Genetic Essentiality, Biochemical and Structural Properties of

Fructose 1,6-bisphosphatases II

BY

HITEN JAYANTI GUTKA
B. Pharm. Sci. University of Mumbai, INDIA, 2002
M. Pharm. Sci. University of Mumbai, INDIA, 2004

DISSERTATION
Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Pharmacognosy
in the Graduate College of the
University of Illinois at Chicago, 2013

Chicago, Illinois

Defense Committee:
Farahnaz Movahedzadeh, Advisor
Scott Franzblau, Chair
Vadim Gaponenko, Biochemistry and Molecular Genetics
Michael Federle
Pavel Petukhov
Cele Abad Zapatero, Center for Pharmaceutical Biotechnology

This thesis is dedicated to my Parents (Mr. Jayanti V. Gutka and Mrs. Sushila J. Gutka), my wife Sejal and our beloved son Rishabh

ACKNOWLEDGEMENTS

During the course of my graduate studies at UIC, I have had the privilege of interacting with a variety of people including graduate students, post-doctoral fellows, research specialists and professors. Without the guidance, support, interactive learning and company of this intellectual and scholarly people this work would not have been possible.

I would like to give a very special thanks to my advisors Dr. Farahnaz Movahedzadeh and Prof. Scott Franzblau for their constant support and help during my years as a graduate student in The Institute for Tuberculosis Research (ITR). Dr. Farah's constructive criticism, encouragement and inspirational discussions were instrumental in accomplishment of my work. I am grateful to Prof. Franzblau for providing me with all the facilities (including a well resourced Biosafety Level 3 laboratory) required to accomplish this work at ITR. I wish to thank both Dr. Farah and Prof. Franzblau for thinking out this unique project for my dissertation. I still remember those tough initial days when finding a lab to continue the PhD research had become extremely difficult. Against all odds Prof. Franzblau accepted me as a graduate student in the ITR and provided me with this unique opportunity. I would like to extend my thanks to my committee members Dr. Michael Federle, Dr. Pavel Petukhov, Dr. Cele Abad Zapatero and Dr. Vadim Gaponenko for their time and comments during the preliminary and final stages of my work.

I extend my special thanks to Dr. Cele Abad Zapatero for mentoring me on the crystallography and structural biology sections of this work. Indeed, working with him was a wonderful learning experience. I am extremely thankful to Dr. Bernard Santarsiero and Dr. Shahila Mehboob from Centre for Pharmaceutical Biotechnology for accommodating my samples during the data collection campaigns at Argone National Labs (ANL). Dr. Santarsiero's help (including logistics for delicate protein crystals to ANL), thoughtful suggestions and guidance helped me crystallize

an additional similar protein; I sincerely wish to thank Dr. Santarsiero for all his help. I am also thankful to Dr. Michael Johnson and members of his lab (in particular Dr. Mehboob) for allowing me to use their FPLC. Prof. Chuan He and Pedro Brugarolas from the University of Chicago are acknowledged for providing facility, equipment and technical assistance to perform molecular weight estimation of the purified proteins.

I am thankful to Dr. Yuehong Wang of ITR for her extraordinary effort and help with the animal studies described in this thesis. Indeed it was Dr. Wang's relentless support that helped me going with performing a 6 month long animal study. I wish to thank my fellow graduate student friends in ITR Geping Cai, Nan Zhang, Wei Gao and seniors Chang Hwa Hwang, Yang Song for their helpful discussions and the wonderful time that I spent at ITR.

I am thankful to Dr. Kamolchanok Ruksree, for her initial help with protein expression. Dr. Paul Wheeler (Tuberculosis Research Group, Veterinary Laboratories Agency) is acknowledged for his valuable discussions on enzyme assay set up and biochemical characterization. I am thankful to Dr. Timothy Larson from the Virginia Technical University for providing me with *Escherichia coli* strains for genetic complementation studies. I would want to thank Maxwell Rutter, Ryan Patwell and Trinidad Gonzalez who made our small lab within the ITR a wonderful place to work. Special thanks to Maxwell Rutter for helping me out with experiments and equipment troubleshooting on several occasions. I would want to thank my friend Pankaj Pandey for his help with visual graphics and structural biology software. I gratefully acknowledge the support of some special individuals – Sumit Sahni, Prahlad Rao, Pulkit Gupta, Chaitanya Agrawal, Hazem Abdelkarim, Hatem El Shabrawy, Tian Zhu, Sambit Mohapatra and Yash Jalundwala. I wish to thank Rohan Dargad, Himanshu Sharma, Romeel Bhargaw, Karan Sarda,

ACKNOWLEDGEMENTS (continued)

Karan Sharma and Gaurav Kundu for being great roommates to me on different occasions during the 4 long years. Indeed the time I spent at Medical Centre Apartments with some of these special individuals, ensured that I never missed out on any festivities or celebrations be it watching a cricket match, celebrating Indian festivals, a weekend party or birthday celebrations.

Nothing can be achieved without the help and support of family, this thesis wouldn't have ever been possible without the relentless support of my beloved wife Sejal. It was the strong will and determination of Sejal that we both could successfully pass through the tough yet memorable 4 long years. My parents deserve a special mention here for their blessings, love and affection and persistent support. My Mother is the one who sincerely raised me with her love and affection and supported me in whatever I intended to do in spite of all difficulties that she had to face. My beloved son Rishabh needs a special mention here for being an extremely well behaved child ever since the time he was born. I wish to thank my sister Bijal and my brother-in-law Bhavin for taking up some of the important responsibilities back home. My maternal Aunt Kanchan Maru, Uncle Ramnik Maru need a special mention here for all their help and support during the difficult days of my undergraduate education. My cousins Kunal and Bhavita and my brother-in-law Paras Gada are acknowledged for all their motivational talks and support. Above all I express my reverence for the ALMIGHTY who gave me this wonderful life transforming opportunity to pursue science and make a career out of it.

H. J. G.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF TABLES

TABLES **PAGE**

LIST OF FIGURES

FIGURES **PAGE**

LIST OF FIGURES (Continued)

FIGURES **PAGE**

LIST OF ABBREVIATIONS

- 1,3BPG 1,3-bisphosphoglycerate
- 3PG 3-phosphoglycerate
- ADP adenosine diphosphate
- Ala alanine
- ALS Advanced Light Source
- AMP Adenosine monophosphate
- APS Advance Photon Source
- Arg arginine
- Asp aspartate
- ATP adenosine triphosphate
- au. Asymmetric unit
- CFU colony forming units
- CIT citrate
- COOT Crystallographic Object-Oriented Toolkit
- Da daltons
- DCO double cross over
- DETANO diethylenetriamine nonoate
- DMSO dimethyl sulfoxide
- DNA deoxyribonucleic acid
- DNAse deoxyribonuclease
- DOTS Directly Observed Therapy Short-course
- DTT dithiothreitol

LIST OF ABBREVIATIONS (Continued)

- EDTA ethylenediaminetetraacetic acid
- EMB ethambutol
- F1,6BP/FBP fructose 1,6-bisphosphate
- F6P fructose 6-phosphate
- FBPase fructose 1,6-bisphosphatase
- FLD first-line drugs
- FPLC Fast protein liquid chromatography
- *Ft*GlpX GlpX protein of *Francisella tularensis*
- G3P glyceraldehyde 3-phosphate
- G6P glucose 6-phosphate
- Glu glutamate
- HBSS Hanks' balanced salt solution
- HCl hydrogen chloride/hydrochloric acid
- hexose-P hexose phosphate
- His histidine
- HTS
- Hyg hygromycin B
- IC_{50} 50 % inhibitory concentration
- IC90 90 % inhibitory concentration
- IMP inositol monophosphate/ inositol 1-phosphate
- IMPase inositol monophosphatase
- INH isoniazid

LIST OF ABBREVIATIONS (Continued)

- IPTG isopropyl β-D-1-thiogalactopyranoside
- ISOCIT isocitrate
- Kan kanamycin
- kcat Catalytic rate of an enzyme
- KCl potassium chloride
- KDa killodaltons
- Km Michaelis constant
- LB Luria-Bertani
- Li^+ lithium
- LiCl lithium chloride
- Lys lysine
- MABA Microtiter alamar blue assay
- MAD multi-wavelength anomalous dispersion
- Mal malic acid
- MBP maltose binding protein
- MDR-TB multi-drug resistant tuberculosis
- $MgCl_2$ magnesium chloride
- MIC Minimum inhibitory concentration
- MR molecular replacement
- *Mtb* *Mycobacterium tuberculosis*
- *Mt*FBPase fructose 1,6-bisphosphatase of *Mycobacterium tuberculosis*
- NaCl sodium chloride

LIST OF ABBREVIATIONS (Continued)

- NaOH sodium hydroxide
- Ni-NTA nickel- nitrilotriacetic acid
- OAA oxaloacetate
- OADC Oleic Albumin Dextrose Catalase
- OD optical density
- OH hydroxyl group
- PBS Phosphate buffered saline
- PCR polymerase chain reaction
- PDB Protein Data Bank
- pentose-P pentose phosphate
- PEP phosphoenolpyruvic acid/ phosphoenolpyruvate
- PG phosphoglycerate
- PGI phosphoglucose isomerase
- PISA protein interactions, surfaces and assemblies
- PZA pyrazinamide
- RMP rifampicin/rifampin
- SAD single-wavelength anomalous dispersion
- SBDD Structure Based Drug Design
- SCO Single cross over
- SDS-PAGE sodium dodecyl sulfate polyacrylamide gel electrophoresis
- SEC Size Exclusion Chromatography
- SLD second-line drugs

LIST OF ABBREVIATIONS (Continued)

- SRF Self Rotation Function
- SUC succinic acid
- SUMO Small Ubiquitin like Modifier
- TB Tuberculosis
- TBSGC Tuberculosis Structural Genomics Consortium
- TCA Tricarboxylic acid
- Thr threonine
- TLD third-line drugs
- TraSH Transposon site hybridization
- Tyr tyrosine
- U Units
- WHO World Health Organization
- WT wild type
- XDR-TB extensively-drug-resistant tuberculosis
- αKG alpha-ketoglutarate
- λmax wavelength maximum

SUMMARY

Tuberculosis (TB) in the 21st century continues to be a global public health problem mainly in underdeveloped or developing nations. This is due to major limitations in the public healthcare infrastructure, poor diagnostic facilities, and lack of necessary health care organization. The current standard of treatment, Directly Observed Therapy Short course (DOTS) is lengthy. Treatment duration is about 6 months and some of the drugs have serious side effects, some of which ultimately result in patient non-compliance and, hence relapse of the disease. Resistance to current anti-TB drugs is manifested as multi drug resistant (MDR) and extensively drug resistant (XDR) strains of *Mycobacterium tuberculosis* (*Mtb*), the etiological agent of tuberculosis (TB).

The large number of *Mtb* infections and the evolution of MDR and XDR strains demand the search for novel anti-tuberculosis agents. The availability of the *Mtb* genome sequence has significantly enhanced the understanding of biology, immunology, and molecular mechanisms of pathogenesis for this dreaded pathogen. Furthermore, the genome sequence coupled with consortium wide structural biology efforts have placed TB drug research in an unprecedented position. It has generated voluminous genetic, bioinformatic, and structural information to probe potential drug targets to combat TB.

The goal of the interdisciplinary studies described herein is to investigate the gene essentiality and perform biochemical and structural characterization of the *glpX* encoded protein in *Mtb*. The genome wide transposon site hybridization (TraSH) experiment suggests the *glpX* gene to be essential for survival in a mice infection model. This method has several limitations however, such as; heterogeneous transposon mutant population. Although the method is high throughput and genome wide, it does not characterize individual transposon mutants. Therefore, a section of

the experimental work is dedicated towards the generation of an unmarked *glpX* gene deletion knock out (KO) in *Mtb* and characterization of its *in vitro* and *in vivo* growth and survival behavior. The results indicate that although the growth profile of Δ*glpX* in the standard *Mtb* 7H9 media (nutrient rich media) is comparable to that of *Mtb*, its growth is dysgonic in a media containing a single gluconeogenic carbon source (e.g. oleic acid, glycerol, acetate). The *in vivo* survival profile in the mouse model indicates that the Δ*glpX* not only fails to achieve the initial dose required for infection, but also maintains bacterial loads similar to the *Mtb* strain at later time points (about 2-3 log lower). Bacterial count lower by about 2-3 logs is considered significant and suggests that the *glpX* gene is essential *in vivo*. The results indicate that the *glpX* gene encodes a functional FBPase and is essential for both *in vitro* and *in vivo* growth and survival of *Mtb*. A section describes the successful heterologous expression and purification of the *glpX* gene-encoded fructose 1,6-bisphosphatase II (FBPase II). The FBPase retained activity post purification because of the presence of bivalent magnesium ions in the purification buffer. Moreover the storage stability of FBPase also improved due to presence of magnesium and glycerol in the purification buffer. Finally, extensive biochemical characterization of the FBPase was performed to determine its key biochemical properties. The crystal structure of FBPase protein was solved in 2 different forms (native and catalytic product fructose 6-phosphate bound form) using a molecular replacement approach, employing *Escherichia coli* GlpX (PDB ID: 3bih) as a model template structure. The functional unit of *Mtb* FBPase was found to be tetrameric in nature, with a possible indication of alternate allosteric regulatory mechanism. The monomeric structure fold (α-β-α-β-α) of *Mtb* FBPase is similar to that of other proteins belonging to the lithium sensitive phosphatase super family. The catalytic site is functionally

conserved with that of *E. coli* GlpX, essentially following the similar and well characterized two-metal sites assisted catalysis model.

The combined genetic, biochemical, and structural studies described herein establish a solid foundation for Structure Based Drug Design (SBDD) for this important *Mtb* target. The *glpX* gene has also been identified to be essential for virulence in *Francisella tularensis*, however its biochemical function was not verified. A section of the experimental work focuses on the verification of the biochemical function of *glpX*-encoded protein in *F. tularensis*. The FBPase activity for *F. tularensis glpX* was verified by *in vitro* genetic complementation. Further preliminary studies including expression, purification, and crystallization were carried out to assess the suitability of this target for Structure Based Drug Design (SBDD) in a manner similar to that of *Mtb* FBPase.

Knowledge acquired from these interdisciplinary studies will help in realizing the ultimate long term goal of identifying new chemical entities inhibiting this important metabolic target in pathogenic bacteria.

CHAPTER I

INTRODUCTION

Tuberculosis (TB) is a global public health problem, with 2 billion people, equal to about one-third of the world's population, infected with *Mycobacterium tuberculosis* (*Mtb*), the microbe that causes TB. In 1993, the World Health Organization (WHO) declared TB a global public health emergency, when an estimated 7–8 million cases and 1.3–1.6 million deaths were found to have been occurring each year. In 2010, there were an estimated 8.5–9.2 million cases and 1.2–1.5 million deaths (including deaths from TB among HIV-positive patients). TB is the second leading cause of death from an infectious disease worldwide after HIV (WHO 2008). TB is a disease of poverty, affecting mostly young adults in their most productive years. The vast majority of TB deaths are in the developing world, and more than half of all deaths occur in Asia (WHO 2008). TB hits the under developed or developing nations due to major limitations in the public healthcare infrastructure, poor diagnostic facilities, lack of necessary health care organization and the inability to provide the long treatment to TB patients. Current control methods of chemotherapy and vaccination are inadequate and treatment is lengthy. With the standard anti–TB therapy (Directly Observed Therapy Short (DOTS) course) taking about 6 months, non-compliance with treatment presents a serious possibility of antibiotic resistance.

According to the WHO, DOTS course treatment regimen/Stop TB strategy has been adopted by most nations for their national tuberculosis control programs (WHO 2008). However the long duration of these regimens and side effects arising from certain drugs of the treatment cocktail results in poor patient compliance. As a consequence of non-compliance there has been the emergence of multidrug-resistant (MDR) and extremely-drug-resistant (XDR) *Mtb* strains that impose a serious threat. The problem of drug-resistance and disease relapse after treatment is

a result of the discontinuation of standard DOTS course treatment. In most cases, there is poor patient compliance because the disease symptoms start to resolve within the first few weeks of treatment and the patient starts feeling better at which point they choose to discontinue the treatment. Current drugs available for TB treatment (according to WHO) have been classified into first, second and third line.

First line drugs: These drugs are the ones included in the current standard TB therapy (DOTS). All first-line drugs (FLDs) are designated by either a standard three-letter or a single-letter abbreviation: ethambutol (EMB or E), isoniazid (INH or H), pyrazinamide (PZA or Z), and rifampicin/rifampin (RMP or R),

Second line: drugs have been classified as second-line due to any of these possible reasons: associated toxic side-effects (e.g., cycloserine); less effective than the first-line drugs (e.g., p-aminosalicylic acid); or not available easily throughout the globe, particularly developing nations (e.g., fluoroquinolones). Following are the second-line drugs (SLDs) used for the treatment of TB: fluoroquinolones: e.g. moxifloxacin, ciprofloxacin, levofloxacin, aminoglycosides: e.g. amikacin, kanamycin; cyclic peptide: eg. caperomycin; thioamides: eg. ethionamide; clofazimine, cycloserine and p-aminosalicylic acid.

Third line: Other drugs that may be useful, but are not in the WHO list of SLDs, are categorized as third line drugs (TLDs). These drugs are considered TLDs because their efficacy has not yet been proven (e.g., linezolid, clarithromycin) and are mostly experimental drugs. The third line drugs are: rifabutin, macrolides: e.g. clarithromycin, linezolid, thioacetazone and thioridazine.

MDR-TB is defined by resistance to the two most commonly used drugs in the current first-line regimen, isoniazid and rifampin. XDR-TB**,** also known as Extensively Drug-Resistant TB, is emerging as an even more ominous threat. XDR-TB is defined as TB that is resistant to

any fluoroquinolone, and at least one of three injectable second-line drugs (capreomycin, kanamycin, and amikacin), in addition to isoniazid and rifampin. This makes XDR-TB treatment extremely complicated, if not impossible, in resource-limited settings. Due to the problems of MDR and XDR-TB, there is an urgent medical need for new drugs and treatment regimens that can better manage the latent tuberculosis infection (Koul, Vranckx et al. 2008; Singh, Manjunatha et al. 2008; Hugonnet, Tremblay et al. 2009).

In addition to the approved FLDs and SLDs for TB chemotherapy, there are at least four promising drugs under clinical development (TMC207, PA-824, OPC-67683 and SQ109) (Guy and Mallampalli 2008). Taken together, these drugs (approved for treatment or in the clinical development) target several different aspects of *Mtb* cellular structural and biological processes (e.g. transcription, protein synthesis, cell wall synthesis, catalase-peroxidase enzyme, ATP synthesis, DNA replication, and cofactor synthesis). Additionally, there is an ever evolving scope for new anti-TB drug discovery via novel anti-TB drug target identification and validation.

1.1 Identification and Validation of novel TB Therapeutic targets in post-genomic era: As described above, the ever-increasing number of *Mtb* infections, together with the evolution of MDR and XDR strains, has stimulated the search for novel anti-tuberculosis agents. The availability of the *Mtb* genome sequence (Cole, Brosch et al. 1998) has placed TB drug research in an unprecedented position. It has generated voluminous genetic and bioinformatic information to probe potential drug targets to combat TB.

