and the carbonyl oxygen of E89 side chain also accepts hydrogen from –NH of R87 side chain at distance 3.3Å (Figure 2-16A). Therefore, the E83 residue is required to maintain the correct conformation of the R81 side chain for substrate binding and the flexibility feature of R-loop might make the role of E83 residue more important for efficient catalysis.

Glycerol binding pocket between the R-loop and Ω-loop of the Mtb PEPCK active site would be explored for the design of species-specific drug

The *Mtb* crystal structures contained a glycerol molecule (from the cryo-solution, 10% glycerol) bound close to the R-loop as shown in Figure 2-17. We found that water-mediated interaction between the -OH group of glycerol and the NH group of the R81 side chain in the R-loop at 3.0 Å and 2.9 Å. In addition, the glycerol -OH groups also H-bond to the backbone amide of P468 and N101, 2.9 Å and 2.8 Å distance, respectively. Also, the other -OH group of glycerol interacts H-bonding to carbonyl oxygen of N100 side chain at 3.0 Å.

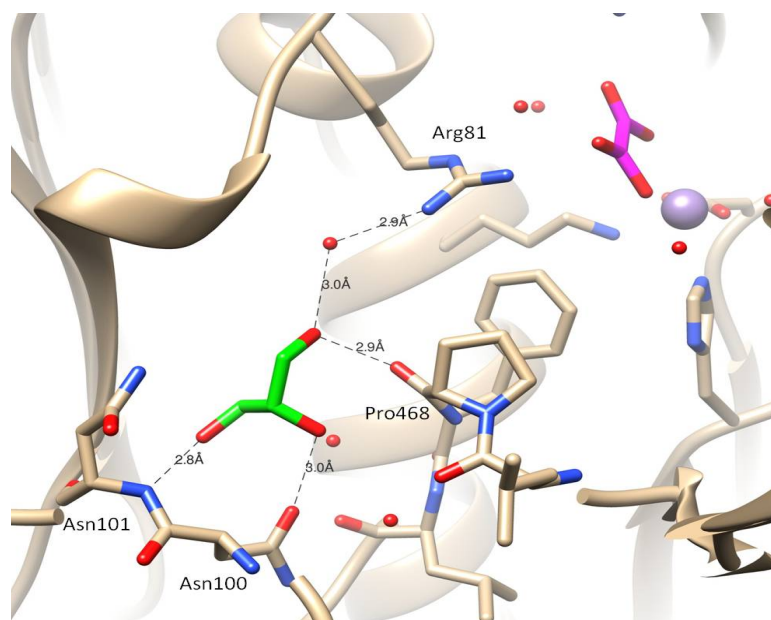


Figure 2-17. Glycerol binding pocket in *Mtb* PEPCCK structure.

At 20%, the concentration of glycerol did not inhibit enzyme activity during the enzyme assay. However, this novel glycerol binding site is not present in the human cPEPCCK crystal structure. The N100 in *Mtb* PEPCCK is substituted with G114 in human cPEPCCK, which is not able to H-bond to the -OH group of glycerol (Figure 2-18). This unique glycerol binding site in *Mtb* PEPCCK could be useful for novel and selective inhibitor design. For example, potent and selective inhibitors of *Mtb* dihydrofolate reductase (DHFR) were designed to take advantage of the glycerol binding pocket in the active site which was absent in human DHFR¹⁶². Therefore, the same logic applied in this situation could provide opportunities for developing new selective inhibitors of *Mtb* PEPCCK.

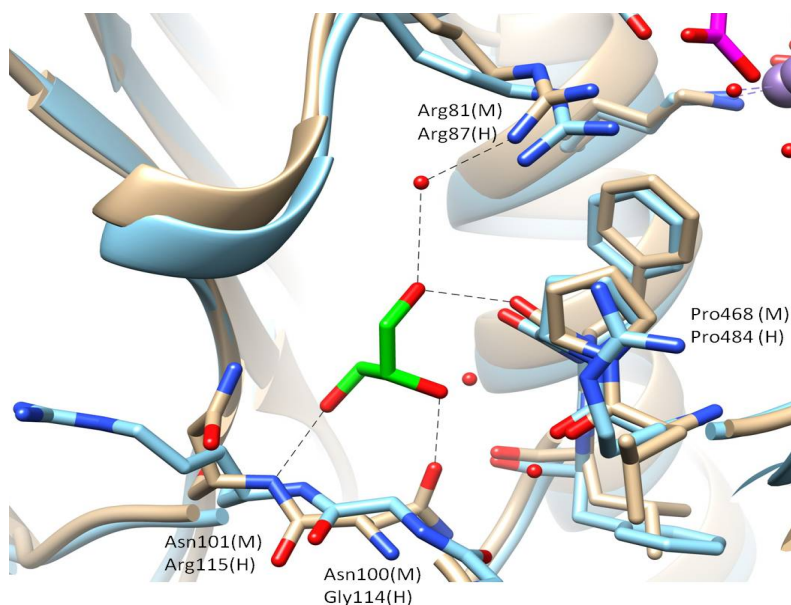


Figure 2-18. Comparison of a glycerol binding pocket of *Mtb* PEPCK and human cPEPCK. Glycerol (green), human cPEPCK (blue), *Mtb* PEPCK (grey).

Conclusion

We obtained and refined a substrate bound crystal of *Mtb* PEPCK at high resolution. We found mobile loops in the active site that show conformational changes upon substrate binding. Comparisons between *Mtb* PEPCK and other GTP-dependent PEPCKs reveal that the flexible loops in *Mtb* PEPCK behave differently because of the differences in the residues adjacent to those involved in loop interactions. Notably, the unique interaction between the P-loop and Ω -loop in *Mtb* PEPCK allows the Ω -loop to close upon GDP binding, which was not observed in other GTP-dependent PEPCKs. The Ω -loop closure helps position PEP in the catalytic site and is required for efficient turnover especially for the anaplerotic reaction. Moreover, the R-loop shows increased flexibility in *Mtb* PEPCK while it is restricted in other GTP-dependent PEPCKs. In

addition, a glycerol binding pocket was found in the *Mtb* PEPCK, which is absent in the human cPEPCK. Despite the conservation of the overall structure and active site between host and *Mtb* PEPCK enzyme, we explored structural differences in *Mtb* PEPCK, which provides an opportunity to develop selective inhibitors for targeting the pathogenic enzyme.

Experimental procedures

Cloning, protein expression, and purification of M. tuberculosis PEPCK (Rv0211)

The plasmid of *M. tuberculosis* PEPCK (*Rv0211*) was given by Sabin Ehrh's group. To overexpress PEPCK, *pckA* (1,838 bp) was amplified by PCR and cloned into the pET28a vector (Novagen) using the restriction sites NdeI and HindIII. Primer sequences are available upon request. The plasmid transformed into *Escherichia coli* BL21(DE3) (Novagen) for expression. The cells were grown to an OD₆₀₀ of 0.7-0.8 at 37 °C in Luria-Bertani medium containing 50 µg ml⁻¹ kanamycin. *Mtb* PEPCK expression was induced with 0.5 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG). Cell growth continued at 291 K for 18 h after IPTG induction, and the cells were harvested by centrifuging at 4200g for 15 min at 277 K. The cell pellet was suspended in ice cold lysis buffer (20 mM Tris-HCl pH 7.0, 500 mM sodium chloride, 5 mM imidazole and 1 mM phenylmethylsulfonyl fluoride (PMSF)) and the cells were lysed by French press. The cell debris was centrifuged at 36000 x g for 30 min at 277 K. Recombinant PEPCK with an N-terminal poly-His tag in the supernatant fraction was purified by nickel affinity chromatography (Ni-NTA resin (Qiagen)) followed by gel filtration (HiLoad XK

16 Superdex 200 prep-grade column (GE Healthcare)). It was then concentrated to 22 mg/ml in 20 mM HEPES (pH 7.0), 50 mM NaCl using a Spin-X UF concentrator (Corning). The protein concentration was estimated by measuring the absorbance at 280 nm, employing the calculated extinction coefficient of 26600 M⁻¹ cm⁻¹ (SWISS-PROT; <http://www.expasy.ch/>).

Site-directed mutagenesis of Q448A, R377A, a double mutant (Q448A, E447A) and D72A, D75A, E83A mutants

Site-directed mutagenesis of the *M. tuberculosis* PEPCK gene was performed with the QuikChange II Site-Directed Mutagenesis kit (Stratagene). Pfu Turbo DNA polymerase (Stratagene) was used for PCR reactions in the course of 30s at 95 °C; 1min at 58 °C; 7 min at 68 °C for 25 cycles. The pET28-pckA was used as template to construct the PEPCK single point mutants (Q448A, and R377A), while plasmid of Q448A mutant was used as template for double mutant (Q448A, E447A). PCR products were digested with Dpn I for 50 min at 37 °C to digest the DNA templates.

Subsequently, plasmids containing the desired mutation were transformed to into One shot[®] TOP10 cells (Invitrogen) for plasmid propagation and extraction, and then transformed into *Escherichia coli* BL21 (DE3) (Novagen) for expression. The sequence of mutant DNA was confirmed by determining the plasmid DNA sequencing at the Gene Technology Lab of Texas A&M University. For the Q448A mutant, the forward primer (5'- CTG GGT AGC GAG GCG ACC GCC GCG G-3') and reverse primer (5'- CCG CGG CGG TCG CCT CGC TAC CCA G-3') were used. For R377A, forward (5'- CAA CGA CTG GTA CTT CGC GGA GAC GGA AAC CAA TG -3'), and reverse (5'- CAT

TGG TTT CCG TCT CCG CGA AGT ACC AGT CGT TG-3') primers were used. For double mutant, Q448A and E447A, forward (5'- CCC TGG GTA GCG CAG CGA CCG CCG C -3') and reverse (5'- GCG GCG GTC GCT GCG CTA CCC AGG G-3') primers were used. The plasmids containing mutants, D72A, D75A, E83A, of the *M. tuberculosis* PEPCK was provided by Sabine's group. Expression and purification of the mutants was performed following the protocol used for the wild-type protein.

Enzyme assay of M. tuberculosis PEPCK (Rv0211)

The gluconeogenic reaction (PEP-formation)

The activity of *Mtb* PEPCK in direction of gluconeogenic reaction was determined using the pyruvate kinase and lactate dehydrogenase (PK/LDH) coupled enzyme assay (Figure 2-19). The PEP produced by PEPCK was taken by pyruvate kinase to produce pyruvate. The LDH enzyme converts pyruvate to lactate with oxidation NADH to NAD⁺. The reduction in absorbance at 340nm can be monitored as a rate of PEP formation by *Mtb* PEPCK. Use extinction coefficient for NADH ($\epsilon_{340} = 6220\text{M}^{-1}\text{cm}^{-1}$) to determine rate of PEP production.

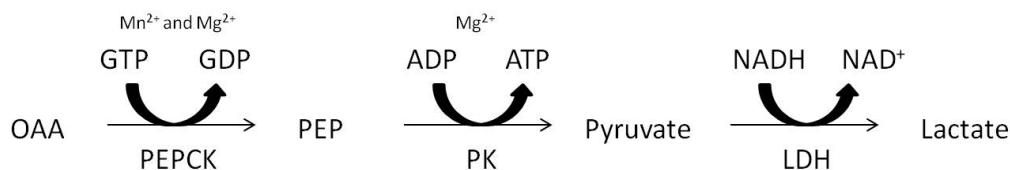


Figure 2-19. The coupled enzyme reactions (PK/LDH) used for the measurement of *Mtb* PEPCK enzyme activity.

The assay mixture contained 0.4 mM GTP, 4 mM MgCl₂, 0.5 mM MnCl₂, 1 mM DTT, 2 mM ADP, 6 unit PK/LDH enzymes, 0.25 mM NADH, and 2 μg/ml *Mtb* PEPCK enzyme in 20 mM Tris-HCl, pH 7.5, with 50 mM NaCl. The enzyme reaction was started by adding varied OAA concentration (0-1 mM). To determine K_{cat} and V_{max} in Table 2-2, GraphPad Prism version 6.00 (GraphPad Software, La Jolla California USA) were used.

The anaplerotic reaction (OAA-formation)

PEPCK converts PEP to OAA by fixing CO₂ in presence of GDP and metals. The produced OAA by PEPCK was consumed by MDH to generate L-malate in order to maintain its equilibrium (Figure 2-20). During this process, MDH converts reduced NADH to NAD⁺, thus enzyme activity can be monitored by detecting decreased NADH absorbance, spectrophotometrically, at 340 nm. The assay mixture contained 20 mM Tris-HCl, 50mM NaCl, 2 units of MDH, 2mM MgCl₂, 0.1 mM MnCl₂, 100 mM NaHCO₃, 5 mM DTT, and 0.25 mM NADH with varied GDP and PEP concentrations depending upon the experiment. To determine K_{cat} and V_{max} in Table 2-2, GraphPad Prism version 6.00 (GraphPad Software, La Jolla California USA) were used.

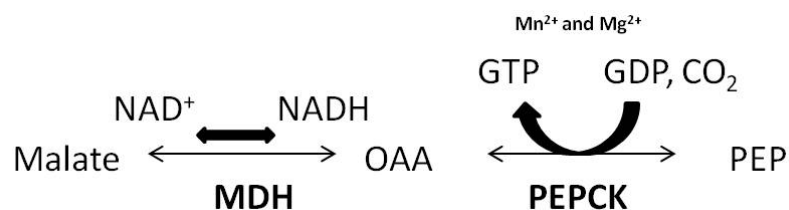


Figure 2-20. The enzyme reaction of coupled malate dehydrogenase (MDH) enzymes. It is used for detecting anaplerotic reaction (OAA-formation, CO₂ fixation) of *Mtb* PEPCK enzyme.

Metal-dependent activity assay

To detect metal activity of enzyme, the PEP forming activity of *Mtb* PEPCK was detected using the malate dehydrogenase (MDH) coupled enzyme (Figure 2-20) due to the activity of PK/LDH coupled enzyme was also metal dependent. The assay mixture contained 20 mM Tris-HCl buffer, 50 mM NaCl, 100 μM GTP, 1 mM malate, 1 mM NAD, and 6 units of MDH. The PEP forming activity of *Mtb* PEPCK was measured using MDH, followed by monitoring the increased NADH absorbance at 340nm. The reaction was started by adding metal for the pH-dependent experiment at room temperature. The effects of Zn²⁺ and Mn²⁺ on the kinetic parameters of *Mtb* PEPCK was performed an activity comparison at 62.5 μM of the metal over a pH range of 5.8 - 8.5.

Crystallization and data collection of M. tuberculosis PEPCK

Mtb PEPCK was crystallized by hanging drop vapor diffusion at 16 °C. Crystallization was achieved in a 4 μl drop consisting of 2 μl well solution [10% -16% (w/v) PEG 3350, 0.2 M KH₂PO₄] and 2 μl of protein. PEPCK in complex with co-purified GDP grew within 24 h. The apo PEPCK was formed after one month in the same drop. Crystals of *Mtb* PEPCK in complex with oxalate, non-hydrolyzable GTP

analog (Guanosine 5'-[β,γ -imido]triphosphate trisodium salt hydrate), and PEP, were obtained by soaking the apo crystal in 2 mM oxalate or PEP in the presence of 5 mM MnCl₂. We were unable to obtain a wild-type PEPCK complex with PEP-GDP because the compound turns over in the active site. The crystallization trial of mutant Q448A was unsuccessful. The mutant E83A PEPCK in complex with co-purified GDP was grown under the same conditions as WT PEPCK, and mutant E83A PEPCK in complex with PEP-GDP was obtained by soaking with the apo E83A crystal in 0.5 mM PEP in the presence of 5 mM MnCl₂ and 0.5 mM GDP. Crystals were cryo-protected by dipping them into a solution [10% (w/v) PEG3350 and 20% glycerol (v/v)]. Diffraction data were collected at a wavelength of 1.542 Å using a Rigaku R-Axix IV ++ at home source and at the Advanced Photon Source beamline 23-ID using a MAR 300 CCD detector (MarMosaic from Marresearch Charged Coupled Device) (Rayonix LLC, Evanston, IL). The HKL2000 program package (HKL Research, Inc., Charlottesville, VA) was used for integration and scaling.

Structure determination and refinement of *M. tuberculosis* PEPCK

The molecular replacement program PHASER (University of Cambridge, Cambridge, U.K.) was used to determine the structure of *Mtb* PEPCK using the crystal structure of human cPEPCK (PDB entry 1KHG) as a search model with water and ions removed. The GDP-bound *Mtb* PEPCK structure represented C2221 space group with a=103.32, b=125.17, c=122.15, α =90.00, β =90.00, and γ =90.00, while apo-*Mtb* PEPCK structure showed C2 space group with a=103.24, b=122.43, c=64.38, α =90.00, β =116.89, and γ =90.00. Except GDP-bound *Mtb* PEPCK structure, all other structures

are scaled into C2 space group. Therefore, the apo structure of *Mtb* PEPCK used as a search model for substrates-bound structures and E83A mutant structure. Rigid body refinement, followed by simulated annealing at 5000 K, was conducted using PHENIX (Python-based Hierarchical Environment for Integrated Xtallography). Subsequently, refinement was conducted in alternating cycles of manual model building in COOT (Crystallographic Object Oriented Toolkit), followed by refinement in PHENIX until the R factors converged. Data collection and refinement statistics for *Mtb* PEPCK in complex with substrates were shown in Table 2-5.

Anomalous difference scattering experiment of M. tuberculosis metal co-purified PEPCK

The *Mtb* PEPCK crystal with co-purified metals was collected at 1.542 Å using Rigaku the R-AXIS IV++ home source; the same crystal was collected again at the wavelength of 0.9199 Å at the Advanced Photon Source. The data were processed in HKL2000 and the Sfall program in CCP4 was used to calculate Fcalc and PHIcalc. The anomalous difference maps were created using the program FFT (Fast Fourier Transform) in CCP4.

ICP-MS analysis of M. tuberculosis metal co-purified PEPCK

Mtb PEPCK co-purified with metals was used for trace metal analysis. Protein samples were digested in 1-4% HNO₃, 20 mM HEPES pH 7.0, 50 mM NaCl at 70 °C for 24 h and then diluted with deionized water. The trace metals in the sample were determined using an inductively coupled plasma mass spectrometer (ICP-MS; ELAN 6100 DRC-2 Perkin-Elmer, Concord, Ontario, Canada), a Meinhardt microcentric

nebulizer, and a quartz cyclonic spray chamber. Rhodium (Rh) 103 was used as an internal standard with external calibration.

Table 2-5. Data collection and refinement statistics for *Mtb* PEPCK in apo-form and complex with substrates.

	<i>Mtb</i> PEPCK- apo	<i>Mtb</i> PEPCK- oxalate-Mn ²⁺	<i>Mtb</i> PEPCK- nhGTP	<i>Mtb</i> PEPCK- Zn-Mn-GDP	<i>Mtb</i> PEPCK- PEP	E83A <i>Mtb</i> PEPCK-PEP- GDP
PDB ID	4WIE	4WIU	4WL8	4WOU	4WPT	4WPU
Wavelength (Å)	1.54	1.54	1.00	0.92	0.95	1.54
Resolution range (Å)	50 - 2.18 (2.22 - 2.18)	50 - 2.02 (2.05 - 2.02)	50 - 1.61 (1.64 - 1.61)	50 - 2.12 (2.16 - 2.12)	50 - 1.13 (1.15 - 1.13)	50 - 2.26 (2.3 - 2.26)
Space group	C 1 2 1	C 1 2 1	C 1 2 1	C 2 2 2 1	C 1 2 1	C 1 2 1
Unit cell	103.958 122.564 64.555 90 117 90	103.969 122.739 64.697 90 116.98 90	105.616 123.009 64.035 90 116.97 90	103.316 125.172 122.147 90 90 90	103.574 122.813 64.556 90 116.66 90	104.636 122.715 63.861 90 116.61 90
Total reflections	157035	173856	368099	382393	708883	123037
Unique reflections	37534	45080	92528	44382	90212	33627
Multiplicity	4.2 (4.1)	3.9 (3.9)	4.0 (2.3)	8.5 (7.0)	1.9 (1.1)	3.7 (3.2)
Completeness (%)	99.98 (99.76)	94.98 (90.31)	98.02 (83.54)	98.47 (93.61)	95.21 (96.46)	99.61 (96.44)
Mean I/sigma(I)	24.8 (5.05)	35.3 (7.6)	21.0 (1.2)	27.8 (3.4)	6.51 (2.67)	17.4 (2.2)
Wilson B- factor	24.9	23.82	20.28	35.48	16.89	33.16
R-merge	0.146 (0.912)	0.107 (0.215)	0.128 (0.77)	0.11 (0.708)	0.087 (0)	0.121 (0.508)
R-work	0.1350 (0.1613)	0.1551 (0.1809)	0.1642 (0.2680)	0.1767 (0.2643)	0.1692 (0.2198)	0.1670 (0.2008)
R-free	0.1888 (0.2472)	0.1953 (0.2243)	0.1869 (0.3058)	0.2107 (0.3064)	0.1946 (0.2735)	0.2231 (0.2803)
Number of atoms	5409	5377	5415	5014	5572	5142
macromolecul es	4791	4787	4770	4657	4787	4769
ligands	7	13	39	41	21	44
water	611	577	606	316	764	329
Protein residues	600	600	599	595	600	599
RMS(bonds)	0.007	0.008	0.008	0.007	0.007	0.008
RMS(angles)	1.04	1.02	1.16	1.11	1.12	1.14
Ramachandran favored (%)	98	97	99	97	98	97
Ramachandran outliers (%)	0	0	0	0.17	0	0.33
Clashscore	2.45	1.28	2.44	3.26	3.18	4.18
Average B- factor	24.6	24.7	23.1	41.6	20.10	34.3
macromolecul es	23.7	23.8	21.8	41.5	18.40	33.9
ligands	28.5	35.3	30.6	37.4	26.50	39.2
solvent	32.1	31.5	33	44.2	30.60	39.3

Additional results

Pyruvate-kinase activity of Mtb PEPCK

PEPCK catalyzes the decarboxylation of oxaloacetate for pyruvate formation in the presence of nucleoside diphosphate. A study of the rat liver PEPCK revealed that Inosine 5'-diphosphate (IDP) and a divalent metal ion are required for the formation of pyruvate from oxaloacetate^{148,163,164}. As shown in Figure 2-21, pyruvate-kinase activity in *Mtb* PEPCK shows better activity (2.5 fold) in the presence of GDP than without GDP. The apparent K_m of OAA in the pyruvate forming reaction is 376 μM in the absence of GDP and 168 μM in the presence of GDP. This activity was observed a decade ago, however, the reason for the disparate activity was unknown. In our structural study of *Mtb* PEPCK, we observed conformational changes of the Ω -loop upon GDP binding; it closed in presence of GDP. Therefore, Ω -loop closure may participate in catalysis of pyruvate production.

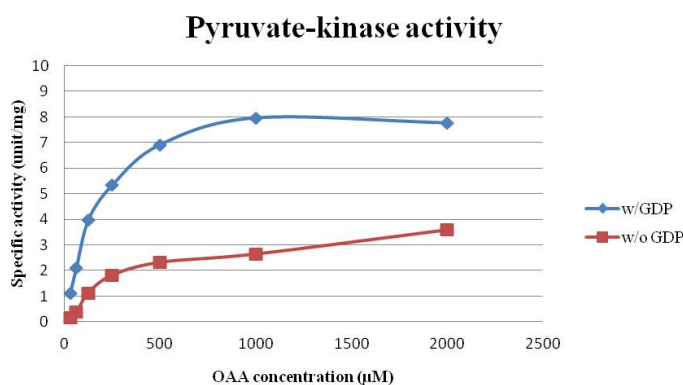


Figure 2-21. Comparison pyruvate-kinase activity of *Mtb* PEPCK in absence and presence of GDP molecule.

Crystal contacts in the Mtb PEPCK structure

The GDP molecules seem to make crystal packing contact that block the active site of *Mtb* PEPCK. The guanidinium group of the R377 side chain on the neighboring molecule forms electrostatic interaction with the phosphate groups of the GDP molecule (distance 3.0 Å), as shown in Figure 2-22. To discriminate between biologically relevant interfaces and artificial crystal packing contacts in the crystal structures, the *Mtb* PEPCK enzyme sample after incubation with the GDP molecule ran the gel-filtration, however the peak was correlated to a monomeric sized peak. Thus, the contacts the between two monomers in the crystal structure are symmetry related crystal packing contacts.

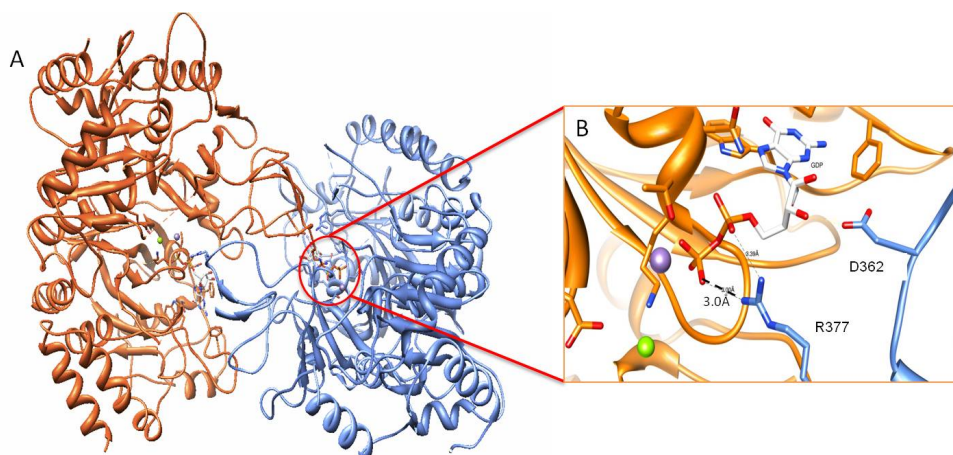


Figure 2-22. Crystal contacts in *Mtb* PEPCK crystal structure. (A) Two monomeric molecules in crystal structure, and (B) GDP molecule is located between two molecules.

CHAPTER III
STRUCTURAL ANALYSIS OF PEPCK INHIBITOR COMPLEX AND
DEVELOPMENT OF SELECTIVE INHIBITORS

Introduction

The *Mtb* PEPCK enzyme is essential for *Mtb* growth on fatty acids in vitro⁴³ and for TB persistence during infection in mice²⁵ making it an attractive target for anti-TB drug development. Despite the conservation of the overall structure and the highly conserved active site, there are structural differences that provide opportunities to develop selective inhibitors targeting *Mtb* PEPCKs over the host enzyme in CHAPTER II.

3-mercaptopicolinic acid (3-MPA) is the classic reversible PEPCK inhibitor, known to inhibit gluconeogenesis by blocking the PEPCK enzyme in the rat and guinea-pig⁶⁴. Its IC₅₀ against human cPEPCK was reported in the 20 μM range⁶⁵. Recently, the 3-MPA binding site was found in a substrate binding site (PEP/OAA) and also in an allosteric site near the nucleotide binding site⁶⁶. Moreover, a variety of substrate analogues (OAA/PEP) exhibited poor enzyme inhibition even though they bind tightly to the substrate-binding site⁶⁷.

Mammalian GTP-dependent PEPCK has a regulatory role in hepatic gluconeogenesis, making it a target for type 2 diabetes drug development. Transcription of the PEPCK gene is inhibited by insulin, resulting in a decreased rate of gluconeogenesis in the liver and in rat hepatoma cells⁶⁸. Overexpressed PEPCK

contributes to an increase in the rate of gluconeogenesis and the development of fasting hyperglycemia in transgenic mice⁶⁹, and PEPCK siRNA treatment of diabetic mice decreased the quantity of hepatic PEPCK enzyme and eliminated hyperglycemia. These data demonstrate the involvement of the PEPCK enzyme in the development of hyperglycemia, and suggest that inhibiting the human cPEPCK enzyme should decrease blood glucose levels in type 2 diabetic patients⁷⁰. Therefore, numerous efforts to find PEPCK inhibitors have been made to treat type 2 diabetes.

A variety of reversible GTP-competitive inhibitors were designed and developed by the Roche research center based on the human cPEPCK nucleotide binding pocket. The first scaffold developed from an HTS hit is 1-allyl-3-butyl-8-methylxanthine ($IC_{50} = 225 \pm 25 \mu M$).

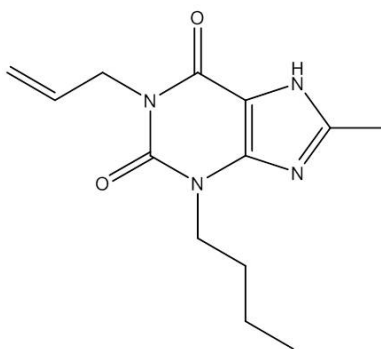


Figure 3-1. The first scaffold from HTS hit is 1-allyl-3-butyl-8-methylxanthine ($IC_{50} = 225 \pm 25 \mu M$) against human cPEPCK.