One of the basic criteria for any chemical compound to be a potential anti-TB drug is its essentiality for bacterial virulence, growth or survival during infection. Identification of such essential targets in infectious disease drug discovery is challenging. Genes encoding the essential

targets are referred to as essential genes, that are absolutely required for bacterial growth or survival under defined conditions (Sassetti, Boyd et al. 2001; Sassetti, Boyd et al. 2003; Sassetti and Rubin 2003) . Many studies described in the literature rely only on bacterial growth as a measure of essentiality. Thus, they actually define genes required for growth and not survival. There is a need to understand several pathophysiological aspects of TB and then perform such essentiality assays. While an essentiality assay that exactly matches/mimics the disease physiology would be an ideal way forward, it still seems like a distant dream. Several assays/screens have been developed which closely mimic the disease physiology of TB. Here several strategies employed to identify and validate such "essential genes" in mycobacteria and understand their advantages and disadvantages will be discussed. In addition, other high throughput strategies like microarray, Transposon Site Hybridization (TraSH) will be mentioned.

1.2 Targeted/ defined Knockouts (gene deletions): Generating knockout mutants is usually considered the first step in determining the function of any particular gene. Preliminary evidence for gene essentiality generally comes up from the inability to delete/knockout a particular gene.

Several methods including the use of non-replicating vectors, linear DNA fragments and incompatible plasmids, have been employed to produce targeted/defined knockouts. Gene replacement can either be performed in a two-step recombination method using plasmids (homologous recombination) or as a single-step mycobacteriophage-mediated delivery method. However, these methods are limited due to the relatively high rate of illegitimate recombination, particularly when attempting to introduce an undesirable mutation (e.g. loss of function).

The simplest method to generate a targeted knockout (gene deletion) is by homologous recombination. In this strategy, the naturally occurring event of homologous recombination is

used to insert a delivery vector into the appropriate chromosomal location by a single crossover event (Parish, Gordhan et al. 1999). A second crossover event (resolution) can either restore back the original structure of the chromosome or result in a deletion of the targeted gene. In the absence of selection, both second crossover events should occur at equal frequency. If the targeted gene is essential, the second crossover event has to preserve the original structure of the chromosome (and hence, the corresponding gene's function). If sufficient number of resolution reactions fail to remove the gene (which can be verified by a simple colony PCR based screening method), it is very likely the gene would be essential. Essentiality can further be confirmed by adding a second copy of the target gene, integrated into the chromosome (Movahedzadeh, Wheeler et al. 2010). If the resolution reactions occur only in the presence of a functional second copy, this strongly suggests gene essentiality. While this method is simple, it is often not persuasive, since the readout can be biased by the specific attributes of the constructs that are employed. For example, inserts might create independent toxicity (Movahedzadeh, Wheeler et al. 2010). However failure to disrupt a gene is indirect evidence of essentiality and does not provide any information on the functional role of the gene product.

Considerable progress has been made in using homologous recombination as a method for generation of deletion mutants (Balasubramanian, Pavelka et al. 1996; Hinds, Mahenthiralingam et al. 1999; Parish, Gordhan et al. 1999; Pavelka and Jacobs 1999). A two-step method utilizing a suicidal delivery construct (Parish and Stoker 2000) has been used to produce defined unmarked mutants of several genes (Parish and Stoker 2000; Parish and Stoker 2002; Movahedzadeh, Smith et al. 2004; Movahedzadeh, Wheeler et al. 2010). Two series of vectors have been described, the first of which, named pNIL, allows manipulation of the target gene sequence . The second series, named pGOAL, contains marker cassettes flanked by PacI

restriction enzyme sites. The suicidal delivery plasmid is constructed by cloning a marker cassette from a pGOAL vector into the single PacI site of the pNIL vector containing the modified gene of interest. This plasmid lacks a mycobacterial origin of replication (oriM), and is thus unable to replicate in *Mtb* (making it a suicidal plasmid). The suicidal plasmid can be pretreated with UV light or an alkaline solution to stimulate homologous recombination. In the two-step strategy, single cross-over events are first selected and then screening for the second crossover is carried out to yield the mutant strains. Construction of the suicidal delivery plasmid has often been problematic as several cloning steps are required to include the relevant markers and finding appropriate restriction sites for inserting these genes has become limiting. Two series of vectors (pNIL/GOAL) have been developed (Parish and Stoker 2000) to overcome the cloning bottleneck. Separating the mutagenesis of the target gene from the inclusion of the required marker genes, makes the construction of the delivery plasmid rapid and flexible.

1.3 Site-specific Recombination system

One of the major limitations of homologous recombination is its low rate of occurrence. One way to overcome this is to use a high frequency method like site-specific recombination. Pashley and Parish describe the mycobacteriophage L5 excisionase-based site specific recombination system where excisionase enzymes efficiently remove DNA integrated into the ΦL5 attB site (Pashley and Parish 2003). A merodiploid strain is designed in which the only copy of the gene in question is integrated at the L5 attB site with the native copy of the gene being deleted. Introduction of another vector carrying the L5 excisionase gene will remove the integrated copy. For non-essential genes, transformation efficiency of the excisionase-carrying plasmid should be the same in all 3 strains (wild type, strain wherein one of the native genes is integrated into attB site, merodiploid strain with the only functional gene copy integrated into

attB site). On the other hand, if the target gene is essential, no transformants will be obtained in the strain where the native gene has been deleted.

1.4 Antisense oligonucleotides

Inhibiting gene expression using antisense oligonucleotides has been an alternative strategy for identifying essential genes, determining gene functions, and identifying potential drug targets in *Mtb* (Harth, Zamecnik et al. 2000; Harth, Horwitz et al. 2002; Harth, Zamecnik et al. 2007; Bai, Xue et al. 2010). A further advantage of the methodology is that it can directly be utilized for therapeutic intervention, although stability and efficient delivery of oligonucleotides remains a challenge. Yuanyuan Li and coworkers describe the effects of phosphorothioate-modified antisense oligodeoxyribonucleotides (PS-ODNs) against the mRNA of inositol-1-phosphate (I-1-P) synthase of *Mtb*, a key enzyme for inositol biosynthesis (Li, Chen et al. 2007). Real-time RT-PCR analysis revealed a significant reduction in the mRNA expression of I-1-P synthase upon addition of 20 μM PS-ODNs. Treatment with antisense PS-ODNs also reduced the level of mycothiol and the proliferation of *Mtb* and enhanced susceptibility to antibiotics. The experiments indicated that not only is I-1-P synthase essential for proliferation of *Mtb* but the antisense PS-ODNs could successfully enter the *Mtb* cytoplasm and inhibit its expression (Li, Chen et al. 2007).

Harth and coworkers described a proof of principle therapeutic strategy against tuberculosis using antisense technology. The study describes the effects of sequence-specific antisense PS-ODNs targeting different regions of each of the 30/32-kDa protein complex (antigen 85 complex) encoding genes involved in the multiplication of Mtb. Addition of single PS-ODNs targeting one of the three mycolyl transferase transcripts, added over a 6-wk observation period, inhibited bacterial growth by one log unit. A combination of three PS-ODNs

specifically targeting all three transcripts inhibited bacterial growth by about two logs. This study further described several critical aspects of PS-ODNs-based inhibition of mycolyl transferase such as the stability of PS-ODNs in culture over a 6-wk period, target specificity and the effect of the target site secondary structure on inhibition efficiency (Harth, Horwitz et al. 2002).

Harth and coworkers further extended their proof of principle to another gene involved in glutamine synthesis. They describe the effects of PS-ODNs against the mRNA of glutamine synthetase of *Mtb*, an enzyme whose export is associated with pathogenicity and with the formation of a poly-l-glutamate/glutamine cell wall structure. Treatment of virulent *Mtb* with 10 μM antisense PS-ODNs reduced glutamine synthetase activity and its expression by 25–50% depending on whether one, two, or three different PS-ODNs were used and the PS-ODN's specific target sites on the mRNA. Treatment with antisense PS-ODNs reduced *Mtb* growth by 0.7 (one PS-ODN) to 1.25 logs (three PS-ODNs). Although the inhibition is marginal and indicating a more potent ODN technology is required, it provides evidence that the antisense PS-ODNs enter the cytoplasm of *Mtb* and bind to their cognate targets. These studies demonstrate the potential of using antisense ODNs against *Mtb* (Harth, Zamecnik et al. 2000; Harth, Zamecnik et al. 2007).

There have been attempts to develop antisense oligonucleotides as a workable therapeutic strategy against *Mtb*. To improve the stability of ODNs, Li and coworkers combined the ODNs with the biodegradable polymer chitosan to form chitosan-ODN nanoparticles. The chitosan-ODN nanoparticles partially protected the encapsulated ODN from nuclease degradation. Chitosan-ODN nanoparticles were also much more effective in inhibiting the proliferation of *Mtb* than free ODNs (Li, Chen et al. 2009).

Das *et al*. describe the development of a novel thiocationic lipid-based formulation of PS-ODNs showing inhibitory activity against *Mtb*. PS-ODNs were designed based on sequences complementary to five essential regions of the mycobacterial genome. These included the superoxide dismutase sodA gene (TBS3), catalase-peroxidase katG gene (TBK1, TBK10), RNA polymerase beta-subunit rpoB gene (TBR5), and diaminopimelate decarboxylase lysA gene (TBL5). The effect of PAOs (TBS3, K1, K10, R5 and L5) alone on *Mtb* was not significant in comparison with the no-drug control. Liposomal formulations of PAOs resulted in statistically significant inhibition compared with PAOs alone, thiocationic liposomal control and the liposomal components (Das, Dattagupta et al. 2003).

1.5 Regulated Promoters (conditional expression)

Generating defined knockouts/mutants is the best way of understanding the function of genes and their importance in pathogenesis. However, mutagenesis of the essential genes is lethal meaning without these genes, the microorganism dies. Also, the significance or essentiality of such genes needs to be studied at various stage of disease progression (acute vs. chronic infection). A number of essential genes have been identified or predicted (Parish and Stoker 2000; Zahrt and Deretic 2000; Sassetti, Boyd et al. 2001; Parish and Stoker 2002; Sareen, Newton et al. 2003; Sassetti and Rubin 2003), thus generating conditional mutants to understand the function of each essential gene is required. The most common way to construct a conditional mutant is by controlling the gene expression with a regulated/inducible promoter.

Two different regulated promoters have been extensively described in mycobacteria. First, the acetamidase promoter that can be induced by acetamide and is derived from *M. smegmatis*. This promoter is most useful in *M. smegmatis*, however this system cannot be tightly

switched off and hence may not be suitable for the study of an essential gene since there is always a basal level of expression. Moreover there are reports of its instability in *Mtb* (Brown and Parish 2006). Second, the tetracycline-regulated promoter and several of its variations, each has distinctive characteristics - the strength of the promoter, degree of inducibility, and background level of transcription in the absence of the regulator - each of which determine how appropriate they are for studying individual genes. The tetracycline–regulated promoter (Ptet) has opened a new era in the study of essential genes in many microorganisms, including *Mtb* (Carroll, Muttucumaru et al. 2005). Use of Ptet systems in mycobacteria has been extensively published and has been used to make a conditionally auxotrophic mutant of *Mtb* trpD (Carroll, Muttucumaru et al. 2005). The Ptet system consists of two divergent promoters, one which drives expression of the TetR repressor and one which drives expression of the target gene. Two operator sites, which are the target of TetR binding, overlap the promoters. In the absence of tetracycline, TetR binds to the operator sites and prevents transcription. In the presence of tetracycline, TetR binds to tetracycline, which induces a conformational change, making it unable to interact with the operator, so that target gene expression can occur. The Ptet system has been used to regulate gene expression both *in vitro* and *in vivo* in mycobacteria (Blokpoel, Murphy et al. 2005; Ehrt, Guo et al. 2005; Ehrt and Schnappinger 2006; Guo, Monteleone et al. 2007; Klotzsche, Ehrt et al. 2009). This inducible promoter can be switched on and off in a controlled manner. One advantage of the tetracycline-inducible system is the ability of the inducer (tetracycline) to enter eukaryotic cells and that it can be administered via drinking water during *in vivo* infections. This allows the control of mycobacterial gene expression *in vivo* as well as *in vitro*.

In conclusion, comprehensive mutational studies, combined with the underlying structural information should identify and validate novel targets, and provide the best opportunity for developing effective new drug therapies, improved diagnostic techniques and better tools to understand host–pathogen interactions in cases of TB infection.

1.6 The Tuberculosis Structural Genomics Consortium and its efforts for TB drug discovery: The Tuberculosis Structural Genomics Consortium (TBSGC, http://www.webtb.org) established in 2000, is an international collaboration of researchers whose primary objective is to determine the three-dimensional structure of *Mtb* proteins with the intent of improving TB diagnosis and treatment (Goulding, Perry et al. 2003; Smith and Sacchettini 2003; Terwilliger, Park et al. 2003; Arcus, Lott et al. 2006; Baker 2007; Holton, Weiss et al. 2007; Ioerger and Sacchettini 2009; Musa, Ioerger et al. 2009; Chim, Habel et al. 2011). TBSCG is one of the seven original initiatives funded by the U.S. National Institutes of Health under their Protein Structure Initiative (PSI). Today, the Consortium comprises of over 100 collaborating research labs spanning more than 15 countries around the world, including structural genomic efforts in Germany, New Zealand and India (Arora, Chandra et al. 2011). TBSGC has applied the concept of a high-throughput pipeline to determining the structures of functionally important proteins in *Mtb*, rather than merely focusing on novel protein folds or structural biology of proteins with unknown functions. TBSGC employs state-of-the-art technologies for high throughput gene cloning, protein expression and structure determination by X-ray crystallography. Most of the relevant ≈ 4000 ORFs annotated in the TB genome have been cloned into the Gateway system (Invitrogen Inc.), which affords a rapid and convenient method for recombination into expression systems (Walhout, Temple et al. 2000; Underwood, Vanderhaeghen et al. 2006; Matsuyama and Yoshida 2009). For high-throughput expression and purification, the consortium has a protein-

production core facility at the Los Alamos National Laboratory which utilizes robotics, cell-free expression systems and high-throughput solubility assays. Purification is most routinely achieved by attaching an N-terminal 6-His-tag with a protease cleavage site, for separation on a nickel column (affinity capture), removal of the fragment via proteolysis and then polishing of the captured protein by size-exclusion chromatography (Rupp, Segelke et al. 2002; Rupp 2003). This methodology suffers from several limitations, with protein insolubility being a major problem. Some groups, with an aim to improve expression of soluble *Mtb* proteins, have utilized traditional techniques for trying to increase the solubility of insoluble proteins, such as making fusions with MBP (maltose binding protein) (Korepanova, Moore et al. 2007; Watkins and Baker 2008), SUMO (Small Ubiquitin like Modifier) or truncations of the target protein (Gutierrez-Lugo, Newton et al. 2006).

While most labs use commercially available screens, some labs also use custom-developed sets of solvents/conditions for optimization of crystal screening. The TBSGC has access to several high-intensity synchrotron beamlines for collecting X-ray diffraction data, including Advance Photon Source (APS), Advanced Light Source (ALS), and Stanford. Most structures have been solved via multi-wavelength anomalous dispersion (MAD) or by molecular replacement (MR) and some by single-wavelength anomalous dispersion (SAD). Advanced crystallographic software systems such as CCP4 and Phenix are used for data manipulation, phasing, automated model building, and refinement (Rupp, Segelke et al. 2002; Goulding, Perry et al. 2003; Musa, Ioerger et al. 2009).

According to a recent review describing the accomplishments of the TBSGC a decade after inception, approximately 250 *Mtb* protein structures have been elucidated by Consortium

members (Chim, Habel et al. 2011), accounting for over one third of the total *Mtb* structures deposited into the Protein Data Bank (PDB, http://www.rcsb.org).

The information from atomic resolution details of proteins, particularly in complex with their substrates, catalytic products or cofactors as well as protein-protein interaction networks, may aid scientists in interpreting their genetic and biochemical data and may ultimately be explored in the rational structure-based design of therapeutics for TB (Arcus, Lott et al. 2006). While the ultimate long term goal of TBSGC is to determine high-resolution structures of new drug targets in *Mtb*, targeting specific proteins of *Mtb* is not as straightforward of a process. The TBSGC uses a bioinformatics based multi-dimensional approach to prioritize drug targets. This approach takes into account all available background information on drug ability, enzyme pathway analysis, corresponding homologous proteins in humans, essentiality (e.g. via gene deletions or transposon mutagenesis), and gene expression under different persistence/dormancy/stress conditions to identify genes whose inhibition might lead to bacterial cell death (or attenuation of virulence).

The valuable information available in terms of high-resolution molecular structures provides the starting-point for many new research avenues and future drug design initiatives. The rich structural information available via the TBSGC initiative calls for more collaborative research efforts in the TB community within the biological, molecular genetics, drug discovery and medicinal chemistry domains to ultimately develop new therapeutic means to combat TB.

1.7 Successful heterologous expression of *Mtb* proteins, a difficult yet achievable task: The formation of insoluble inclusion bodies, associated with recombinant protein over-expression in *E. coli* is a particularly pronounced obstacle with *Mtb* proteins (Bellinzoni and Riccardi 2003). One of the strategies adopted to overcome the formation of inclusion bodies is to use an

expression host like *Mycobacterium smegmatis* that is more closely related to *Mtb*. Bashiri and coworkers describe the expression of FGD1, an F420-dependent glucose-6-phosphate dehydrogenase from *Mtb* (FGD1 is essential for activation of the anti-TB compound PA-824), using an *M. smegmatis*-based expression system. Initial attempts to produce recombinant FGD1 using *E. coli* as a host were unsuccessful; however, soluble protein yields of seven mg/L of culture were achieved when the alternative host *M. smegmatis* was used (Bashiri, Squire et al. 2007). The study used the pYUB1049 plasmid, a T7 promoter–based vector which expresses recombinant FGD1 by IPTG or autoinduction. The study described purification of both native and selenomethionine-substituted FGD1 by culturing *M. smegmatis* in autoinduction media protocols. Using the autoinduction media offered reduced handling time, as cultures did not require monitoring of optical density and induction, and reduced cost by removing the need for expensive ADC enrichment normally used in mycobacterial cultures. Selenomethionine was efficiently incorporated at levels required for multi-wavelength anomalous diffraction (MAD) experiments for crystal structure determination. Native and selenomethionine-labeled FGD1 were successfully crystallized by vapor diffusion, with the crystals diffracting to 2.1A˚ resolution (Bashiri, Squire et al. 2008).

Goldstone and coworkers have successfully adapted the *M. smegmatis* expression vector pYUB1049 (Bashiri, Squire et al. 2007) into the Gateway cloning system by the addition of *att* recombination recognition sequences. The resulting vector, designated pDESTsmg, is compatible with Gateway methods for *E. coli* expression. A target can be subcloned into pDESTsmg by a simple LR reaction using an entry clone previously generated for *E. coli* expression, removing the need to design new primers and re-clone target DNA. Furthermore, the method describes the use of the *M. smegmatis* strain, $mc^2$4517, as host for expression of *Mtb* proteins using the auto-

induction media. The described expression method was applied to a set of *Mtb* proteins that otherwise form inclusion bodies when expressed in *E. coli*. Five of the eight insoluble proteins become soluble when expressed in *M. smegmatis*, demonstrating this method as an efficient alternative salvage strategy (Goldstone, Moreland et al. 2008).