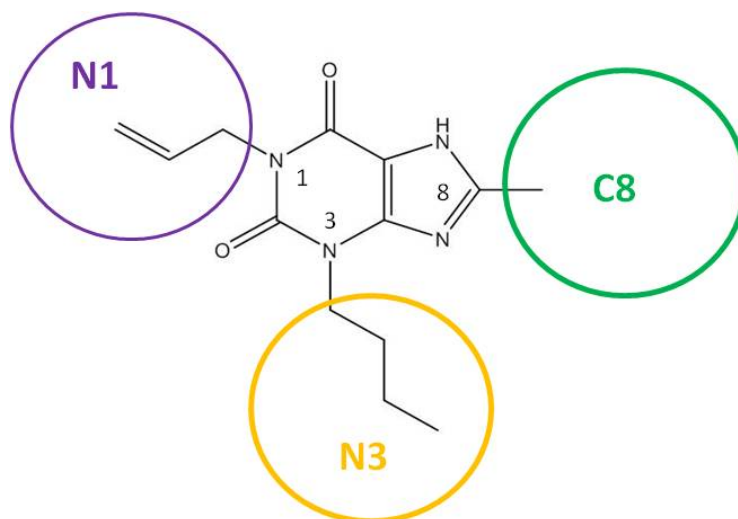


Figure 3-2. The modifications of Roche compounds targeting the human cPEPCK. Based on a xanthine ring, the series of compounds contain diversity on N1, N3, and C8.

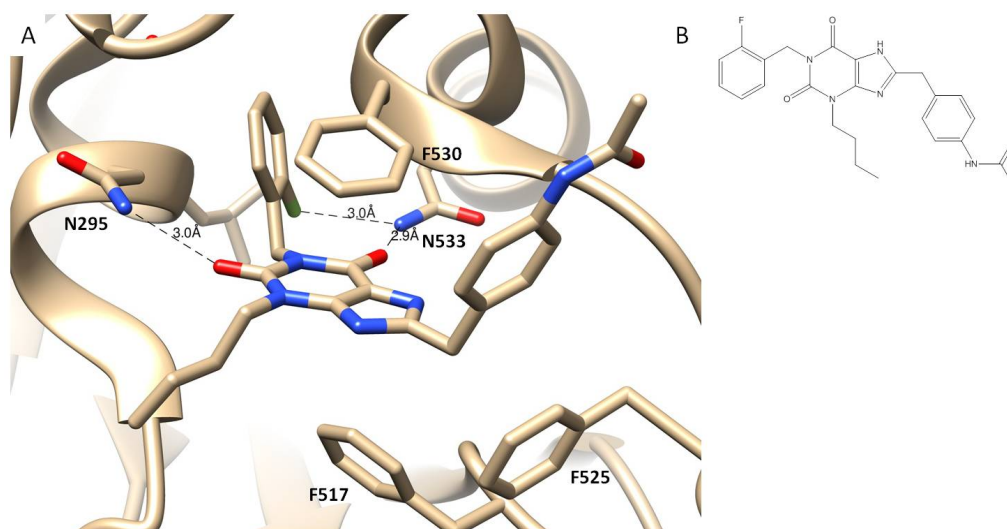


Figure 3-3. The binding mode of a GTP-competitive PEPCK inhibitor. (A) (N-{4-[1-(2-FLUOROBENZYL)-3-BUTYL-2,6-DIOXO-2,3,6,7-TETRAHYDRO-1H-PURIN-8-YLMETHYL]-PHENYL}-ACETAMIDE or Axon1165) in human cPEPCK (PDB code: 1NHX), and (B) chemical structure of Axon1165.

Roche made a variety of xanthine analogues with modifications on the N1, N3, and C8 positions (Figure 3-2). Potency from the initial HTS hit was improved about 100 fold ($IC_{50} = 2.11 \pm 0.60 \mu\text{M}$) through the substituent of 2-fluorobenzyl at N1 and N-Acetyl-4-aminobenzyl at C8 (chemical structure shown in Figure 3-3). This compound, Axon1165, is commercially available from the Axon Medchem Company. Figure 3-3A shows how the human PEPCK-Axon1165 structure (PDB: 1NHX) explains the interaction mode of this compound. The xanthine ring makes a π -stacking interaction with F530 at 3.4 Å and with F517 at 3.3 Å in a sandwich configuration, and there is another π - π interaction between the benzyl ring on the C-8 xanthine ring with F525 in a parallel offset configuration. The favorable edge-to-face interaction between F530 and the phenyl ring of the C-8 benzyl unit is also seen in the X-ray structure. The C-2 carbonyl of the xanthine ring hydrogen bonds to the amide of the N295 side chain at a distance of 3.0 Å. The 2-fluorobenzyl group makes π -interactions with the peptide bond between N295 and L293 and the fluoro group hydrogen bonds to the N533 side chain amide at 3.1 Å. In this chapter, the structural differences between the complex structures of Axon1165 with human and *Mtb* PEPCK are addressed in detail and used for the identification of a lead compound that shows selective inhibition against the pathogenic enzyme. In addition, the Roche GTP-competitive human cPEPCK inhibitor library (xanthine analogues) was intensively tested against *Mtb* PEPCK to find potent and selective inhibitors.

In addition, to find new molecule scaffolds that inhibit the *Mtb* PEPCK enzyme, 800 fragments (low molecular weight, <300 Da, compounds) and SRI hit 1120 anti-

tuberculosis active compounds ⁷¹ were screened against the *Mtb* PEPCK enzyme using affinity testing with differential scanning fluorimetry (DSF) as initial screening method. Also, medically relevant compounds were screened against *Mtb* PEPCK to search for molecular scaffolds via enzyme inhibition assay against the *Mtb* PEPCK enzyme. For this, 4700 anti-tuberculosis active compounds were selected for inhibiting *Mtb* growth on acetate, as a sole carbon source (from among 51K compounds).

Result and discussion

Crystal structure of Mtb PEPCK bound to Axon1165 (GTP-competitive inhibitors for human cPEPCK) and comparison with the human cPEPCK structure (PDB:1NHX)

Axon1165 (CAS_628279-07-2) is a GTP-competitive human cPEPCK inhibitor that is commercially available. The complex structure of human cPEPCK with Axon1165 is available in the PDB (PDB:1NHX) and its IC₅₀ against human cPEPCK is $2.11 \pm 0.60 \mu\text{M}$ ⁶⁵. The interaction mode is as described in the beginning of this chapter. This compound has an IC₅₀ of $3.54 \mu\text{M}$ against *Mtb* PEPCK and we successfully created a complex structure of *Mtb* PEPCK with Axon1165. The information from this structure allows us to investigate the difference in the binding interactions between the human and *Mtb* PEPCK enzymes.

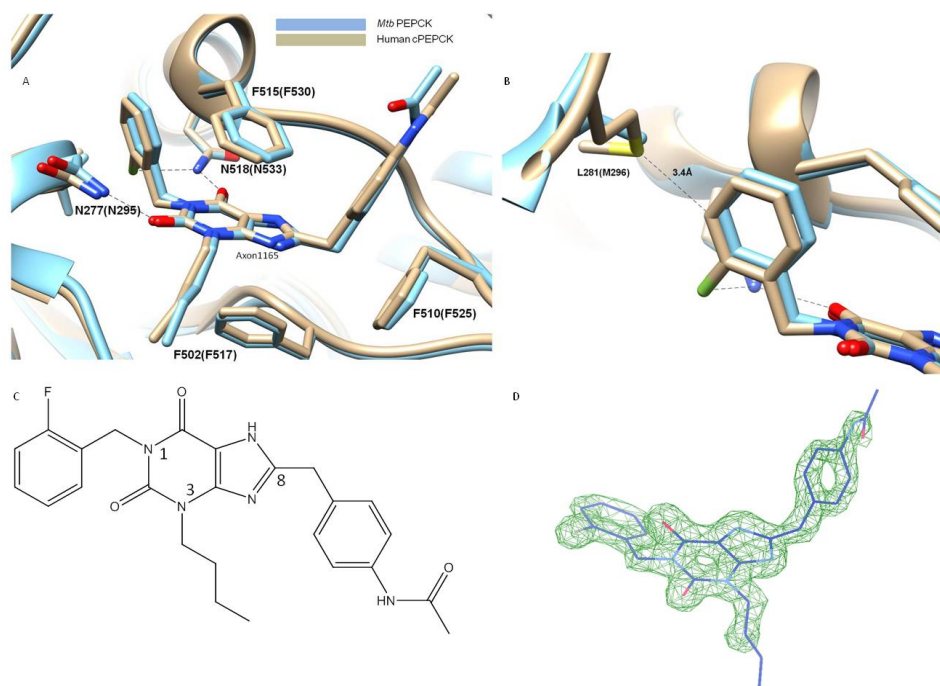


Figure 3-4. The binding site of Axon1165 in *Mtb* PEPCK. (A) Superimposed complex structure of Axon1165 in *Mtb* PEPCK (blue) and human cPEPCK (grey) (PDB:1NHX), (B) different residues placed near the Axon1165 binding area between *Mtb* and human PEPCK enzyme, (C) the structure of Axon1165, and (D) the Fo-Fc omit map contoured at 5.0 σ of the Axon1165 ligand in the *Mtb* PEPCK structure.

In Figure 3-4, a superimposed complex structure of the *Mtb* PEPCK and human Axon1165 structures reveals that the binding mode of Axon1165 is highly conserved between two structures. In the *Mtb* PEPCK-Axon1165 structure, the xanthine ring makes a π -stacking interaction with the benzyl ring of the F515 side chain at 3.5 Å and with that of the F502 residue at 3.7 Å. There is another π - π interaction between the benzyl ring on the C-8 xanthine ring and the ring of F510 side chain in a parallel offset configuration. The favorable edge-to-face interaction between the benzyl ring of the F515 side chain and the phenyl ring of the C-8 benzyl unit is also seen in the X-ray

structure. The C-2 carbonyl of the xanthine ring accepts a hydrogen from the amide group of the N277 side chain at a distance of 2.9 Å. The 2-fluorobenzyl ring makes π -interactions with the peptide bond between the N277 and L278 residue and the fluoro atom accepts hydrogen from the N518 side chain amide for a H-bonding interaction (distance 2.9 Å). These interactions are identical to those found in the human cPEPCK structure.

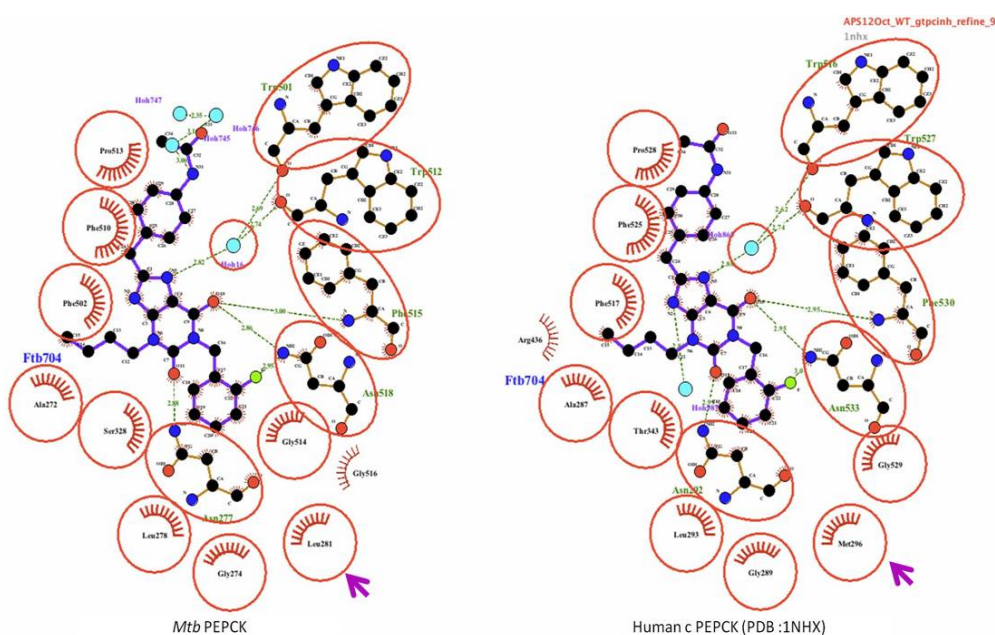


Figure 3-5. Comparison of binding interaction with Axon1165 in *Mtb* PEPCK and human cPEPCK (PDB ID: 1NHX). This figure was generated by using Ligplot program. The different residues are pointed with arrows.

Despite the highly conserved binding interactions, there are two interesting differences in the xanthine analog binding area which could be exploited for the design of selective *Mtb* PEPCK inhibitors. First, the 2-fluorobenzyl ring of Axon1165 is 3.4 Å

from M296 in the human cPEPCK (Figure 3-4 B, Figure 3-5); the equivalent residue in the *Mtb* PEPCK complex structure is L281. Both L281 in *Mtb* PEPCK and M296 in human cPEPCK are hydrophobic residues, but modification of the fluorobenzyl group of Axon1165 could extend near the L281 residue and create potent and selective *Mtb* PEPCK inhibitors.

Another important structural difference between the *Mtb* and human PEPCK occurs close to the acetamide of the C8 benzyl ring, in an area denoted the small binding pocket (pocket size is about 10 Å). As shown in Figure 3-6, the electrostatic surface model of the *Mtb* PEPCK represents a number of polar residues in this small pocket and is sufficiently large to accommodate the larger inhibitors that are extended from the C8 benzyl ring. However, this pocket in human cPEPCK structure contains fewer polar residues and is smaller than that of *Mtb* PEPCK.

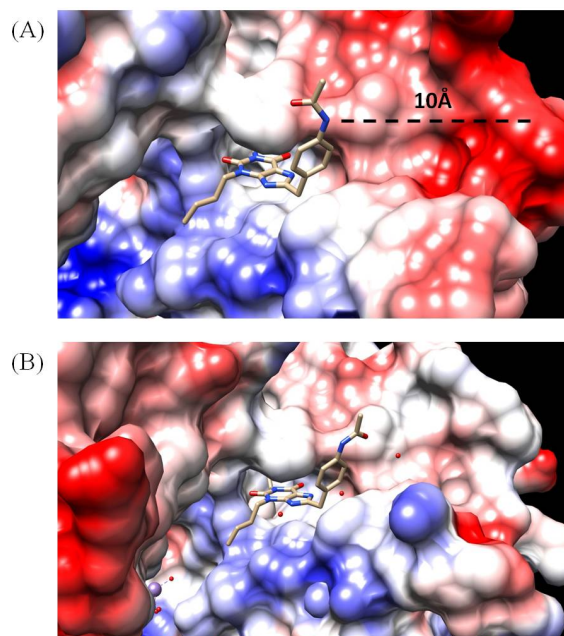


Figure 3-6. Comparison of the electrostatic surface of the small pocket near the inhibitor binding site. (A) *Mtb* PEPCK structure in complex with Axon1165, (B) human cPEPCK (PDB:1NHX).

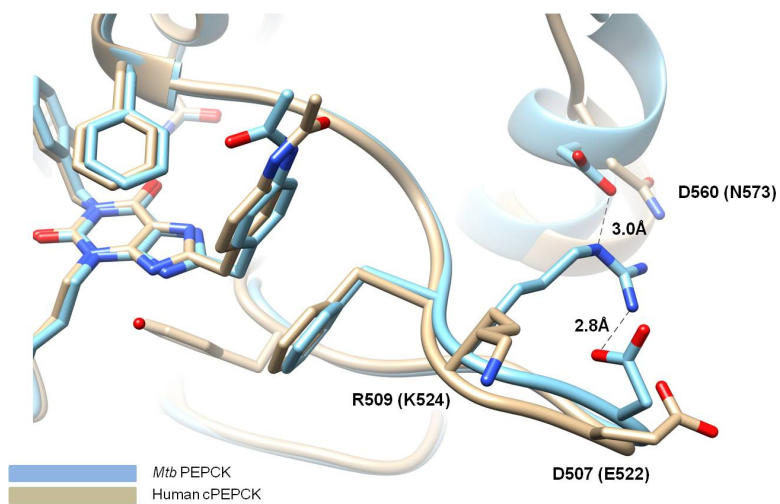


Figure 3-7. Superimposed inhibitor binding pocket of the human cPEPCK and *Mtb* PEPCK structures.

The superimposed structure of human cPEPCK (PDB:1NHX) and *Mtb* PEPCK in complex with Axon1165 (Figure 3-7) highlights several key residues that create the larger pocket in *Mtb* PEPCK, but not in human cPEPCK. The side chain of the *Mtb* PEPCK R509 residue faces the outer pocket and hydrogen bonds with the side chains of D560 (3.0 Å) and D507 (2.8 Å). The equivalent residue in human cPEPCK is K524 which does not interact with adjacent residues and seems to have free rotation. These differences may offer opportunities to develop selective *Mtb* PEPCK inhibitors. To find a lead molecule that binds to this area, we screened a variety of xanthine analogs with extended groups at the C-8 position.

Screening Roche 104 compounds (1st set of Roche library)

104 human cPEPCK inhibitors (GTP-competitive inhibitors), mostly with modifications on the C8 position were screened against *Mtb* PEPCK at a final concentration of 10 µM, resulting in 80 compounds that showed inhibition in the range of 50-90%, results are shown in Table 3-1.

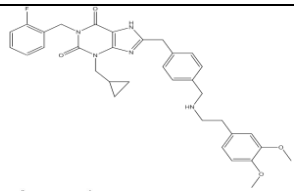
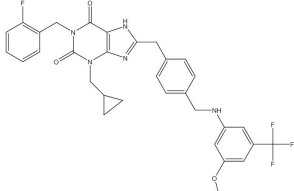
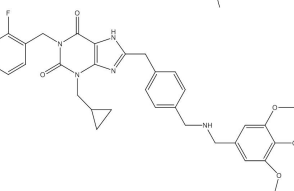
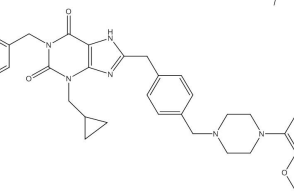
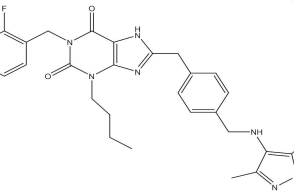
Table 3-1. 104 compound screen against the *Mtb* PEPCK enzyme.

Enzyme Inhibition at 10 µM	Number of compounds
> 90% inhibition	5
>70% inhibition	30
>50% inhibition	45

A group of 104 compounds were tested against *Mycobacterium tuberculosis* H37Rv for whole cell inhibition by the tuberculosis research lab in University Illinois Chicago (UIC). The data related to the compounds (sulfonamide, amide, and benzylamine linkage analogues with a xanthine core) are summarized in Table 3-2 and Table 3-3. Screening of the first set of Roche compound libraries suggested that sulfonamide derivatives are effective against the enzyme at 10 μM , but they have poor activity against TB whole cells, probably because the polarity of compounds prevents entry to the bacterium. The N-methyl amide derivatives shows increased *Mtb* PEPCK activity compared to the amide analogues. The benzylamine derivatives are most potent against the cell and enzymes with a low μM range of activity for MIC and IC_{50} , as shown Table 3-2. It suggests that benzylamine derivatives possess better permeability against cell. However, though RO4355651 has poor inhibition against the enzyme (no enzyme inhibition at 10 μM), it shows MABA MIC is 25 μM , thus the possibility of nonselective target inhibition cannot be ruled out.

We selected five compounds that were potent against both the enzyme (IC_{50} of 4 of 5 compounds in range of 0.5-1.5 μM) and whole cells (99-100% at 50 μM) (Table 3-2). Those five compounds were taken to X-ray crystallography to figure out binding modes, resulted in four compounds, RO4355622, RO4355668, RO4355703, and RO505244, PEPCK-inhibitor complex structures were solved for all of them at 2.0 Å resolutions.

Table 3-2. Chemical structures, IC₅₀, and MIC value of selected five compounds.

Compound ID	Structure Xanthine analogs - Benzylamine	ENZYME IC ₅₀ (μ M)	MABA MIC (μ M)
RO4355622		0.626	22.7
RO4355651		No inhibition <10 μ M	25
RO4355668		0.752	16.4
RO4355703		1.26	7.6
RO0505244		1.21	28.3

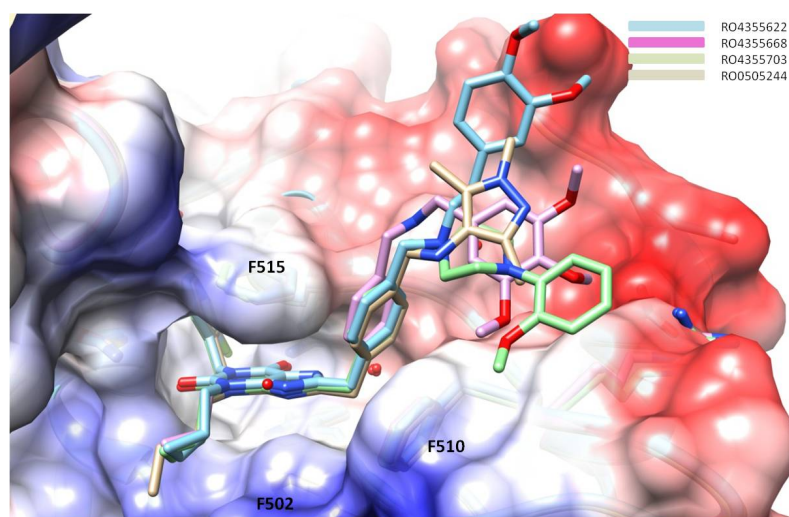


Figure 3-8. Superimposition of PEPCK-inhibitor complex structures with RO4355622, RO4355668, RO4355703, and RO505244.

All four compounds bind to the GTP-binding site in a similar fashion as Axon 1165, the π - π interactions of xanthine core and C-8 benzyl ring with the phenyl ring of the F530, F517, and F525 side chains are conserved. However, interestingly, the extended ring systems of C-benzyl ring do not interact with any enzyme residues (Figure 3-8). Instead, the human cPEPCK is in complex with a sulfonamide compound (PDB:2GMV) and shows that the imidazole ring extends from the sulfonamide linkage positioned on top of the phenyl ring from the F530 side chain for a π - π interaction (Figure 3-9).

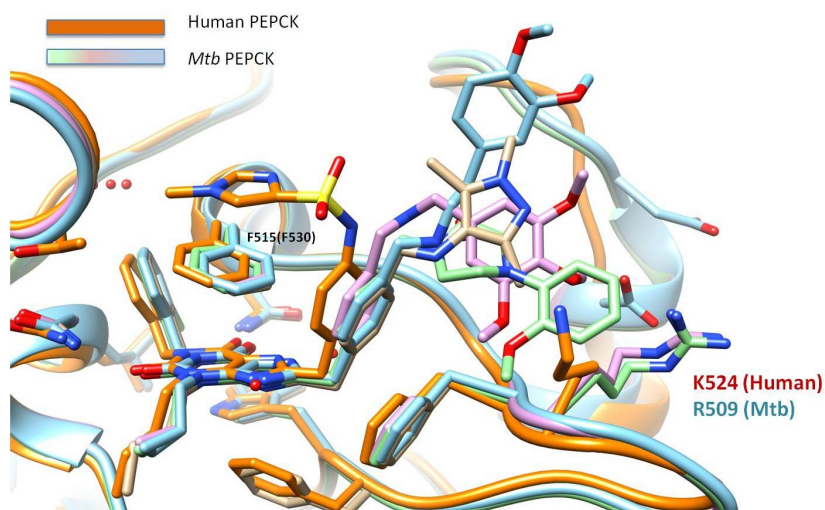


Figure 3-9. Superimposed inhibitor complex structures of *Mtb* PEPCK with RO4355622, RO4355668, RO4355703, and RO505244 and human cPEPCK with N-(4-{[3-butyl-1-(2-fluorobenzyl)-2,6-dioxo-2,3,6,7-tetrahydro-1H-purin-8-yl]methyl}phenyl)-1-methyl-1H-imidazole-4-sulfonamide (orange color) (PDB code:2GMV).

The difference between conformation of K524 (human cPEPCK) and R509 (*Mtb* PEPCK) may induce the differences in the binding mode of the extended ring from the C-8 benzyl on R8. As shown in Figure 3-9, the different placement of a sulfonamide derivative in the human PEPCK structure might be due to the steric hindrance caused by the K524 residue. It is possible that the steric hindrance caused by K524 facilitates this turn, while R509 of *Mtb* PEPCK can accommodate any extended ring system near the small binding pocket, which represents a different environment compared to the human one. Therefore, the bulky group of the sulfonamide linkage in certain compounds could be critical to confer selectivity against the *Mtb* PEPCK enzyme. The xanthine analogs, sulfonamide, and amide derivatives with MABA MIC and percent inhibition at 10 μ M of compounds against the *Mtb* PEPCK enzyme are represented in Table 3-3. Despite the

compounds have inhibitory effect on enzyme at 10 μM , whole cell activity is very poor. Probably, the polarity of sulfonamides and amides makes compounds difficult to enter into cell.

Table 3-3. Chemical structure of xanthine analogs, sulfonamide, and amide derivatives with MABA MIC and percent inhibition at 10 μM of compounds against the *Mtb* PEPCK enzyme.

^a data from TB research center at UIC

Compound ID	Structure Xanthine analogs - Sulfonamides	MABA MIC (μM) ^a	% Enzyme Inhibition at 10 μM
RO0504005		> 50	91
RO0503982		> 50	87
RO0505928		> 50	4
RO0282693		41.8	80
RO0505787		49.2	92

Table 3-3. Continued.

Compound ID	Structure Xanthine analogs - Amides	MABA MIC (μM) ^a	% Enzyme Inhibition at 10 μM
RO0282340		> 50	40
RO0503883		> 50	77
RO0506405		49.9	53
RO0506263		48.2	81
RO0506834		>50	90
RO0507332		46.4	89

Screening of 2nd set of Roche libraries selected based on structural information of complex structure of Mtb PEPCK with inhibitors, revealed the improved activity with respect to both cell and enzyme inhibition

The complex structures of *Mtb* PEPCK with compounds from the 1st set of Roche compounds revealed the small binding site with different residues between human and *Mtb* PEPCK, and provided opportunities for the development of selective inhibitors for anti-TB drugs. The 2nd set of Roche libraries contained selected 472 compounds with a variety of modifications on the C8 position and could potentially bind to the small binding site. The results of screening 2nd of Roche inhibitors revealed the improved activity with respect to both cell and enzyme inhibition. Of these compounds, 65 showed over 70% enzyme inhibition at 10 μ M. These 65 compounds were retested in dose-response for IC₅₀ value. The 20 compounds had IC₅₀ values ranging from 0.2 μ M to 2.5 μ M and MIC values were less than 10 μ M. Benzylamines continue to show very good activity in both the enzyme assay and MABA, and more potent N-methyl amides were identified. The IC₅₀ and MABA MIC values of the hit compounds are summarized in Table 3-4.

Table 3-4. The IC₅₀ and MABA MIC values of the hit compounds with chemical structure.

Compound ID	Structure Xanthine analogs – 2 nd Roche compounds	MABA MIC (μM) ^a	ENZYME IC ₅₀ (μM)
RO4355653		1.9	5
RO4355629		4.3	0.6
RO4355660		2.8	2.5
RO4355664		2.8	3
RO4355707		0.38	1.8
RO4355630		6.1	2.3
RO0504732		4.8	0.2
RO0506952		4.6	1.8
RO0507002		9.8	0.8
RO0507694		5.1	0.7

Structural study of 2nd Roche compound library hits revealed important interactions for binding to C-8 binding pocket

We identified a small pocket in *Mtb* PEPCK whose amino acid composition differs from the human one, providing us an opportunity to find selective inhibitors against pathogenic enzyme. Thus, we screened several Roche 2nd compounds using X-ray crystallography and identified inhibitors that interact with the binding pocket.

Structural analysis of *Mtb* PEPCK with RO4355707, which has the most potent activity against the cell (MIC = 0.38 μ M), reveals that the benzo-d-imidazole ring of RO4355707 contacts a hydrophobic interaction with the guanidinium group of the R509 residue in *Mtb* PEPCK (distance 3.6 Å) (Figure 3-10). The piperidine ring might make a better linker for entering the cell than the sulfonamides, and extending the benzo-d-imidazole ring successfully captured the interaction of the R509 side chain. Further modification of this compound could improve potency and selectivity against the enzyme.