Bashiri et al. describe the generation of two different vectors with additional features, designed and prepared from the parental pYUB1049 vector. The pYUB28b vector can be used for restriction/ligation cloning of single genes with the capability of expressing N– and C–terminal His–tags, whereas the pYUBDuet vector is a co–expression vector for simultaneous expression of two genes in an *M. smegmatis* host. The pYUB28b and pYUBDuet vectors, together with the previously described pYUB1049 and pDESTsmg vectors, represent a repertoire of T7 promoter–based vectors which can be routinely used for expression of a wide range of ORFs in a mycobacterial host (Bashiri, Rehan et al. 2010).

Unfortunately, the Hsp60 chaperone GroEL1, which is relatively highly expressed, is often co-purified with polyhistidine-tagged recombinant proteins as a major contaminant when using the *M. smegmatis* expression system (Bashiri, Squire et al. 2007). This is due to a histidine-rich C-terminus in GroEL1.

In order to improve purification efficiency and yield of polyhistidine-tagged mycobacterial target proteins, Noens et al. created a mutant version of GroEL1 by removing the coding sequence for the histidine-rich C-terminus, termed GroEL1ΔC. GroEL1ΔC, which is a functional protein, cannot bind to nickel affinity beads. Using a selection of challenging test proteins, Noens et al. demonstrate that GroEL1ΔC is no longer present in protein samples purified from the GroEL1ΔC *M. smegmatis* expression strain. They also demonstrate the feasibility and advantages of purifying and characterizing proteins produced using this strain.

This GroEL1ΔC *M. smegmatis* expression strain allows efficient expression and purification of mycobacterial proteins while concomitantly removing the troublesome contaminant GroEL1 and increasing the speed and efficiency of protein purification (Noens, Williams et al. 2011).

Korepanova et al. cloned and expressed 70 integral membrane proteins from *Mtb* in *E. coli*. A combination of T7 promoter-based vectors with 6-His affinity tags and BL21 *E. coli* strains with additional tRNA genes to supplement sparsely used *E. coli* codons have been most successful. The authors demonstrate that the expressed proteins have a wide range of molecular weights and number of transmembrane helices. Expression of these proteins was observed in the membrane, the insoluble fraction of *E. coli* cell lysates; and in some cases, the soluble fraction. The highest expression levels in the membrane fraction were restricted to a narrow range of molecular weights and relatively few transmembrane helices (Korepanova, Gao et al. 2005).

Korepanova et al. further expressed 16 of 22 low molecular weight integral membrane proteins from *Mtb* in *E. coli* as fusions with both maltose binding protein (MBP) and His8-tag. There was a high yield (130mg/L) expression of 68% of targeted proteins in soluble and/or inclusion body forms. Thrombin cleavage of the MBP fusion protein was successful for 10 of 13 proteins expressed in soluble form and for three proteins expressed only as inclusion bodies. The use of autoinduction growth media increased yields over Luria-Bertani (LB) growth media in 75% of the expressed proteins. Expressing integral membrane proteins with yields suitable for structural studies from a set of previously low and non-expressing proteins has proved highly successful upon attachment of the maltose binding protein as a fusion tag (Korepanova, Moore et al. 2007).

Qin et al. describes a series of ligation independent cloning based vectors, constructed to express and purify 41 putative membrane proteins from *M. tuberculosis*. The efficiency for direct cloning of these target genes from PCR products was 95% (39/41). Over 40% of cloned genes were overexpressed in *E. coli* BL21 (DE3)-RP codon plus strain in the first round of expression screening. For those proteins which showed no expression, three protein fusion partners were prepared and it was found that each of the target proteins could be overexpressed by at least one of these fusions, resulting in the overexpression of two-thirds of the cloned genes (Qin, Hu et al. 2008).

1.8 Central Carbon Metabolism and Gluconeogenesis in *Mtb*: *Mtb* grows on a variety of substrates *in vitro* but mounting evidence indicates that during infection most of its energy comes from fatty acids (Boshoff and Barry 2005; Russell, Cardona et al. 2009). When bacterial metabolism is fueled by fatty acids, synthesis of sugars from intermediates of the TCA cycle (particularly the glyoxylate shunt) become important for growth and persistence (McKinney, Honer zu Bentrup et al. 2000; Sharma, Sharma et al. 2000; Munoz-Elias and McKinney 2005; Dunn, Ramirez-Trujillo et al. 2009). Hence, the glyoxylate shunt enzymes (malate synthase and isocitrate lyase) are considered potential targets for the development of new antibacterial agents (Sharma, Sharma et al. 2000; Smith, Huang et al. 2003). Phosphoenolpyruvate carboxykinase (PEPCK), the enzyme connecting the TCA cycle and gluconeogenesis, catalyses the reversible decarboxylation and phosphorylation of oxaloacetate (OAA) to form phosphoenolpyruvate (PEP). The PEPCK-encoding gene *pckA* is up-regulated by acetate or palmitate, but down-regulated by glucose. Deletion of the *pckA* gene of *Mycobacterium bovis* BCG led to a reduction in the capacity of the bacteria to infect and survive in macrophages (Liu, Yu et al. 2003). A study demonstrated that PEPCK plays a pivotal role in the pathogenesis of *Mtb*, as it is essential for

growth and survival of this pathogen during infections in mice and that *Mtb* relies primarily on gluconeogenic substrates for *in vivo* growth and persistence (Marrero, Rhee et al. 2010). Except for these recent studies, the role of gluconeogenesis in *Mtb* pathogenesis has remained largely unaddressed. Therefore, understanding the key structural and functional aspects of enzymes in the gluconeogenic pathway becomes important.

1.9 TrasH experiment and the *glpX (Rv1099c)* gene: Sasseti and Rubin describe the use of a method termed Transposon Site Hybridization (TraSH) to generate probes from a pool of transposon mutants that would hybridize to chromosomal regions adjacent to the site of transposon insertion on microarrays (Sassetti, Boyd et al. 2001; Sassetti and Rubin 2003). This method allows determining the positions of pooled probes in the chromosome and rapid identification of genes that lack insertions. Since the entire pool is analyzed simultaneously, the experiment can be performed under several different conditions (growth on different defined media, in macrophages with different activation states and at various points during murine infection) to define the various sets of required genes (Sassetti, Boyd et al. 2001; Sassetti, Boyd et al. 2003; Sassetti and Rubin 2003; Murry, Sassetti et al. 2008). A total of 194 genes that are specifically required for mycobacterial growth *in vivo* were identified. The behavior of these mutants provided a detailed view of the changing environment that the bacterium would encounter as infection progressed.

Sassetti and Rubin assessed the predictive value of the TraSH data. They isolated several mutant strains which were predicted to have varying *in vivo*-growth phenotypes. Mutants were isolated either by allelic exchange (Δ*bioF* and Δ*yrbE4A*) or by identifying strains with transposon insertions in appropriate genes (Tn::*fadD10* and Tn::*Rv1099c*). The *in vivo* growth rate of each strain was then determined relative to wild-type bacteria in mixed infections. Each of

the strains that were predicted to be attenuated for growth showed a significant defect compared with control strains. TraSH has much less resolution than sequencing and is more susceptible to false positive and negative results. However, it is far easier to perform multiple experiments. These findings warrant a logical investigation into the role of *glpX* gene in *Mtb* and also understanding the biochemistry of its translational product.

1.10 *glpX* gene in *Mtb* encodes a Class II fructose 1,6 bisphosphatase: Using genetic and biochemical methods, Movahedzadeh *et al*. have identified *Rv1099c* gene as a *glpX* gene that encodes a class II FBPase (Movahedzadeh, Rison et al. 2004). Fructose-1,6-bisphosphatase (FBPase, EC 3.1.3.11), a key enzyme of gluconeogenesis, catalyzes the hydrolysis of fructose-1, 6-bisphosphate to form fructose 6-phosphate and orthophosphate. This reaction is the opposite of that catalyzed by phosphofructokinase in glycolysis. The catalytic product of fructose 1,6-bisphosphate (i.e. fructose 6-phosphate) is an important precursor in various biosynthetic pathways generating important structural components of the cell wall and glycolipids in mycobacteria. In all organisms, gluconeogenesis is an important metabolic pathway that allows the cells to synthesize glucose from noncarbohydrate precursors such as organic acids, amino acids, and glycerol. FBPases are members of the large superfamily of lithium-sensitive phosphatases, which includes both the inositol phosphatases and FBPases. These enzymes demonstrate bivalent metal-dependent and lithium-sensitive phosphatase activity (Movahedzadeh, Rison et al. 2004).

Until recently, five different classes of FBPases have been identified based on their amino acid sequences (FBPases I to V) (Donahue, Bownas et al. 2000; Nishimasu, Fushinobu et al. 2004; Hines, Fromm et al. 2006; Hines, Chen et al. 2007). Eukaryotes possess only the FBPase I-type enzyme, but all five types exist in various prokaryotes. Many organisms have

more than one FBPase, mostly the combination of types I and II. The type I FBPase is the most widely distributed among living organisms and is the primary FBPase in *Escherichia coli*. An additional class II FBPase is encoded by the *glpX* gene in *E. coli*, which is part of the glycerol 3-phosphate regulon (Donahue, Bownas et al. 2000). The completion of the genome sequence of *Mtb* allowed the identification of genes that were predicted to encode enzymes for most central metabolic pathways (Cole, Brosch et al. 1998); however, no FBPase was initially assigned. Results from genetic and biochemical analyses revealed that the *Rv1099c* gene of *Mtb* encodes the missing mycobacterial FBPase (II) (Movahedzadeh, Rison et al. 2004). The protein encoded by the *Mtb glpX* (*Rv1099c*) gene is identical to other class II FBPase from *E. coli* (GlpX) (42% identity) and *Corynebacterium glutamicum* FBPase II (65% identity) (Rittmann, Schaffer et al. 2003). The outcomes of the Transposon Site Hybridization (TraSH) experiment further make investigation extremely essential using a rational approach (targeted gene deletion). Our results indicate a probable role of this enzyme in mycobacterial pathogenesis.

1.11 *glpX* gene in *Francisella tularensis* is essential for virulence in mice: The virulence of several *F. tularensis* SCHU S4 strain mutants was assessed by following the outcome of infection after intradermal infection. In this study the virulence of 20 in-frame deletion mutants and 37 transposon mutants was assessed (Kadzhaev, Zingmark et al. 2009). The majority of the mutants did not show an increase in prolonged time to death. However, mutations in six unique targets, *tolC, rep,* FTT0609, FTT1149c, *ahpC*, and *hfq* resulted in significantly prolonged time to death and mutations in nine targets, *rplA, wbtI, iglB, iglD, purL, purF, ggt, kdtA,* and *glpX* led to marked attenuation with an LD_{50} of 10^3 CFU. While the LD_{50} of wt strain was one CFU, the latter seven mutants showed a very marked attenuation with an LD_{50} of 10^7 CFU. The extreme

attenuation of Δ*glpX* mutant which showed an LD_{50} of 10^7 CFU, suggests that *glpX* is required for the virulence of *F. tularensis in vivo* (Kadzhaev, Zingmark et al. 2009).

1.12 Statement of intent

The goal(s) of the studies described in the subsequent sections can be divided into 3 important sections:

i. Investigate the genetic essentiality of the *glpX* gene in *Mycobacterium tuberculosis*. While the genome wide TraSH experiment indicates the *glpX* gene to be essential for survival in a mouse infection model, the method suffers several limitations, a heterogeneous transposon mutant population being one of them. Also, the method does not characterize individual transposon mutants. Hence, a section of the experimental work focuses on generation of an unmarked *glpX* gene deletion knock out (KO) in *Mtb*, and extensive characterization of this deletion KO for its *in vitro* and *in vivo* growth and survival behavior.

ii. Biochemical and structural characterization of the *glpX*-encoded FBPase (essential target) to evaluate its suitability for a structure based drug discovery (SBDD) project. A section focuses on the purification and biochemical characterization of this enzyme. Additionally, a section describes the structural biology of *Mt*FBPase.

iii. Verification of the biochemical activity/function of the *glpX* gene encoded protein of *F. tularensis*. A section of the experimental work focuses on verification of the biochemical function of the *glpX*-encoded protein of *F. tularensis*. The *glpX*-encoded protein has been purified using standard heterologous protein purification methods. Further preliminary studies have been carried out to evaluate the suitability of this target for SBDD.

Knowledge acquired from these studies will help in realizing our ultimate goal of identifying new chemical entities inhibiting this important metabolic target in pathogenic bacteria.

CHAPTER 2

GLPX GENE OF MYCOBACTERIUM TUBERCULOSIS IS REQUIRED FOR EUGONIC GROWTH ON GLUCONEOGENIC SUBSTRATES (*IN VITRO*) AND SURVIVAL IN MICE (*IN VIVO*)

2.1 RATIONALE AND HYPOTHESIS: *Rv1099c/glpX* is the major FBPase of *Mtb* (Movahedzadeh, Rison et al. 2004). The *glpX* transposon mutant of *Mtb* was predicted to be highly attenuated in Transposon Site Hybridization (TraSH) experiments (Sassetti and Rubin 2003). However, since the TraSH experiments can only predict essentiality and provide a mutant library as opposed to a single targeted gene knock-out (KO), the individual phenotypes of TraSH-identified essential gene(s) remain to be confirmed experimentally on a case by case basis. Also a deletion mutant provides the opportunity to test the growth and survival profiles both *in vitro* and *in vivo*. This chapter describes the generation of the *glpX* deletion mutant in *Mtb* and the effect of its deletion on growth and survival of *Mtb* (both *in vitro* and *in vivo*).

2.2 MATERIALS AND METHODS:

A. Cloning and mutagenesis of *Mtb glpX* gene

Plasmids and strains used in this work are shown in Table I and II. Mutagenesis was essentially carried out as previously described (Figure 1 and 2) (Parish and Stoker 2000). Briefly, the coding sequence of *M. tuberculosis* H37Rv *glpX* (Rv1099c) with flanking DNA, 893bp upstream and 970bp downstream of the gene, was amplified by PCR using *Pfu* Turbo DNA polymerase enzyme (Stratagene). The primers used, each at 300 nM, were GlpX_1 and GlpX_2 (Table II), and cycling conditions were: an initial 3 min at 95°C, then 35 cycles of 45 sec at 95°C, 30 sec at 52°C and 3 min at 72°C, and a final extension step of 72°C for 10 min to complete primer extension. PCR was carried out using 5% DMSO to compensate for the high

guanine-cytosine content of mycobacterial DNA. The PCR product was phosphorylated using polynucleotide kinase and cloned into the PmlI site of plasmid p2NIL, producing pFM143. To create a deletion in pFM143, 943bp were deleted in the *glpX* gene by inverse PCR, using Pfu Turbo DNA polymerase enzyme (Stratagene). The primers used, each at 300 nM, were glpX_Rev1 and glpX_Rev2 (Table II), and cycling conditions were: an initial 3 min at 95°C, then 30 cycles of 45 sec at 95°C, 30 sec at 63°C and 7 min at 68°C, and a final extension step of 68°C for 10 min to complete primer extension. PCR was carried out using 6% DMSO. The PCR product was Klenow polymerase-repaired and phosphorylated using polynucleotide kinase followed by re-ligation producing FM147. The deletion created was confirmed by sequencing. Following insertion of a gene cassette carrying the *lacZ* and *sacB* genes from pGOAL19 into the vector's PacI site (producing FM152); the DNA was introduced into *M. tuberculosis* H37Rv by electroporation. Cells carrying single-crossovers (SCOs) were isolated by selection for blue hygromycin resistant (hyg^R) and kanamycin resistant (kan^R) colonies on Middlebrook 7H11 agar supplemented with OADC (BBL Middlebrook, BD). One SCO colony was plated onto agar containing sucrose (2%) to isolate bacteria with a second crossover, which will lead to mutant or wild-type cells depending on the location of the recombination event.

In order to screen for *glpX* mutants, DNA was extracted from sucroseS kanS white colonies (obtained from plating *M. tuberculosis* onto sucrose medium) and analyzed by PCR using primers that flank the *glpX* gene (GlpX_1: CCCCTGGGGTATCATCATC and GlpX_2: GATCTGGCAGTTGGATCTCG). The mutant and wild type strains were distinguished by performing colony PCR; further the *glpX* deletion mutation was confirmed by Southern blotting.

B. Complementation of Δ*glpX*

A fragment containing the *glpX* gene together with 445 bp of upstream sequence was produced by PCR of *M. tuberculosis* genomic DNA using *Pfu* Turbo DNA polymerase enzyme (Stratagene). The primers used, each at 300 nM, were tbglpX_up and tbglpX_end (Table II), and cycling conditions were: an initial 3 min at 95°C, then 35 cycles of 45 sec at 95°C, 30 sec at 54°C and 2 min at 72°C, and a final extension step of 72°C for 10 min to complete primer extension. PCR was carried out using 5% DMSO. The PCR product was phosphorylated using polynucleotide kinase and cloned into the SmaI site of pBluescript-SK+ to produce pFM158. The *Hin*dIII cassette of pUC-Gm-int, carrying the *int* and *gm* genes was cloned into the *Hin*dIII site of pFM158 to produce pFM163. The plasmid was electroporated into HG1, yielding HG2.

TABLE I: STRAINS AND PLASMIDS USED FOR THE GENERATION OF ΔGLPX AND GLPX COMPLEMENT STRAINS.

Strains	**Characteristics**	**Source**
*M. tuberculosis*H37Rv	wild-type laboratory strain	ATCC 27294
HG1	*M. tuberculosis ΔglpX*	This study
HG2	*M. tuberculosis ΔglpX*::pFM163	This study
E.coli DH5α		Invitrogen
p2NIL	manipulation vector	(Parish and Stoker 2000)
pGOAL19	delivery cassette vector	(Parish and Stoker 2000)
pBluescript II SK+		Stratagene
pFM143	p2NIL::*glpx*	Movahedzadeh, unpublished data
pFM147	pFM143::*glpXΔ*	Movahedzadeh, unpublished data
pFM152	pFM147/with PacI cassette of pGoal19	Movahedzadeh, unpublished data
pFM158	pBluescript SK+::*glpX* (445/bp upstream)	Movahedzadeh, unpublished data
pFM163	pFM158::*intgm*	Movahedzadeh, unpublished data

TABLE II: LIST OF PRIMERS USED FOR THE GENERATION OF ΔGLPX.

GlpX_1	CCCCTGGGGTATCATCATC
GlpX_2	GATCTGGCAGTTGGATCTCG
Glpx_Rev1	AGCGCCGTGTACCCATTGCC
Glpx_Rev2	GTCACCCGGACCAGCTCCAT
tbglpX_up	GCTCTGGGTCAAGCTCAGAT
tbglpX_end	GGGCAATGGGTACACGGC

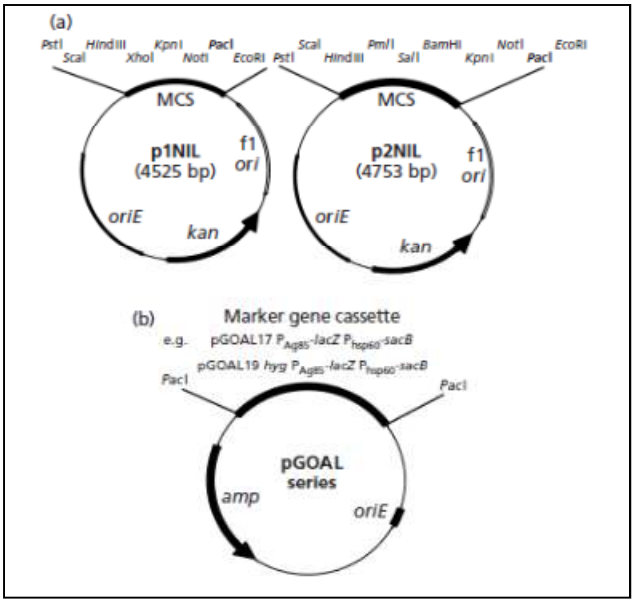


Figure 1. General features of the pNIL, pGOAL vector series. (a) p1NIL (not used in this study) and p2NIL cloning vectors. Each vector carries a *kan* gene for selection in both *E. coli* and mycobacteria, and oriE for replication in *E. coli* and a small MCS and a single PacI site. (b) pGOAL vectors carry different combinations of marker genes flanked by two PacI sites. Each vector also carries an *oriE* and *amp* for selection in *E. coli*. (Adapted with permission from (Parish and Stoker 2000))

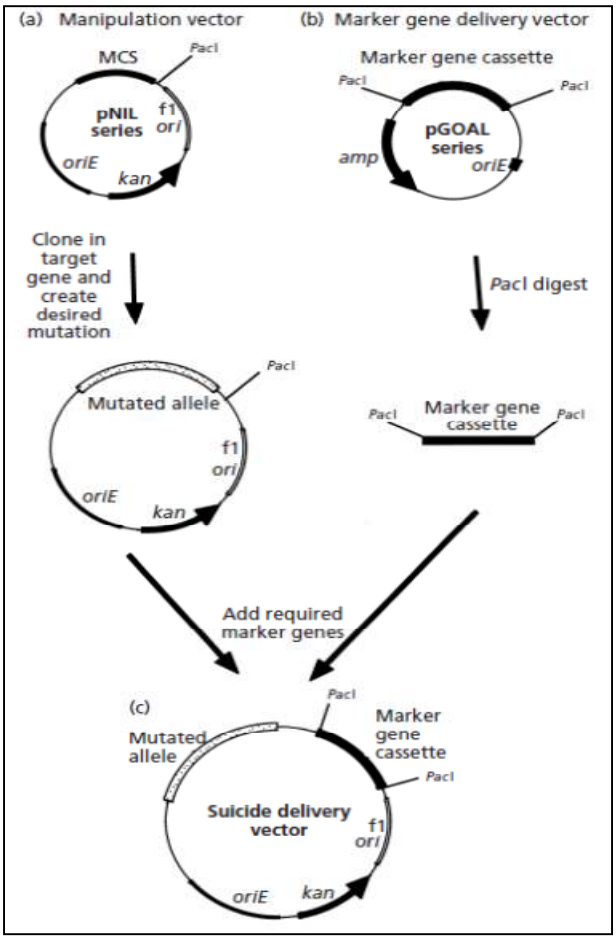


Figure 2. Cloning strategy for generating suicide delivery vector. (a) The target gene is cloned into p2NIL vector and deletion made using inverse PCR). (b) The pacI cassette containing the desired marker genes is then excised from the pGOAL vector and cloned into the single PacI site of the p2NIL harboring the mutated gene resulting in the final suicide delivery vector. (c) This final vector contains *oriE*, the kanamycin (*kan*) resistance and f1 origin. (Adapted with permission from (Parish and Stoker 2000).

C. Southern Blot analysis (confirmation of the *glpX* gene deletion):

1. DNA extraction: Mycobacterial DNA extraction was performed as described (Belisle, Mahaffey et al. 2009).

2. Neutral Transfer of DNA Nytran® SuPerCharge Nylon Membranes:

Neutral Transfer Buffers

Denaturing Buffer

0.5 M Sodium hydroxide (NaOH)

1.5 M Sodium chloride (NaCl) pH was adjusted to 7.0

20x SSC Transfer Buffer

3 M Sodium chloride (NaCl)

0.3 M Sodium citrate

Neutralizing Buffer

0.5 M Tris-Hydrochloride pH 7.0

1.5 M Sodium chloride (NaCl) pH was adjusted to 7.0 with concentrated HCl.

3. Denaturation: After digesting the genomic DNA with BamH1 (following the restriction enzyme manufacturer's instructions), the digested DNA was run on an agarose gel and denatured with denaturing buffer for 30 min. at room temperature with gentle shaking . Following this, the gel was rinsed with distilled water, and transferred to neutralizing buffer. The gel was then shaken slowly again for 30 min at room temperature. Finally, the gel was soaked in 20x SSC transfer buffer for 30 min with gentle shaking.

4. Transfer: A nylon membrane (Nytran SuPerCharge) is wetted and immersed in distilled water, followed by soaking in 20x SSC transfer buffer for 5 min. The "stack tray" of the transfer device is then placed on a level surface and 20 sheets of dry GB004 blotting paper (thick)

(Whatman) are placed in stack tray. Additionally 4 sheets of dry 3MM Chr blotting paper (thin) (Whatman) are placed on top of the stack. Finally one sheet of 3MM Chr blotting paper, pre-wet in transfer buffer, is placed on the stack. The transfer membrane is placed on the stack. The membrane is then covered with an agarose gel; which is previously cut to the size of the membrane, making sure there are no air bubbles between the gel and the membrane.

The top surface of the gel is wetted with transfer buffer, and 3 sheets of 3MM Chr blotting paper, presoaked in transfer buffer are placed on top of the gel. The "buffer tray" of the transfer device is attached to the bottom tray, using the circular alignment buttons to align both trays. The buffer tray is filled with transfer buffer. Transfer is started by connecting the gel stack with the buffer tray, using the precut "buffer wick" (included in each blotter stack), pre-soaked in transfer buffer. The wick was placed across the stack so that the short dimension of the wick completely covered the blotting stack and both ends of the long dimension extended into the buffer tray. Finally the "wick cover" is placed on top of the stack to prevent evaporation, making sure that the edges of the wick are immersed in the transfer buffer. Transfer was continued for 3 hrs to ensure complete transfer of the DNA onto the membrane. Following transfer, the transfer membrane was gently washed in 2x SSC for 5 min.

5. Drying (Immobilization): Post neutralization, the membrane was placed on a fresh sheet of dry 3MM Chr blotting paper to remove any excess 2x SSC buffer. As a precautionary measure, it was ensured that the membrane was damp enough and did not dry out. The membrane was placed in a UV chamber for cross linking and immobilization (UVP HL-2000 HybriLinker). The DNA could be easily covalently bonded to the membrane by cross-linking the molecule to the matrix in the presence of UV light. The blot was exposed to UV light (254 nm) in the chamber for a total dose of 120 mJ/cm^2.

6. Hybridization and detection: Hybridization and detection was carried out following the manufacturer's protocol for The Gene Images™ AlkPhos Direct™ labelling and detection system (GE Healthcare).

D. Bacterial Strains, media and growth conditions: Δ*glpX*, *glpX* complement and WT *Mtb* H37Rv strains were routinely grown using shake flask method (culture volume of 100ml in 500ml capacity flask) in a side arm flask at 37 °C in Middlebrook 7H9 liquid medium (Difco) containing 0.2% glycerol, 10% BBL Middlebrook OADC enrichment (Becton Dickinson) and 0.05% Tween 80. For growth with defined carbon sources, 7H9 medium with 0.05% Tyloxapol and a carbon substrate at 0.1% or 0.2% (wt/vol). Gentamycin was added at 10 µg/mL when required for selection. The optical density (OD) was measured using a Klett-Summerson Photoelectric Colorimeter; the optical density was reported in Klett Units. 1 Klett Unit (KU) = OD/0.002 or 500 KU = 1.0 (OD600).

E. Experimental methods for determination of *in vivo* growth and survival profile of Δ*glpX* strain: 108 female BALB/c mice (approximately eight week old) were used for the purpose of evaluating the effect of *glpX* gene knockout on *in vivo* growth. A total of 36 mice per strain were used for infection.

1. Bacterial culture preparation: The respective strains Δ*glpX*, Δ*glpX* complement and WT *Mtb* were grown in 7H9 + OADC medium for about 14 days until the optical density was 70 – 110 KU. The bacterial cultures were harvested by centrifugation at 3000 g for 10 minutes. The cell pellet was washed and re-suspended twice in phosphate buffered saline (PBS) with Tween 80 (0.01% w/v). The re-suspended cell culture was filtered through a Whatman polyethersulfone membrane filter, 47mm diameter, 0.8 µm. The cell stock was stored at - 80°C until further use,

Colony forming units (CFU) were estimated in a similar manner as described in subsequent sections of this chapter.

2. Infection of mice: BALB/c mice were infected by aerosol delivery as follows. An aliquot of the respective *Mtb* strain (or gene knock-out strain) was thawed, sonicated at 50% of the maximal power settings for 15 sec in a sonic cup horn (Ultrasonic, Inc.) to disperse any clumps, and diluted in PBS to an approximate bacterial count of 2 x 10^6 CFU/ml. The organisms were transferred to the nebulizer of a Middlebrook Aerosol Generation Device (Glas-Col Inhalation Exposure System, Model A4212, Terre Haute, IN), and an aerosol was generated with compressed air pressure of 15 cubic feet/hr (CFH) and a vacuum of 70 CFH for 30 min. These parameters effectively result in an infection of approximately 50-100 bacilli per mouse. Up to 100 mice can be infected in a single run (based on historical experimental data in our laboratory and the specifications provided by the manufacturer of aerosol generation device).

3. Enumeration of bacteria in mouse tissues: At indicated time points 6 mice from each group (infected with the respective strain) were sacrificed via carbon dioxide asphyxiation and the lungs from individual mice were aseptically removed, homogenized in Hank's balanced salt solution (HBSS) containing 0.05% Tween 80 (HBSS-Tw) using a motorized tissue homogenizer (ProScientific, Monroe, CT), and serially diluted (10-fold) in HBSS-Tw. The ideal dilution suitable for CFU enumeration is never fixed, hence, dilutions between 1:10 to 1: 1,000,000 are plated on 6 well plates with Middlebrook 7H11 solid medium supplemented with 0.5% glycerol and 10% OADC as the growth media. Plates were incubated at 37°C for about 17-21 days, CFU were enumerated thereafter. The animal protocol was approved by the animal care committee: Protocol number 12 – 074.

F. Measurement of cellular FBPase activity in *Mtb* strains: Bacteria were grown to mid-log phase in 7H9 + OADC enrichment, and collected by centrifugation. Each pellet was washed once in distilled water followed by re-suspension in 2.5 ml (2 mM dithiothreitol in 50 mM Tris-Cl, pH8). The suspended bacteria were disrupted in a laboratory bead beater for 3 cycles of 60 sec in Lysing Matrix B (0.1 mm silica beads), with cooling on ice between cycles for 120 sec. The resulting cell-extracts were then clarified at 4000 g for 10 min using a bench centrifuge and filter-sterilized through 0.22 µm pore Millipore milex filter (PVDF membrane). FBPase assays were conducted by incubating 2.0 mM substrate fructose 1,6-bisphosphate at 37°C for 10 min, as described previously (Gutka, Rukseree et al. 2011). The phosphate released thereby was estimated using the method described by Baykov *et al.* (Baykov, Evtushenko et al. 1988) with appropriate modifications for use in 96-well plates in a final reaction volume of 125 ul..

1. Color reagent. Concentrated sulfuric acid (60 ml) was slowly added to 300 ml of water. The solution was then cooled to room temperature and supplemented with 0.44 g of malachite green (Sigma Aldrich). The resulting orange solution is stable for at least 1 year at room temperature. On the day of use, 2.5 ml of 7.5% ammonium molybdate (Sigma Aldrich) was added to 10 ml of the dye solution followed by 0.2 ml of 11% Tween 20 (Fisher Scientific); this mixture is referred to as “color reagent”.

Important note: It is important to check that all enzyme preparations and assay buffers do not contain free phosphate. This was conveniently verified by adding 25 µL of the color reagent to 100 µL sample solution. The blank OD values at 630 nm were found to be consistently less than 0.1.

2. Preparation of phosphate standards: A Premix solution containing 40 μM phosphate was prepared by pipetting 40 μL 1 mM phosphate standard to 960 μL distilled water or enzyme reaction buffer.

3. Color formation and its measurement: 100 μL of standard solutions were transferred (in triplicate) into wells of a clear-bottom 96-well plate. Appropriate blank controls containing water or reaction buffer only, were also incorporated in the assay. 100 μL test samples (appropriately diluted) were transferred into separate wells of the plate. 25 μL of color reagent was added and mixed in each well by tapping the plate. Plates were then incubated for 30 min at room temperature for color development. Absorbance was measured at (630 nm) on a plate reader. The green color complex has a λ_{max} in the range of 600 nm - 660nm.

4. Phosphate determination in samples: The enzyme reaction was terminated directly by the addition of the color reagent. One volume (25 μL) of the color reagent was mixed with four volumes of the solution to be analyzed (100 μL), the mixture was allowed to stand for 10 min, and the absorbance at 630 nm was measured. Enzyme activity in *Mtb* cells is then calculated by subtracting the phosphate formed in wells with cell extract and substrate, from phosphate formed in corresponding wells with cell extract but without substrate.

G. Additional phenotypic characterization of Δ*glpX*:

1. Hydrogen peroxide sensitivity: *Mtb* cultures (Δ*glpX,* Δ*glpX* complement and *M. tuberculosis* H37Rv WT) were grown in 7H9 medium + OADC (100ml) until late log or early stationary phase (Klett units approximately = 150) and then treated with 1ml of 500mM hydrogen peroxide solution (effective peroxide concentration = 5mM). 1 ml samples were taken as controls before

hydrogen peroxide treatment (referred to as 0 hr /input sample). Additional samples (1 ml) were taken at 2 hrs, 6 hrs and 12 hrs after hydrogen peroxide treatment.

2. pH sensitivity: Late log or early stationary phase *Mtb* cultures were treated with 6ml of sterile 2 mM acetic acid solution (effective pH ≈ 4.5). 1 ml sample was taken as control before addition of acetic acid (referred to as 0 hr /input sample). Additional 1ml samples were taken at 2 days, 4 days and 6 days after pH change to ≈ 4.5 units.

3. Nitrosative stress: Late log or early stationary phase *Mtb* cultures were treated with 10ml of sterile 2.2 mM DETANO solution (effective DETANO concentration = 0.2 mM). 1ml sample was taken as control before DETANO treatment (referred to as 0 Hr /input sample). Additional 1ml samples were taken at 1 day, 2 days and 3 days after DETANO treatment.

CFU determination: The samples were diluted 10, 10^2, 10^3, 10^4 and 10^5 fold, respectively, using PBS as a diluent. The respective dilutions were plated (in duplicate) on plates containing 7H11 + OADC enrichment.

2.3. RESULTS AND DISCUSSION:

A. Successful generation of unmarked deletion mutant of *glpX* gene in *Mtb*: The mutant and wild type strains were distinguished by performing colony PCR, and the *glpX* deletion mutation was confirmed by Southern blotting. Strains carrying the *glpX* wild-type allele would produce a band of 2841 bp (lane 4, Figure 3a) whereas strains carrying a *glpX* deletion mutation would produce a smaller band of 1898 bp, accounting for deletion of 943bp (lanes 2,3 Figure 3a.). Mutant candidates and a wild-type control were digested with BamH1 and subjected to Southern blot analysis using a 2841 bp *glpX* probe (*glpX* plus ≈1kb flanking regions). The wild-type strain showed 3 bands of 6 kb, 4 kb and ≈ 600bp band whilst the mutant showed 2 bands of 6 kb and 4 kb respectively on the Southern blot, accounting for 1 lost BamH1 site (Figure 3b). The Southern blot results can be better understood by the cartoon representation of the genomic region of Rv1099c and the methodology used for design of the Δ*glpX* (Figure 4).

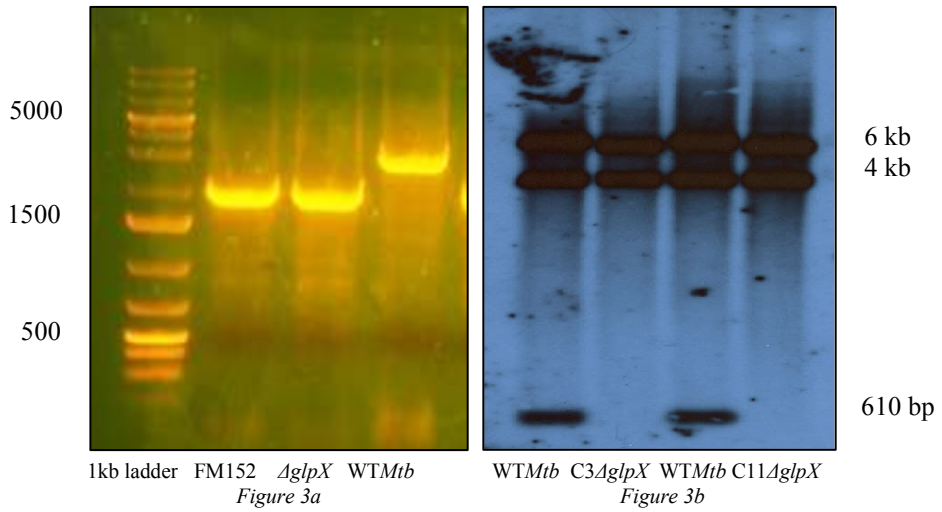


Figure 3 Colony PCR and Southern Blot confirming the deletion of *glpX* gene. A. Colony PCR results of the potential double cross overs (DCOs): Lane 1: 1kb ladder, Lane 2: PCR amplification product of suicidal delivery vector FM152 is about 2kb, Lane 3: PCR product of one such potential DCO (Colony 11) is 2kb, Lane 4: PCR product of a WT *Mtb* is about 3kb. **B. Southern blot:** Lane 1: WT genomic DNA digest with *Bam*H1 which gives a fragment of about 610 bp (lowermost band in lane 1 and 3), Lane 2: Genomic DNA digest of a potential DCO (Colony 3) with the 610 bp fragment missing, Lane 3: Same as Lane 1, Lane 4: Genomic DNA digest of a potential DCO (Colony 11) with the 600 bp fragment missing.