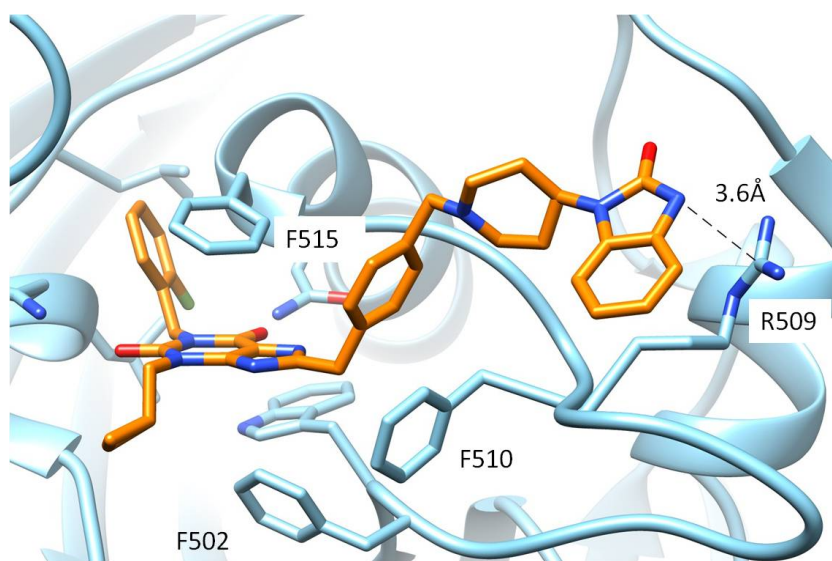


Figure 3-10. The binding mode of RO4355707 in *Mtb* PEPCK-RO4355707 complex structure.

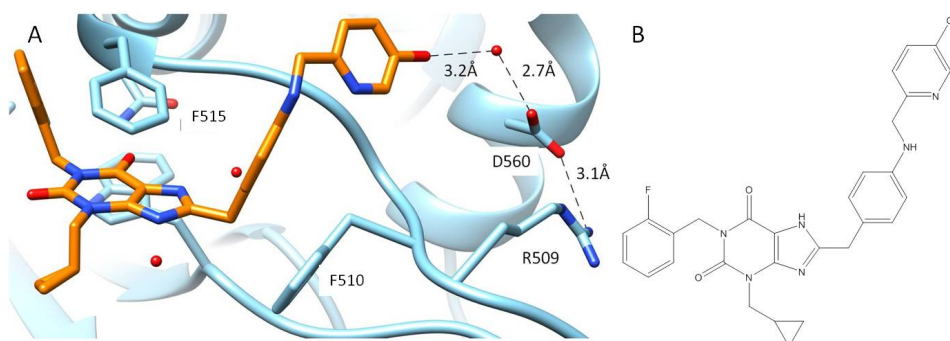


Figure 3-11. The binding mode of RO0506473 in *Mtb* PEPCK and chemical structure of RO0506473.

In the crystal complex structure of *Mtb* PEPCK-RO0506473, the molecule binds to the small pocket by interacting with the D560 residue through water mediated interactions as shown in Figure 3-11. The hydroxyl group interacts with the water

molecule through hydrogen bonds, and the water interacts with the carbonyl group of the D560 side chain, at 3.2 Å and 2.7 Å, respectively. This compound shows an IC_{50} of 2.5 μ M against *Mtb* PEPCK, but cell data is not available. Binding small pocket through water interaction shows no improved and selective activity on *Mtb* PEPCK. The water mediated interaction with D560 may not support a strong interaction to improve inhibitory effect. This molecule probably interacts with phenylring F530 of human cPEPCK to show an increased activity against human PEPCK (IC_{50} = 1.48 μ M).

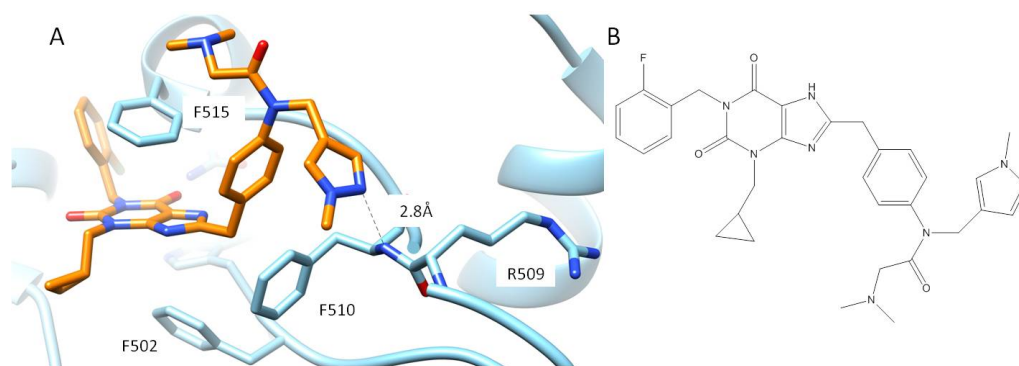


Figure 3-12. The binding mode of RO0506333 in *Mtb* PEPCK-RO0506333 complex structure and chemical structure of RO0506333.

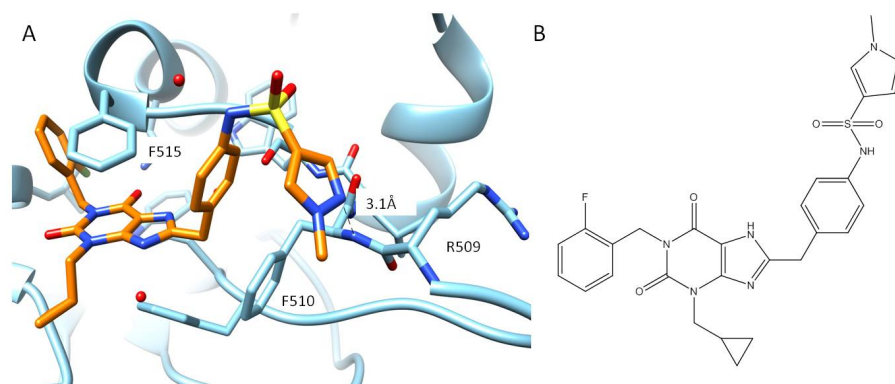


Figure 3-13. The binding mode of RO0505787 in *Mtb* PEPCK-RO0505787 complex structure and chemical structure of RO0505787.

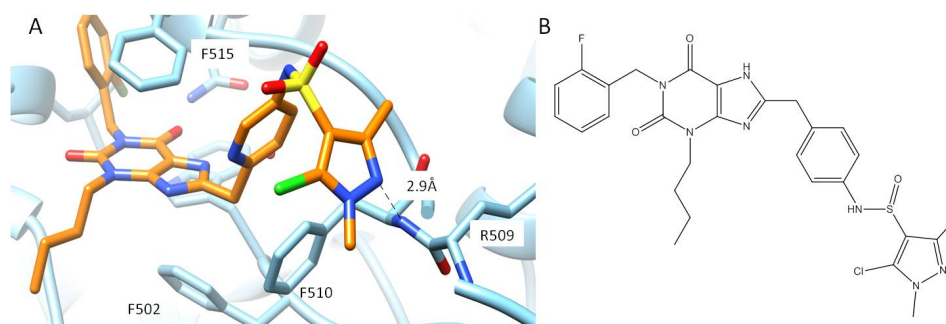


Figure 3-14. The binding mode of RO504449 in *Mtb* PEPCK-RO504449 complex structure and chemical structure of RO504449.

The complex structures of RO506333, 0505787, and 0504449 inhibitors (Figures 12, 13, and 14) with *Mtb* PEPCK reveal that the pyrazole ring is positioned near the R509 residue in C-8 binding pocket, not on the top of the benzyl ring of the F515 side chain. The 2-N atom of the pyrazole ring functions as a hydrogen bond acceptor interacting with the backbone of amide of the F510 residue. Crystallographic evidence suggests that the H-bond interaction between the molecule and the backbone of amide of

F510 is critical to position the extended ring near R509. These pyrazole derivatives showed very potent inhibition on *Mtb* PEPCK, RO506333 ($IC_{50}=0.30 \mu\text{M}$), RO050787 ($IC_{50}=0.96 \mu\text{M}$), RO0504449 ($IC_{50}=0.80 \mu\text{M}$). Therefore, this H-bond interaction is important to improve enzyme inhibitory activity against *Mtb* PEPCK. However, they showed no selective inhibition. It is possible that these compounds can still interact with phenyl ring of F530 side chain in human PEPCK to have π - π interaction showing strong inhibition.

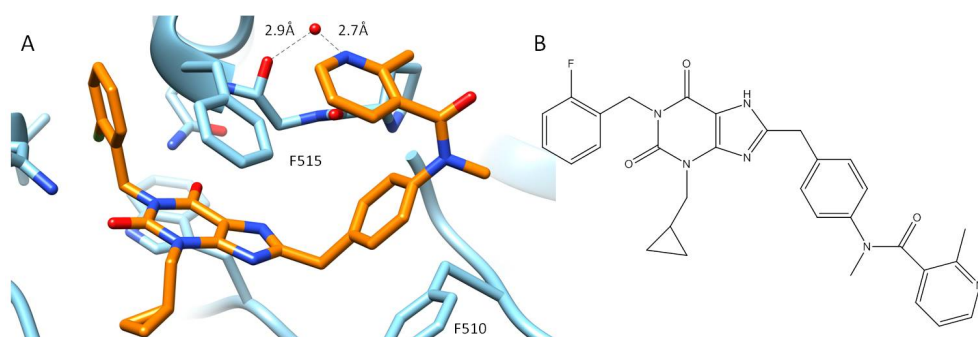


Figure 3-15. *Mtb* PEPCK-RO507000 complex structure and chemical structure of RO507000.

The nitrogen atom on the pyridine ring of RO507000 hydrogen bonds to the backbone carbonyl group of the G514 residue through a water molecule (Figure 3-15). The carbonyl group of the G514 backbone interacts with water (distance 2.9 Å) and the water interacts with the nitrogen of the pyridine ring (distance 2.7 Å) for a H-bonding interaction, this makes the pyridine ring turn closer to the F515 residue, unlike the pyrazole ring. The IC_{50} value of RO507000 is 0.76 μM , indicating not much different

activity were shown between the interaction closer to the R509 and the interaction closer to the F515 residue.

The crystal complex structures of *Mtb* PEPCK with the Roche inhibitors demonstrate clearly the importance of the H-bonding interaction between the compound and the backbone amide group of F510 to position the extended ring near the R509 residue. This localization of inhibitors near R509 is expected to increase selective activity against *Mtb* PEPCK because the lysine residue at the equivalent position may cause steric hindrance and thus would exhibit less activity against the human enzyme. This interaction also improved inhibitory efficacy on *Mtb* PEPCK. We tested the inhibitors against a cloned human cPEPCK enzyme using identical conditions to those used for the *Mtb* PEPCK enzyme. None of compounds showed selectivity against *Mtb* PEPCK, they all represent better activity on human enzyme (Table 3-5). It is plausible that the compounds interact with the benzyl group of the F530 side chain via π - π interactions as discussed in Figure 3-9, with a similar degree of enzyme inhibition. Therefore, further direction for the design of selective inhibitors against *Mtb* PEPCK: (1) keep H-bonding interaction between the compound and the backbone amide group of F510 to position the extended ring near the R509 residue, (2) have bulky groups that can reside near R509 residue in *Mtb* PEPCK, but induce steric hindrance with K524 in human cPEPCK, (3) avoid a π - π interaction with phenylring of F530 side chain in human cPEPCK would be valued strategies.

Table 3-5. The chemical structure of Roche compounds with the IC₅₀ value against human cPEPCK and *Mtb* PEPCK.

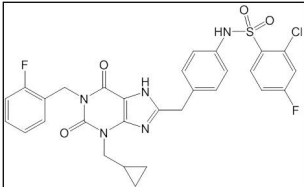
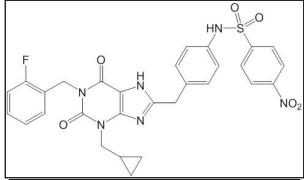
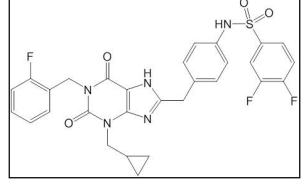
Compound ID	Structure	<i>Mtb</i> PEPCK IC ₅₀ (μM)	Human cPEPCK IC ₅₀ (μM)
RO4355707		2.29	1.40
RO0506473		2.50	1.48
RO0506333		0.30	0.22
RO0505787		0.96	0.16
RO0504449		0.80	0.80
RO0507000		0.76	0.70

Sulfonamides derivatives show selective cell growth inhibition depending on carbon sources, indicating that the compounds may be targeting on Mtb PEPCK

In order to figure out whether growth inhibition is via PEPCK inhibition or not, newly synthesized of sulfonamides analogs (by Wuxi AppTec) were tested against H37Rv for whole cell inhibition by the tuberculosis research lab in University Illinois Chicago (UIC), using different carbon source, palmitate and glycerol containing media (Table 3-6). The compounds which show better cell inhibition on 7H12 (in presence palmitic acid as a sole carbon source) were presented here with IC₅₀ value of human and *Mtb* PEPCK enzyme.

Table 3-6. Xanthine analogs – sulfonamides are tested on mc²7000 under different carbon source.

^{a,b} Cell data were provided by the tuberculosis research lab in University Illinois Chicago (UIC).

Compound ID	Structure Xanthine analogs - Sulfonamides	7H12 ^a	7H12 without palmitic acid add 0.2% glycerol ^b	<i>Mtb</i> PEPCK IC ₅₀ (μM)	Human cPEPCK IC ₅₀ (μM)
ES2888-133-P1		3.9	13.1	0.44	0.20
ES652-550-P1		9.6	21.0	0.24	0.26
ES2888-132-P1		5.6	12.2	0.58	0.22

In vitro, *Mtb* PEPCK enzyme is essential for TB growth in fatty acids condition as a carbon source²⁵. Therefore, inhibitors targeting on *Mtb* PEPCK would be specific to palmitic acid medium rather than specific to glucose. In Table 3-6, three potent inhibitors inhibit growth very efficiently on palmitic acid medium. This suggests that their growth inhibition could be via PEPCK inhibition. However, they are also active on glycerol condition, thus, the possibility of multiple target inhibition of those compounds cannot be ruled out.

Modification to the N1 binding pocket of xanthine analogues shows selectivity against the Mtb PEPCK enzyme over the human counterpart

Our collaborators, Schrödinger and the TB alliance modified the N1 position of the xanthine analogues to design PEPCK-0057 (N-(4-((3-(cyclopropylmethyl)-1-(2-fluoro-4-(methylsulfonyl)benzyl)-2,6-dioxo-2,3,6,7-tetrahydro-1H-purin-8-yl)methyl)phenyl)methanesulfonamide) and PEPCK-0058 (N-(4-((3-(cyclopropylmethyl)-1-(2-fluoro-4-(2-oxooxazolidin-3-yl)benzyl)-2,6-dioxo-2,3,6,7-tetrahydro-1H-purin-8-yl)methyl)phenyl)methanesulfonamide) (Figure 3-16). There is a second pocket that can be explored for selective inhibitor development, located at L281 in *Mtb* PEPCK, and M296 in the human PEPCK enzyme (Figure 3-17). In *Mtb* PEPCK two water molecules hydrogen bond with residues adjacent to L281. One water molecule interacts with the amide backbone of Q282 at a distance of 3.0 Å and the other water molecule interacts with the carbonyl backbone of N277 at a distance of 2.8 Å. Those two water molecules are connected via a hydrogen bond (2.7 Å). PEPCK-0057 and PEPCK-0058 contain a methyl sulfonyl and oxooxazolidinyl group in the N1 position, respectively, which should act to replace the two water molecules.

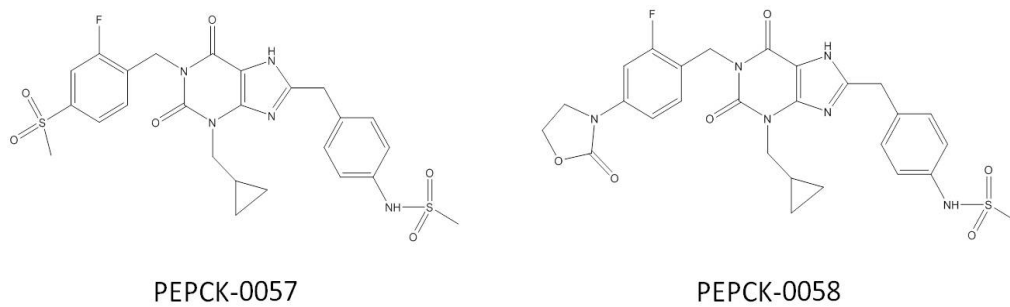


Figure 3-16. The chemical structures of PEPCK-0057 and PEPCK-0058.

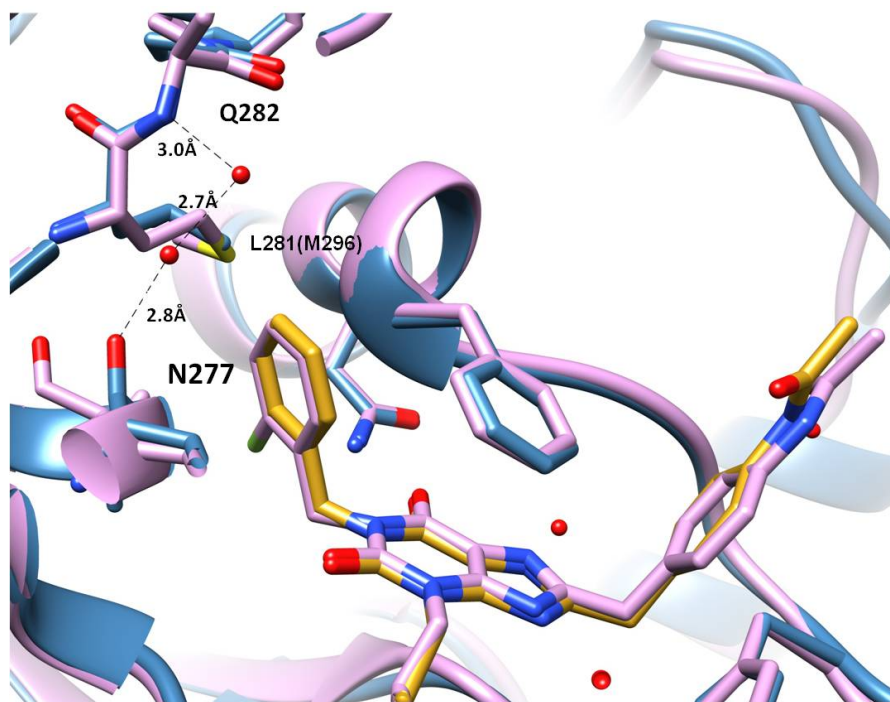


Figure 3-17. Superimposed structure of *Mtb* PEPCK (blue) and human PEPCK (pink) complex with Axon1165.

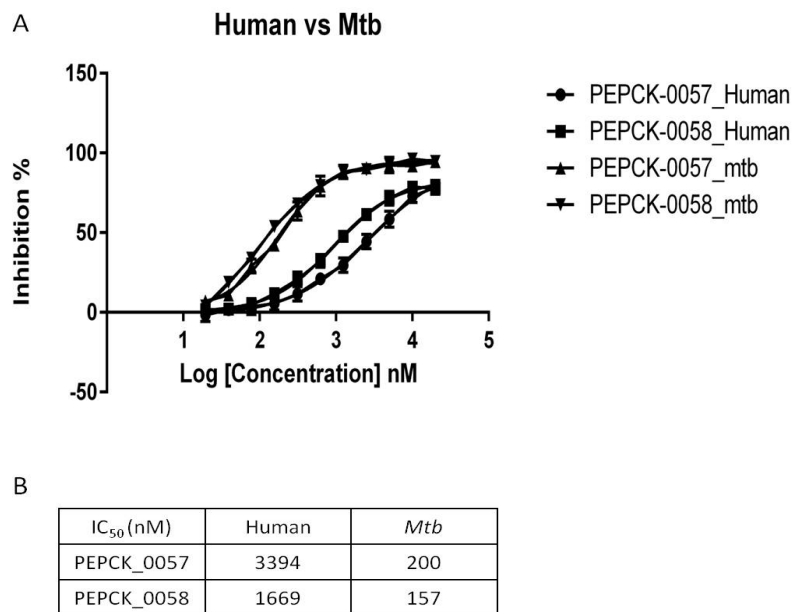


Figure 3-18. The selectivity of PEPCK-0057 and PEPCK-0058 against *Mtb* PEPCK over human cPEPCK. (A) Dose-response graph for PEPCK-0057 and PEPCK-0058 against both human and *Mtb* PEPCK enzymes. (B) IC₅₀ of PEPCK-0057 and PEPCK-0058 on human and *Mtb* PEPCK enzymes.

PEPCK-0057 and PEPCK-0058 exhibited enzyme inhibition with IC₅₀s of 200 nM and 157 nM, respectively, which shows about a 10-fold greater selectivity for the *Mtb* PEPCK enzyme than the human analog (Figure 3-18). As shown Figure 3-19, the complex structure with PEPCK-0057 indicates that the oxygen of the methyl sulfonyl group from the fluorobenzyl ring on xanthine replaced a water molecule by forming a hydrogen bond to the backbone amide of Q282 (distance 2.9 Å). This interaction induces a minor shift, 0.4 Å, of the C α atoms of Q282, whereas the remaining residues interacting with the inhibitor remained in their original positions. This minor shift to interact with the methyl sulfonyl group of the inhibitor is not favorable in human

PEPCK. These differences can be explained by the presence of the longer M296 side chain in the human PEPCK compared to the shorter L281 in *Mtb* PEPCK. In addition, a high degree of flexibility in M296 may cause steric hinderance to the methyl sulfonyl group of the inhibitor (Figure 3-20). In Figure 3-21, the PEPCK-0058 complex of the *Mtb* PEPCK structure indicates that the oxygen of the oxazolidone ring carbonyl group interacts with the backbone amide of Q282 at a distance of 2.9 Å. Also, the side chain of L281 supports hydrophobic interaction with the oxazolidine ring at a distance of 3.4 Å. The superimposed structure of *Mtb* PEPCK and human PEPCK reveals that the side chain of M296 in human PEPCK is too close to have this hydrophobic interaction at 2.1 Å. The co-crystal structures provide an explanation of the selectivity observed and these data suggest that the site binding the methyl sulfonyl group and oxazolidine ring of the inhibitors is a promising one for the design and development of selective inhibitors against the *Mtb* PEPCK enzyme.

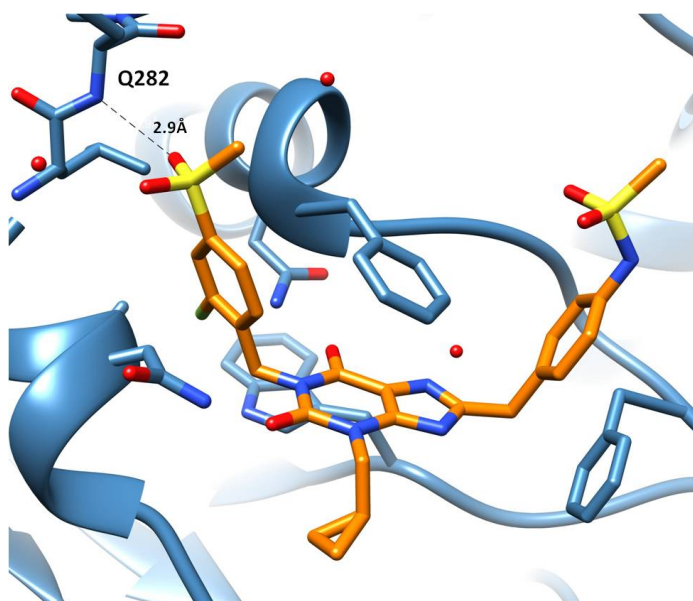


Figure 3-19. The complex structure of *Mtb* PEPCK with PEPCK-0057.

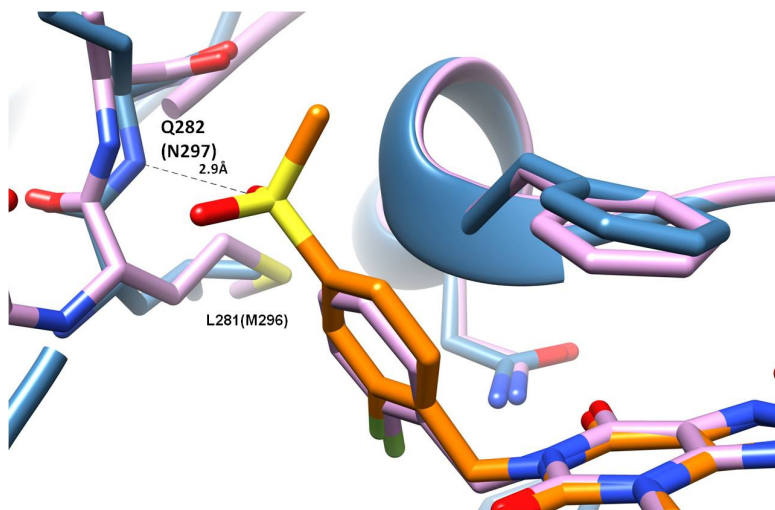


Figure 3-20. The superimposed structures of the PEPCK-0057-*Mtb* PEPCK complex (blue) and the Axon 1165-human PEPCK complex (pink). PEPCK-0057 is shown in orange and Axon 1165 in light pink.

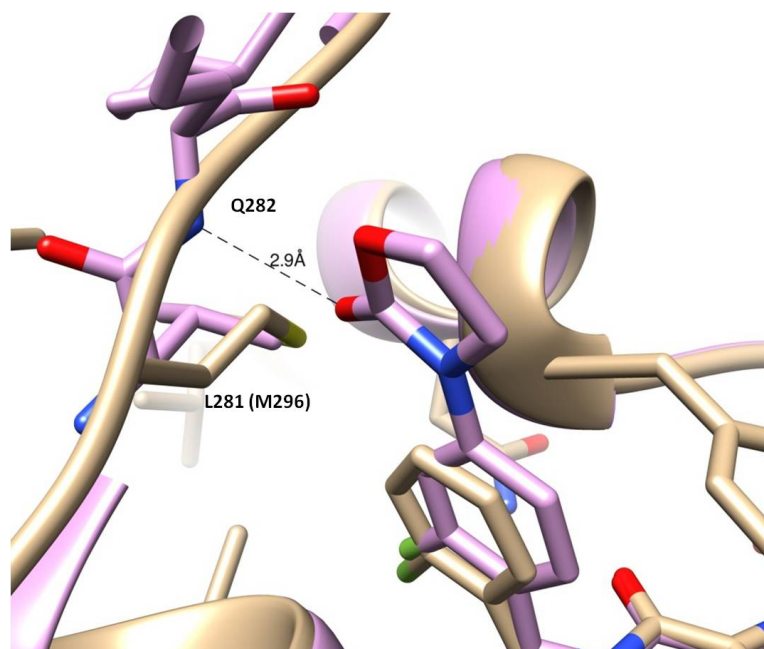


Figure 3-21. The superimposed structures of the PEPCK-0058- *Mtb* PEPCK complex (pink) and the Axon 1165-human PEPCK complex (tan).

Conclusion

The Roche inhibitors that bind near the R509 residue in C-8 binding pocket and near L281 in N-1 binding pocket have been identified through X-ray structure studies. Moreover, the IC_{50} values against the human and *Mtb* PEPCK enzymes have been compared to determine the selectivity of the inhibitors. We discovered an important interaction which resulted in improved activity and led us to place a ring near the R509 residue. We also identified potent whole-cell inhibitors, specific to growth on fatty acids. These compounds could be used for proof of concept (POC) for the development of an anti-TB drug targeting *Mtb* PEPCK. The xanthine inhibitors with modification to the N1 position (PEPCK-0057 and PEPCK-0058) were designed by Schrödinger and the TB

alliance to bind near the L281 residue. They show both potency and selectivity against *Mtb* PEPCK. The co-crystal structure of PEPCK-0057 and -00058 with *Mtb* PEPCK revealed that the minor shift created through the hydrogen bond might be not favorable in the human enzyme. The longer methionine residue in the human PEPCK instead of the leucine in *Mtb* PEPCK, may cause steric hindrance upon binding these inhibitors. Therefore, this site near the L281 residue is a promising site that can be explored for the development of selective and potent inhibitors against *Mtb* PEPCK. The detailed characterization of the inhibitors presented here provides further insight into design strategies for pathogen-specific inhibitors.

Experimental procedures

Enzyme assay for high-throughput screening for identification of M. tuberculosis PEPCK inhibitors

To identify inhibitors through high-throughput screening (HTS), PK/LDH coupled assay was used. The standard assay mixture contained 0.3 mM GTP, 2 mM MgCl₂, 0.1 mM MnCl₂, 1 mM DTT, 1 mM ADP, 2 unit PK/LDH enzyme, 0.25 mM NADH, and 2 μg/ml *Mtb* PEPCK (28nM) enzyme in 20mM Tris-HCl, pH 7.5, with 50mM NaCl. The enzyme reaction was started by adding 0.2 mM OAA after 20 minute incubation with inhibitors. The enzyme assay was monitored the decreased NADH absorbance at 340 nm during 5min. The % inhibition was calculated from this equation, %inhibition = (1- (decreased absorbance slope)/control slope (no inhibitor, only DMSO))*100. The value of IC₅₀ was calculated by using Prism. Activity assays were

performed with a Thermo Scientific *Multiskan GO* spectrometer. The plate used in the enzyme assay was a 96-well, half area, UV-star plate, (Greiner bio-one).