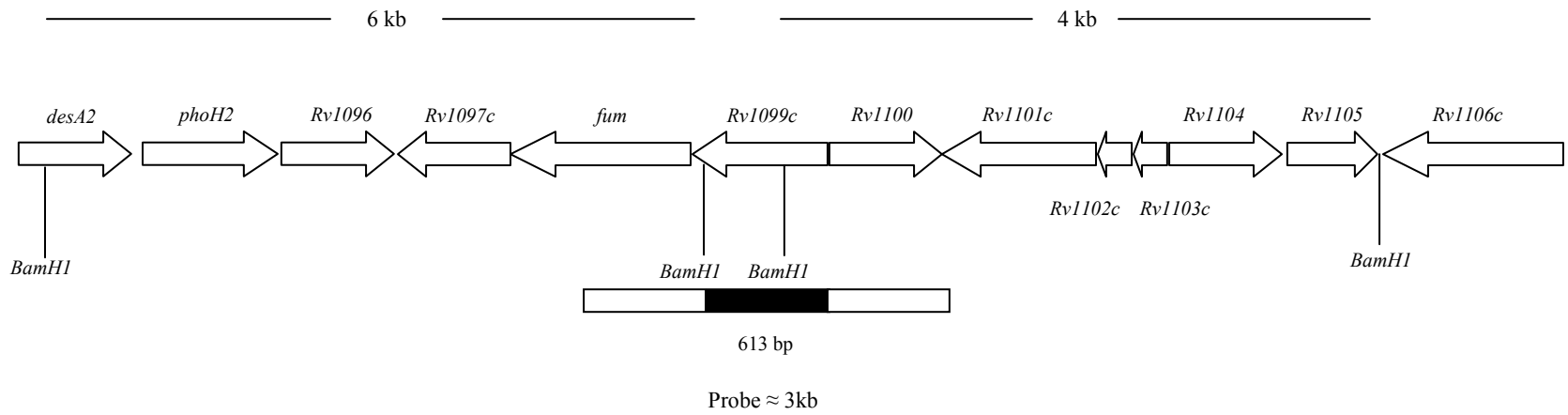


Figure 4. Genomic region of wild-type *Rv1099c* and design of the disrupted allele. Vertical lines indicate the 4 BamH1 restriction sites throughout the genomic region. Probe size for Southern blot ≈ 3kb. Since BamH1 site 3 (3rd from left in the above figure) is lost in the *glpX* KO, the southern blot for KO displays only 2 bands of 4 kb and 6 kb respectively (lanes 2 and 4).

B. *In vitro* growth profile of Δ*glpX* strain (glycerol, dextrose and acetate as carbon sources):

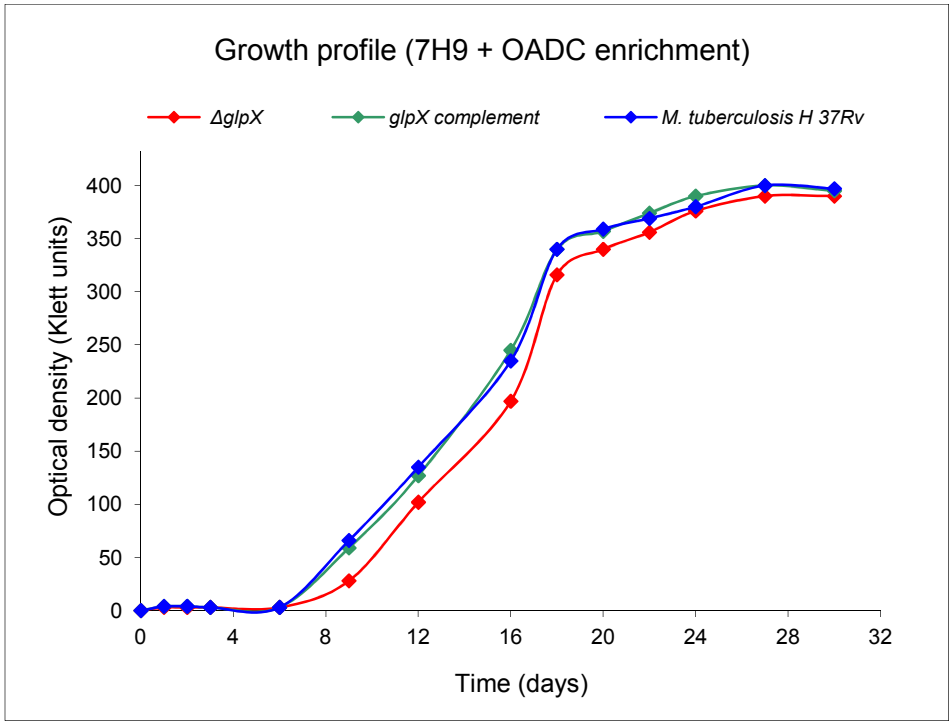


Figure 5. Growth profile of Δ*glpX* on 7H9 medium + OADC enrichment. The shown growth profile is representative of a triplicate data set.

The growth profile of Δ*glpX* is similar to that of WT *Mtb* when grown in 7H9 medium with OADC enrichment (Figure 5). This indicates that the disruption of *glpX* does not significantly affect the growth profile of *Mtb*, in enriched media. The growth medium used in this study is highly complex and contains OADC (Oleic acid, Albumin, Dextrose and Catalase), plus glycerol and Tween 80, a surfactant which can furnish *Mtb* with oleic acid (a fatty acid carbon source) through de-esterification (Dubos, Davis et al. 1946). For studies involving growth of Δ*glpX* on defined carbon sources, defined media

was used containing equal concentrations of dextrose, acetate or glycerol as single carbon sources and replacing Tween 80 with a non-hydrolysable detergent, Tyloxapol.

Studies (Youmans and Youmans 1953; Youmans and Youmans 1953; Youmans and Youmans 1953; Youmans and Youmans 1954; Youmans and Youmans 1954; Munoz-Elias and McKinney 2005) indicate that *Mtb* grows fastest *in vitro* on glycerol, less quickly on dextrose, and least quickly on acetate.

As expected, in aerated batch cultures, WT *Mtb* grew fastest and achieved the highest growth rate on glycerol, followed by dextrose and acetate, similar to the earlier studies (de Carvalho, Fischer et al. 2010). The ability of Δ*glpX* to grow on glycerol or acetate as a sole carbon source was severely compromised as compared to WT *Mtb* (Figures 6 and 7). However, Δ*glpX* grew twice as fast as WT *Mtb* on dextrose as a sole carbon source. The slower growth profile of Δ*glpX* on glycerol or acetate was expected since *glpX* encodes a functional fructose 1,6-bisphosphatase (FBPase) required for growth on gluconeogenic substrates (glycerol/fatty acids/acetate) (Movahedzadeh, Rison et al. 2004; Gutka, Rukseree et al. 2011). Also the slow growth of Δ*glpX* on glycerol as a sole carbon source was demonstrated by the phenotype observed on solid 7H11 media (Figure 8).

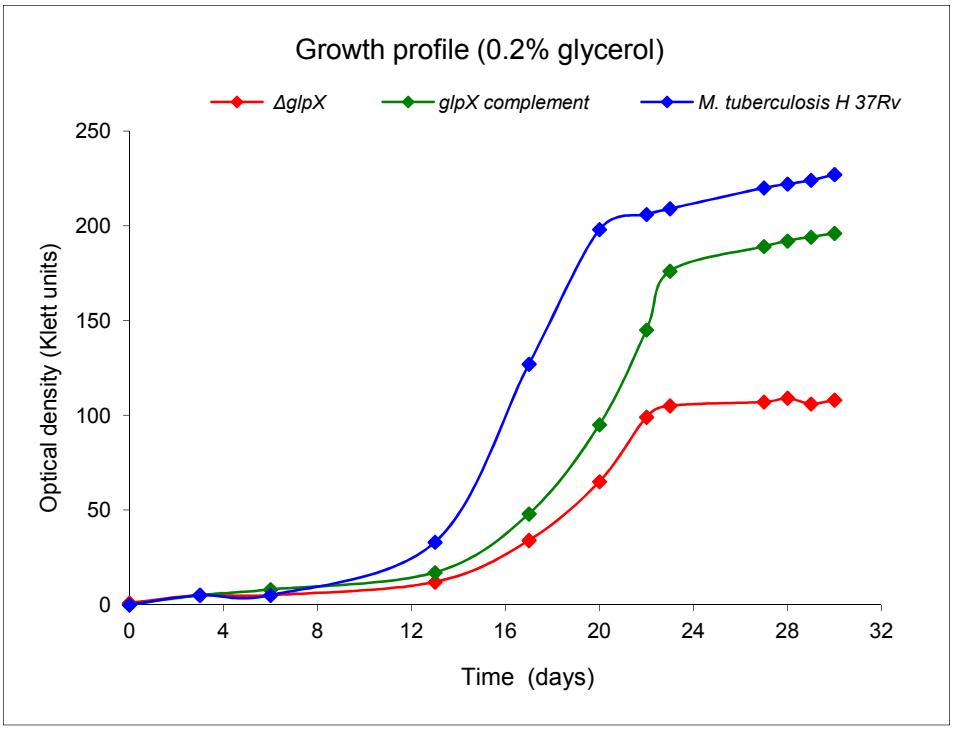


Figure 6. Growth profile of *ΔglpX* in 7H9 medium + 0.2% glycerol. The shown growth profile is representative of a triplicate data set.

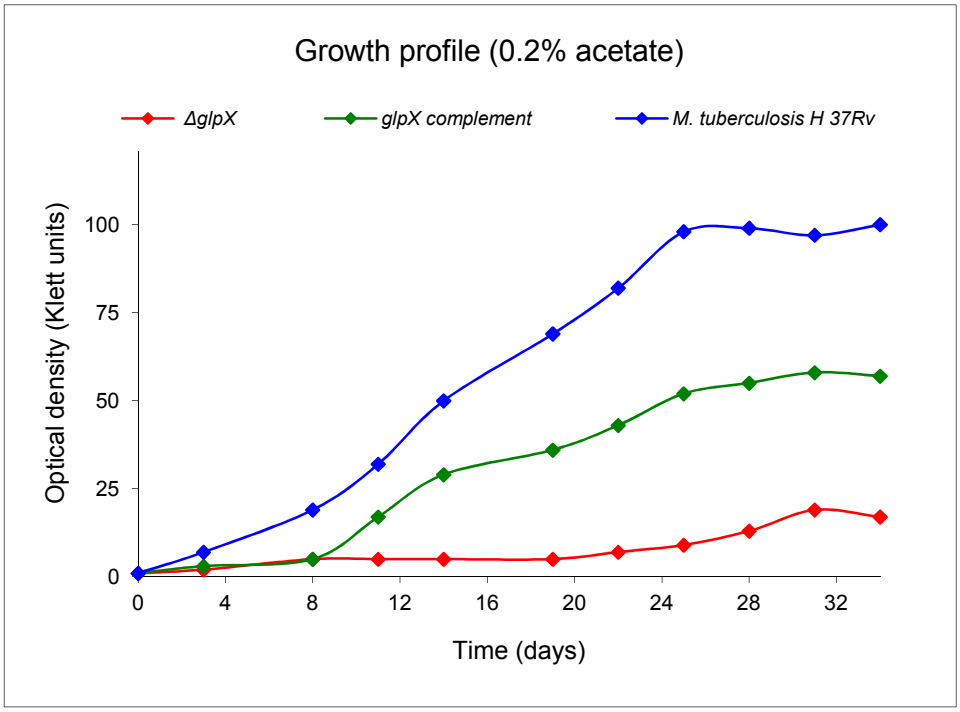


Figure 7. Growth profile of Δ*glpX* in 7H9 medium + 0.2% Acetate. The shown growth profile is representative of a triplicate data set.

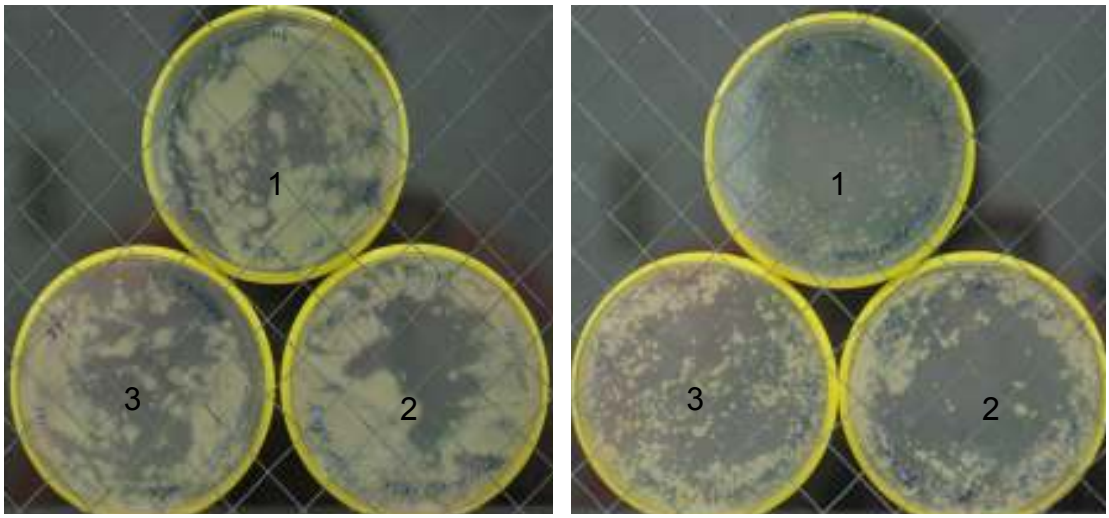


Figure 8. Growth characteristics of *ΔglpX* on solid medium. Growth of the *ΔglpX* (1); *Δ glpX* + complement (2) and *Mtb* H37Rv (3) on 7H11 + OADC enrichment (left panel) and 7H11 + 0.5% glycerol solid medium (right panel). (Images taken 21 days post-incubation of the plates at 37°C).

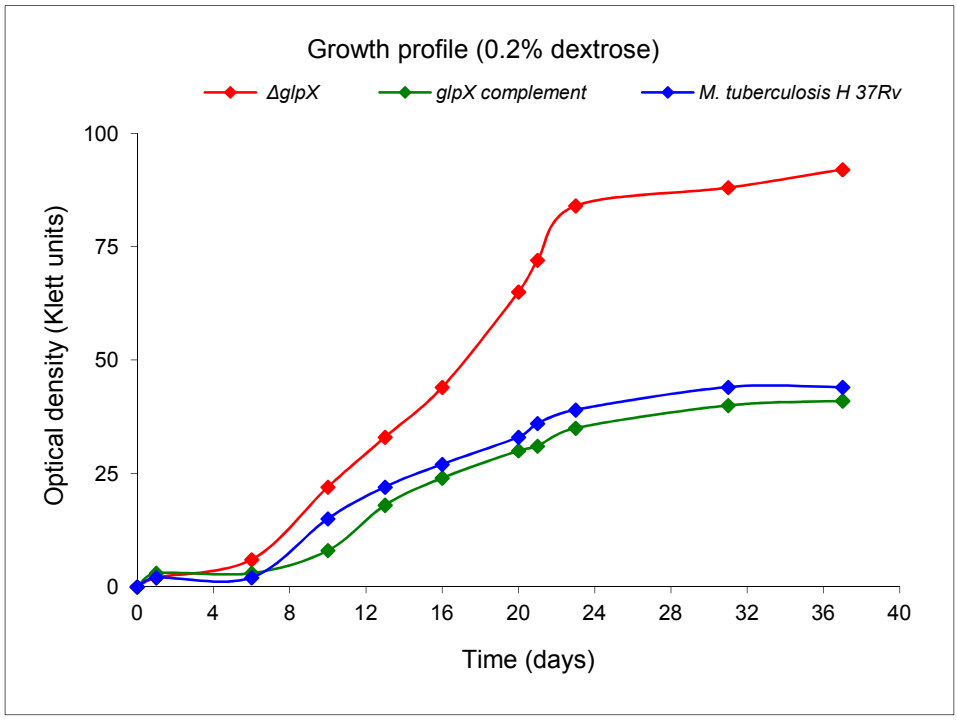


Figure 9. Growth profile of Δ*glpX* in 7H9 medium + 0.2% dextrose. The shown growth profile is representative of a triplicate data set.

The fast growth of Δ*glpX* on dextrose (Figure 9) is most likely a result of the disruption of a possible regulatory mechanism of FBPase with the phosphofructokinase (encoded by *pfk1* and *pfk2* in *Mtb*) enzyme (which catalyzes the reverse reaction converting fructose 6-phosphate to fructose 1,6-bisphosphate). However the presence of such a regulatory mechanism for FBPase is not experimentally verified in *Mtb*.

Both *ΔglpX* and WT *Mtb* failed to grow significantly in the absence of any external carbon source (glycolytic/gluconeogenic) as indicated in the growth profile (Figure 10). Figure 11 depicts the metabolic pathway diagram of central carbon metabolism in *Mtb*.

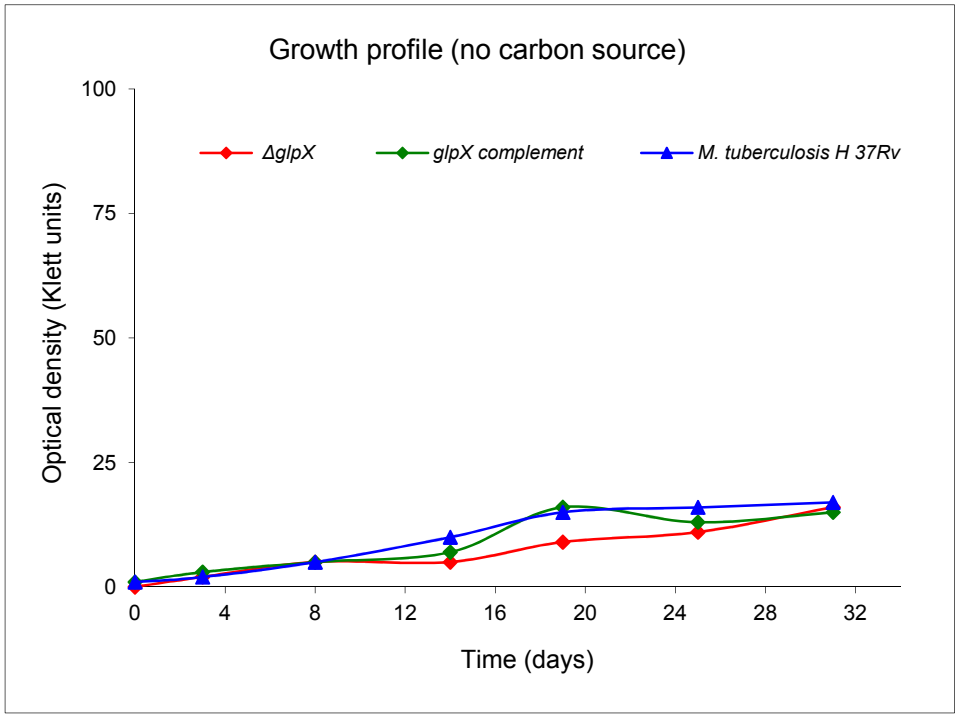


Figure 10. Growth profile of Δ*glpX* in 7H9 medium with no additional carbon source. Medium contains albumin (0.5 g/100ml), growth profile is representative of a triplicate data set.