Differential Scanning Fluorimetry (DSF) assay of M. tuberculosis PEPCK

The real time (RT)-PCR machine (Stratagene Mx3005P) was set to equilibrate samples at 25°C for 5 minutes and then gradually increase temperature to 95°C at a rate of 1°C/minute. The excitation and emission of SYPRO orange can be monitored using filters: FAM (492nm) and ROX (610nm). The RT-PCR instrument (MxPro software) calculated first derivative plot, indicating the lowest point is melting point of the protein^{105, 106}. The master mix contained 40µg/ml protein in buffer (20mM Tris-HCl pH 7.5, 50mM NaCl) and 5X SYPRO orange (diluted from 5000X concentrate). Each well of a 96-well plate has 10µl solution contained 9µl master mix and 1µl of 1mM inhibitors stock (final concentration 100 µM). For the fragment screening, final 5mM inhibitors were used. It is not recommended exceed DMSO amount not over 2% (v/v), but this case, 10% DMSO does not influence the results.

Crystallization of Mtb PEPCK in complex with inhibitors and data collection

Crystallization was accomplished by the hanging-drop vapor diffusion method. Co-crystal of inhibitors was obtained by soaking the apo crystal of the mutant E83A PEPCK in 125µM inhibitors for 2 days in the presence of 10% PEG3350, 20%Glycerol, and 0.5mM MnCl₂. Crystals were cryo-protected by dipping them into a solution [10% (w/v) PEG3350 and 20% glycerol (v/v)]. Diffraction data were collected at a wavelength of 1.542 Å using a Rigaku R-Axix IV ++ at home source and at the Advanced Photon Source beamline 23-ID using a MAR 300 CCD detector (MarMosaic from Marresearch

Charged Coupled Device) (Rayonix LLC, Evanston,IL). The HKL2000 program package (HKL Research,Inc., Charlottesville,VA) was used for integration and scaling.

Structure determination and refinement of M. tuberculosis PEPCK in complex with inhibitors

To determine the co-crystal structure of Mtb PEPCK with inhibitors, the apo structure of the mutant E83A Mtb PEPCK used as a template for refinement by using PHENIX (Python-based Hierarchical Environment for Integrated Xtallography). Rigid body refinement, followed by simulated annealing at 5000 K, was conducted. Subsequently, refinement was conducted in alternating cycles of manual model building in COOT (Crystallographic Object Oriented Toolkit), followed by refinement in PHENIX until the R factors converged. Inhibitors were manually built in the active site of the model by examining the F_o-F_c map in Coot (Paul Emsley and Kevin Cowtan, 2004). The full data collection and refinement statistics are given in Table 3-7.

Cloning, expression and purification of human cytosolic PEPCK (cPEPCK)

cDNA of human cPEPCK purchased from Origene Technologies was amplified by PCR (30s at 94°C; 30s at 55°C ; 4min at 72°C; 10min at 72°C) and then digested with NdeI and BamHI enzymes. Digested PCR products ligated into pET19b (+) vector contained Nde and BamH restriction sites. The ligate product was transformed into ONE Shot Top10 cells (Invitrogen). After confirmed DNA sequence, isolated the plasmid DNA was transformed into BL21 (DE3) (Novagene) for over-expression. The expression and purification of the human cPEPCK was performed following the protocol used for the *M. tuberculosis* PEPCK.

Enzyme assay of human cytosolic PEPCK (cPEPCK)

The enzyme assay of human cPEPCK was conducted under identical condition used in Mtb PEPCK.

Whole cell assay for MIC determination on mc²7000 strain

mc² 7000 starting culture in 7H9 containing OADC, Tyloxapol, Pantothenic acid, Malachite green was used to grow until OD₆₀₀ reaches to 1.0. The growing cell from starting culture was added into testing media 7H9 with dextrose and M9 with acetate separately. For the dextrose condition, it consists 0.5% dextrose, 0.085% NaCl, 0.05% tyloxapol, 25µg/ml Pantothenic acid, and 0.25µg/ml malachite green. For acetate condition, it consists 0.5% acetate, 2mM MgSO₄, 0.1 mM CaCl₂, 0.05% tyloxapol, 25µg/ml Pantothenic acid, and 0.25µg/ml malachite green. Incubation with inhibitors (50, 25, 12.5, 6.25, 3.12, 1.56, 0.78 µM final concentration) in 96-well plate lasts for 7 days at 37 °C for dextrose plate, and 14 days for acetate plate. The each well was stained with either alamar blue (MABA assay) or Resazurin (5 µg/ml final) for MIC determination.

Table 3-7. Data collection and refinement statistics for *Mtb* PEPCK in complex with inhibitors.

	<i>Mtb</i> PEPCK - Axon1165	<i>Mtb</i> PEPCK - RO4355622	<i>Mtb</i> PEPCK - RO4355668	<i>Mtb</i> PEPCK - RO4355703	<i>Mtb</i> PEPCK - RO505244	<i>Mtb</i> PEPCK - RO4355707
Wavelength (Å)	0.98	0.98	0.98	0.98	0.98	0.98
Resolution range (Å)	50 - 1.7 (1.77-1.71)	50 - 1.73 (1.79 - 1.73)	50 - 1.57 (1.62 - 1.57)	50 - 1.57 (1.60 - 1.57)	50 - 1.37 (1.39 - 1.37)	50 - 2.05 (2.09 - 2.05)
Space group	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1
Unit cell	104.667 122.246 64.245 90 117.15 90	105.680 122.117 64.869 90 117.70 90	105.655 122.353 64.663 90 117.43 90	105.770 122.314 64.868 90 117.53 90	105.616 122.655 64.788 90 117.33 90	106.297 121.798 64.587 90 117.56 90
Total reflections	327317	75172	99639	390909	532432	151036
Unique reflections	76938	73671	98029	98212	136717	43328
Multiplicity	4.3 (4.0)	1.0 (1.0)	1.0 (1.0)	4.0 (3.0)	3.9 (2.6)	3.5 (2.5)
Completeness (%)	99.75(97.76)	96.71 (75.87)	96.36 (85.14)	96.85 (77.73)	89.38 (61.95)	94.48 (69.29)
Mean I/sigma(I)	27.03 (9.66)	21.54 (0.957)	31.590 (1.459)	25.560 (1.559)	18.204 (0.934)	14.475 (1.262)
Wilson B-factor	16.33	25.84	22.17	19.91	15.08	34.62
R-merge	0.053 (0.161)	0.096 (0.838)	0.076 (0.617)	0.090 (0.509)	0.081 (0.758)	0.134 (0.673)
R-work	0.1536 (0.1573)	0.1638 (0.2819)	0.1604 (0.2542)	0.1639 (0.2359)	0.1692 (0.2747)	0.1632 (0.2581)
R-free	0.1810 (0.2056)	0.1921 (0.3243)	0.1742 (0.2832)	0.1908 (0.2650)	0.1897 (0.3110)	0.2019 (0.3132)
Number of atoms	5551	5557	5635	5604	5679	5355
Macromolecules	4773	4767	4775	4766	4766	4766
Ligands	51	51	51	51	46	53
Water	727	739	809	787	867	536
Protein residues	598	598	599	598	598	598
RMS (bonds)	0.007	0.009	0.008	0.008	0.008	0.010
RMS (angles)	1.21	1.37	1.37	1.38	1.23	1.38
Ramachandran favored (%)	98	98	97	97	98	97
Ramachandran outliers (%)	0	0	0	0	0.16	0
Clash score	2.78	3.49	4.32	3.97	3.60	3.32
Average B-factor	18.8	27.40	25.30	22.70	16.90	29.40
Macromolecules	17.0	25.80	23.30	20.90	14.80	28.50
Ligands	22.4	32.80	34.80	30.60	21.70	32.50
Solvent	29.9	37.80	36.50	33.30	27.90	36.90

Table 3-7. Continued.

	<i>Mtb</i> PEPCK- RO0506473	<i>Mtb</i> PEPCK- RO0506333	<i>Mtb</i> PEPCK- RO0505787	<i>Mtb</i> PEPCK- RO0504449	<i>Mtb</i> PEPCK- RO0507000
Wavelength (Å)	0.98674	0.98674	0.97924	0.97949	0.97949
Resolution range (Å)	50 - 1.63 (1.66 - 1.63)	50 - 1.63 (1.66 - 1.63)	50 - 1.62 (1.65 - 1.62)	50 - 1.84 (1.87 - 1.84)	50 - 1.84 (1.87 - 1.84)
Space group	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1
Unit cell	105.273 121.783 65.14 90 117.91 90	106.149 121.754 65.031 90 117.81 90	105.829 121.362 64.952 90 117.94 90	106.548 122.574 64.882 90 117.69 90	105.894 121.84 65.265 90 118.04 90
Total reflections	366063	372794	374128	233435	236775
Unique reflections	88592	89009	91578	63380	58388
Multiplicity	4.1 (3.4)	4.2 (3.4)	4.1 (3.2)	3.7 (3.6)	4.1 (4.2)
Completeness (%)	99.46 (94.96)	98.40 (94.12)	99.53 (96.17)	99.48 (96.83)	92.87 (83.79)
Mean I/sigma(I)	30.775 (3.11)	22.878 (2.225)	31.221 (3.129)	17.780 (2.108)	43.518 (19.362)
Wilson B-factor	20.28	15.51	20.02	18.24	14.52
R-merge	0.073 (0.336)	0.066 (0.387)	0.104 (0.393)	0.117(0.726)	0.073 (0.138)
R-work	0.1535 (0.2077)	0.1490 (0.2196)	0.1683 (0.2028)	0.1633 (0.2356)	0.1469 (0.1559)
R-free	0.1730 (0.2666)	0.1676 (0.2488)	0.1805 (0.2345)	0.1988 (0.2967)	0.1767 (0.1974)
Number of atoms	5678	5724	5494	5612	5702
Macromolecules	4782	4782	4782	4782	4766
Ligands	51	56	46	54	49
Water	845	886	666	785	887
Protein residues	600	600	600	600	598
RMS (bonds)	0.017	0.009	0.008	0.008	0.008
RMS (angles)	2.08	1.39	1.53	1.24	1.36
Ramachandran favored (%)	98	98	97	98	98
Ramachandran outliers (%)	0	0	0	0	0
Clash score	4.57	4.05	3.20	3.52	3.75
Average B-factor	24	19.40	22.90	18.50	16.60
Macromolecules	21.90	17.20	21.40	16.80	14.30
Ligands	28.60	20.00	24.10	20.90	15.90
Solvent	35.70	31.60	33.60	28.80	28.90

Table 3-7. Continued.

	<i>Mtb</i> PEPCK- PEPCK-0057	<i>Mtb</i> PEPCK- PEPCK-0058
Wavelength (Å)	1.542	0.979
Resolution range (Å)	23.89 - 1.90 (1.97 - 1.90)	31.44 - 2.1 (2.05 - 1.98)
Space group	C 1 2 1	C 1 2 1
Unit cell	104.01 121.69 63.95 90 116.66 90	104.65 121.74 63.92 90 116.59 90
Total reflections	105612	205874
Unique reflections	48108	45970
Multiplicity	2.2 (2.1)	4.5 (2.0)
Completeness (%)	86.30 (82.96)	91.70 (72.15)
Mean I/sigma(I)	28.00 (3.04)	10.91 (0.17)
Wilson B-factor	28.59	45.11
R-merge	0.086 (0.350)	0.197 (3.39)
R-work	0.1732 (0.2118)	0.1851 (0.3891)
R-free	0.2106 (0.2668)	0.2449 (0.4005)
Number of atoms	5190	5062
Macromolecules	4782	4782
Ligands	50	41
Water	357	239
Protein residues	600	600
RMS (bonds)	0.009	0.010
RMS (angles)	1.31	1.37
Ramachandran favored (%)	97	97
Ramachandran outliers (%)	0	0
Clash score	2.93	3.74
Average B-factor	26.00	43.40
Macromolecules	25.50	43.00
Ligands	27.40	45.40
Solvent	33.50	49.80

Additional results

HT screening results from chemical and fragments libraries

1,131 SRI anti-tuberculosis active compounds were screened against *Mtb* PEPCK. The DSF method is used as an initial screen against *Mtb* PEPCK to identify any molecule that has affinity to the active site. The hits were confirmed using the enzyme assay. The summary of the hits, IC₅₀ and ΔT_m, are shown in Table 3-8 and their chemical structures are shown in Figure 3-22.

Table 3-8. IC₅₀ and ΔT_m value of seven SRI hits against *Mtb* PEPCK.

Compound ID	Name	IC ₅₀ of <i>Mtb</i> PEPCK (μM)	Δ T _m (C°)
333612	1-(4-(2-bromo-4-chlorophenoxy)butyl)-1H-imidazole	7.19	5.0
333496	1-(4-(4-bromo-2-chlorophenoxy)butyl)-1H-imidazole	7.26	5.0
333360	1-(4-(4-isopropyl-3-methylphenoxy)butyl)-1H-imidazole	4.94	5.5
332991	1-(4-(3-isopropyl-5-methylphenoxy)butyl)-1H-imidazole	5.91	5.5
332702	1-(4-(5-isopropyl-2-methylphenoxy)butyl)-1H-imidazole	6.19	5.5
332657	1-(3-(3,4-dichlorophenoxy)propyl)-1H-imidazole	4.70	5.6
309287	1-(5-(2,4-dichlorophenoxy)pentyl)-1H-imidazole	4.09	5.6

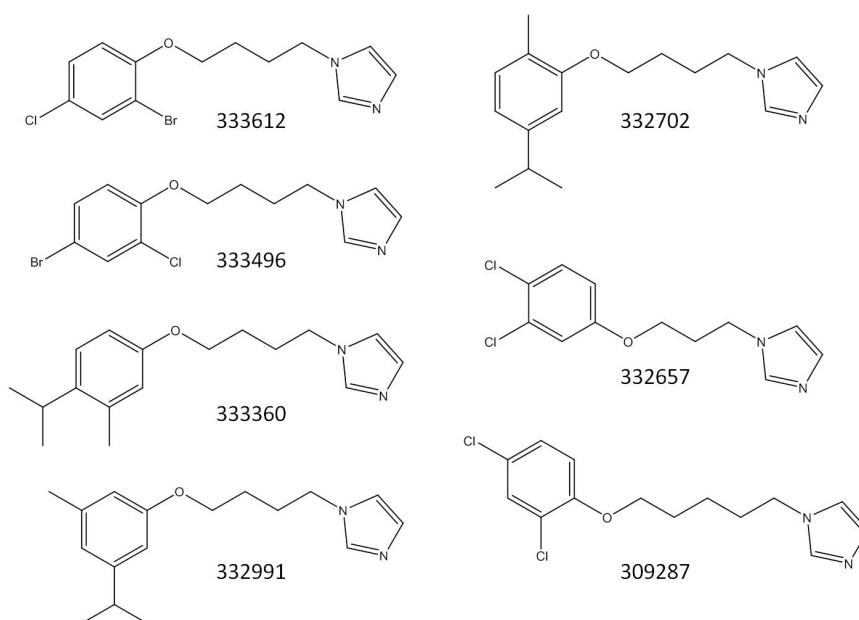


Figure 3-22. The chemical structure of the SRI hits from the HT screen against *Mtb* PEPCK by DSF and enzyme inhibition assay.

The series of phenoxyimidazoles from the SRI compound library show a consistent thermal shift of about 5 degrees from the melting point of *Mtb* PEPCK. These compounds are known to possess have about 1-2 $\mu\text{g/mL}$ of anti-tubercular activity, but its mechanism of action is unclear⁷¹. As shown in Table 3-1, these compounds show IC_{50} s in the low micro molar range against *Mtb* PEPCK. However, the purchased series of phenoxyimidazoles from ChemBridge Corporation (Figure 3-23) shows no thermal shift and no inhibition against the *Mtb* PEPCK enzyme.

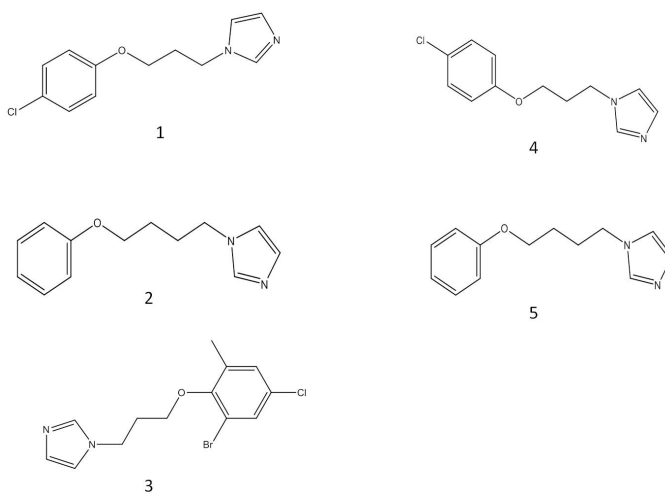


Figure 3-23. The chemical structure of the phenoxyalkylimidazole compounds purchased from ChemBridge Corporation.

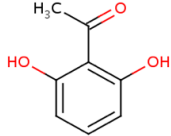
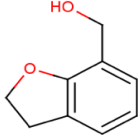
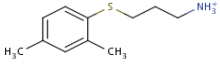
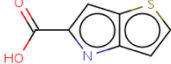
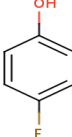
Despite the high structural similarity of compounds between the SRI libraries and the purchased one, the experimental results are not the same. We examined the difference between the purchased compounds and SRI compounds in the libraries, and found that the SRI compounds, especially the phenoxyalkylimidazole series, contained oxalate salt. Comparing the DSF results from the apo enzyme, the apo enzyme with Mn^{2+} , the apo enzyme with inhibitors, and the apo enzyme with Mn^{2+} and inhibitor sample (Table 3-9) showed that the inhibitor plus metal causes an approximately 6 degree thermal shift, while the inhibitor without metal shows no thermal shift, indicating that it is the complex of oxalate and metal binding to the active site that causes the thermal shift. This is in agreement with the *Mtb* PEPCK-oxalate complex structure discussed in chapter 2, which indicates that the oxalate coordinates directly with manganese and binds to the catalytic site (Figure 2-3B). Moreover, it has been reported

that oxalate inhibits the PEPCK enzyme in the micromolar range ⁶⁷. Taken together, the contamination of oxalate salt in the phenoxyalkylimidazole compounds in the SRI libraries resulted in false positives in the DSF study and enzyme inhibition assay. While screening these diverse libraries we learned that the presence of any acidic salt may cause false positive results, especially for the enzymes of the TCA cycles due to their intermediate structures resemble to the structure of acetate or oxalate.

Table 3-9. Melting temperature and thermal shift of samples, apo enzyme, apo enzyme in the presence of Mn^{2+} , apo enzyme with inhibitors, and apo enzyme with inhibitors in presence of Mn^{2+} . The DSF results suggest that the combination of metal and inhibitors cause a greater temperature shift.

Sample	Tm	ΔTm
Apo	50	-
Apo + Mn^{2+}	52	+2
Apo+ Inhibitors	50	0
Apo + Mn^{2+} + Inhibitors	56	+6

Table 3-10. Chemical structure and thermal shift (ΔT_m with Mn) of fragment hits.

Molecule name	Structure	ΔT_m with Mn
Cambridge Frag Lib10-E2		+ 20
Cambridge Frag Lib10-F6		+ 2.5
EN:T5267544		+ 3.0
CB:4012315		+ 2.0
Cambridge Frag Lib 15-G9		+ 3.0

A fragment library of 591 small molecules was screened against *Mtb* PEPCK, and five hits with thermal shifts with a range of 2.0 - 20 degrees were found (Table 3-10), but only one exhibited enzyme inhibition activity, EN:T5267544 ($IC_{50} = 50\mu M$). We were unable to conduct complex crystal trials for these hits due to the weak binding characteristics of the fragments.

HT screening of Sac1 diverse libraries (1667 compounds)

Our lab selected anti-tuberculosis active molecules from among 51K drug-like molecules, based on inhibition of TB growth on different carbon sources (such as dextrose vs. acetate). Since *Mtb* PEPCK is essential for TB growth on acetate as a carbon source in *in vitro* studies, 1667 anti-TB compounds active on acetate conditions were screened against *Mtb* PEPCK using an enzyme inhibition assay. Two compounds, EN: T5251690 and 5292830, exhibited enzyme inhibition with IC₅₀s of 58 μM and 38 μM, respectively (Figure 3-24). We could not obtain a complex structure of *Mtb* PEPCK with either of these inhibitors, likely due to low potency. However, these compounds could provide opportunities for further lead design and optimization.

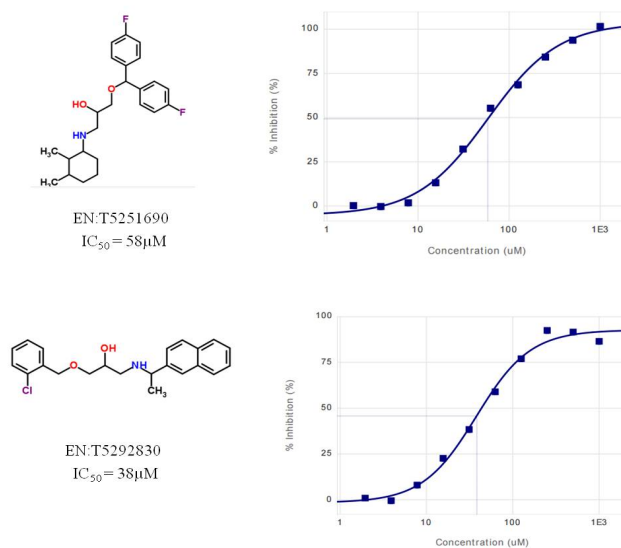


Figure 3-24. The chemical structure of two hits (EN:T5251690, EN:T5292830) from Sac1 libraries, and IC₅₀ value and dose-response curve.

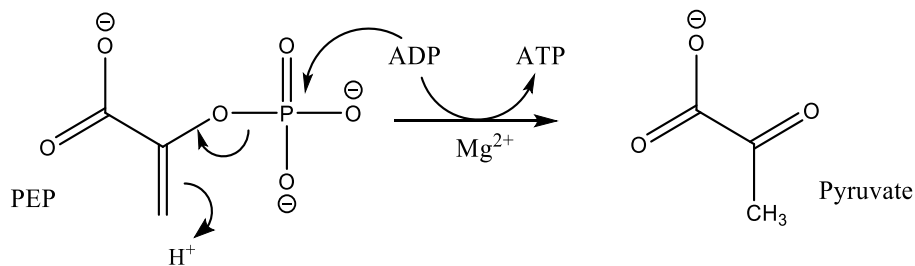
CHAPTER IV
STRUCTURAL STUDY OF PYRUVATE KINASE FROM *MYCOBACTERIUM*
TUBERCULOSIS

Introduction

Pyruvate kinase (PykA) has a regulatory role in central carbon metabolism as the rate-limiting enzyme in the final step of the glycolysis pathway. Its kinetic and regulatory properties have been well characterized with three-dimensional structures in divergent species, but little is known about the PykA from *Mycobacterium tuberculosis*. TraSH analyses have shown that *Mtb* PykA is essential for in vitro growth of *Mtb* on glycerol and fatty acids³⁰, indicating it may be a potential target for anti-TB chemotherapy.

PykA reaction

Pyruvate kinase (EC 2.7.1.40) (PykA) is involved in the junction of the gluconeogenic and glycolytic pathways. It controls the metabolic flux from fructose-1,6-bisphosphate to pyruvate, by acting as a rate-limiting enzyme in the last step of the glycolytic pathway. It catalyzes an irreversible reaction ($K_{eq} = 10^3$ to 10^4) under physiological conditions; the production of ATP from ADP by converting phosphoenolpyruvate (PEP) into pyruvate in the presence of either monovalent or divalent cations⁷² (Scheme 4-1).



Scheme 4-1. Pyruvate kinase (PykA) reaction.

PykA regulation by allosteric effector

PykA's position at the junction of the gluconeogenesis and glycolysis pathways makes it play a role in changing metabolic flux ⁷³. Thus, the intracellular level of ATP, GTP, and glycolytic intermediates are controlled by PykA ⁷⁴. The enzyme is regulated in multiple ways, via gene expression, allosteric regulation by small metabolites, and post-translational modification. Specially, levels of intracellular intermediates are commonly regulated by non-genetic mechanisms ⁷⁵. Most PykA enzymes are allosteric enzymes that are regulated by glycolytic intermediates (Figure 4-1). PykA enzymes from diverse species display different allosteric regulatory properties. For example, type 1 pyruvate kinase (PykF) from *Escherichia coli* (*E. coli*) is allosterically regulated by fructose 1,6-bisphosphate (F1,6BP) and type 2 *E. coli* PykA is activated by adenosine monophosphate (AMP). In addition, *L. mexicana* and *Trypanosoma cruzi* are regulated by fructose 2,6-bisphosphate (F2,6-BP), and both *Saccharomyces carlsbergensis* and yeast pyruvate kinase are regulated by F1,6BP ^{76,77, 78, 79}. In addition, mammalian pyruvate kinase enzymes existed as four different isoforms, representing different regulatory properties. The R and L isoforms are found in erythrocytes and the liver, the

M1 isoform is found in skeletal muscle, and the M2 form is widely spread with a low level of expression under normal cellular conditions, but is highly expressed in tumors. The M1 isoform is an unregulated enzyme, but R, L and M2 are allosterically regulated by F1,6-BP⁸⁰. Some bacilli including *M. tuberculosis*, *Geobacillus stearothermophilus*, and *B. stearothermophilus* are thought to be regulated by AMP and pentose monophosphates, such as ribose 5-phosphate⁸¹⁻⁸³. The variety of allosteric regulators for the pyruvate kinase enzyme across species implies that it has different regulatory mechanisms and different roles in different species. Moreover, the different allosteric molecule binding sites could offer an opportunity for the design of organism-specific drugs. A putative AMP-binding site has been proposed based on sequence alignment and mutation studies⁸³, but direct crystallographic evidence for the binding site of AMP has yet to be established.

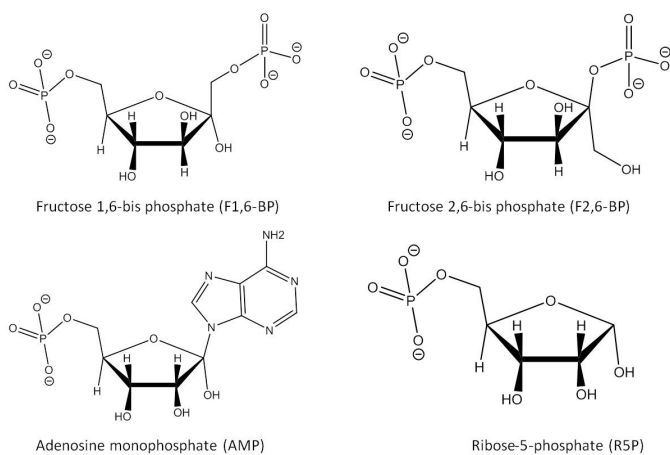


Figure 4-1. Chemical structure of common allosteric effectors for pyruvate kinase from diverse species.

Structural study of pyruvate kinase from divergent species

The three-dimensional structure of pyruvate kinase has been studied in many species including the human M2^{84,85} and R isozymes⁸⁴, as well as in *E. coli*⁸⁶, *Saccharomyces cerevisiae*⁷⁴, *Leishmania mexicana*⁷⁷, and *Trypanosoma brucei*⁸⁷. Most pyruvate kinase enzymes are homo-tetramers (each monomer contains about 500 residues) under physiological conditions, and have also been shown to be tetramers in the crystal lattice. Most of the pyruvate kinase structure has a highly conserved amino acid sequence in the active site and possess similar overall folds. Each subunit has three or four distinct domains, depending on the organism. In the human and yeast pyruvate kinase enzyme, each monomer consists of four domains, N (α -helical stretch N-terminal domain), A (the common parallel α/β_8 barrel motive), B (nine-stranded β barrel motif), and C (regulatory domain, α/β open-sheet motif)^{74,85}. Two A-domains contact to form a dimer and then tetramerization occurs between interfaces of the C-domains of the dimers. In the *E. coli* pyruvate kinase enzyme the N domain is not present it consists of only three domain, the A- ($(\alpha/\beta)_8$ barrel), B- (a flexible β barrel), and C-terminal domains⁸⁶. The catalytic site is located between the A- and B- domains, and residues from both domains are involved in substrate binding. The B-domain displays flexibility in many crystal structures, and is involved in the catalytic reaction. Upon PEP and ADP binding to the catalytic site, the B-domain moves like a hinge to start catalysis. The C-domain includes an allosteric effector binding site, that is 40 Å from the active site^{86,74}.

Pyruvate kinase is one of the housekeeping metabolic enzymes which is present in humans, thus a detailed structural comparison between the host and pathogen pyruvate

kinase structures is necessary to develop selective anti-TB drugs. As discussed previously, human pyruvate kinase has 4 different isozymes, M1, M2, L, and R. The sequence identity between *Mtb* PEPCK and the human isozyme is in the range of 35-37%. In Figure 4-2, multiple sequence alignment between *Mtb* pyruvate kinase and human isozymes reveals that the α -helical stretch of the N-terminal domain found in human pyruvate kinase is absent in *Mtb*.