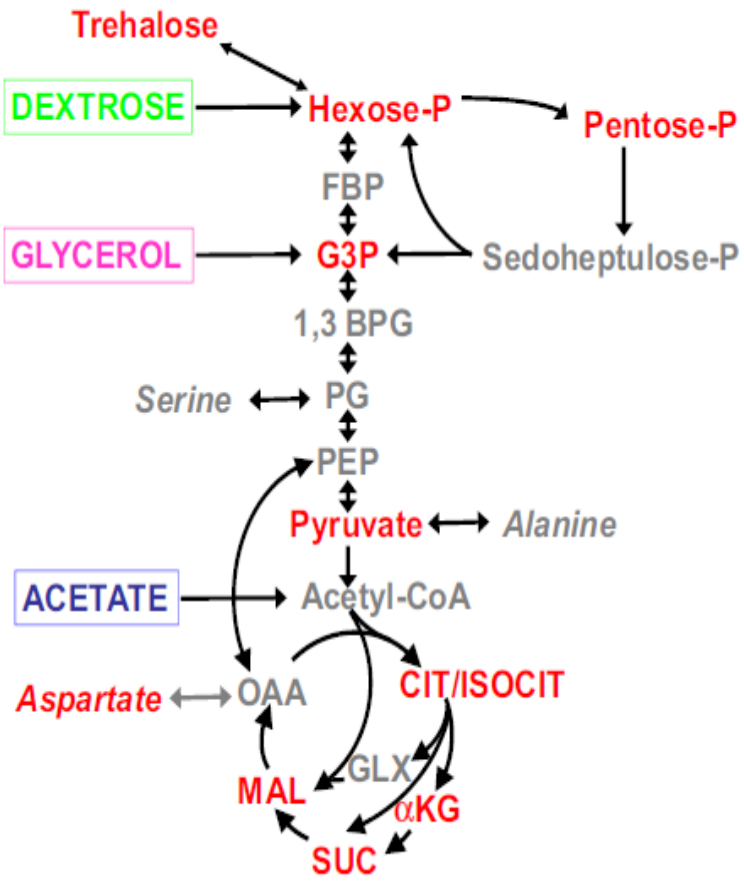


Figure 11. Metabolic pathway diagram of central carbon metabolism in *Mtb*. Exemplary metabolite reporters of glycolysis, pentose phosphate pathway and tricarboxylic acid (TCA) cycle are highlighted in red. Experimental carbon sources used in this study are boxed and their expected entry points into metabolism noted with arrows. In *Mtb*, the path from α-ketoglutarate to succinate is incompletely defined. αKG: alpha-ketoglutarate, 1,3 BPG: 1,3-bisphosphoglycerate, CIT: citrate, FBP: fructose bisphosphate, G3P: glyceraldehyde phosphate, hexose-P: hexose phosphate, ISOCIT: isocitrate, mal: malic acid, OAA: oxaloacetate, pentose-P: pentose phosphate, PEP: phosphoenolpyruvate, suc: succinic acid, PG: phosphoglycerate. (Adapted with permissions from (de Carvalho, Fischer et al. 2010).

C. *In vitro* growth profile of Δ*glpX* strain on fatty acids as a sole carbon source: As expected (similar to acetate and glycerol based growth profiles), the growth of *ΔglpX* was severely compromised on oleic acid (C18) or valeric acid (C5) as a sole carbon source (Figure 12, 13), further substantiating the observation with glycerol. Taken together the observations indicate the *ΔglpX* does not display an eugonic growth phenotype on gluconeogenic substrates (acetate, glycerol, fatty acids), rather it is dysgonic. This observation is concordant with the expected disruption of FBPase activity in *Mtb* by deletion of the *glpX* gene.

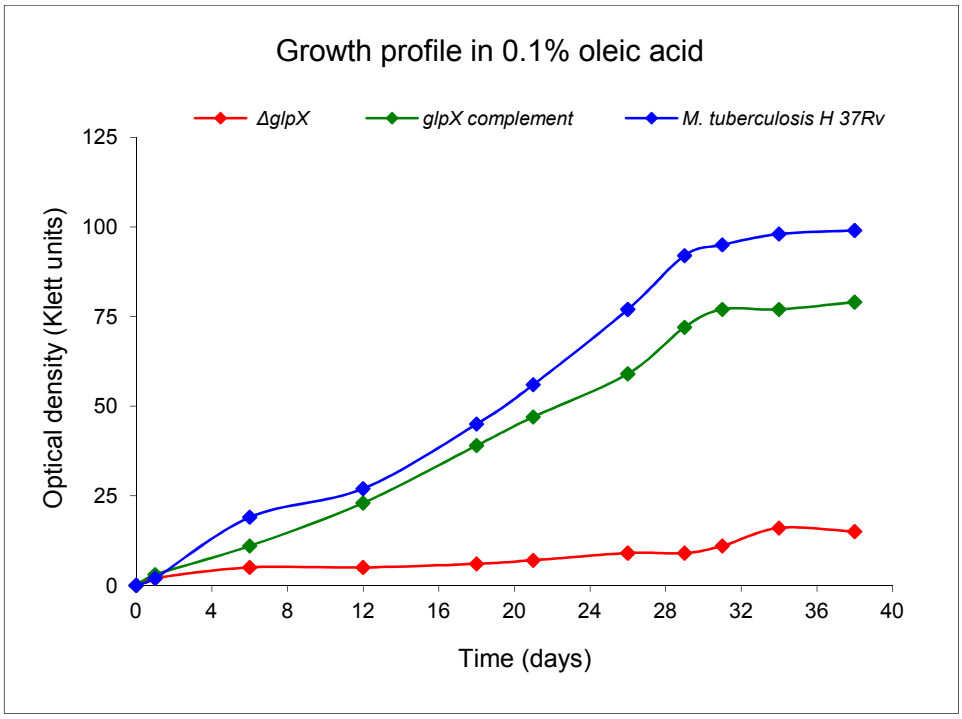


Figure 12. Growth profile of Δ*glpX* in 7H9 medium + 0.1% oleic acid. shown growth profile is representative of a triplicate data set.

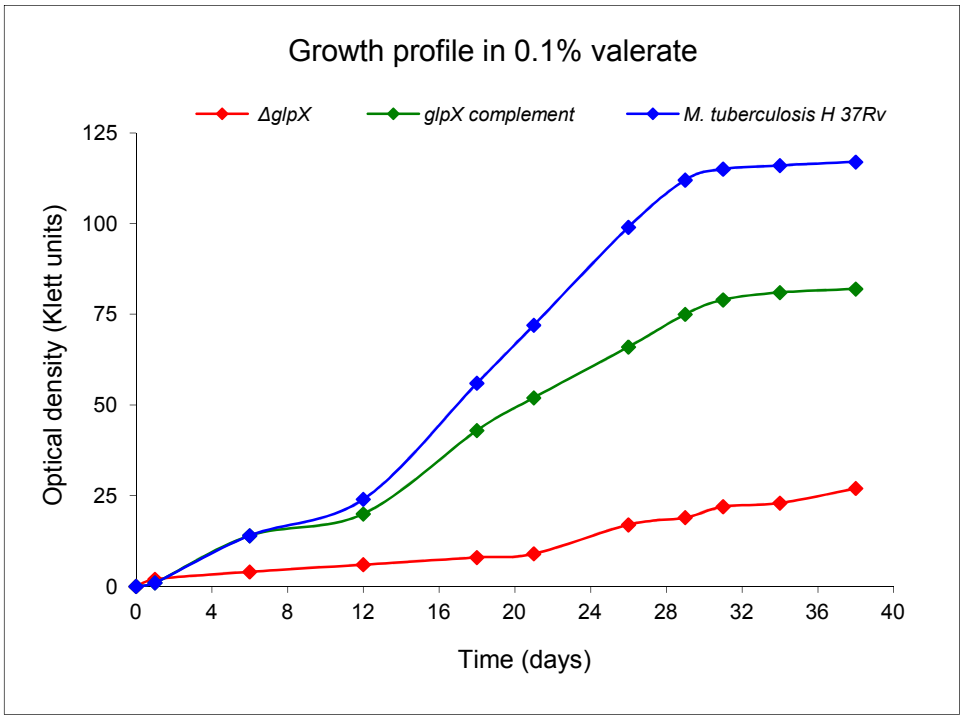


Figure 13. Growth profile of Δ*glpX* in 7H9 medium + 0.1% valeric acid. Shown growth profile is representative of a triplicate data set.

D. *In vitro* growth profile of Δ*glpX* strain on combination of carbon sources: Based on the observed growth profiles in glycerol, acetate and dextrose as individual carbon sources, growth of these strains was subsequently monitored using combinations of two such carbon sources at a time. The growth of WT *Mtb* on an equimolar mixture of glycerol and acetate exceeded that achieved with either constituent alone (Figures 14, 6, 7). Similar effects were observed using a mixture of dextrose and glycerol and a mixture of acetate and dextrose (Figure 15, 16). The compromised growth of *ΔglpX* on glycerol or acetate was not completely rescued by their combination either (Figure 16). However, with the presence of dextrose in the medium (dextrose + glycerol or dextrose + acetate), the growth profile was almost similar to WT *Mtb* (Figures 15, 16). A reason for such a phenotype is that both glycerol and acetate are gluconeogenic in nature and therefore

cannot be effectively utilized by the *ΔglpX*. However, dextrose being a preferred glycolytic nutrient for *ΔglpX* (Figures 9, 15, 16) easily rescues the *glpX* mutant strain and the combinations involving dextrose grew in a manner similar to WT *Mtb*.

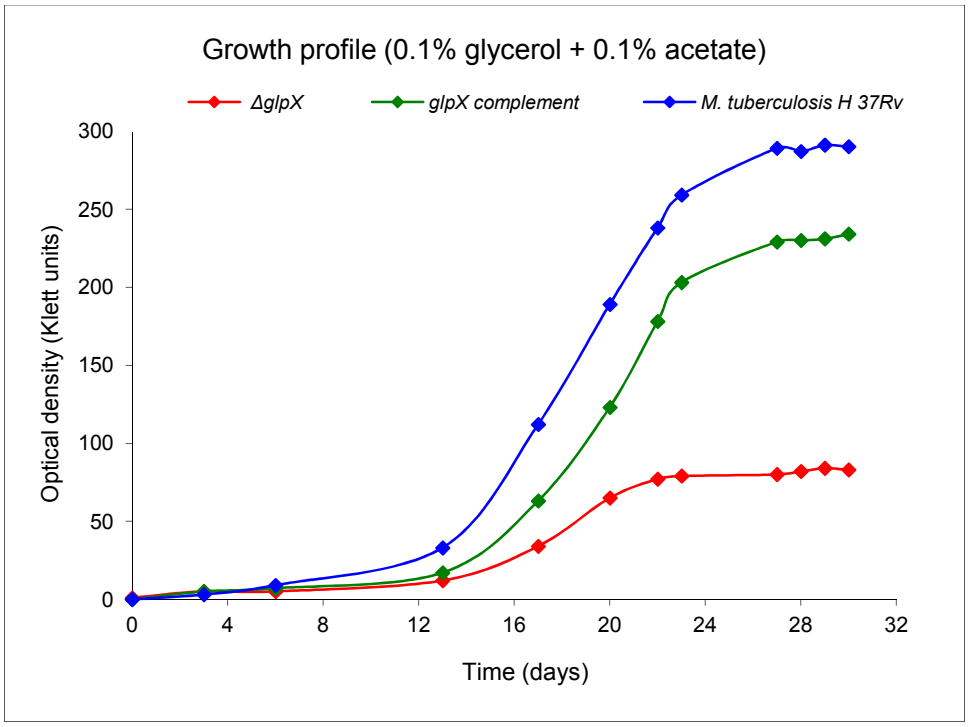


Figure 14. Growth profile of Δ*glpX* in 7H9 medium + 0.1% each of glycerol and acetate. Growth profile is representative of a triplicate data set.

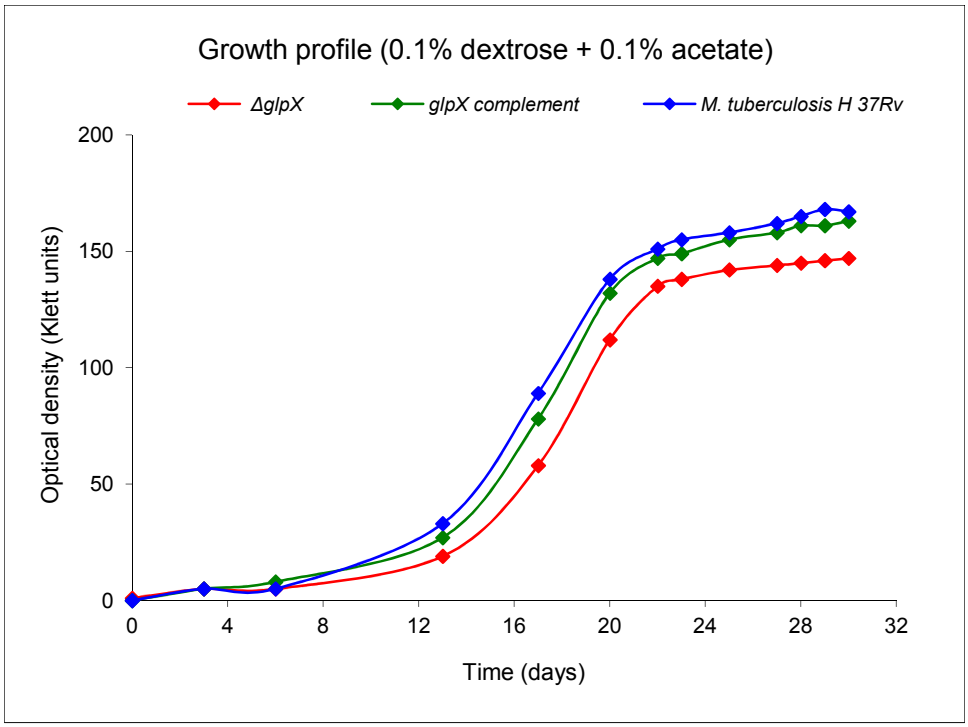


Figure 15. Growth profile of Δ*glpX* in 7H9 medium + 0.1% each of dextrose and acetate. Growth profile is representative of a triplicate data set.

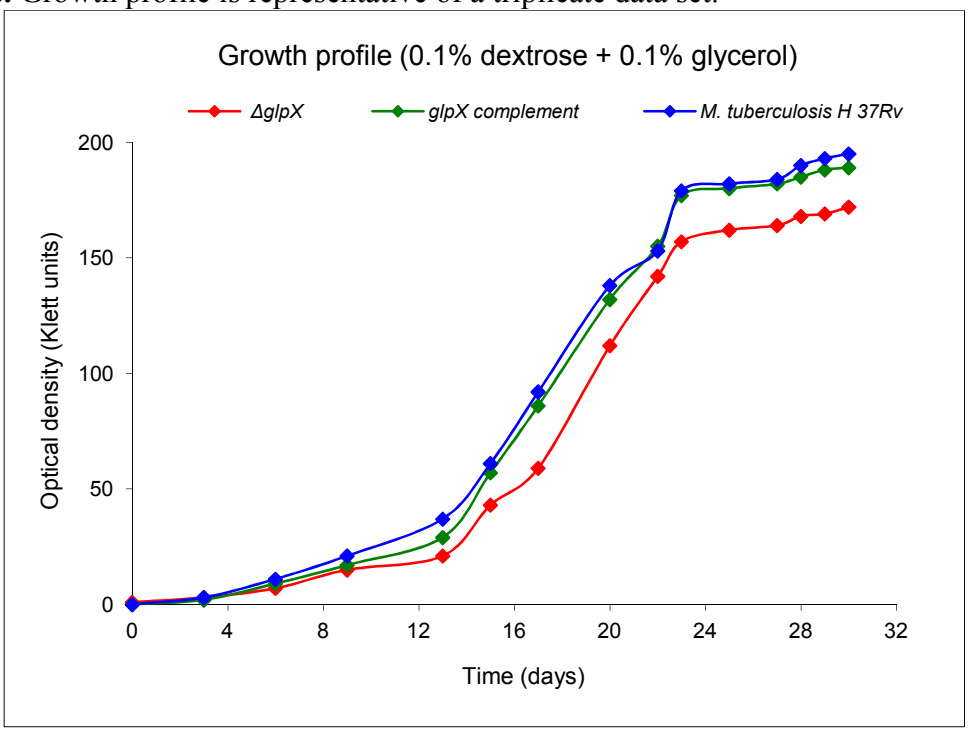


Figure 16. Growth profile of Δ*glpX* in 7H9 medium + 0.1% each of dextrose and glycerol. Growth profile is representative of a triplicate data set.

E. Compartmentalized central carbon metabolism in *Mtb*: Central carbon metabolism in *Mtb* is a peculiar case involving compartmentalization of metabolic pathways and thereby the metabolites. As understood from our results and the model of compartmentalized central carbon metabolism (de Carvalho, Fischer et al. 2010), *Mtb* preferentially utilizes glycerol over dextrose and acetate. As shown in the figure below (Figure 17C.), the predominant pattern of distribution for glycerol and acetate combination as carbon source requires a functional *glpX*/FBPase. However, since Δ*glpX* does not possess a fully functional FBPase, its growth is severely compromised on gluconeogenic substrates like glycerol and acetate.

The reduced growth/growth-compromised phenotype of Δ*glpX* is rescued to some extent by dextrose in combination growth media since with dextrose as a carbon source, FBPase activity becomes dispensable (predominant pattern of distribution is not gluconeogenic but rather glycolytic) (Figure 17A. and B.).