We have determined the crystal structure of *Mtb* PykA in complex with a co-purified allosteric effector, AMP. Structural evidence for the AMP binding site has not been reported, all the available allosteric binding sites are for the F1,6BP and F2,6BP. The structural explanation of the activator specificity of *Mtb* PykA and the activator binding site compared with human M2 pyruvate kinase are discussed in this chapter. The difference between the effector binding sites provides another therapeutic approach for the design of selective anti-TB drugs targeting *Mtb* PykA.

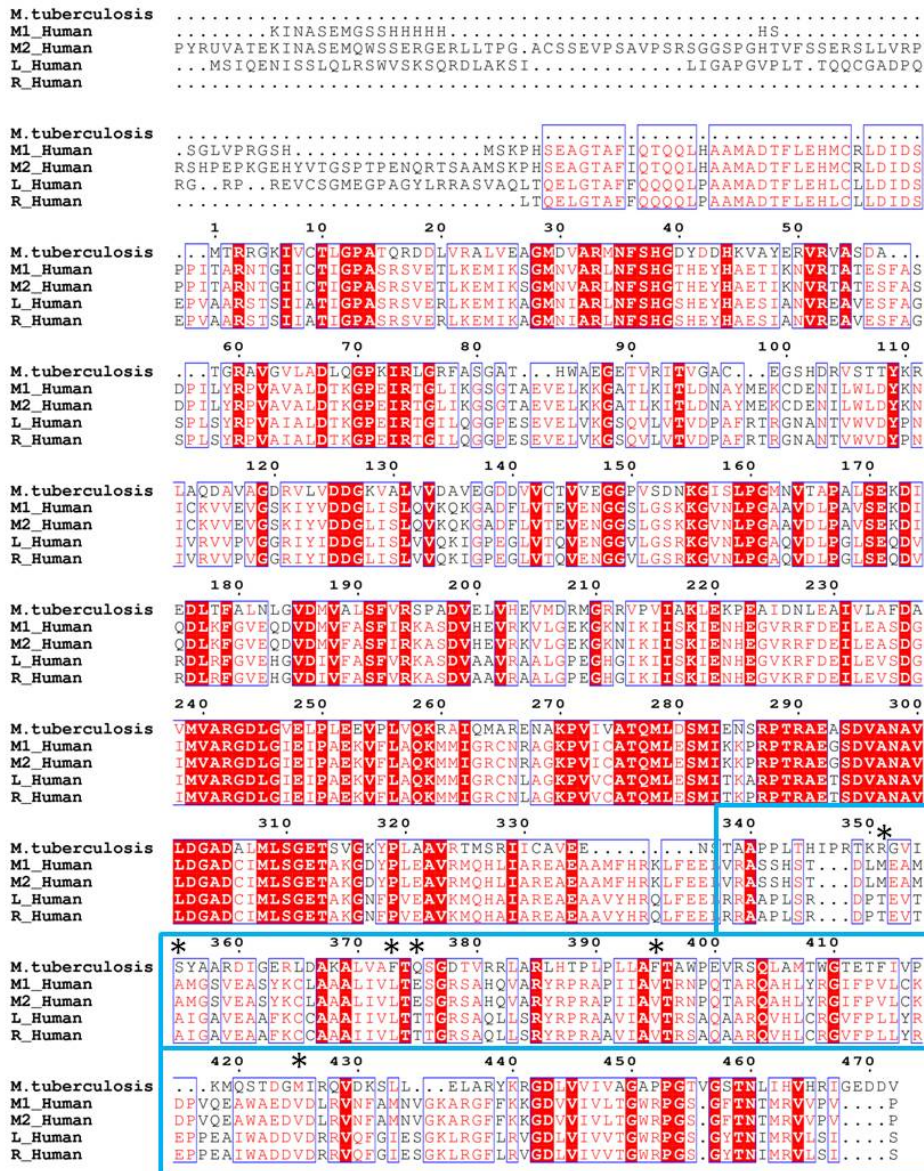


Figure 4-2. Multiple sequence alignment of *Mtb* PykA with all four human isozyms (M1, M2, L, R). The stars mark the directly interacting residues with AMP in *Mtb* PykA, indicating the differences between human orthologs and *Mtb* PykA.

Result and discussion

Overall structure of Mtb PykA

To determine the *Mtb* PykA structure, we cloned and expressed *Mtb* PykA (Rv1617) from *E. coli*, then purified it using affinity chromatography (Ni-column) and size exclusion chromatography (Superdex 200). Size exclusion chromatography confirmed that *Mtb* PykA (472 residues) is a functional tetramer in the solution. *Mtb* PykA was crystallized in 15-18% PEG3350 and 0.15-0.30 M ammonium acetate, and the resulting crystal belonged to space group P1 with unit cell parameters $a = 93.7$, $b = 110.5$, $c = 135.1$, $\alpha = 85.9$, $\beta = 87.9$, $\gamma = 83.7$. It was impossible to solve the structure with a higher symmetry space group than P1, the molecular replacement solution only resulted in a correct structure solution in that space group. The crystal structure of *Mtb* PykA was solved at 2.1 Å, and refined to an overall crystallographic R-factor of 19% and a free R-factor of 23%. Data collection and refinement statistics for *Mtb* PykA are shown in Table 4-1. The asymmetric unit contains eight monomeric molecules, forming two homotetramers as shown in Figure 4-3. The inter-subunit interactions define two (large and small) contact areas. The A/A' interface is constructed by $A\alpha_6$, $A\alpha_7$, and $A\alpha_8$ helices of the $(\alpha/\beta)_8$ barrel of adjacent subunits with a vertical twofold axis. The C/C' interface is maintained by hydrogen bonds from the adjacent $C\alpha_1$ helices and $C\beta_5$ strands of the C-domains.

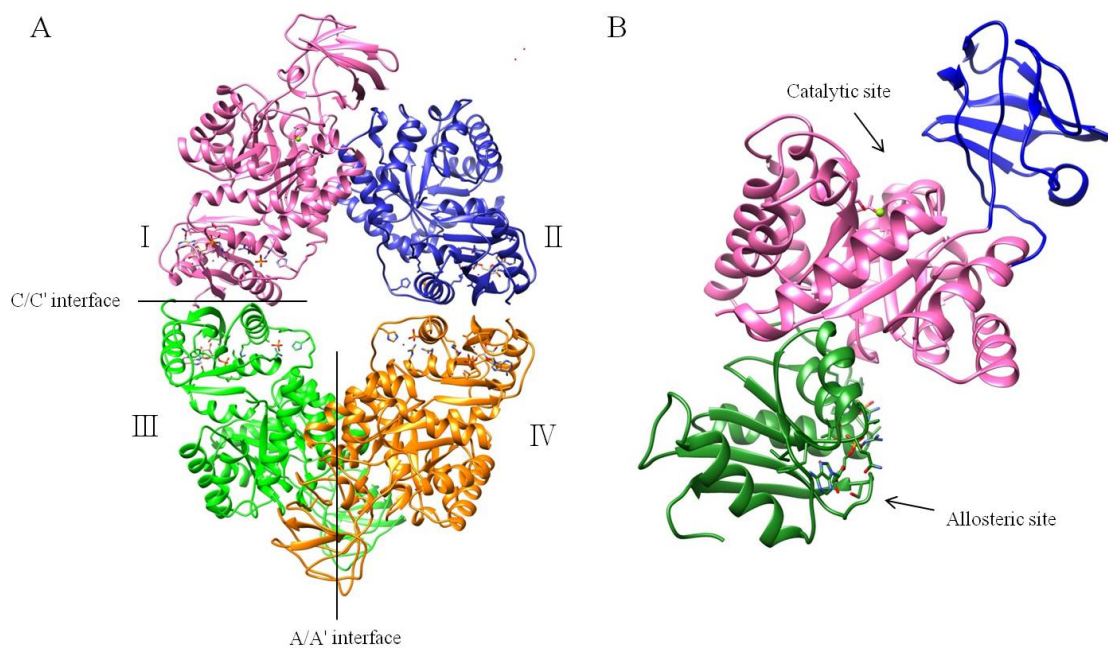


Figure 4-3. *Mtb* PykA structure. (A) homo-tetramer, B-domain was not built due to lack of electron density in subunit II. A/A' and C/C' interfaces between monomers are shown as lines, and (B) Each subunit has an A- (pink), B- (blue), and C- (green) domain. The catalytic site is located between the A and B domains, and the allosteric effector binding site is located in the C-domain.

The overall *Mtb* PykA architecture was similar to reported other species pyruvate kinase structures, having each a subunit composed of three A-, B-, and C-domains (Figure 4-3). The A-domain (residues 2-70 and 168-338) constitutes the central eight stranded α/β barrels. The B-domain (residues 71-167) consists of antiparallel β -sheets, and is known as a flexible domain. While electron density was fully observed for the A- and C-domains, density for the B-domain was completely disordered in the subunit II and we were unable to build it. Similar observation of the flexibility of the B-domain has been reported in many pyruvate kinase crystal structures, including *E. coli*, human and *L.*

mexicana. This flexibility of the B-domain is important for catalysis^{85,86,88}. The C-domain (residues 339-472) is composed of five β -strands surrounded by four α -helices.

The allosteric activator binding site is located in the C-domain. In the *Mtb* PykA structure, a co-purified AMP molecule and a PO_4^{2-} ion were observed in the allosteric effector binding pocket of each subunit, according to the Fo-Fc electronic density map (Figure 4-4). The *Mtb* PykA -ADP bound structure was determined in this study to explore the catalytic site. The complex crystal was obtained by soaking the apo-crystal into 10 mM ADP and Mg^{2+} in the mother liquid (15-18% PEG 3350 and 0.15-0.30 M ammonium acetate). The crystal space group is P1 with cell parameters $a = 93.6$, $b = 110.7$, $c = 136.1$, $\alpha = 85.6$, $\beta = 81.5$, $\gamma = 83.5$. The electron density of ADP and Mg^{2+} only appears in subunit G, which was used as a 'reference subunit' for the analysis of the ADP complex *Mtb* PykA structure in this chapter. The data collection and refinement statistics for AMP- PO_4^{2-} -bound complex structure and ADP- Mg^{2+} -AMP- PO_4^{2-} complex structure are shown in Table 4-1. They did not show notable differences as similar cell parameters, indicating that conformational changes do not occur between subunits.

Table 4-1. Data collection and refinement statistics for *Mtb* pyruvate kinase (PykA) in complex with AMP-PO₄²⁻ and AMP-ADP-PO₄²⁻.

	<i>Mtb</i> PykA-AMP-PO ₄ ²⁻ - complex	<i>Mtb</i> PykA-AMP-ADP- PO ₄ ²⁻ -complex
Wavelength (Å)	1.54	0.98
Resolution range (Å) ¹	50 - 2.1 (2.14 - 2.1)	42.69 - 2.13 (2.21 - 2.13)
Space group	P1	P 1
Unit cell	93.714 110.496 135.075 85.87 81.49 83.65	93.634 110.753 136.071 85.59 81.52 83.47
Total reflections	913375	881730
Unique reflections	203637	288351 (24036)
Multiplicity ¹	3.3 (2.0)	3.9(3.6)
Completeness (%) ¹	97.88 (91.76)	96.33 (80.58)
Mean I/sigma(I)	5.58 (0.42)	7.56 (0.65)
Wilson B-factor	49.53	39.36
R-merge ¹	0.129 (0.709)	0.159 (0.00)
R-work ¹	0.1891 (0.2886)	0.2167 (0.3375)
R-free ¹	0.2291 (0.3407)	0.2624 (0.3592)
Number of atoms	28051	28998
Macromolecules	26471	27155
Ligands	41	67
Water	1539	1776
Protein residues	3486	3580
RMS (bonds)	0.009	0.009
RMS (angles)	1.40	1.40
Ramachandran favored (%)	95	96
Ramachandran outliers (%)	1.8	1.6
Clashscore	8.74	8.97
Average B-factor	43.80	48.30
Macromolecules	43.80	48.30
Ligands	46.60	62.90
Solvent	43.60	47.30

¹. Numbers in parenthesis indicates data for highest-resolution shell.

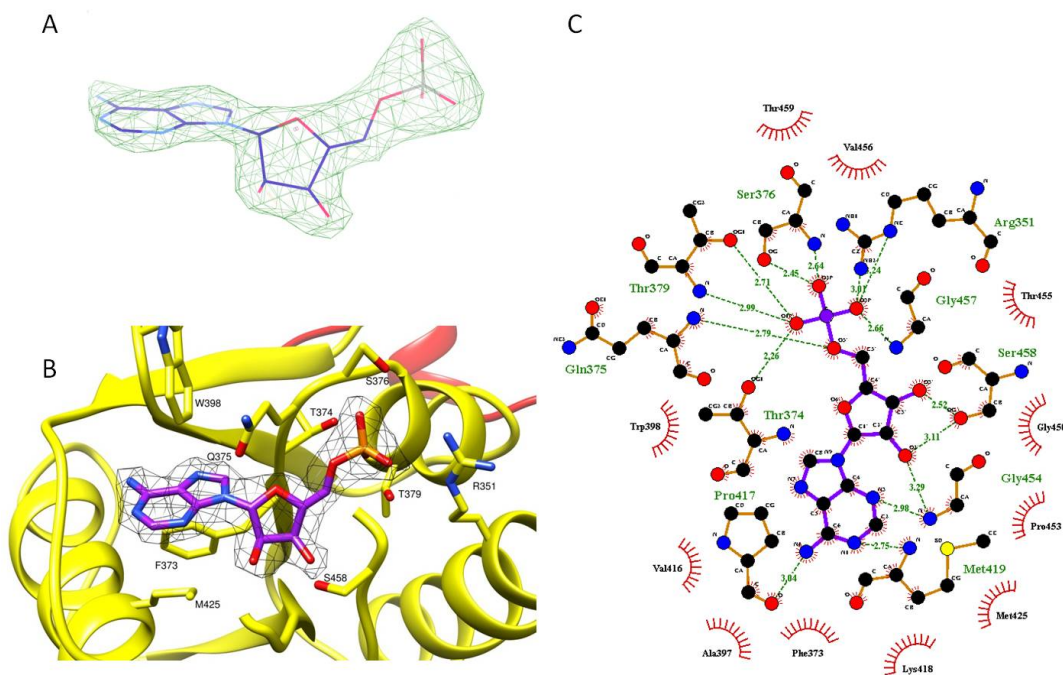


Figure 4-4. AMP molecule found in the allosteric effector binding site. (A) Refinement without ligand shows Fo-Fc density (green) contoured at 4.2σ for the AMP (blue) molecule. The same quality maps are observed in all four allosteric activator sites of the tetramer in the asymmetric unit. (B) The $2Fo-Fc$ map of AMP contoured at 3.0σ with residues involved in direct interaction. (C) Detailed interaction of AMP in allosteric binding site.

Allosteric effector site in Mtb PykA

The Fo-Fc difference electron density map shown in Figure 4-4 produced strong positive features that were indicative of an adenosine monophosphate (AMP) molecule in the allosteric site in the *Mtb* PykA structure. The crystal structure of *Mtb* PykA revealed the AMP binding site for the first time, opening windows for designing organism-specific inhibitors targeting *Mtb* PykA enzyme. All eight subunits have one AMP molecule in the same location in each C-domain. The detailed interactions between

AMP and the *Mtb* PykA enzyme are shown in Figure 4-4C. The adenine ring lies on several hydrophobic residues pocket including W398, F373, and M425 for π - π interactions at 4.0, 3.5, and 4.1 Å, respectively. The C(6)NH₂ forms a hydrogen bond with the backbone carbonyl oxygen of P417. The 1N of adenine interacts with the backbone amide nitrogen of M419 (bond distance of 2.9 Å) and 3N hydrogen bonds to the backbone amide nitrogen of G454 at 2.9 Å. The ribose ring 2'-OH hydrogen bonds to the backbone carbonyl oxygen of G450 and the oxygen of the 3'-OH group forms a hydrogen bond with the hydroxyl group of side chain S458 at 2.5 Å. The phosphate C(5)O picks hydrogen from the backbone amide of Q375 at 2.9 Å. The oxygen from the phosphate group forms a hydrogen bond with the backbone amide nitrogen of G457, and the hydroxyl group of side chains of S376, T379, and T374 at 2.6, 2.6, 2.5, and 2.5 Å, respectively. Interestingly, it has been proposed that the W398 residue is not essential for binding allosteric activator, ribose-5 phosphate in *E.coli* type II PK⁸², but our structure reveals that the aromatic ring of the W398 side chain interacts with the adenosine ring of AMP through π - π interactions in a face-to-edge configuration, indicating it is likely essential for AMP binding.

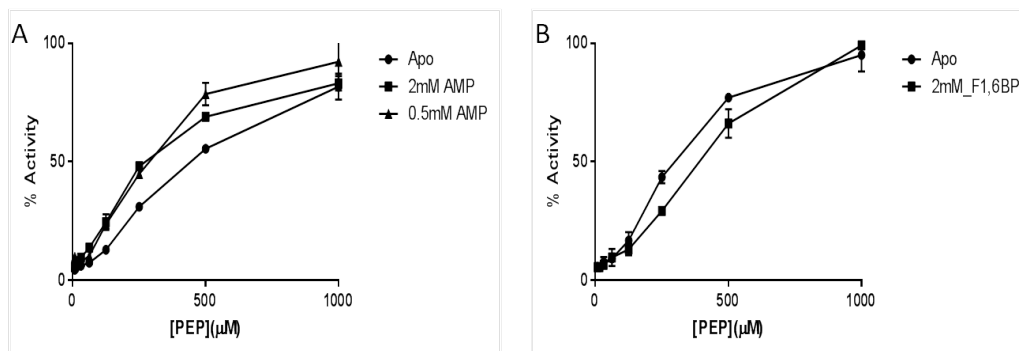


Figure 4-5. The effect of AMP and FBP on *Mtb* PykA. Graph of percent enzyme activity against PEP concentration. (A) 0.5 mM and 2 mM AMP treated in *Mtb* PykA enzyme reaction and compared to untreated enzyme (apo), resulting in enzyme activation by AMP, (B) 2 mM F1, 6BP treated in the reaction, it did not activate the enzyme.

The effect of AMP on the *Mtb* PykA enzymatic activity was shown in Figure 4-5.

The *Mtb* PykA enzyme is activated by AMP, but not FBP. *Mtb* PykA exhibits a sigmodal kinetic response to the binding of PEP in the apo sample, indicating the enzyme is allosterically regulated. In the presence of AMP, the enzyme reaction was activated at a low concentration of PEP. However, F1,6BP did not increase enzyme activity, it showed decreased enzyme activity compared to the apo sample in the absence of F1,6BP. Therefore, the structural evidence of AMP binding in allosteric effector site and the kinetic data are in agreement with the AMP as an allosteric activator of *Mtb* PykA enzyme.

The allosteric activation mechanism of PykA proposes that the activator induces conformational changes to shift the conformational equilibrium of this tetrameric enzyme from the T state (inactive) to the R state (active)⁸⁹. Unfortunately, the apo

structure of *Mtb* PykA was not obtained in this study, thus the T and R states were not compared.

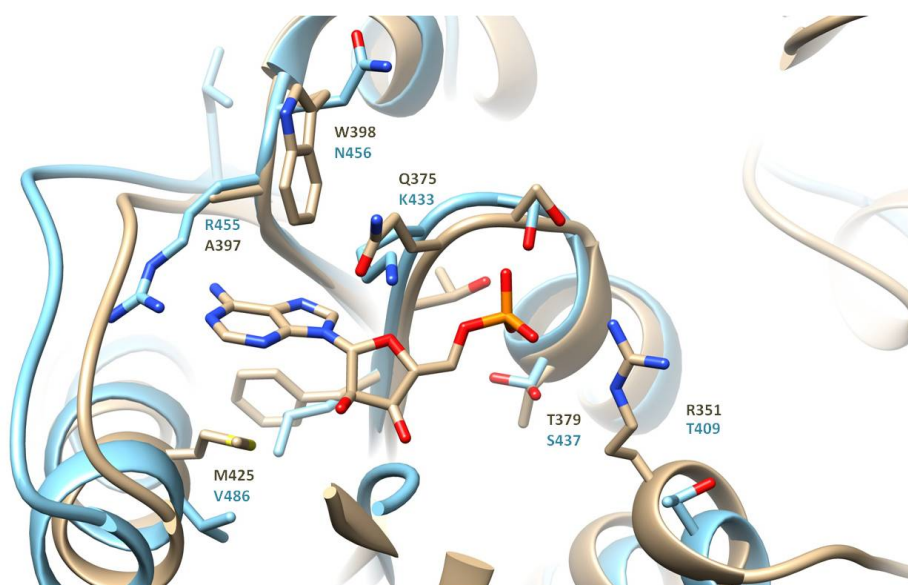


Figure 4-6. Comparison of the allosteric binding site in the *Mtb* PykA (gray) structure with those of the human M2 pyruvate kinase structure (sky blue) (PDB: 4G1N). It reveals the site is not conserved.

Comparison of the allosteric activator binding pocket between human M2 and Mtb pyruvate kinase

The difference in allosteric effectors used by the human (F1,6BP) and *Mtb* (AMP) PykA implies that there is a difference between the effector binding sites. Superimposing the effector binding site of *Mtb* PykA with human M2 shows that several different residues support direct interactions. In *Mtb* PykA the adenine ring binding pocket is surrounded by hydrophobic residues such as M425, A397, and W398, while

the site in human pyruvate kinase contains charged residues like R455 and N453. The AMP phosphate binding pocket also contains different residues, R351 and T379, but they are substituted with the T409 and S437 residues in human M2. A structural alignment of *Mtb* PykA with all four human isozymes (M1, M2, L, R) is shown in Figure 4-2. The blue box represents the sequence of the effector binding site and the stars indicate the residues that directly interact with AMP in *Mtb* PykA, indicating a difference between the human orthologs and *Mtb* enzymes. The differences in the regulatory properties between the pathogen and human create another potential avenue for the development of selective drugs, the unique effector binding pocket of *Mtb* PykA should be explored in future studies.

PO₄²⁻ binding site

Phosphate ions are bound at identical pockets near to the effector binding sites in all eight subunits of the asymmetric unit and have peak heights of greater than 6σ in Fo-Fc map. The phosphate ion interacts with the R385 (2.8 Å), R382 (3.0 Å), H345 (2.7 Å), T349 (2.6 Å) side chain and the backbone amide of the R348 residue (2.7 Å) (Figure 4-7). This phosphate ion binding site is not observed in other species, but is unique to *Mtb* PykA. The site is located 13.2 Å from the phosphate of AMP. The residues, T349, R348, and H345 lie in the loop between the A-domain and the C α 1 helices, which maintains the C/C' interface, and may possibly affect the regulatory mechanism. Unfortunately, due to the lack of an apo *Mtb* PykA structure, we could not further explore the role of the site in this study.

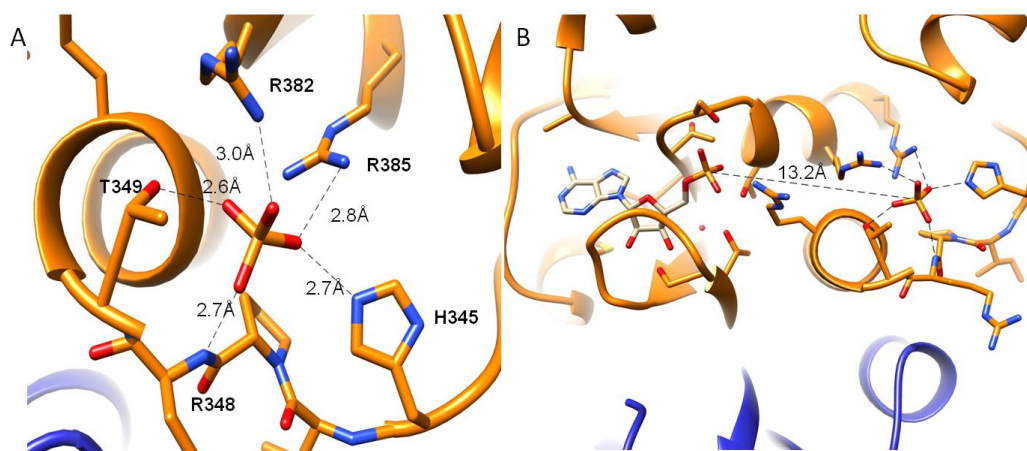


Figure 4-7. Phosphate binding pocket in *Mtb* PykA. (A) Phosphate ion interacting with R382, R385, H345, and T349 side chains and the backbone amide of R348. (B) The distance between the phosphate binding site and allosteric binding site is 13.2Å.

Catalytic site – ADP-Mg²⁺ binding site

The active site in each subunit is located in the cavity between the A- and B-domains. The electron density of the ADP molecule appeared only in the G subunit among eight monomers. In Figure 4-8, the Fo-Fc density map shows the ADP molecule contoured at 3.1 σ . The adenine ring is surrounded by hydrophobic residues H38, P13, and V315. The imidazole ring of H38 residue makes a π - π interaction with the adenine ring of ADP at a distance of 3.3 Å in a face-to-edge configuration. The P13 and V315 side chains provide hydrophobic interactions with the adenine ring. The C-6 amine group of the adenine ring has a water-mediated interaction with the backbone carbonyl oxygen at a distance of 2.7 Å. The N-7 atom of the adenine ring functions as a hydrogen bond acceptor interacting with a water molecule (2.7 Å), and the water interacts with the carbonyl group of the N35 side chain (3.1 Å). The ADP phosphate oxygen forms a hydrogen bond with the carbonyl group of N35 (3.2 Å). Both divalent and monovalent

cations are required for pyruvate kinase activity ⁷². Though the soaking solution contained 10 mM Mg²⁺, no electron density was observed in the active site. The ADP binding geometry in *Mtb* PykA is identical to that observed in the *L. mexicana* pyruvate kinase-ATP complex (PDB code: 3HQP), as the active sites are highly conserved among the pyruvate kinase enzymes from different species. It has been reported that the B-domain moves closer to the A-domain upon the binding of the two substrates, oxalate and ATP in pyruvate kinase enzyme from rabbit muscle ⁹⁰, but the ADP-*Mtb* PykA crystal structure showed ADP binding does not cause conformational changes in the B-domain, it remains in the same position as it is seen in the structure without ADP.

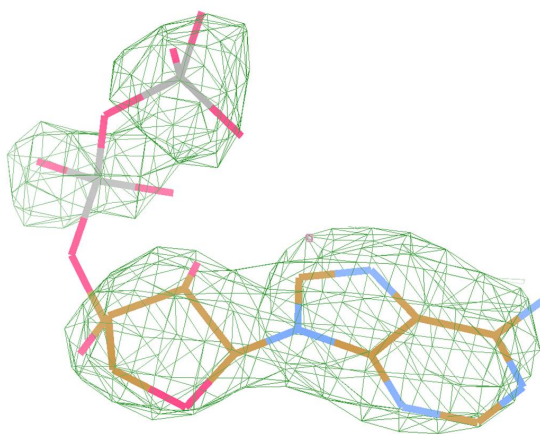


Figure 4-8. The Fo-Fc density map (green) for the ADP bound in the active site of subunit G is shown contoured at 3.1 σ .

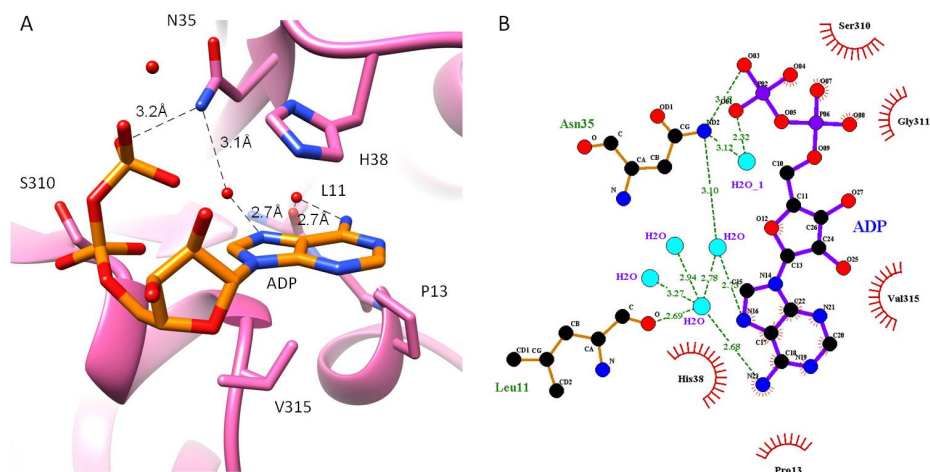


Figure 4-9. ADP bound to the active site of the *Mtb* PykA enzyme. (A) ADP interacting residues. (B) Ligplot of ADP binding with water molecules.

Conclusion

The complex *Mtb* PykA structure with ADP as a substrate and AMP as an allosteric effector were determined. Structural and biochemical study of *Mtb* PykA demonstrated that its activity is regulated by AMP molecule. Despite of highly conserved active site, detailed structural studies of *Mtb* PykA reveals considerable difference in an allosteric effector binding site between host and pathogen pyruvate kinase enzyme. The binding mode of AMP molecule was identified for the first time, structural evidence of *Mtb* PykA suggests optimistic avenues for the design species-specific drug of new TB chemotherapy.

Experimental procedures

Cloning, protein expression, and purification of M. tuberculosis PykA (Rv1617)

The *M. tuberculosis* Pyruvate Kinase (PykA) (Rv1617) gene was amplified by PCR, ligated into restriction site NdeI and HindIII of the pET28a vector and then transformed into *Escherichia coli* BL21 (DE3) cells (Novagen) for over-expression. The cells were grown to an OD₆₀₀ of 0.7-0.8 at 37 °C in Luria-Bertani medium containing 50 µg ml⁻¹ kanamycin. *Mtb* PykA expression was induced with 0.5 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG). Cell growth continued at 293 K for 18 h after IPTG induction, and the cells were harvested by centrifugation at 4200 g for 15 min at 277 K. The cell pellet was resuspended with lysis buffer (20 mM Tris-HCl pH 7.0, 500 mM sodium chloride, 5 mM imidazole and 1 mM phenylmethylsulfonyl fluoride (PMSF)) and the cells were lysed by French press. The cell debris was centrifuged at 36000 g for 50 min at 277 K. The *M. tuberculosis* PykA with an N-terminal poly-His tag in the supernatant fraction was purified by nickel affinity chromatography (Ni-NTA resin (Qiagen)) followed by gel filtration (HiLoad 26/60 Superdex 200 prep-grade column (GE Healthcare)). The buffer used during gel filtration was 20 mM Potassium phosphate (pH 7.0), 50 mM KCl, 2mM DTT, 5% glycerol and 0.1mM EDTA. The protein was then concentrated to 12 mg/ml in using a Spin-X UF concentrator (Corning). The protein concentration was estimated by measuring absorbance at 280 nm by using the Nano drop (Thermo Science), employing the calculated extinction coefficient of 26600 M⁻¹ cm⁻¹ (SWISS-PROT; <http://www.expasy.ch/>).

Crystallization and data collection of M. tuberculosis PykA

Mtb PykA was crystallized by hanging drop vapor diffusion at 16 °C.

Crystallization was achieved in a 4 µl drop consisting of 2 µl well solution [15% (w/v) PEG 3350, 0.2-0.4 M Ammonium Acetate] and 2 µl of protein. *Mtb* PykA -ADP complex was co-crystallized in condition of 10 mM ADP and 10 mM MgCl₂ with protein sample, after 30 min incubation on ice. Diffraction quality crystals were obtained in a 4 µl drop consisting of 2 µl well solution [15% (w/v) PEG 3350, 0.2-0.4 M Ammonium Acetate] and 2 µl of protein. Crystals were cryo-protected through soaking for a second in solution [10% (w/v) PEG3350, 0.1M Ammonium Acetate] containing 20% glycerol (v/v). Diffraction data were collected at a wavelength of 1.542 Å using a Rigaku R-Axix IV ++ at home source and at the Advanced Photon Source beamline 23-ID using a MAR 300 CCD detector (MarMosaic from Marresearch Charged Coupled Device) (Rayonix LLC, Evanston, IL). The HKL2000 program package (HKL Research, Inc., Charlottesville, VA) was used for integration and scaling.

Structure determination and refinement of M. tuberculosis PykA

The structure of *Mtb* PykA was solved by molecular replacement (MR) with PHASER program (University of Cambridge, Cambridge, U.K.). The coordinates of pyruvate kinase from E.coli (PDB entry 1PKY) were used as a search model. Rigid body refinement, followed by simulated annealing refinement at 5000 K, was conducted using PHENIX (Python-based Hierarchical Environment for Integrated Xtallography). Subsequently, refinement was conducted in alternating cycles of manual model building

in COOT (Crystallographic Object Orientated Toolkit), followed by refinement in PHENIX until the R factors converged.

In vitro acetylation assay of M. tuberculosis PykA

Acetylation assay was performed in buffer condition of 25 mM Tris-HCl pH 7.5, 100 mM NaCl contained 100 μ M Acetyl CoA and 0.67 mg/ml *Mtb*-PatA in the presence of 200 μ M cAMP. Reactions were started by the addition of 20 μ M *Mtb* PykA and incubated for 10 min at 25 °C. After the reaction was quenched by boiling in SDS loading dye, it was analyzed in parallel with SDS-PAGE and western blotting with antibody to acetylated lysine (Cell Signaling Technology, cat. no. 9441S; 1:2,000 dilution) detected by chemiluminescence.

Trypsin digestion in solution

Trypsin was used for the protein digestion of *M. tuberculosis* PykA to find out acetylated peptides. The protein was diluted in the buffer of 50mM Ammonium Bicarbonate and incubated 5min at 95°C. Then 0.1 μ g/ μ l trypsin was added to the buffer contained *Mtb* PK and incubated overnight at 37°C.

Mass-spectroscopy of M. tuberculosis pyruvate kinase to identify acetylated peptides

Digested samples were concentrated and desalted by using ziptip with C18 resin (EMD Millipore). Samples were spotted on MALDI plate with matrix (α -Cyano-4-hydroxycinnamic acid 7mg/ml 70% acetonitrile and 0.1% Trifluoroacetic acid). MS and MS/MS spectra were collected by AB SCIEX 4800 Plus MALDI TOF/TOF™ Analyzer. The analysis of searching acetylated peptides were performed using Protein Prospector (<http://prospector.ucsf.edu/prospector/>)

CD spectroscopy

The CD spectra of *Mtb* PK proteins, wild-type, K110Q, K110R, and K110A mutants were analyzed between 180-260nm at 0.5-nm intervals by using a Chirascan™ CD Spectrometer (Applied photophysics). The solution contains 8μM proteins and distilled water.

Enzyme assay of M. tuberculosis PykA used for HTS

The pyruvate kinase activity can be measured with coupled the lactate dehydrogenase (LDH) reaction. *Mtb* PykA catalyzes pyruvate production from PEP by converting ADP to ATP in presence of divalent metal, Mg²⁺. Then lactate dehydrogenase (LDH) produces lactate from pyruvate by converting NADH to NAD⁺. Therefore, a decrease in NADH absorbance was monitored at 340nm for PykA enzyme activity (Figure 4-10).

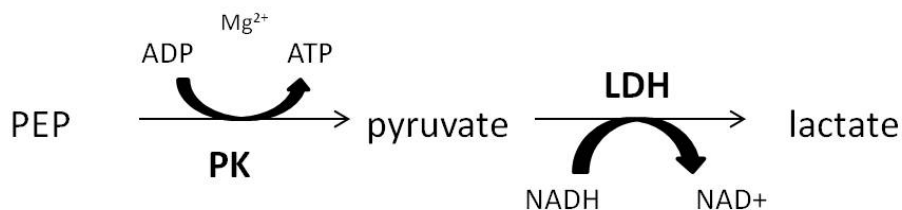


Figure 4-10. Schematic of the *Mtb* PykA enzyme assay coupled LDH enzyme reaction. The PykA activity was measured in decrease of NADH absorbance at 340nm.

The 4nM purified *Mtb* PykA enzyme was used for measuring enzyme assay in solution contained 0.5mM ADP, 0.25 mM NADH, 2.5 mM MgCl₂ and 1unit of LDH.

The reaction was initiated by addition of 0.5mM PEP after pre-incubated with inhibitors for 20min at room temperature.

Additional results

Lysine acetylation of Mtb PykA – identification of acetylated lysine

The pyruvate kinase enzyme is regulated in multiple ways, including gene expression, allosteric regulation by small metabolites, and post-translational modification. Major post-translational modification, lysine acetylation, is commonly found in metabolic enzymes⁹¹. Rv0998, is a cAMP-regulated PAT of *Mtb* (*Mtb* PatA) that is a key enzyme in the adaptation to different carbon sources and shifting to slow growth^{75,92}. It has been proposed that acetyl coenzyme A (acetyl-CoA) synthetase (ACS) is acetylated by *Mtb* PatA⁹³, but the true substrates of *Mtb* PatA have not yet been determined. Christopher M. Sasseti's group identified acetylated peptide (VSTTYk(ac)QLAQDAAR) on *M. smegmatis* pyruvate kinase (*M. smeg* PykA) when incubating with recombinant *Mtb* PatA (data not published).

We attempted to investigate whether *Mtb* PykA is the substrate for *Mtb* PAT, we incubated *Mtb* PykA with *Mtb* PAT (Rv0998) (Method in chapter x) and digested the sample with trypsin. We analyzed the resulting peptide fragments using LC/MS/MS. However, the expected lysine acetylated peptide fragment (VSTTYK*R) was not found in the digested sample though it had been in the unmodified sample, as shown in figure 4-11. Trypsin preferentially cleaves positively charged residues such as lysine or arginine. However, trypsin digestion is limited to acetylated lysines due to both steric

effects and the loss of positive charge on the ϵ -amine group of lysine, which disrupts interaction with the aspartate residue of the trypsin catalytic center⁹⁴. Therefore, the VSTTYKR fragment could have been acetylated resulting in the missing peak in the modified sample. Many trials to identify the acetylated fragment were unsuccessful, further analysis is required for this study.

	Unmodified sample PK enzyme + AcetylCoA + cAMP	Modified sample PK enzyme + AcetylCoA + cAMP + <i>Mtb</i> PAT
	Peptide R.DIGER.L R.DDLVR.A R.VFVIK.L R.IGEDDV.- K110 → R.VSTTYK.R R.VLVDDGK.V K.SLLELAR.Y R.LAQDAVAGDR.V K.MQSTDGMIR.Q R.ALVEAGMDVAR.M R.AVGVLADLQGP.K.I K.ALVAFTQSODTVR.R K.ALVAFTQSODTVR.R R.MNFSHGDYDDHK.V R.SPADVELVHEVMDR.I R.SPADVELVHEVMDR.I R.FASGATHWAEGETVR.I R.FASGATHWAEGETVR.I K418 → R.SQLAMTWGTSTFIVPK.M R.GDLGVLPLEEVPVQK.R R.GDLGVLPLEEVPVQK.R R.LHTPLPLAFTAWPEVR.S K.DIEDLTFALNLGVDNVALSEFVR.S R.GDLVVIVAGAPPQTVGSTNLIHVHR.I R.GDLVVIVAGAPPQTVGSTNLIHVHR.I	Peptide R.DIGER.L R.DDLVR.A R.VFVIK.L R.IGEDDV.- R.VLVDDGK.V R.LAQDAVAGDR.V K.MQSTDGMIR.Q R.ALVEAGMDVAR.M R.AVGVLADLQGP.K.I K.ALVAFTQSODTVR.R R.SPADVELVHEVMDR.I R.SPADVELVHEVMDR.I R.FASGATHWAEGETVR.I R.FASGATHWAEGETVR.I K.GISLPGMNVTAPALSEK.D R.GDLGVLPLEEVPVQK.R R.GDLGVLPLEEVPVQK.R R.LHTPLPLAFTAWPEVR.S R.GDLVVIVAGAPPQTVGSTNLIHVHR.I R.GDLVVIVAGAPPQTVGSTNLIHVHR.I R.AEASDVANAVLDGADALMLSGETSVGK.Y K.LEKPEAIDNLEAIVLAFDAVMVAR.G K.LEKPEAIDNLEAIVLAFDAVMVAR.G

Figure 4-11. Trypsin digested peptide sequence identified from an LC/MS/MS comparison of unmodified and modified samples. Unmodified sample ran in the absence of acetylase (*Mtb* PAT), while modified sample ran in the presence of acetylase (*Mtb* PAT). The peptide fragments, VSTTYK(110), and SQLAMTWGTSTFIVPK(418) are not present in the LC/MS/MS result of modified sample.

Lysine acetylation of Mtb PykA – site-directed mutagenesis at the K110 site

K112 in *M. smegmatis* is identical to the K110 residue in the B-domain of *Mtb* PykA. To establish the role of K110 in *Mtb* PykA, we performed site-directed mutagenesis to make K110A, K110Q, and K110R mutants. As shown in figure 4-12, the K110A mutant exhibits about 25% decreased activity and the K110Q and K110R mutants show 60% decreased activity compared to the wild-type. The CD analysis of each mutant revealed that no significant structural changes occurred (Figure 4-13). In figure 4-14, K110 was located in B-domain and pointed to solvent area without interaction with vicinity residues. This suggests that acetylated K110 might affect B-domain flexibility inducing loss of substrate affinity. However, the mutants did not completely impair enzyme activity, thus, further study is required to conclude whether *Mtb* PykA is one of the substrates for *Mtb* PAT or not.

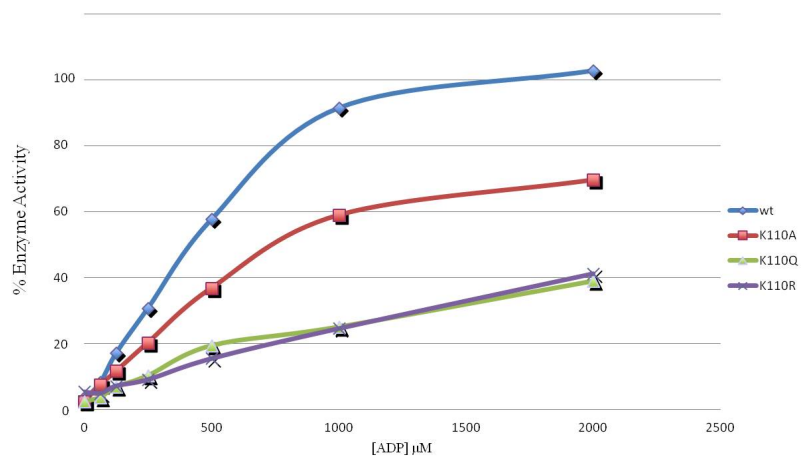


Figure 4-12. The enzyme activity of wild-type (blue) and mutants of *Mtb* PykA, K110A (red), K110Q (green), and K110R (purple). K110A mutant exhibits about 25% decreased activity and the K110Q and K110R mutants show 60% decreased activity compared to the wild-type.

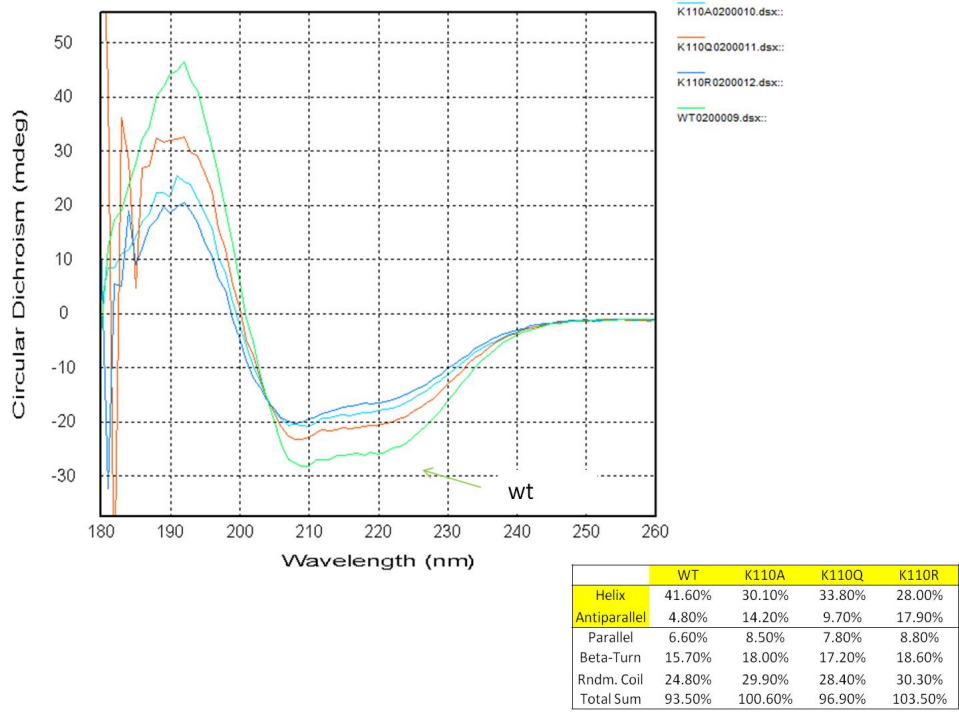


Figure 4-13. CD result of wild-type and mutants of *Mtb* PykA, K110Q, K110R, and K110A. The table shows no major conformational changes occur in mutants.

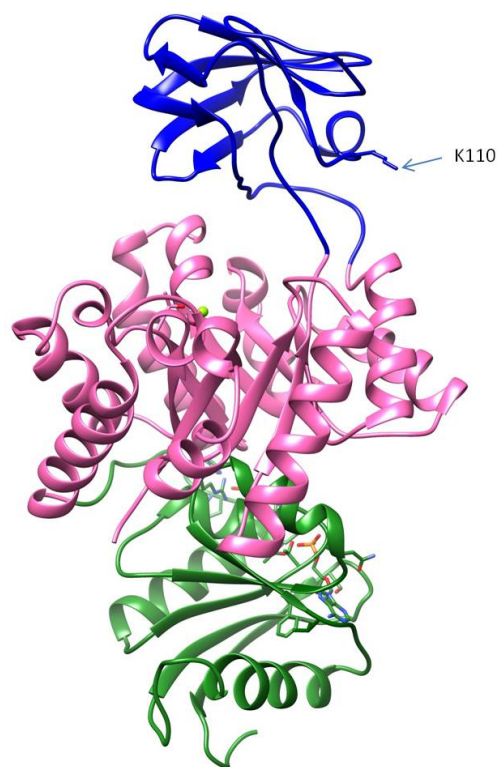


Figure 4-14. Location of K110 in the B-domain. K110 residue is pointing to solvent area, no major interaction found with adjacent residues.

CHAPTER V

IDENTIFICATION OF POTENTIAL INHIBITORS TARGETING PYRUVATE

KINASE FROM *MYCOBACTERIUM TUBERCULOSIS*

Introduction

The pyruvate kinase enzyme is ubiquitous to most of living cells as it is a critical enzyme in the glycolytic pathway. Pyruvate kinase is an essential gene in methicillin-resistant *Staphylococcus aureus* (MRSA), *Haemophilus influenzae*, and *Streptococcus pneumoniae*^{95,96,97} and has been studied as a potential drug target against *Trypanosoma brucei* and *Leishmania Mexicana* pyruvate kinase enzyme⁹⁸ due to they rely on glycolytic pathway to produce catabolic source of ATP. The pyruvate kinase isoenzyme M2 (PKM2) is an anti-cancer drug target due to its role in controlling glycolysis speed and reprogramming the glycolytic flux to provide the special metabolic needs of proliferating cancer cells^{99,100}.

The successful examples in the development of species-specific inhibitors against pyruvate kinase enzyme have been reported to date. The C-C' interface site of the pyruvate kinase enzyme from methicillin-resistant *Staphylococcus aureus* (MRSA) has been targeted as a druggable site. The C-C' interface site is formed by the interaction of the C α 1 helices and the C β 5 strands in each subunit's C-domain (Figure 4-3). Bis-indole alkaloid inhibitors show nano-molar IC₅₀ values for MRSA pyruvate kinase, where they bind to the C-C' interface causing tetramer rigidity and ineffective catalysis¹⁰¹. Another example, dye-like molecules, suramin and the trypan blue mimetics, inhibit *Leishmania*

mexicana (*Lm*) pyruvate kinase through binding in ATP binding site, these inhibitors selectively inhibits PKs from humans¹⁰². These previous results indicate positive view of the development of selective inhibitors against *Mtb* PykA. The sequence alignment (Figure 5-1) compares the C α 1 helices and C β 5 strands in the C-domain of the *Mtb* and human pyruvate kinase isoforms, indicating which amino acids in this region are not highly conserved among species. Therefore, the C- C' interface can also be exploited for the design of pathogen-specific inhibitors. In addition, as discussed in Chapter 4, we highlighted the sequence divergence of pyruvate kinase between the bacterial and mammalian enzymes, especially the allosteric regulator binding site. This also provides an opportunity for the design of *Mtb* PykA specific inhibitors which would not disturb human glycolysis.

		349		366
	<i>M.tb</i> PK	T	K R G V I S Y A A R D I G E R S D	
	M1_Human	D	L M E A M A M G S V E A S Y K C L	
C α 1 helices	M2_Human	D	L M E A M A M G S V E A S Y K C L	
	L_Human	D	P T E V T A I G A V E A A F K C C	
	R_Human	D	P T E V T A I G A V E A A F K C C	
		458		466
	<i>M.tb</i> PK	S	T N L I H V H R I	
	M1_Human	F	T N T M R V V P V	
C β 5 strands	M2_Human	F	T N T M R V V P V	
	L_Human	Y	T N I M R V L S I	
	R_Human	Y	T N I M R V L S I	

Figure 5-1. Sequence alignment of the C- C' interface composed of C α 1 helices (349-366, numbering in *Mtb* PykA (or PK) and C β 5 strands (458-466 residues) of each subunit in the *Mtb* and human pyruvate kinase isozymes.

Transposon mutagenesis of the H37Rv strain showed that *Mtb* PykA (Rv1617) is an essential gene for *in vitro* growth of *Mtb* on glycerol and fatty acids³⁰, indicating that the *Mtb* PykA enzyme might be related to TB pathogenesis. A knock-out mutant of the *Mtb* PykA gene failed to grow *in vitro* on glycerol as the sole carbon source¹⁰³. Despite the importance of *Mtb* PykA for TB growth and numerous inhibitor studies for the pyruvate kinase enzyme across divergent species, no inhibitor studies have been done against *Mtb* PykA. As is common for multi-species housekeeping enzymes, pyruvate kinase exhibits a conserved architecture and active site. The presence of highly conserved pyruvate kinase homologue in host may hamper development of inhibitors targeting on *Mtb* PykA. To search for *Mtb* PykA inhibitors, a chemical library of 1428 compounds (Sac2 hits library) that inhibit *Mtb* (mc² 7000), and the predicted pyruvate kinase inhibitors from *in silico* screening were tested against *Mtb* PykA. Several hits were found to inhibit *M. tuberculosis* growth with low micromolar range MICs and with IC₅₀s in the nanomolar range. One of the hits bound to the ATP/ADP binding site, this result provides potential scaffolds for the development of more potent inhibitor targeting on *Mtb* PykA.

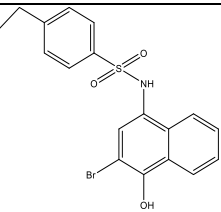
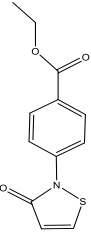
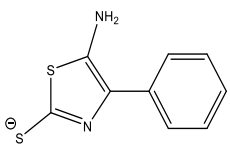
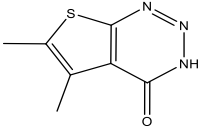
Result and discussion

In vitro HT screening of the anti-tuberculosis active compounds library

To identify lead molecules inhibiting *Mtb* PykA, 1428 anti-tuberculosis active compounds (TB growth inhibition on dextrose and acetate as dual carbon sources) were screened against the enzyme from our libraries (Sac2 hit libraries). Twelve hits exhibited

more than 60% enzyme inhibition at 20 μM . Four compounds among 12 hits showed an inhibitory activity against both whole-cell and enzyme. Binding test against enzyme for these four hits was tested via DSF. The chemical structures of compounds, IC_{50} , MIC, and ΔTm values are listed in Table 5-1.

Table 5-1. Distinct structural classes of four hits were identified that exhibit inhibition of both the *Mtb* PykA enzyme and the whole-cell (mc^27000).

ID	Structure	ΔTm	IC_{50} (μM)	MIC (mc^27000) (μM)
LC:F0808-1900		2	0.76	10
MB_P:KM04416		-3	0.04	5-10
LC:F2147-0563		1	0.23	5-10
LC:F3284-4918		2	0.06	10

LC:F3284-4918 (5,6-dimethylthieno[2,3-d][1,2,3]triazin-4(3H)-one) caused a two degree shift in the DSF assay with an IC₅₀ value of 0.06 μM and an MIC value of 10 μM. LC:F2147-0563 (5-amino-4-phenylthiazole-2-thiolate) showed a one degree shift in the thermal shift assay, and an IC₅₀ value of 0.23 μM. MB_P:KM04416 (ethyl 4-(3-oxoisothiazol-2(3H)-yl)benzoate) caused a negative shift of about three degrees, with an IC₅₀ of 0.04 μM. LC:F0808-1900(N-(3-bromo-4-hydroxynaphthalen-1-yl)-4-ethylbenzenesulfonamide) resulted in a two degree shift with an IC₅₀ value of 0.76 μM and an MIC value of 10 μM. Two compounds, KM01446 and F0808-1900 compounds were tested on human fibroblasts cell for testing cytotoxicity with compound concentration in range of from 0 to 50μM. While KM01446 compound showed toxicity at 12.5μM of compound, F0808-1900 is non-toxic to cells up to 25μM range. However, trials (soaking and co-crystallization) in obtaining a complex structure to determine the binding mode of these compounds were unsuccessful.

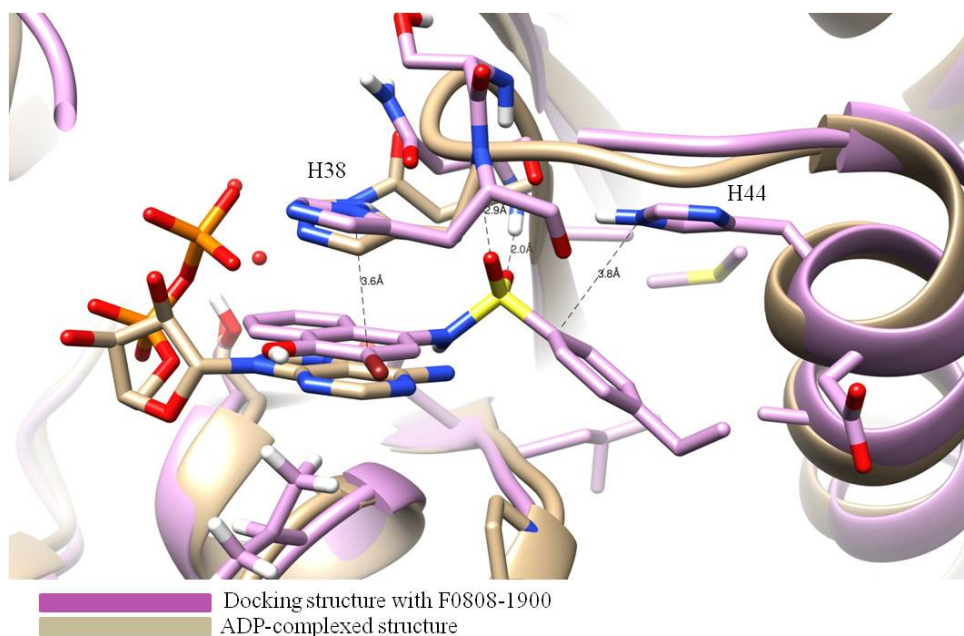


Figure 5-2. Predicted binding mode of F0808-1900 (N-(3-bromo-4-hydroxynaphthalen-1-yl)-4-ethylbenzenesulfonamide) with *Mtb* PykA discovered in docking. Docking structure with F0808-1900 was colored in light purple, and ADP complex structure of *Mtb* PykA was colored in grey. F0808-1900 binds to ADP/ATP binding pocket in the active site.

F0808-1900 bound to the ATP-binding site

The interactions of F0808-1900 with the *Mtb* PykA enzyme were analyzed using a molecular docking approach, Glide (grid-based ligand docking)¹⁰⁴, was done by Dr. Charles Hutchins from Abbvie. The docking result shows the F0808-1900 molecule binds to the ATP binding site, the naphthalene ring has a π - π interaction with the imidazole ring of the H38 side chain (3.6 Å), the site overlaps with the adenine ring as shown in the ADP-complex in the *Mtb* PykA structure (Figure 5-2). The two oxygens of the sulfonamide linker accept hydrogen atoms from the backbone amide of the N35 and H38 residues to form hydrogen bonds at 2.0 Å and 2.9 Å, respectively. The benzyl ring

interacts with the H44 side chain at a distance of 3.8 Å. The interactions of the F0808-1900 ethyl benzyl ring with enzyme are supported by hydrophobic interactions, such as the side chain of the H44 residue (3.8 Å) and the peptide bonds of G12 and P13 (3.2 Å). Four compounds having similar structure with F0808-1900 purchased from ChemBrige vendor were tested against *Mtb* PykA to check which parts of compounds are important for interaction. The results revealed that the naphthalene ring of this compound is critical to inhibit enzyme activity (Figure 5-3).