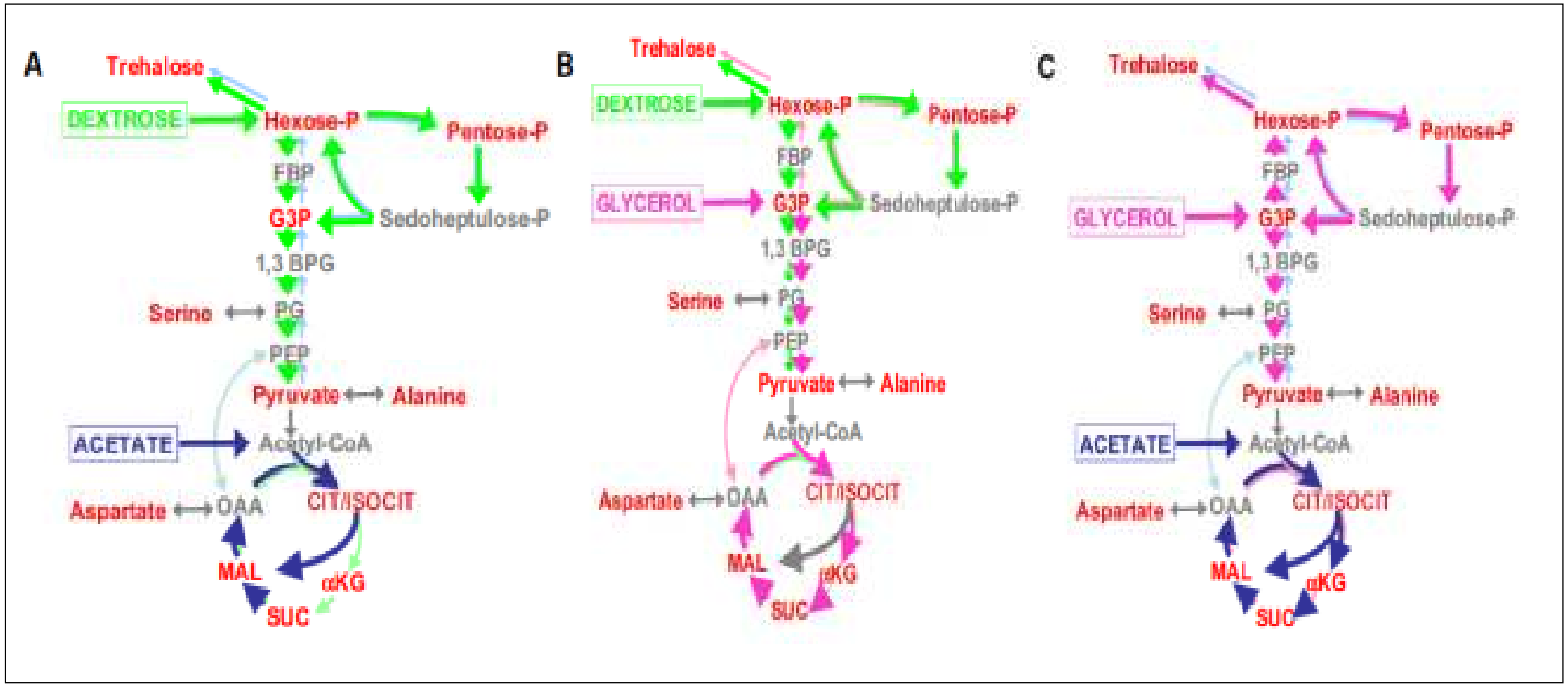


Figure 17. Metabolic pathway of central carbon metabolism in *Mtb*: Pathway schematic of glycolysis/gluconeogenesis, pentose phosphate shunt, and TCA cycle indicating distinct metabolic fates of dextrose (green), acetate (blue), and glycerol (pink) during growth on carbon substrate mixtures of: (A) dextrose and acetate, (B) dextrose and glycerol, and (C) acetate and glycerol. Exemplary metabolite reporters of glycolysis, pentose phosphate pathway, and TCA cycle are highlighted in red. The color and thickness of the arrows connecting metabolites indicates the **predominant pattern of distribution** of the corresponding carbon source. Abbreviations are as in legend to Figure 11 (Adapted with permission from (de Carvalho, Fischer et al. 2010)

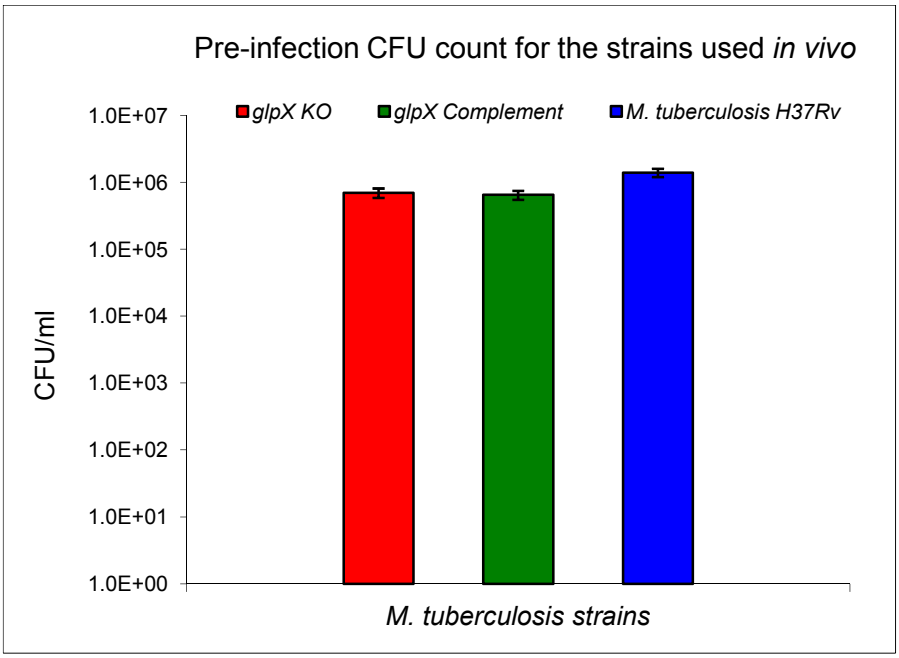
F. *glpX* is essential for growth in acute phase and survival during chronic phase of *Mtb* infection in mice (*in vivo* essentiality):

To determine the role of *glpX* encoded FBPase in a model of pulmonary tuberculosis, immune-competent BALB/c mice were infected with aerosolized WT *Mtb*, Δ*glpX*, and the complemented strains. The aerosolization parameters were set and validated such that they ensured an initial pulmonary bacterial load of about 25 – 75 CFU. Such an initial instillation of CFU was achieved for the WT *Mtb* and the complemented strain but not for the Δ*glpX* strain, indicating a possible attenuation with regard to infectivity. There are several reports wherein the initial bacterial loads of certain gene deletion strains of *Mtb* are significantly different than the control WT *Mtb* strain (Sun, Converse et al. 2004; Awasthy, Ambady et al. 2010; Venugopal, Bryk et al. 2011). In the study by Venugopal *et al*., in describing the phenotype of Δ*dlaT*Δ*pdhC,* the deletion strain was not recovered post-day one despite equal input into the aerosolizer. However, by day 7, the Δ*dlaT*Δ*pdhC* and Δ*lpdC* mutants were both recovered at 100 CFU, suggesting that the Δ*dlaT*Δ*pdhC* mutant was viable but non-culturable at day one. We expect a similar case for Δ*glpX*, where the deletion of *glpX* has limited the viability of the bacteria. Further it is clearly evident that this difference in initial instillation bacterial count was observed irrespective of the fact that the CFU of the strains used for infection were almost similar at about 1 X 10^6 CFU/ml (Figure 18a).

It is clear that the Δ*glpX* strain failed to efficiently replicate during the acute phase of infection as compared to WT *Mtb* (first 30 days of *in vivo* study). The CFU for the Δ*glpX* strain are consistently lower by about 2-3 logs than those for WT *Mtb*. The CFU for Δ*glpX* in lungs started to decline rapidly after day 57 (Figure 18b.) and continued to remain low thereafter (time points 120 and 180 days). Introduction of *glpX* expressed from its native promoter (*glpX*

complement strain) restored survival and replication through both acute and chronic phases, although the bacterial load in the lungs did not reach the WT level (Figure 18b.).

Similar to the *in vitro* growth profile of the *glpX* complemented strain on fatty acids or gluconeogenic growth substrates, growth *in vivo* was not fully restored to WT levels (Figure 18b.) Such a behavior has been observed for complemented strains of several mutants (Movahedzadeh, Smith et al. 2004; Awasthy, Ambady et al. 2010; Marrero, Rhee et al. 2010). It is noteworthy that a *glpX* gene deletion does not completely abolish FBPase activity in *Mtb* (discussed in sufficient detail in subsequent section). These experiments demonstrate that *glpX* is not only essential for *Mtb* to establish an infection and to grow during the acute phase of infection, but is equally important for survival of *Mtb* during the chronic phase of infection.



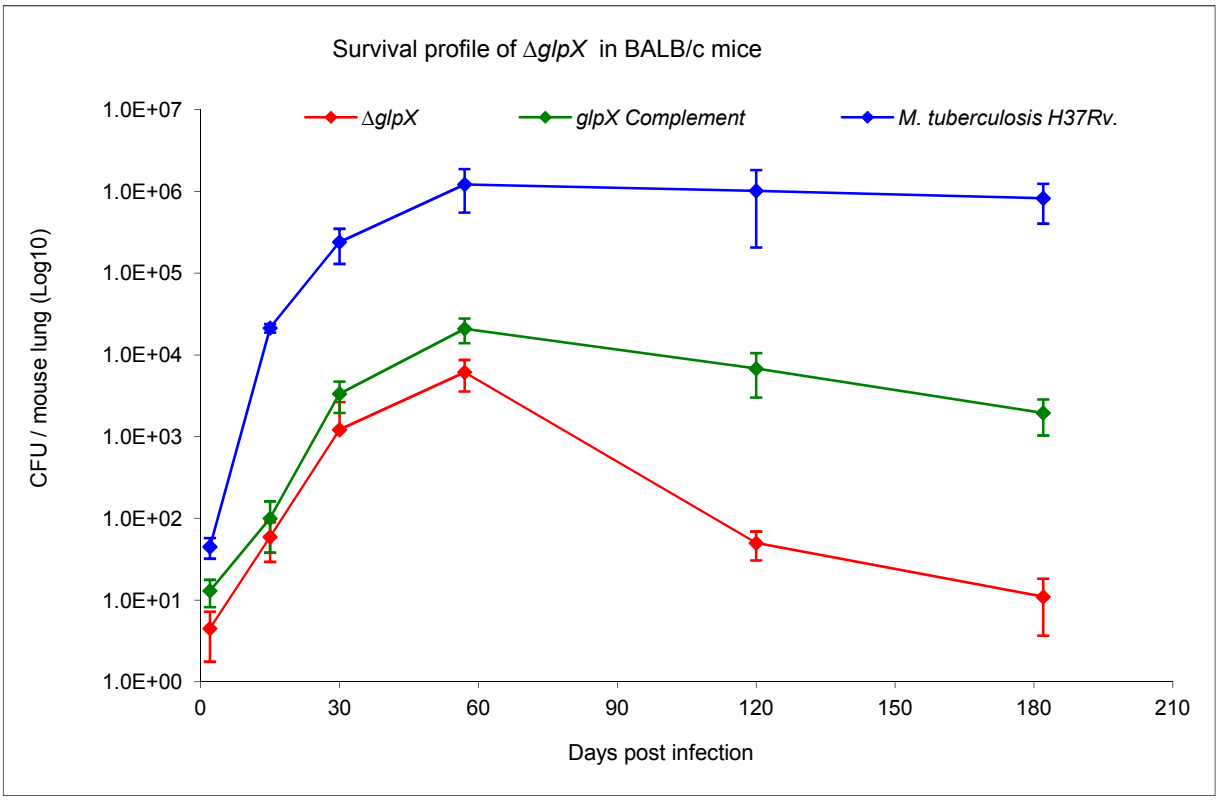


Figure 18. *glpX* is essential for growth in acute phase and survival during the chronic phase of *Mtb* infection in mice. a. Pre-infection colony forming unit count for the Δ*glpX, glpX* complement and the *Mtb* strains. b. *In vivo* Survival curve for the 3 strains in BALB/c mice infection model.

G. *glpX* gene deletion does not completely abolish the FBPase activity in *Mtb*:

As anticipated, *glpX* gene deletion adversely affected the FBPase activity levels in Δ*glpX*. However, FBPase activity was not completely abolished in the Δ*glpX* indicating a possible compensation by other genes/alternate proteins having FBPase activity (Figure19, Table III).

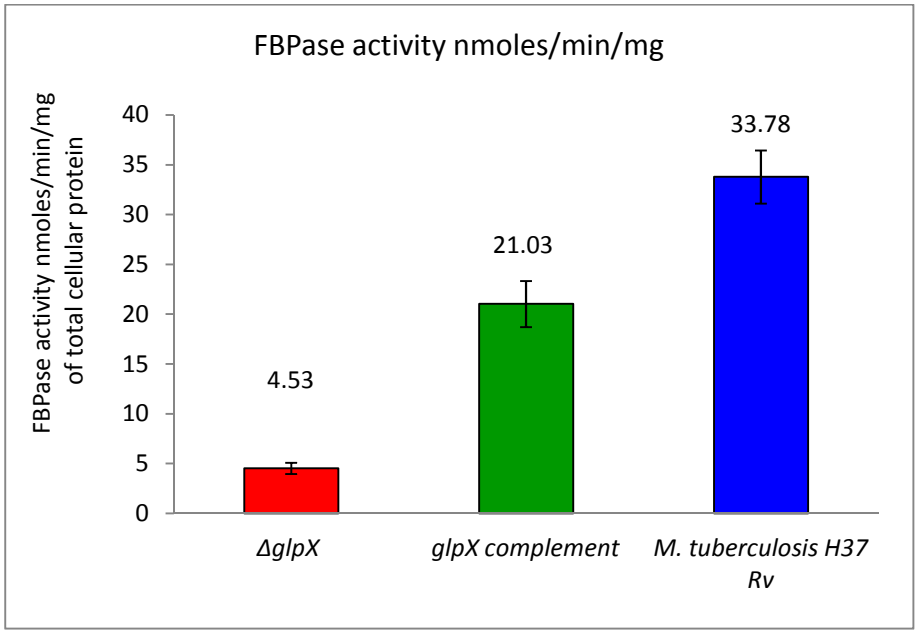


Figure 19. Fructose 1,6-bisphosphatase activity in cell extracts of different strains of *Mtb*.

Strain	FBPase activity (2 mM F1,6BP as substrate in the assay)[a]
Mtb H37Rv	33.78 ± 2.67
△glpX.	4.53 ± 0.57
glpX complement	21.03 ± 2.31

[a] **FBPase Activity** in nmol/min/mg protein in crude extracts, mean of two determinations, limit of detection = 0.4. All values are with a substrate-free control (no FBP) subtracted. The reported readings are an average of total 6 measurements coming from duplicate protein samples (n = 3X2 = 6).

TABLE III. FRUCTOSE 1,6-BISPHOSPHATASE ACTIVITY IN CELL EXTRACTS OF DIFFERENT STRAINS OF MTB.

The total cellular FBPase activity results highlight two important findings:

i. Deletion of *glpX* gene does not completely abolish FBPase activity (Δ*glpX* activity = 4.53 ± 0.57 nmol/min/mg).

ii. Although the *glpX* complement strain was constructed to complement the Δ*glpX* and restore the FBPase activity levels (FBPase activity = 21.03 ± 2.31 nmol/min/mg), it does not restore it to the WT *Mtb* levels (FBPase activity = 33.78 ± 2.67 nmol/min/mg)

The residual FBPase activity in Δ*glpX* can be attributed to several other gene products/ other classes of enzymes such as inositol monophosphatases (IMPase). The *M. tuberculosis* genome encodes four IMPase like genes, ImpA, SuhB (*Rv2701c*), ImpC (*Rv3137*) and CysQ (*Rv2131c*) (Cole, Brosch et al. 1998; Nigou, Dover et al. 2002; Movahedzadeh, Wheeler et al. 2010). The enzyme most closely related to human IMPase (25% identity), *M. tuberculosis* SuhB (*Rv2701c*), has been annotated as a 'putative extragenic suppressor protein' according to its homology with *E. coli* SuhB. The *E. coli* enzyme has been implicated in post-transcriptional control of gene expression (Chen and Roberts 2000). The SuhB of *Mtb* has been extensively characterized both structurally and biochemically (Nigou *et. al*. 2002, Brown *et al*. 2007). SuhB appears to be a bona fide IMPase with no activity towards fructose-1,6-bisphosphate. Based on the transposon mutagenesis study (Sassetti and Rubin 2003), and the gene deletion study (Movahedzadeh, Wheeler et al. 2010), impC (*Rv3137*) is identified as essential gene. SuhB and ImpA are dispensable for *in vitro* growth since their deletion knock-outs were obtainable following a standard gene deletion method (Movahedzadeh, Wheeler et al. 2010).

Deletion of CysQ was possible only when a porin-like gene of *M. smegmatis* was expressed, and also by gene switching in the merodiploid strain. *In vitro* analyses showed that *Mtb* CysQ hydrolyses inositol-1-phosphate (IMP), adenosine monophosphate (AMP) and F1,6BP

(Hatzios, Iavarone et al. 2008). Gu *et al*. identified the *Rv2131c* gene product as an unconventional enzyme that has both inositol monophosphatase (IMPase) and FBPase (Gu, Chen et al. 2006) activities. The purified *Rv2131c* gene product, in addition to having IMPase and FBPase activity, showed substrate specificity that was broader than those of several bacterial and eukaryotic IMPase. The dimeric enzyme exhibited dual activities of IMPase and FBPase, with K_m of 0.22 ± 0.03 mM for IMP and K_m of 0.45 ± 0.05 mM for FBPase. However these K_m values are at least 10 fold higher that those reported for conventional FBPases, indicating that *Rv2131* is not a major FBPase or has some other enzymatic activity. Hatzios *et al*, identified the *Rv2131*gene product as a CysQ, a 3'-phosphoadenosine-5'-phosphatase (PAP Phosphatase). It was found that the protein dephosphorylates PAP in a magnesium-dependent manner, with optimal activity observed between pH 8.5 and pH 9.5 using 0.5 mM $MgCl_2$ (Hatzios, Iavarone et al. 2008). The kinetic parameters for PAP were reported as k_m (8.1 ± 3.1 μM) and k_{cat} (5.4 ± 1.1 s^{-1}) comparable to those reported for other CysQ enzymes. The second-order rate constant for PAP was determined to be over 3 orders of magnitude greater than those determined for IMP and F1,6BP, otherwise considered to be the primary substrates of this enzyme. The ability of the *Rv2131c*-encoded enzyme to dephosphorylate PAP and PAPS *in vivo* was confirmed by functional complementation of an *E. coli* ΔCysQ mutant. In a separate study it was found that CysQ disruption attenuates *Mtb* growth *in vitro* and decreases the biosynthesis of sulfated glycolipids but not major thiols, suggesting that CysQ specifically regulates mycobacterial sulfation (Hatzios, Schelle et al. 2011).

The CysQ mutants had significantly less FBPase activity than the parent strain, ($P < 0.05$; t-test). However, the FBPase activity in the *Mtb* H37Rv control for the cysQ mutants was significantly less than in *Mtb* H37Rv control used for impA and suhB mutants ($P < 0.05$; t-test)

suggesting that the small but significant differences reported in this study may be due to batch to-batch variation rather than a result of any mutations (Movahedzadeh, Wheeler et al. 2010). Also, this reduced FBPase activity in the cysQ mutant could be attributed to the FBPase activity (although very low) of CysQ as reported by Gu *et al*. Finally, all four IMPases in *Mtb* share very limited sequence homology with each other (<30% identity with the best being SuhB which is 30% identical to ImpC (*Rv3137*)). Both ImpA and ImpC share similar sequence homology (20 – 25 % sequence identity) with TM1415 IMPase/FBPase – *Thermotoga maritime* (PDB id: 2P3N) (Chen and Roberts 1999) and MJ0109 IMPase/FBPase – *Methanococcus jannaschii* (PDB id: 1G0H) (Chen and Roberts 1998). Also it is noteworthy that both ImpA and ImpC have not been characterized structurally or biochemically yet to rule out the possibility of a dual function IMPase/FBPase. Taken together the residual FBPase activity in Δ*glpX* is attributable to CysQ or ImpA and ImpC.

H. Δ*glpX* is sensitive to stress factors (low pH, hydrogen peroxide and nitric oxide) in a similar manner to that observed for WT *Mtb*:

To identify stresses that might be responsible for killing of Δ*glpX in vivo*, the survival of the knockout under various *in vitro* stress conditions was measured. Δ*glpX* was not significantly more sensitive than *Mtb* WT to stresses likely to be encountered inside the host, such as low pH, hydrogen peroxide, nitric oxide (NO), (Figure 20, 21 and 22). These findings further indicate that the observed Δ*glpX* phenotype is primarily due to loss of fructose 1,6-bisphosphatase activity and not due to any alternate stress related effect.