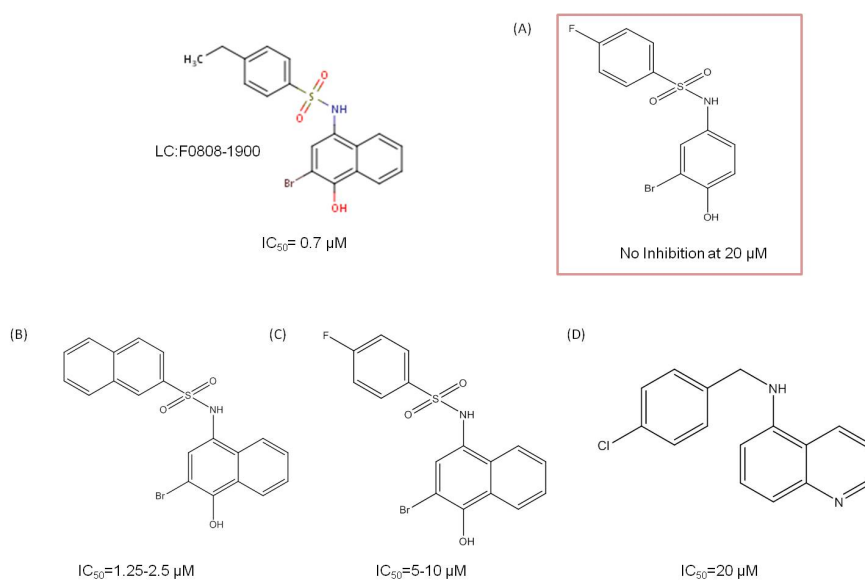


Figure 5-3. Four purchased compounds which are similar to “hit” structure (F0808-1900) with IC_{50} values. These suggest the naphthalene ring is important to maintain enzyme inhibition. (A) N-(3-bromo-4-hydroxy-1-naphthyl)-4-fluorobenzenesulfonamide (B) N-(3-bromo-4-hydroxy-1-naphthyl)-2-naphthalenesulfonamide (C) N-(3-bromo-4-hydroxyphenyl)-4-fluorobenzenesulfonamide (D) N-(4-chlorobenzyl)quinolin-5-amine

Comparison ADP binding site of Mtb and human pyruvate kinase

Despite the conservation of the ADP binding area among species, the ethylbenzyl ring sulfonamide of F0808-1900 binding site is different in the host enzyme as shown in the superposed structure of human and *Mtb* pyruvate kinase. Notably, as shown in Figure 5-4, Y83 in human M2 pyruvate kinase will cause steric hindrance with the ethylbenzyl ring, while D43 residue in *Mtb* PykA does not, indicating that this compound may not inhibit the human enzyme. Adjacent to that, there are two more residue differences; M34 and A47 residues in *Mtb* PykA are substituted with L74 and T87 residues in human M2 PK. This structural evidence of docked F0808-1900 with *Mtb* PykA implies that this adjacent pocket can be exploited for the design of selective inhibitors against *Mtb* PykA.

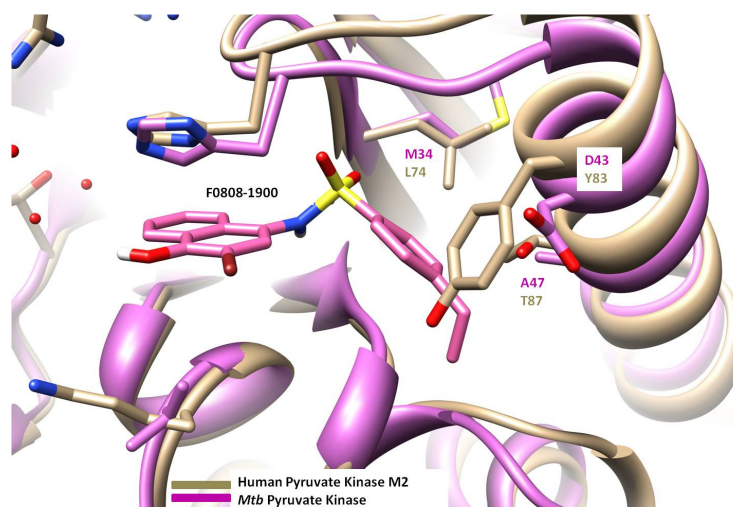


Figure 5-4. Comparison of the docked F0808-1900 compound in the *Mtb* PykA structure (pink) with the human M2 PK (tan; PDB 3GR4). The inhibitor binding site indicates different residues between the human and *Mtb* PykA enzymes, indicating a potential site for selective inhibitor development.

Virtual screening hits were experimentally validated against Mtb PykA

The Abbvie Company conducted virtual screening to identify molecules with inhibition potential against *Mtb* PykA using computational drug discovery tools to generate 531 potential *Mtb* PykA inhibitors. Virtual screening (VS) was based on the ADP binding site in the *Mtb* PykA -ADP complex crystal structure. Glide was applied to predict which molecules would potentially bind to the active site¹⁰⁴. The 531 compounds were tested against the *Mtb* PykA enzyme, resulting in 10 hits showing a more than 75% inhibition. We eliminated six hits due to conflicts with the other institute's projects, leaving us with four hits for further study. The chemical structure of the four hits, the IC₅₀, and MIC₉₀ values are represented in Figure 5-5. A-1261670 (5-(2,5-dimethyl-1H-pyrrol-1-yl)-2-hydroxybenzoic acid) exhibited potent activity against the enzyme (IC₅₀ = 1.2 μM), but it was not effective in the whole-cell assay (MIC₉₀ > 100 μM). Interestingly, this compound shares a notable structural similarity with MB_P:KM04416 and LC:F2147-0563 (a benzene and a five membered ring) (Table 5-1), which were previous hits. These results suggest that a scaffold including benzene and a five membered ring could be modified to increase potency against both the enzyme and the cell.

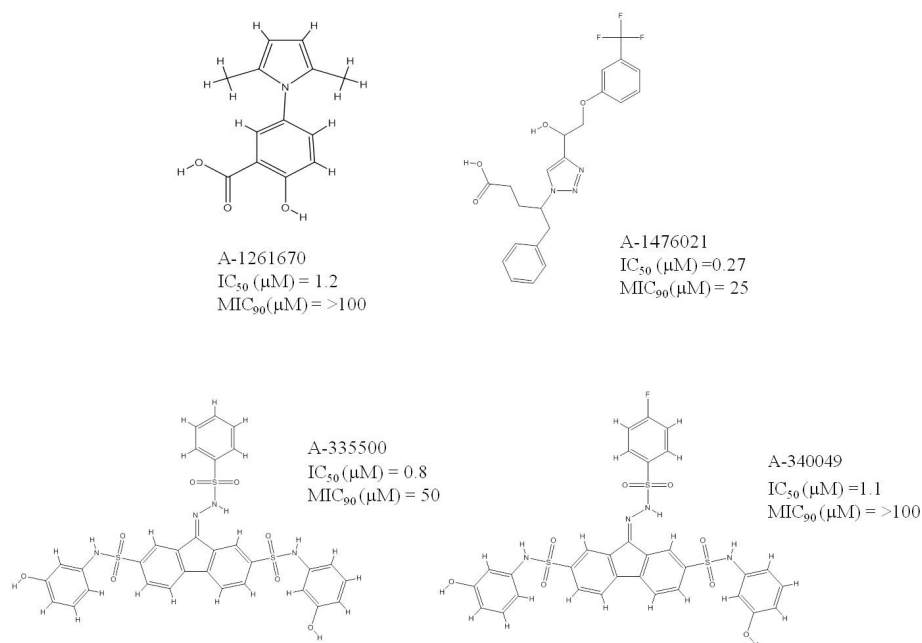


Figure 5-5. The chemical structure of the four identified active compounds from the virtual screen done by AbbVie with IC_{50} and MIC_{90} values for each compound.

A-1476021, 4-(4-(1-hydroxy-2-(3-(trifluoromethyl)phenoxy)ethyl)-1H-1,2,3-triazol-1-yl)-5-phenylpentanoic acid, has a very potent enzyme inhibition, showing an IC_{50} value of 0.27 μM and mild activity against the cell with an MIC_{90} value of 25 μM . Two hits with a close structural resemblance, A-335500 and A-340049, had IC_{50} values of 0.8 and 1.1 μM , respectively, but their whole-cell activities were not quite as strong as A-1476021, perhaps because its larger size might make it difficult to enter the cell. In this study, we used virtual screening and anti-*Mtb* active compounds to search for novel *Mtb* PykA inhibitors. We are conducting trials to obtain complex structures for all the hits to determine their binding mode. Binding orientations and optimization should be

investigated further. These results will provide potential scaffolds for further development of more potent and selective *Mtb* PykA inhibitors.

Conclusion

The screening of our chemical compound libraries (Sac2) resulted in identify four hits which showed potent activity on enzyme at nano-molar range and whole cell at micro-molar range. Docking of F0808-1900 hit into the active site of *Mtb* PykA revealed its binding site was slightly overlapped with the ADP/ATP binding site, but different residues were involved in contact of F0808-1900. Thus, this site could be explored further for the design of selective inhibitors. Also, the naphthalene ring of F0808-1900 is critical for having potent inhibitory activity against enzyme. Hits from virtual screening using *Mtb* PykA structure were screened *Mtb* PykA enzyme, and four hits show very potent inhibitory effect on enzyme. These works provides promising directions for the lead modification and design selective inhibitor strategy.

Experimental procedures

Enzyme assay of M. tuberculosis pyruvate kinase used for HTS

The pyruvate kinase activity can be measured with coupled the lactate dehydrogenase (LDH) reaction. *Mtb* PykA catalyzes pyruvate production from PEP by converting ADP to ATP in presence of divalent metal, Mg^{2+} . Then lactate dehydrogenase (LDH) produces lactate from pyruvate by converting NADH to NAD⁺.

Therefore, a decrease in NADH absorbance was monitored at 340nm for PykA enzyme activity.

The 4nM purified *Mtb* PykA enzyme was used for measuring enzyme assay in solution contained 0.5mM ADP, 0.25 mM NADH, 2.5 mM MgCl₂ and 1unit of LDH. The reaction was initiated by addition of 0.5mM PEP after pre-incubated with inhibitors for 20min at room temperature. The enzyme assay was monitored the decreased NADH absorbance at 340 nm during 5min. The % inhibition was calculated from this equation, %inhibition = (1- (decreased absorbance slope)/control slope (no inhibitor, only DMSO))*100. The value of IC₅₀ was calculated by using Collaborative Drug Discovery (CDD) database. Activity assays were performed with a Thermo Scientific *Multiskan GO* spectrometer. The plate used in the enzyme assay was a 96-well, half area, UV-star plate, (Greiner bio-one).

Differential Scanning Fluorimetry (DSF) assay of M. tuberculosis PykA

The real time (RT)-PCR machine (Stratagene Mx3005P) was set to equilibrate samples at 25°C for 5minutes and then gradually increase temperature to 95°C at a rate of 1°C/minute. The excitation and emission of SYPRO orange can be monitored using filters: FAM (492nm) and ROX (610nm). The RT-PCR instrument (MxPro software) calculated first derivative plot, indicating the lowest point is melting point of the protein^{105 106}. The master mix contained 20µg/ml protein in buffer (20mM Tris-HCl pH 7.5, 50mM NaCl) and 5X SYPRO orange (diluted from 5000X concentrate). Each well of a 96-well plate has 10µl solution contained 9µl master mix and 1µl of 1mM inhibitors

stock (final concentration 100 μ M). It is not recommended exceed DMSO amount not over 2% (v/v), but this case, 10% DMSO does not influence the results.

Whole cell assay for MIC determination on mc²7000 strain

mc² 7000 starting culture in 7H9 containing OADC, Tyloxapol, Pantothenic acid, Malachite green was used to grow until OD₆₀₀ reaches to 1.0. The growing cell from starting culture was added into testing media 7H9 with dextrose. The media consists 0.5% dextrose, 0.085% NaCl, 0.05% tyloxapol, 25 μ g/ml Pantothenic acid, and 0.25 μ g/ml malachite green. Incubation with inhibitors in 96-well plate lasts for 7 days at 37 °C and then the each well was stained with resazurin (5 μ g/ml final) for viability testing.

CHAPTER VI

SUMMARY

New anti-TB drugs are urgently needed to eradicate *M. tuberculosis* from the human population. The *Mtb* PEPCK and PykA enzymes are involved in the CCM, which is a major determinant of the pathogenicity of *Mtb*, but drug development has traditionally shied away from targeting enzymes with human orthologs. I have dedicated the years of my Ph.D to the challenging topic of making progressive effort in targeting the *Mtb* PEPCK and PykA enzymes for the development of efficient anti-TB drugs. In CHAPTER 1, the literature review on TB drug discovery, and the problem of currently used TB drugs were discussed. Also, the prospect of the *Mtb* CCM as target for anti-tubercular drug design was addressed. In CHAPTER 2, I presented the crystal structures of *Mtb* PEPCK in apo and complex forms. Despite a highly conserved active site between the human and *Mtb* enzymes, distinct movements of the flexible loops were observed. A functional study of *Mtb* PEPCK revealed that a unique interaction between the P-loop and the Ω -loop in *Mtb* PEPCK is required to allow efficacious the anaplerotic reaction (PEP to OAA) of PEPCK, which is important to *Mtb*'s pathogenicity. Comparing the structural differences between the human enzyme and *Mtb* PEPCK supported its viability as a target for the design of selective inhibitors. In CHAPTER 3, GTP-competitive inhibitors were introduced and optimized based on structural evidence for inhibitor interaction with *Mtb* PEPCK. This work has created a positive outlook on the success of designing a pathogen-specific anti-tuberculosis

drug. We developed selective inhibitors against *Mtb* PEPCK enzyme. In CHAPTER 4, crystal structures of the *Mtb* PykA enzyme were determined, and the binding mode of an allosteric effecotor, AMP was identified for the first time. Comparison of the allosteric effector binding site between *Mtb* and human revealed important differences in the residues present in the site that could be exploited for the development of selective inhibitors. In CHAPTER 5, identification of inhibitors from HT screening of diverse compound libraries was addressed. The hits were active against both the enzyme and whole-cell in nanomolar and low micromolar ranges, respectively. This works provide potential scaffolds that can be used for further rounds of inhibitor optimization. Structural studies of *Mtb* PEPCK and PykA have provided a better understanding of their role in the gluconeogenesis/glycolysis pathway. The inhibitor studies presented in this dissertation of two important enzymes in the *Mtb* CCM not only reinforce the structure guided inhibitor design strategy but also provide better direction for lead optimization.

REFERENCES

- 1 World Health Organization. Global tuberculosis report 2015, <http://www.who.int/tb/publications/global_report/en/> (2015).
- 2 Keshavjee, S. & Farmer, P. E. Tuberculosis, drug resistance, and the history of modern medicine. *N Engl J Med* **367**, 931-936 (2012).
- 3 Brennan, P. J. & Nikaido, H. The envelope of mycobacteria. *Annu Rev Biochem* **64**, 29-63 (1995).
- 4 Wolucka, B. A., McNeil, M. R., de Hoffmann, E., Chojnacki, T. & Brennan, P. J. Recognition of the lipid intermediate for arabinogalactan/arabinomannan biosynthesis and its relation to the mode of action of ethambutol on mycobacteria. *J Biol Chem* **269**, 23328-23335 (1994).
- 5 Brennan, P. J. Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)* **83**, 91-97 (2003).
- 6 Barry, C. E., 3rd. Interpreting cell wall 'virulence factors' of *Mycobacterium tuberculosis*. *Trends Microbiol* **9**, 237-241 (2001).
- 7 Middlebrook, G., Cohn, M. L. & Schaefer, W. B. Studies on isoniazid and tubercle bacilli. III. The isolation, drug-susceptibility, and catalase-testing of tubercle bacilli from isoniazid-treated patients. *Am Rev Tuberc* **70**, 852-872 (1954).
- 8 Zhang, Y., Heym, B., Allen, B., Young, D. & Cole, S. The catalase-peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature* **358**, 591-593 (1992).
- 9 Rozwarski, D. A., Grant, G. A., Barton, D. H., Jacobs, W. R., Jr. & Sacchettini, J. C. Modification of the NADH of the isoniazid target (InhA) from *Mycobacterium tuberculosis*. *Science* **279**, 98-102 (1998).
- 10 Manjunatha, U. H. *et al.* Direct inhibitors of InhA are active against *Mycobacterium tuberculosis*. *Sci Transl Med* **7**, 269ra263 (2015).
- 11 TB ALLIANCE. InhA inhibitors, <<http://www.tballiance.org/portfolio/compound/inha-inhibitors>> (2015).

- 12 Pajk, S. *et al.* New direct inhibitors of InhA with antimycobacterial activity based on a tetrahydropyran scaffold. *Eur J Med Chem* **112**, 252-257 (2016).
- 13 Sink, R. *et al.* Design, synthesis, and evaluation of new thiadiazole-based direct inhibitors of enoyl acyl carrier protein reductase (InhA) for the treatment of tuberculosis. *J Med Chem* **58**, 613-624 (2015).
- 14 Zumla, A., Nahid, P. & Cole, S. T. Advances in the development of new tuberculosis drugs and treatment regimens. *Nat Rev Drug Discov* **12**, 388-404 (2013).
- 15 Takayama, K., Armstrong, E. L., Kunugi, K. A. & Kilburn, J. O. Inhibition by ethambutol of mycolic acid transfer into the cell wall of *Mycobacterium smegmatis*. *Antimicrob Agents Chemother* **16**, 240-242 (1979).
- 16 Belanger, A. E. *et al.* The embAB genes of *Mycobacterium avium* encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. *Proc Natl Acad Sci U S A* **93**, 11919-11924 (1996).
- 17 Lety, M. A., Nair, S., Berche, P. & Escuyer, V. A single point mutation in the embB gene is responsible for resistance to ethambutol in *Mycobacterium smegmatis*. *Antimicrob Agents Chemother* **41**, 2629-2633 (1997).
- 18 Thomas, J. P., Baughn, C. O., Wilkinson, R. G. & Shepherd, R. G. A new synthetic compound with antituberculous activity in mice: ethambutol (dextro-2,2'-(ethylenediimino)-di-l-butanol). *Am Rev Respir Dis* **83**, 891-893 (1961).
- 19 Lee, R. E. *et al.* Combinatorial lead optimization of [1,2]-diamines based on ethambutol as potential antituberculosis preclinical candidates. *J Comb Chem* **5**, 172-187 (2003).
- 20 Tahlan, K. *et al.* SQ109 targets MmpL3, a membrane transporter of trehalose monomycolate involved in mycolic acid donation to the cell wall core of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **56**, 1797-1809 (2012).
- 21 Morlock, G. P., Metchock, B., Sikes, D., Crawford, J. T. & Cooksey, R. C. ethA, inhA, and katG loci of ethionamide-resistant clinical *Mycobacterium tuberculosis* isolates. *Antimicrob Agents Chemother* **47**, 3799-3805 (2003).
- 22 Wang, F. *et al.* Mechanism of thioamide drug action against tuberculosis and leprosy. *J Exp Med* **204**, 73-78 (2007).

- 23 DeBarber, A. E., Mdluli, K., Bosman, M., Bekker, L. G. & Barry, C. E., 3rd. Ethionamide activation and sensitivity in multidrug-resistant Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* **97**, 9677-82 (2000).
- 24 Baulard, A. R. *et al.* Activation of the pro-drug ethionamide is regulated in mycobacteria. *J Biol Chem* **275**, 28326-28331 (2000).
- 25 Willand, N. *et al.* Synthetic EthR inhibitors boost antituberculous activity of ethionamide. *Nat Med* **15**, 537-544 (2009).
- 26 Lambert, M. P. & Neuhaus, F. C. Mechanism of D-cycloserine action: alanine racemase from Escherichia coli W. *J Bacteriol* **110**, 978-987 (1972).
- 27 Ciak, J. & Hahn, F. E. Mechanisms of action of antibiotics. II. Studies on the modes of action of cycloserine and its L-stereoisomer. *Antibiot Chemother (Northfield)* **9**, 47-54 (1959).
- 28 Lynch, J. L. & Neuhaus, F. C. On the mechanism of action of the antibiotic O-carbamyld-serine in Streptococcus faecalis. *J Bacteriol* **91**, 449-460 (1966).
- 29 Zygmunt, W. A. Antagonism of D-Cycloserine inhibition of Mycobacterial growth by D-alanine. *J Bacteriol* **85**, 1217-1220 (1963).
- 30 Zygmunt, W. A. Reversal of D-cycloserine inhibition of bacterial growth by alanine. *J Bacteriol* **84**, 154-156 (1962).
- 31 Caceres, N. E. *et al.* Overexpression of the D-alanine racemase gene confers resistance to D-cycloserine in Mycobacterium smegmatis. *J Bacteriol* **179**, 5046-5055 (1997).
- 32 Watson, G. B., Bolanowski, M. A., Baganoff, M. P., Deppeler, C. L. & Lanthorn, T. H. D-cycloserine acts as a partial agonist at the glycine modulatory site of the NMDA receptor expressed in Xenopus oocytes. *Brain Res* **510**, 158-160 (1990).
- 33 Caminero, J. A., Sotgiu, G., Zumla, A. & Migliori, G. B. Best drug treatment for multidrug-resistant and extensively drug-resistant tuberculosis. *The Lancet Infectious Diseases* **10**, 621-629 (2010).
- 34 Feng, Z. & Barletta, R. G. Roles of Mycobacterium smegmatis D-alanine:D-alanine ligase and D-alanine racemase in the mechanisms of action of and resistance to the peptidoglycan inhibitor D-cycloserine. *Antimicrob Agents Chemother* **47**, 283-291 (2003).

- 35 Prosser, G. A. & de Carvalho, L. P. Metabolomics Reveal d-Alanine:d-Alanine Ligase As the Target of d-Cycloserine in. *ACS Med Chem Lett* **4**, 1233-1237 (2013).
- 36 Lin, Y. *et al.* Identification of antituberculosis agents that target ribosomal protein interactions using a yeast two-hybrid system. *Proc Natl Acad Sci U S A* **109**, 17412-17417 (2012).
- 37 Carter, A. P. *et al.* Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* **407**, 340-348 (2000).
- 38 Grigg, E. R. The arcana of tuberculosis with a brief epidemiologic history of the disease in the U.S.A. *Am Rev Tuberc* **78**, 151-172 (1958).
- 39 Sreevatsan, S. *et al.* Characterization of rpsL and rrs mutations in streptomycin-resistant Mycobacterium tuberculosis isolates from diverse geographic localities. *Antimicrob Agents Chemother* **40**, 1024-1026 (1996).
- 40 Shi, R., Zhang, J., Li, C., Kazumi, Y. & Sugawara, I. Detection of streptomycin resistance in Mycobacterium tuberculosis clinical isolates from China as determined by denaturing HPLC analysis and DNA sequencing. *Microbes Infect* **9**, 1538-1544 (2007).
- 41 Ruusala, T. & Kurland, C. G. Streptomycin preferentially perturbs ribosomal proofreading. *Mol Gen Genet* **198**, 100-104 (1984).
- 42 World Health Organization. Guidelines for the programmatic management of drug-resistant tuberculosis, <http://www.who.int/tb/challenges/mdr/programmatic_guidelines_for_mdrtb/en/> (2008).
- 43 Reeves, A. Z. *et al.* Aminoglycoside cross-resistance in Mycobacterium tuberculosis due to mutations in the 5' untranslated region of whiB7. *Antimicrob Agents Chemother* **57**, 1857-1865 (2013).
- 44 Sowajassatakul, A., Prammananan, T., Chaiprasert, A. & Phunpruch, S. Molecular characterization of amikacin, kanamycin and capreomycin resistance in M/XDR-TB strains isolated in Thailand. *BMC Microbiol* **14**, 165 (2014).
- 45 Begg, E. J. & Barclay, M. L. Aminoglycosides--50 years on. *Br J Clin Pharmacol* **39**, 597-603 (1995).

- 46 Podobnik, M., McInerney, P., O'Donnell, M. & Kuriyan, J. A TOPRIM domain in the crystal structure of the catalytic core of Escherichia coli primase confirms a structural link to DNA topoisomerases. *J Mol Biol* **300**, 353-362 (2000).
- 47 Aravind, L., Leipe, D. D. & Koonin, E. V. Toprim--a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins. *Nucleic Acids Res* **26**, 4205-4213 (1998).
- 48 Neu, H. C. Quinolone antimicrobial agents. *Annu Rev Med* **43**, 465-486 (1992).
- 49 Drlica, K., Malik, M., Kerns, R. J. & Zhao, X. Quinolone-mediated bacterial death. *Antimicrob Agents Chemother* **52**, 385-392 (2008).
- 50 Tse-Dinh, Y. C. Exploring DNA topoisomerases as targets of novel therapeutic agents in the treatment of infectious diseases. *Infect Disord Drug Targets* **7**, 3-9 (2007).
- 51 Kocagoz, T. *et al.* Gyrase mutations in laboratory-selected, fluoroquinolone-resistant mutants of Mycobacterium tuberculosis H37Ra. *Antimicrob Agents Chemother* **40**, 1768-1774 (1996).
- 52 Nuermberger, E. L. *et al.* Moxifloxacin-containing regimen greatly reduces time to culture conversion in murine tuberculosis. *Am J Respir Crit Care Med* **169**, 421-426 (2004).
- 53 Losick, R. & Pero, J. Bacillus subtilis RNA polymerase and its modification in sporulating and phage-infected bacteria. *Adv Enzymol Relat Areas Mol Biol* **44**, 165-185 (1976).
- 54 Burgess, R. R. & Travers, A. A. Escherichia coli RNA polymerase: purification, subunit structure, and factor requirements. *Fed Proc* **29**, 1164-1169 (1970).
- 55 Campbell, E. A. *et al.* Structural mechanism for rifampicin inhibition of bacterial rna polymerase. *Cell* **104**, 901-912 (2001).
- 56 Williams, D. L. *et al.* Contribution of rpoB mutations to development of rifamycin cross-resistance in Mycobacterium tuberculosis. *Antimicrob Agents Chemother* **42**, 1853-1857 (1998).
- 57 Artsimovitch, I. *et al.* Allosteric modulation of the RNA polymerase catalytic reaction is an essential component of transcription control by rifamycins. *Cell* **122**, 351-363 (2005).

- 58 Niemi, M., Backman, J. T., Fromm, M. F., Neuvonen, P. J. & Kivisto, K. T. Pharmacokinetic interactions with rifampicin : clinical relevance. *Clin Pharmacokinet* **42**, 819-850 (2003).
- 59 Burman, W. J., Gallicano, K. & Peloquin, C. Comparative pharmacokinetics and pharmacodynamics of the rifamycin antibacterials. *Clin Pharmacokinet* **40**, 327-341 (2001).
- 60 Nijland, H. M. *et al.* Rifampicin reduces plasma concentrations of moxifloxacin in patients with tuberculosis. *Clin Infect Dis* **45**, 1001-1007 (2007).
- 61 Svensson, E. M., Murray, S., Karlsson, M. O. & Dooley, K. E. Rifampicin and rifapentine significantly reduce concentrations of bedaquiline, a new anti-TB drug. *J Antimicrob Chemother* **70**, 1106-1114 (2015).
- 62 Boyer, P. D. The ATP synthase--a splendid molecular machine. *Annu Rev Biochem* **66**, 717-749, doi:10.1146/annurev.biochem.66.1.717 (1997).
- 63 Lu, P., Lill, H. & Bald, D. ATP synthase in mycobacteria: special features and implications for a function as drug target. *Biochim Biophys Acta* **1837**, 1208-1218 (2014).
- 64 Rao, S. P., Alonso, S., Rand, L., Dick, T. & Pethe, K. The protonmotive force is required for maintaining ATP homeostasis and viability of hypoxic, nonreplicating Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* **105**, 11945-11950 (2008).
- 65 Tran, S. L. & Cook, G. M. The F1Fo-ATP synthase of Mycobacterium smegmatis is essential for growth. *J Bacteriol* **187**, 5023-5028 (2005).
- 66 Deoghare, S. Bedaquiline: a new drug approved for treatment of multidrug-resistant tuberculosis. *Indian J Pharmacol* **45**, 536-537 (2013).
- 67 Andries, K. *et al.* A diarylquinoline drug active on the ATP synthase of Mycobacterium tuberculosis. *Science* **307**, 223-227 (2005).
- 68 Huitric, E. *et al.* Rates and mechanisms of resistance development in Mycobacterium tuberculosis to a novel diarylquinoline ATP synthase inhibitor. *Antimicrob Agents Chemother* **54**, 1022-1028 (2010).
- 69 Andries, K. *et al.* Acquired resistance of Mycobacterium tuberculosis to bedaquiline. *PLoS One* **9**, e102135 (2014).

- 70 Lehmann, J. Para-aminosalicylic acid in the treatment of tuberculosis. *Lancet* **1**, 15 (1946).
- 71 Rengarajan, J. *et al.* The folate pathway is a target for resistance to the drug para-aminosalicylic acid (PAS) in mycobacteria. *Mol Microbiol* **53**, 275-282 (2004).
- 72 Zheng, J. *et al.* para-Aminosalicylic acid is a prodrug targeting dihydrofolate reductase in *Mycobacterium tuberculosis*. *J Biol Chem* **288**, 23447-23456 (2013).
- 73 Chakraborty, S., Gruber, T., Barry, C. E., 3rd, Boshoff, H. I. & Rhee, K. Y. Para-aminosalicylic acid acts as an alternative substrate of folate metabolism in *Mycobacterium tuberculosis*. *Science* **339**, 88-91 (2013).
- 74 Ashtekar, D. R. *et al.* In vitro and in vivo activities of the nitroimidazole CGI 17341 against *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **37**, 183-186 (1993).
- 75 Stover, C. K. *et al.* A small-molecule nitroimidazopyran drug candidate for the treatment of tuberculosis. *Nature* **405**, 962-966 (2000).
- 76 Manjunatha, U. H. *et al.* Identification of a nitroimidazo-oxazine-specific protein involved in PA-824 resistance in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* **103**, 431-436 (2006).
- 77 Boshoff, H. Is the mycobacterial cell wall a hopeless drug target for latent tuberculosis? *Drug Discovery Today: Disease Mechanisms* **3**, 237-245 (2006).
- 78 Singh, R. *et al.* PA-824 kills nonreplicating *Mycobacterium tuberculosis* by intracellular NO release. *Science* **322**, 1392-1395 (2008).
- 79 Cellitti, S. E. *et al.* Structure of Ddn, the deazaflavin-dependent nitroreductase from *Mycobacterium tuberculosis* involved in bioreductive activation of PA-824. *Structure* **20**, 101-112 (2012).
- 80 Somasundaram, S., Anand, R. S., Venkatesan, P. & Paramasivan, C. N. Bactericidal activity of PA-824 against *Mycobacterium tuberculosis* under anaerobic conditions and computational analysis of its novel analogues against mutant Ddn receptor. *BMC Microbiol* **13**, 218 (2013).
- 81 Xavier, A. S. & Lakshmanan, M. Delamanid: A new armor in combating drug-resistant tuberculosis. *J Pharmacol Pharmacother* **5**, 222-224 (2014).
- 82 World Health Organization. The use of delamanid in the treatment of multidrug-resistant tuberculosis. *Interim Policy Guidance*,

- < http://www.who.int/tb/features_archive/delamanid/en/ > (2015).
- 83 Bloemberg, G. V. *et al.* Acquired Resistance to Bedaquiline and Delamanid in Therapy for Tuberculosis. *N Engl J Med* **373**, 1986-1988 (2015).
- 84 Yeager, R. L., Munroe, W. G. & Dessau, F. I. Pyrazinamide (aldinamide) in the treatment of pulmonary tuberculosis. *Am Rev Tuberc* **65**, 523-546 (1952).
- 85 Mitchison, D. A. The action of antituberculosis drugs in short-course chemotherapy. *Tubercle* **66**, 219-225 (1985).
- 86 Scorpio, A. & Zhang, Y. Mutations in *pncA*, a gene encoding pyrazinamidase/nicotinamidase, cause resistance to the antituberculous drug pyrazinamide in tubercle bacillus. *Nat Med* **2**, 662-667 (1996).
- 87 Konno, K., Feldmann, F. M. & McDermott, W. Pyrazinamide susceptibility and amidase activity of tubercle bacilli. *Am Rev Respir Dis* **95**, 461-469 (1967).
- 88 Zhang, Y., Wade, M. M., Scorpio, A., Zhang, H. & Sun, Z. Mode of action of pyrazinamide: disruption of Mycobacterium tuberculosis membrane transport and energetics by pyrazinoic acid. *J Antimicrob Chemother* **52**, 790-795 (2003).
- 89 Lu, P. *et al.* Pyrazinoic acid decreases the proton motive force, respiratory ATP synthesis activity, and cellular ATP levels. *Antimicrob Agents Chemother* **55**, 5354-5357 (2011).
- 90 Shpanchenko, O. V., Golovin, A. V., Bugaeva, E. Y., Isaksson, L. A. & Dontsova, O. A. Structural aspects of trans-translation. *IUBMB Life* **62**, 120-124 (2010).
- 91 Shi, W. *et al.* Pyrazinamide inhibits trans-translation in Mycobacterium tuberculosis. *Science* **333**, 1630-1632 (2011).
- 92 Yang, J. *et al.* Structural basis for targeting the ribosomal protein S1 of Mycobacterium tuberculosis by pyrazinamide. *Mol Microbiol* **95**, 791-803 (2015).
- 93 Zhang, S. *et al.* Mutations in *panD* encoding aspartate decarboxylase are associated with pyrazinamide resistance in Mycobacterium tuberculosis. *Emerg Microbes Infect* **2**, e34 (2013).
- 94 Shi, W. *et al.* Aspartate decarboxylase (PanD) as a new target of pyrazinamide in Mycobacterium tuberculosis. *Emerg Microbes Infect* **3**, e58 (2014).

- 95 Monteiro, D. C. *et al.* The structure of the PanD/PanZ protein complex reveals negative feedback regulation of pantothenate biosynthesis by coenzyme A. *Chem Biol* **22**, 492-503 (2015).
- 96 World Health Organization. *Tuberculosis*, <<http://www.who.int/mediacentre/factsheets/fs104/en/>> (2015).
- 97 Gandhi, N. R. *et al.* Multidrug-resistant and extensively drug-resistant tuberculosis: a threat to global control of tuberculosis. *Lancet* **375**, 1830-1843 (2010).
- 98 Mitnick, C. *et al.* Community-based therapy for multidrug-resistant tuberculosis in Lima, Peru. *N Engl J Med* **348**, 119-128 (2003).
- 99 Iseman, M. D. Treatment of multidrug-resistant tuberculosis. *N Engl J Med* **329**, 784-791 (1993).
- 100 Torun, T. *et al.* Side effects associated with the treatment of multidrug-resistant tuberculosis. *Int J Tuberc Lung Dis* **9**, 1373-1377 (2005).
- 101 Kwara, A., Flanigan, T. P. & Carter, E. J. Highly active antiretroviral therapy (HAART) in adults with tuberculosis: current status. *Int J Tuberc Lung Dis* **9**, 248-257 (2005).
- 102 Mitchison, D. A. Basic mechanisms of chemotherapy. *Chest* **76**, 771-781 (1979).
- 103 Prideaux, B. *et al.* The association between sterilizing activity and drug distribution into tuberculosis lesions. *Nat Med* **21**, 1223-1227 (2015).
- 104 van den Boogaard, J., Kibiki, G. S., Kisanga, E. R., Boeree, M. J. & Aarnoutse, R. E. New drugs against tuberculosis: problems, progress, and evaluation of agents in clinical development. *Antimicrob Agents Chemother* **53**, 849-862 (2009).
- 105 Dawson, R. *et al.* Efficiency and safety of the combination of moxifloxacin, pretomanid (PA-824), and pyrazinamide during the first 8 weeks of antituberculosis treatment: a phase 2b, open-label, partly randomised trial in patients with drug-susceptible or drug-resistant pulmonary tuberculosis. *Lancet* **385**, 1738-1747 (2015).
- 106 Rhee, K. Y. *et al.* Central carbon metabolism in *Mycobacterium tuberculosis*: an unexpected frontier. *Trends Microbiol* **19**, 307-314 (2011).

- 107 Baughn, A. D. & Rhee, K. Y. Metabolomics of Central Carbon Metabolism in *Mycobacterium tuberculosis*. *Microbiol Spectr* **2** (2014).
- 108 Gorke, B. & Stulke, J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat Rev Microbiol* **6**, 613-624 (2008).
- 109 de Carvalho, L. P. *et al.* Metabolomics of *Mycobacterium tuberculosis* reveals compartmentalized co-catabolism of carbon substrates. *Chem Biol* **17**, 1122-1131 (2010).
- 110 Baker, J. J., Johnson, B. K. & Abramovitch, R. B. Slow growth of *Mycobacterium tuberculosis* at acidic pH is regulated by *phoPR* and host-associated carbon sources. *Mol Microbiol* **94**, 56-69 (2014).
- 111 Galagan, J. E. *et al.* The *Mycobacterium tuberculosis* regulatory network and hypoxia. *Nature* **499**, 178-183 (2013).
- 112 Munoz-Elias, E. J. & McKinney, J. D. *Mycobacterium tuberculosis* isocitrate lyases 1 and 2 are jointly required for in vivo growth and virulence. *Nat Med* **11**, 638-644 (2005).
- 113 Kornberg, H. L. & Krebs, H. A. Synthesis of cell constituents from C₂-units by a modified tricarboxylic acid cycle. *Nature* **179**, 988-991 (1957).
- 114 Timm, J. *et al.* Differential expression of iron-, carbon-, and oxygen-responsive mycobacterial genes in the lungs of chronically infected mice and tuberculosis patients. *Proc Natl Acad Sci U S A* **100**, 14321-14326 (2003).
- 115 Gould, T. A., van de Langemheen, H., Munoz-Elias, E. J., McKinney, J. D. & Sacchettini, J. C. Dual role of isocitrate lyase 1 in the glyoxylate and methylcitrate cycles in *Mycobacterium tuberculosis*. *Mol Microbiol* **61**, 940-947 (2006).
- 116 Sharma, V. *et al.* Structure of isocitrate lyase, a persistence factor of *Mycobacterium tuberculosis*. *Nat Struct Biol* **7**, 663-668 (2000).
- 117 Kratky, M. *et al.* Salicylanilide derivatives block *Mycobacterium tuberculosis* through inhibition of isocitrate lyase and methionine aminopeptidase. *Tuberculosis (Edinb)* **92**, 434-439 (2012).
- 118 Krieger, I. V. *et al.* Structure-guided discovery of phenyl-diketo acids as potent inhibitors of *M. tuberculosis* malate synthase. *Chem Biol* **19**, 1556-1567 (2012).

- 119 Liu, K., Yu, J. & Russell, D. G. pckA-deficient *Mycobacterium bovis* BCG shows attenuated virulence in mice and in macrophages. *Microbiology* **149**, 1829-1835 (2003).
- 120 Marrero, J., Rhee, K. Y., Schnappinger, D., Pethe, K. & Ehrt, S. 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 (2010).
- 121 Ganapathy, U. *et al.* Two enzymes with redundant fructose bisphosphatase activity sustain gluconeogenesis and virulence in *Mycobacterium tuberculosis*. *Nat Commun* **6**, 7912 (2015).
- 122 Schnappinger, D. *et al.* Transcriptional Adaptation of *Mycobacterium tuberculosis* within Macrophages: Insights into the Phagosomal Environment. *J Exp Med* **198**, 693-704 (2003).
- 123 Sauer, U. & Eikmanns, B. J. The PEP-pyruvate-oxaloacetate node as the switch point for carbon flux distribution in bacteria. *FEMS Microbiol Rev* **29**, 765-794 (2005).
- 124 Awasthy, D. *et al.* Inactivation of the *ilvB1* gene in *Mycobacterium tuberculosis* leads to branched-chain amino acid auxotrophy and attenuation of virulence in mice. *Microbiology* **155**, 2978-2987 (2009).
- 125 Bellamy, D. The endogenous citric acid-cycle intermediates and amino acids of mitochondria. *Biochem J* **82**, 218-224 (1962).
- 126 Beste, D. J. *et al.* (1)(3)C metabolic flux analysis identifies an unusual route for pyruvate dissimilation in mycobacteria which requires isocitrate lyase and carbon dioxide fixation. *PLoS Pathogens* **7**, e1002091 (2011).
- 127 Marrero, J., Trujillo, C., Rhee, K. Y. & Ehrt, S. Glucose phosphorylation is required for *Mycobacterium tuberculosis* persistence in mice. *PLoS Pathogens* **9**, e1003116 (2013).
- 128 Watanabe, S. *et al.* Fumarate reductase activity maintains an energized membrane in anaerobic *Mycobacterium tuberculosis*. *PLoS Pathogens* **7**, e1002287 (2011).
- 129 Sasseti, C. M. & Rubin, E. J. Genetic requirements for mycobacterial survival during infection. *Proc Natl Acad Sci U S A* **100**, 12989-12994 (2003).

- 130 Chavadi, S. *et al.* Global effects of inactivation of the pyruvate kinase gene in the *Mycobacterium tuberculosis* complex. *J Bacteriol* **191**, 7545-7553 (2009).
- 131 Opperdoes, F. R. & Michels, P. A. Enzymes of carbohydrate metabolism as potential drug targets. *Int J Parasitol* **31**, 482-490 (2001).
- 132 Gengenbacher, M., Rao, S. P., Pethe, K. & Dick, T. Nutrient-starved, non-replicating *Mycobacterium tuberculosis* requires respiration, ATP synthase and isocitrate lyase for maintenance of ATP homeostasis and viability. *Microbiology* **156**, 81-87 (2010).
- 133 Lorenz, M. C. & Fink, G. R. Life and death in a macrophage: role of the glyoxylate cycle in virulence. *Eukaryot Cell* **1**, 657-662 (2002).
- 134 Matte, A., Tari, L. W., Goldie, H. & Delbaere, L. T. J. Minireview: Structure and mechanism of phosphoenolpyruvate carboxykinase (vol 272, pg 8105, 1997). *Journal of Biological Chemistry* **272**, 20962-20962 (1997).
- 135 Beste, D. J. *et al.* ¹³C-flux spectral analysis of host-pathogen metabolism reveals a mixed diet for intracellular *Mycobacterium tuberculosis*. *Chem Biol* **20**, 1012-1021 (2013).
- 136 Zelle, R. M., Trueheart, J., Harrison, J. C., Pronk, J. T. & van Maris, A. J. Phosphoenolpyruvate carboxykinase as the sole anaplerotic enzyme in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **76**, 5383-5389 (2010).
- 137 Petersen, S. *et al.* Metabolic consequences of altered phosphoenolpyruvate carboxykinase activity in *Corynebacterium glutamicum* reveal anaplerotic regulation mechanisms in vivo. *Metab Eng* **3**, 344-361 (2001).
- 138 Utter, M. F. & Kurahashi, K. Purification of oxalacetic carboxylase from chicken liver. *J Biol Chem* **207**, 787-802 (1954).
- 139 Hanson, R. W. & Garber, A. J. Phosphoenolpyruvate carboxykinase. I. Its role in gluconeogenesis. *Am J Clin Nutr* **25**, 1010-1021 (1972).
- 140 Dunten, P. *et al.* Crystal structure of human cytosolic phosphoenolpyruvate carboxykinase reveals a new GTP-binding site. *J Mol Biol* **316**, 257-264 (2002).
- 141 Mukhopadhyay, B., Concar, E. M. & Wolfe, R. S. A GTP-dependent vertebrate-type phosphoenolpyruvate carboxykinase from *Mycobacterium smegmatis*. *J Biol Chem* **276**, 16137-16145 (2001).

- 142 Park, E. A., Jerden, D. C. & Bahouth, S. W. Regulation of phosphoenolpyruvate carboxykinase gene transcription by thyroid hormone involves two distinct binding sites in the promoter. *Biochem J* **309** (Pt 3), 913-919 (1995).
- 143 Modaresi, S., Brechtel, K., Christ, B. & Jungermann, K. Human mitochondrial phosphoenolpyruvate carboxykinase 2 gene. Structure, chromosomal localization and tissue-specific expression. *Biochem J* **333** (Pt 2), 359-366 (1998).
- 144 Jiang, W. *et al.* Acetylation regulates gluconeogenesis by promoting PEPCK1 degradation via recruiting the UBR5 ubiquitin ligase. *Mol Cell* **43**, 33-44 (2011).
- 145 Machova, I. *et al.* Mycobacterium tuberculosis phosphoenolpyruvate carboxykinase is regulated by redox mechanisms and interaction with thioredoxin. *The Journal of Biological Chemistry* **289**, 13066-13078 (2014).
- 146 Machova, I. *et al.* Structural and functional studies of phosphoenolpyruvate carboxykinase from Mycobacterium tuberculosis. *PLoS One* **10**, e0120682 (2015).
- 147 Chang, H. C. & Lane, M. D. The enzymatic carboxylation of phosphoenolpyruvate. II. Purification and properties of liver mitochondrial phosphoenolpyruvate carboxykinase. *J Biol Chem* **241**, 2413-2420 (1966).
- 148 Noce, P. S. & Utter, M. F. Decarboxylation of oxalacetate to pyruvate by purified avian liver phosphoenolpyruvate carboxykinase. *J Biol Chem* **250**, 9099-9105 (1975).
- 149 Hammond, K. D. & Balinsky, D. Activities of key gluconeogenic enzymes and glycogen synthase in rat and human livers, hepatomas, and hepatoma cell cultures. *Cancer Res* **38**, 1317-1322 (1978).
- 150 Barns, R. J. & Keech, D. B. Sheep kidney phosphoenolpyruvate carboxylase. Purification and properties. *Biochim Biophys Acta* **276**, 284-296 (1972).
- 151 Aich, S., Imabayashi, F. & Delbaere, L. T. Expression, purification, and characterization of a bacterial GTP-dependent PEP carboxykinase. *Protein Expr Purif* **31**, 298-304 (2003).
- 152 Foster, D. O., Lardy, H. A., Ray, P. D. & Johnston, J. B. Alteration of rat liver phosphoenolpyruvate carboxykinase activity by L-tryptophan in vivo and metals in vitro. *Biochemistry* **6**, 2120-2128 (1967).

- 153 Tari, L. W., Matte, A., Goldie, H. & Delbaere, L. T. Mg(2+)-Mn2+ clusters in enzyme-catalyzed phosphoryl-transfer reactions. *Nat Struct Biol* **4**, 990-994 (1997).
- 154 Sullivan, S. M. & Holyoak, T. Enzymes with lid-gated active sites must operate by an induced fit mechanism instead of conformational selection. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 13829-13834 (2008).
- 155 Kinoshita, K., Sadanami, K., Kidera, A. & Go, N. Structural motif of phosphate-binding site common to various protein superfamilies: all-against-all structural comparison of protein-mononucleotide complexes. *Protein Eng* **12**, 11-14 (1999).
- 156 Johnson, T. A. & Holyoak, T. The Omega-loop lid domain of phosphoenolpyruvate carboxykinase is essential for catalytic function. *Biochemistry* **51**, 9547-9559 (2012).
- 157 Case, C. L., Concar, E. M., Boswell, K. L. & Mukhopadhyay, B. Roles of Asp75, Asp78, and Glu83 of GTP-dependent phosphoenolpyruvate carboxykinase from *Mycobacterium smegmatis*. *The Journal of Biological Chemistry* **281**, 39262-39272 (2006).
- 158 Colombo, G., Carlson, G. M. & Lardy, H. A. Phosphoenolpyruvate carboxykinase (guanosine triphosphate) from rat liver cytosol. Separation of homogeneous forms of the enzyme with high and low activity by chromatography on agarose-hexane-guanosine triphosphate. *Biochemistry* **17**, 5321-5329 (1978).
- 159 Merritt, E. A. X-ray Anomalous Scattering, <<http://skuld.bmsc.washington.edu/scatter/>> (2015).
- 160 Holyoak, T., Sullivan, S. M. & Nowak, T. Structural insights into the mechanism of PEPCK catalysis. *Biochemistry* **45**, 8254-8263 (2006).
- 161 Yang, J., Kalhan, S. C. & Hanson, R. W. What is the metabolic role of phosphoenolpyruvate carboxykinase? *J Biol Chem* **284**, 27025-27029 (2009).
- 162 El-Hamamsy, M. H., Smith, A. W., Thompson, A. S. & Threadgill, M. D. Structure-based design, synthesis and preliminary evaluation of selective inhibitors of dihydrofolate reductase from *Mycobacterium tuberculosis*. *Bioorg Med Chem* **15**, 4552-4576 (2007).

- 163 Jabalquinto, A. M., Laivenieks, M., Zeikus, J. G. & Cardemil, E. Characterization of the oxaloacetate decarboxylase and pyruvate kinase-like activities of *Saccharomyces cerevisiae* and *Anaerobiospirillum succiniciproducens* phosphoenolpyruvate carboxykinases. *J Protein Chem* **18**, 659-664 (1999).
- 164 Colombo, G. & Lardy, H. A. Phosphoenolpyruvate carboxykinase (guanosine 5'-triphosphate) from rat liver cytosol. Divalent cation involvement in the decarboxylation reactions. *Biochemistry* **20**, 2758-2767 (1981).
- 165 DiTullio, N. W. *et al.* 3-mercaptopycolinic acid, an inhibitor of gluconeogenesis. *Biochem J* **138**, 387-394 (1974).
- 166 Foley, L. H. *et al.* Modified 3-alkyl-1,8-dibenzylxanthines as GTP-competitive inhibitors of phosphoenolpyruvate carboxykinase. *Bioorganic & Medicinal Chemistry Letters* **13**, 3607-3610 (2003).
- 167 Balan, M. D., McLeod, M. J., Lotosky, W. R., Ghaly, M. & Holyoak, T. Inhibition and Allosteric Regulation of Monomeric Phosphoenolpyruvate Carboxykinase by 3-Mercaptopycolinic Acid. *Biochemistry* **54**, 5878-5887 (2015).
- 168 Stiffin, R. M., Sullivan, S. M., Carlson, G. M. & Holyoak, T. Differential inhibition of cytosolic PEPCK by substrate analogues. Kinetic and structural characterization of inhibitor recognition. *Biochemistry* **47**, 2099-2109 (2008).
- 169 Sun, Y. *et al.* Phosphoenolpyruvate carboxykinase overexpression selectively attenuates insulin signaling and hepatic insulin sensitivity in transgenic mice. *J Biol Chem* **277**, 23301-23307 (2002).
- 170 Valera, A., Pujol, A., Pelegrin, M. & Bosch, F. Transgenic mice overexpressing phosphoenolpyruvate carboxykinase develop non-insulin-dependent diabetes mellitus. *Proc Natl Acad Sci U S A* **91**, 9151-9154 (1994).
- 171 Rondinone, C. M. Therapeutic potential of RNAi in metabolic diseases. *Biotechniques* Suppl, 31-36 (2006).
- 172 Ananthan, S. *et al.* High-throughput screening for inhibitors of *Mycobacterium tuberculosis* H37Rv. *Tuberculosis (Edinb)* **89**, 334-353 (2009).
- 173 Kayne, F. J. & Price, N. C. Amino acid effector binding to rabbit muscle pyruvate kinase. *Arch Biochem Biophys* **159**, 292-296 (1973).

- 174 Jurica, M. S. *et al.* The allosteric regulation of pyruvate kinase by fructose-1,6-bisphosphate. *Structure* **6**, 195-210 (1998).
- 175 Saghatelian, A. & Cravatt, B. F. Global strategies to integrate the proteome and metabolome. *Curr Opin Chem Biol* **9**, 62-68 (2005).
- 176 Haeckel, R., Hess, B., Lauterborn, W. & Wuster, K. H. Purification and allosteric properties of yeast pyruvate kinase. *Hoppe Seylers Z Physiol Chem* **349**, 699-714 (1968).
- 177 Ernest, I., Callens, M., Opperdoes, F. R. & Michels, P. A. Pyruvate kinase of *Leishmania mexicana mexicana*. Cloning and analysis of the gene, overexpression in *Escherichia coli* and characterization of the enzyme. *Mol Biochem Parasitol* **64**, 43-54 (1994).
- 178 Cazzulo, J. J., Cazzulo Franke, M. C. & Franke de Cazzulo, B. M. On the regulatory properties of the pyruvate kinase from *Trypanosoma cruzi* epimastigotes. *FEMS Microbiol Lett* **50**, 259-263 (1989).
- 179 Johannes, K. J. & Hess, B. Allosteric kinetics of pyruvate kinase of *Saccharomyces carlsbergensis*. *J Mol Biol* **76**, 181-205 (1973).
- 180 Noguchi, T., Inoue, H. & Tanaka, T. The M1- and M2-type isozymes of rat pyruvate kinase are produced from the same gene by alternative RNA splicing. *J Biol Chem* **261**, 13807-13812 (1986).
- 181 Tanaka, K., Sakai, H., Ohta, T. & Matsuzawa, H. Molecular cloning of the genes for pyruvate kinase of two bacilli, *Bacillus psychrophilus* and *Bacillus licheniformis*, and comparison of the properties of the enzymes produced in *Escherichia coli*. *Biosci Biotechnol Biochem* **59**, 1536-1542 (1995).
- 182 Munoz, M. E. & Ponce, E. Pyruvate kinase: current status of regulatory and functional properties. *Comp Biochem Physiol B Biochem Mol Biol* **135**, 197-218 (2003).
- 183 Suzuki, K., Ito, S., Shimizu-Ibuka, A. & Sakai, H. Crystal structure of pyruvate kinase from *Geobacillus stearothermophilus*. *J Biochem* **144**, 305-312 (2008).
- 184 Valentini, G. *et al.* Structure and function of human erythrocyte pyruvate kinase. Molecular basis of nonspherocytic hemolytic anemia. *J Biol Chem* **277**, 23807-23814 (2002).

- 185 Dombrauckas, J. D., Santarsiero, B. D. & Mesecar, A. D. Structural basis for tumor pyruvate kinase M2 allosteric regulation and catalysis. *Biochemistry* **44**, 9417-94293 (2005).
- 186 Mattevi, A. *et al.* Crystal structure of Escherichia coli pyruvate kinase type I: molecular basis of the allosteric transition. *Structure* **3**, 729-741 (1995).
- 187 Zhong, W. *et al.* 'In crystallo' substrate binding triggers major domain movements and reveals magnesium as a co-activator of Trypanosoma brucei pyruvate kinase. *Acta Crystallogr D Biol Crystallogr* **69**, 1768-1779 (2013).
- 188 Fothergill-Gillmore, L. A., Rigden, D. J., Michels, P. A. & Phillips, S. E. Leishmania pyruvate kinase: the crystal structure reveals the structural basis of its unique regulatory properties. *Biochem Soc Trans* **28**, 186-190 (2000).
- 189 Morgan, H. P. *et al.* Allosteric mechanism of pyruvate kinase from Leishmania mexicana uses a rock and lock model. *J Biol Chem* **285**, 12892-12898 (2010).
- 190 Larsen, T. M., Benning, M. M., Rayment, I. & Reed, G. H. Structure of the bis(Mg²⁺)-ATP-oxalate complex of the rabbit muscle pyruvate kinase at 2.1 Å resolution: ATP binding over a barrel. *Biochemistry* **37**, 6247-6255 (1998).
- 191 Wang, Q. *et al.* Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science* **327**, 1004-1007 (2010).
- 192 Lee, H. J., Lang, P. T., Fortune, S. M., Sasseti, C. M. & Alber, T. Cyclic AMP regulation of protein lysine acetylation in Mycobacterium tuberculosis. *Nat Struct Mol Biol* **19**, 811-818 (2012).
- 193 Xu, H., Hegde, S. S. & Blanchard, J. S. Reversible acetylation and inactivation of Mycobacterium tuberculosis acetyl-CoA synthetase is dependent on cAMP. *Biochemistry* **50**, 5883-5892 (2011).
- 194 Zee, B. M. & Garcia, B. A. Discovery of lysine post-translational modifications through mass spectrometric detection. *Essays Biochem* **52**, 147-163 (2012).
- 195 Cherkasov, A. *et al.* Mapping the protein interaction network in methicillin-resistant Staphylococcus aureus. *J Proteome Res* **10**, 1139-1150 (2011).
- 196 Akerley, B. J. *et al.* A genome-scale analysis for identification of genes required for growth or survival of Haemophilus influenzae. *Proc Natl Acad Sci U S A* **99**, 966-971 (2002).

- 197 Song, J. H. *et al.* Identification of essential genes in *Streptococcus pneumoniae* by allelic replacement mutagenesis. *Mol Cells* **19**, 365-374 (2005).
- 198 Nowicki, M. W. *et al.* Design, synthesis and trypanocidal activity of lead compounds based on inhibitors of parasite glycolysis. *Bioorg Med Chem* **16**, 5050-5061 (2008).
- 199 Christofk, H. R. *et al.* The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature* **452**, 230-233 (2008).
- 200 Iqbal, M. A., Gupta, V., Gopinath, P., Mazurek, S. & Bamezai, R. N. Pyruvate kinase M2 and cancer: an updated assessment. *FEBS Lett* **588**, 2685-2692 (2014).
- 201 Zoraghi, R. *et al.* Methicillin-resistant *Staphylococcus aureus* (MRSA) pyruvate kinase as a target for bis-indole alkaloids with antibacterial activities. *J Biol Chem* **286**, 44716-44725 (2011).
- 202 Morgan, H. P. *et al.* The trypanocidal drug suramin and other trypan blue mimetics are inhibitors of pyruvate kinases and bind to the adenosine site. *J Biol Chem* **286**, 31232-31240 (2011).
- 203 Friesner, R. A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* **47**, 1739-1749 (2004).
- 204 DeSantis, K., Reed, A., Rahhal, R. & Reinking, J. Use of differential scanning fluorimetry as a high-throughput assay to identify nuclear receptor ligands. *Nucl Recept Signal* **10**, e002 (2012).
- 205 Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2**, 2212-2221 (2007).