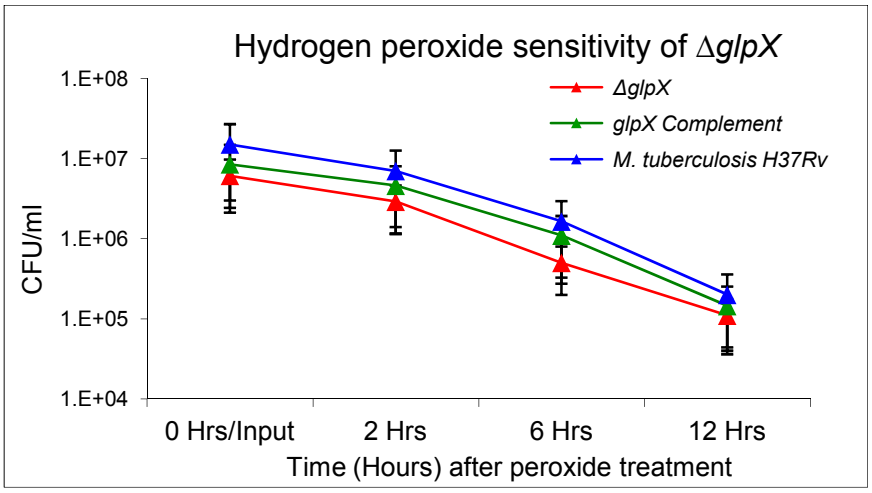


Figure 20. Hydrogen peroxide sensitivity of Δ*glpX*. The figure indicates that Δ*glpX* is affected by peroxide stress in a similar manner to *Mtb*.

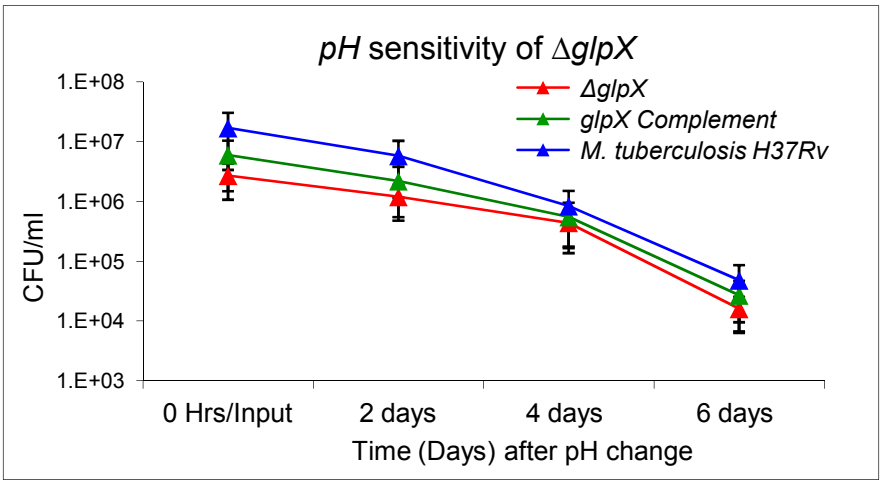


Figure 21. pH sensitivity of Δ*glpX*. The figure indicates that Δ*glpX* is affected by low pH in a similar manner to *Mtb*.

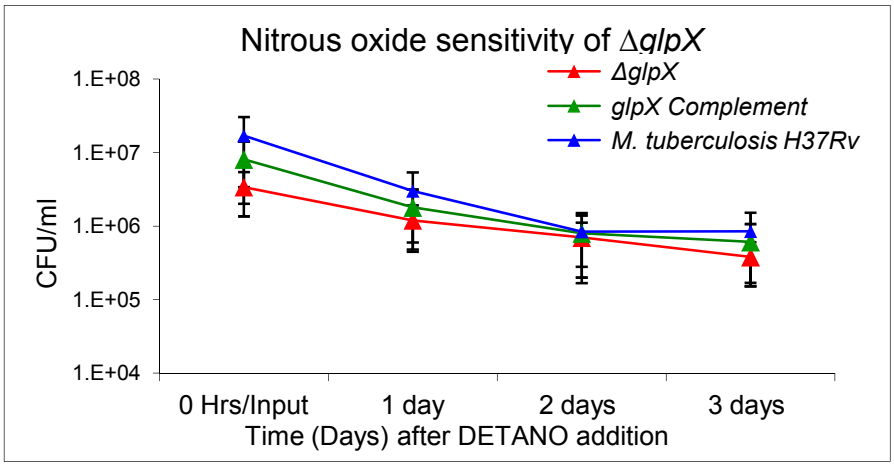


Figure 22. Nitrosative stress response of Δ*glpX*. The figure indicates that both WT *Mtb* and Δ*glpX* are almost equally affected by nitrous oxide as a stress factor.

2.4 CONCLUSION

Δ*glpX* strain was successfully generated. The *glpX* gene is required for eugonic growth on selective gluconeogenic substrates and for survival *in vivo*. The results indicate that although the growth profile of Δ*glpX* is comparable to that of WT *Mtb* when grown on the standard *Mtb* 7H9 media with OADC enrichment, its growth is dysgonic in a media containing gluconeogenic substrates such as oleic acid, glycerol and acetate. *In vivo* survival profile (in mouse model) indicates that the Δ*glpX* not only fails to achieve initial instillation dose required for infection, but also maintain such bacterial loads similar to the *Mtb* strain (about 2-3 log lower) at later time points. A 2-3 log lower bacterial count is considered significant and indicative of *glpX* gene's essentiality *in vivo*. The results indicate that *glpX* gene encodes a functional FBPase and is essential for both *in vitro* and *in vivo* growth and survival of *Mtb*. Moreover the loss of this gene results in drastic reduction of FBPase activity but not complete abolition. The Δ*glpX* strain like WT *Mtb*, is sensitive to stress factors such as low pH, oxidative and nitrosative stress, indicating that the different phenotype (growth suppressed) *in vivo* is primarily due to a drastic reduction of FBPase activity and not any other underlying effects.

CHAPTER 3

SENSITIVITY OF Δ*GLPX* TOWARDS ANTIBACTERIAL DRUGS: EXPERIMENTAL DETERMINATION OF MINIMUM INHIBITORY CONCENTRATION (MIC)

3.1 RATIONALE AND HYPOTHESIS

Understanding the differences in sensitivity to known anti-mycobacterial drugs provides valuable information for rational design of drugs inhibiting the *glpX* encoded target. As a part of the phenotypic characterization of Δ*glpX*, antibacterial activities of a battery of standard anti-mycobacterial drugs and drugs in the clinical phase of development were measured against Δ*glpX***.** MIC of a drug is referred to as the micromolar concentration at which it exhibits 90% inhibition of the mycobacterial species. The MIC obtained by Microtiter Alamar Blue Assay (MABA) was additionally verified by a cellular ATP based assay for TMC207. Further this study also helps understand plausible synergism or antagonism amongst standard anti-mycobacterial drugs and novel drug candidates.

3.2 MATERIAL AND METHODS

A. Microtiter Alamar Blue Assay:

Middlebrook 7H12 broth (preparation described below), Clear bottom 96 well plates (BD Falcon), 20% Tween 80, Alamar Blue reagent (Invitrogen).

Standard control drugs:

Standard Drugs	Molecular Weight	Maximum drug concentration in assay (µM)
Rifampin (RMP)	822.94	4
Isoniazid (INH)	137.14	8
Metronidazole (MET)	171.16	512
Caperomycin sulfate (CAP)	766.78	16
Streptomycin sulfate (SM)	1457.41	16
PA-824	359.3	8
OPC-67683	534.5	0.1
TMC207	555.51	8

TABLE IV. STANDARD ANTI-TUBERCULOSIS DRUGS IN CLINIC OR UNDER DEVELOPMENT. Indicated are their respective molecular weights and the maximum concentrations employed (in µM) in inhibitory assay.

Cell culture preparation: Cell stocks for both Δ*glpX* and WT *Mtb* strains were prepared following the protocol as described in below. The cell stocks were stored at -80°C for routine assays.

Seed stock preparation: Starter cultures for respective strains were generated by inoculating them into 10 ml of 7H9 media with OADC enrichment. After about 10 days of growth, the cultures (1:100 dilution) were transferred to 100 ml media (7H9+OADC enrichment) in a 300 ml nephelo flask (side-arm flask). A larger head space ensured sufficient aeration, and thus growth, for the *Mtb* culture. Attainment of log phase (60-90 Klett units) required 5-7 days incubation and was determined by monitoring growth using a Klett meter. The bacterial cultures were harvested by centrifugation at 3000 g for about 10 min, at 4°C. Three subsequent washes were given with 30 ml of cold sterile PBS after carefully discarding the supernatant. The cell pellet was finally re-suspended in 1 ml of PBS. Aliquots of 200 µl of suspended cells were taken into 1.5 ml screw

capped tubes. Sterility check and CFU count were performed and the cell stocks stored at -80°C. These cell stocks served as seeds for any future MABA working cell stock preparations.

MABA working cell stock preparation: MABA working cell stocks are prepared in a similar manner as the seed stocks with minor modifications. After 3 subsequent washes with 30 ml of cold, sterile PBS, and carefully discarding the supernatant, the cell pellet was re-suspended in PBS with 0.05% Tween 80 (PBST). An additional 17.5 ml of PBST was added to resuspend the cell pellet and make a homogeneous bacterial suspension. This suspension was filtered through an 8 µm pore-size-membrane filter ensuring no spillages or extraneous contamination. 1mL aliquots of the filtrate were stored in screw cap tubes at -80°C for long term use. CFU were measured per ml of the culture.

Middlebrook 7H12 broth preparation: Middlebrook 7H12 broth is essentially Middlebrook 7H9 broth containing 1 mg/ml Casitone, 5.6 µg/ml palmitic acid, 5 mg/ml bovine serum albumin, and 4 µg/ml filter-sterilized catalase (Cho, Warit et al. 2007). Each of these ingredients in their appropriate concentrations were dissolved in deionized water and filter sterilized using a 0.22 um pore-size-filter with the aid of a vacuum. Middlebrook 7H12 broth was prepared and stored under refrigeration for up to three weeks for routine use.

Assay plate set up: 100 µl of Middlebrook 7H12 broth was added to all testing wells except column 2 (column 1, 3-12; rows A-H) of a clear 96 well plate. An additional 98 µl of Middlebrook 7H12 broth was added to columns 1, 2 and 3. An additional 100 µl Middlebrook 7H12 broth was also added to rows A, H and column 12. This addition prevented the drying out/evaporation effect that happens over a 7 day incubation period. 2 µl of standard drug solution was added to columns 1, 2 and 3. After careful mixing, 100 µl of the mixture from column 3 (wells B3 – G3) was transferred to column 4 (wells B4- G4). This step was repeated for all

subsequent columns (columns 4 – 10). Finally, 100 µl of mixture from column 10 was discarded (Figure 23).

Inoculation: The MABA working cell stock is diluted to approximately 1 X 10^6 CFU/ml with Middlebrook 7H12 broth. 100 µl of bacterial cell inoculum is added to all wells apart from the all edge wells (rows B-G, columns 2-11). Final inoculum cell concentration was approximately 5 X 10^5 CFU/ml in culture media. The plates were incubated for 7 days at 37°C, in a 5% carbon dioxide incubator at 95% humidity. After 7 days of incubation, 12.5 uL 20% Tween 80 and 20 uL of Alamar Blue solution were added to all wells. Plates were incubated for an additional 24 hours at 37°C, 5% CO_2 and 95% humidity.

Media	Media	Media	Media	Media	Media	Media	Media	Media	Media	Media	Media
Drug control	Drug 1 level 1	Drug 1 level 2	Drug 1 level 3	Drug 1 level 4	Drug 1 level 5	Drug 1 level 6	Drug 1 level 7	Drug 1 level 8	Drug 1 level 9	Bacteria control	Media
Drug control	Drug 2 level 1	Drug 2 level 2	Drug 2 level 3	Drug 2 level 4	Drug 2 level 5	Drug 2 level 6	Drug 2 level 7	Drug 2 level 8	Drug 2 level 9	Bacteria control	Media
Drug control	Drug 3 level 1	Drug 3 level 2	Drug 3 level 3	Drug 3 level 4	Drug 3 level 5	Drug 3 level 6	Drug 3 level 7	Drug 3 level 8	Drug 3 level 9	Bacteria control	Media
Drug control	Drug 4 level 1	Drug 4 level 2	Drug 4 level 3	Drug 4 level 4	Drug 4 level 5	Drug 4 level 6	Drug 4 level 7	Drug 4 level 8	Drug 4 level 9	Bacteria control	Media
Drug control	Drug 5 level 1	Drug 5 level 2	Drug 5 level 3	Drug 5 level 4	Drug 5 level 5	Drug 5 level 6	Drug 5 level 7	Drug 5 level 8	Drug 5 level 9	Bacteria control	Media
Drug control	Drug 6 level 1	Drug 6 level 2	Drug 6 level 3	Drug 6 level 4	Drug 6 level 5	Drug 6 level 6	Drug 6 level 7	Drug 6 level 8	Drug 6 level 9	Bacteria control	Media
Media	Media	Media	Media	Media	Media	Media	Media	Media	Media	Media	Media

Figure 23. Typical assay plate set up for MIC determination of 6 standard drugs using the 96 well plate format.

Measuring inhibition: Fluorescence reading was performed at 530 nm excitation and 590 nm emission.

Percentage inhibition was calculated as follows:

Percentage Inhibition = (1 – (test well fluorescence units (FU)/ mean FU for bacteria controls)) X 100

Wells in Column 1 represent blank/drug control; wells in Column 12 represent bacterial control. The mean value of the media control was subtracted from all other wells. MIC was calculated by measuring the percentage inhibition over 2-fold serial dilutions (nine concentration levels). The minimum concentration that exhibited 90% inhibition was referred to as the MIC.

B. Cellular ATP assay: The MIC of anti-TB drug TMC207, as obtained with MABA assay, is verified with an additional readout (i.e. the cellular ATP assay method).

ATP Assay CellTiter-Glo® Luminescent Cell Viability Assay (Promega): The CellTiter-Glo® Luminescent Cell Viability Assay is a method to determine the number of viable cells in culture based on quantitation of the ATP present, which signals the presence of metabolically active cells. The homogeneous "add-mix-measure" assay format results in cell lysis or significant leakage and generation of a luminescent signal proportional to the amount of intracellular ATP present. The amount of ATP is directly proportional to the number of cells present in culture. The CellTiter-Glo® Assay works on the principal of a proprietary thermostable luciferase (Ultra-Glo™ Recombinant Luciferase), which generates a stable "glow-type" luminescent signal and improves performance across a wide range of assay conditions.

Materials: Middlebrook 7H12 broth (same as that used for MABA assay), CellTiter-Glo®

Luminescent Cell Viability Assay (Promega), clear bottom 96 well plates (BD, Falcon), white 96 well microplate(NUNC).

Methods:

Assay plate set up: 100 µl of Middlebrook 7H12 broth was added to all testing wells except column 2 (columns 1, 3-12; rows A-H) of a clear 96 well plate. An additional 98 µl of Middlebrook 7H12 broth was added to columns 1, 2 and 3. An additional 100 µl Middlebrook 7H12 broth was also added to rows A, H and column 12. This addition prevented the drying out/evaporation effect that happens over a 7 day incubation period. 2 µl of standard drug solution was added to columns 1, 2 and 3. After careful mixing, 100 µl of the mixture from column 3 (wells B3 – G3) was transferred to column 4 (wells B4- G4). This step was repeated for subsequent columns (columns 4 – 10). Finally, 100 µl of mixture from column 10 was discarded.

Inoculation: The MABA working cell stock is diluted to approximately 1 X 10^6 CFU/ml with Middlebrook 7H12 broth. 100 µl of bacterial cell inoculum is added to all wells apart from all edge wells (rows B-G, columns 2-11). Final inoculum cell concentration was approximately 5 X 10^5 CFU/ml in culture media. The plates were incubated for 7 days at 37°C, in a 5% carbon dioxide incubator at 95% humidity.

Measuring inhibition: CellTiter-Glo® substrate was reconstituted in CellTiter-Glo® buffer (CellTiter-Glo® reagent). 25µl volumes from each of the wells from the clear well plate are transferred to the corresponding wells in the white plate. 25µl of CellTiter-Glo® reagent was added to each of the wells and ingredients were mixed. Luminescence was read on a Victor 3

plate reader following the standard program for "Luminescence" after 2 minutes. Percentage inhibition is calculated in a similar manner as described for MABA assay.

MIC calculation: MIC is calculated by measuring the percentage inhibition over 9 concentration levels (2-fold serial dilutions). The minimum concentration that exhibits 90% inhibition is referred to as the MIC.

3.3 RESULTS AND DISCUSSION

A. MIC for Δ*glpX* strain is comparable to that for WT *Mtb* for all the standard drugs tested except for PA-824 and TMC207: MIC values for standard drugs were obtained using the standard MABA assay, the inhibitory concentrations for both strains were not strikingly different except for PA-824 and TMC-207(Table V).

Standard drugs	MIC (μM)[a]	
	ΔglpX	WT *Mtb* H37Rv
RMP	0.023 ± 0.005	0.053 ± 0.007
INH	0.479 ± 0.010	0.476 ± 0.009
MET	> 512	> 512
CAP	1.722 ± 0.061	1.895 ± 0.095
SM	0.486 ± 0.097	0.827 ± 0.083
PA-824	**0.483 ± 0.027**	**0.161 ± 0.007**
OPC-67683	0.006 ± 0.001	0.009 ± 0.001
TMC207	**< 0.031**	**0.062 ± 0.005**

[a] Reported MIC values are an average (± standard deviation) of 6 independent assays.

TABLE V. MICs FOR STANDARD ANTIBACTERIAL DRUGS AS DETERMINED BY THE MABA ASSAY.

B. *glpX* gene deletion alters the inhibitory activity of PA-824 but not OPC-67683 (a structurally similar nitroimidazole): PA-824 belongs to the 'next generation' nitroimidazole class of compounds that are currently under development for TB treatment. This nitroimidazo-oxazine has been developed by the TB Alliance for use in novel combinations, potentially for

treatment of both drug sensitive TB and MDR-TB patients. Another very similar molecule OPC-67683, a nitroimidazo-oxazole, is also being tested in MDR-TB patients as an add-on to standardized MDR-TB treatment. PA-824 is a prodrug that is converted into active form in the presence of glucose-6-phosphate dehydrogenase (FDG1) or its cofactor, coenzyme F420 from *Mtb*, (Stover, Warrener et al. 2000; Manjunatha, Boshoff et al. 2006).

Three different components have been described to be essential for the intracellular activation of PA-824 in *Mtb*. A mutation in any of these components leads to resistance to this compound: *Rv0407* encoding a nonessential F420 dependent glucose-6-phosphate dehydrogenase, genes in the F420 biosynthetic pathway, as well as *Rv3547* (Stover, Warrener et al. 2000; Manjunatha, Boshoff et al. 2006). Mutations in the mycobacterial genes *fbiA*, *fbiB*, and *fbiC* lead to impaired coenzyme F420 synthesis and therefore resistance to PA-824 (Choi, Bair et al. 2001; Choi, Kendrick et al. 2002). Mutations in the *Rv3547* gene have been described in PA-824-resistant strains (Manjunatha, Boshoff et al. 2006; Singh, Manjunatha et al. 2008). Complementing these mutants with WT *Rv3547* restores the ability of the mutants to metabolize PA-824. *Rv3547*, encoding a 151 amino acid protein with no similarity to any protein with a known function, was characterized as a F420-dependent nitroreductase (Manjunatha, Boshoff et al. 2006; Singh, Manjunatha et al. 2008). F420-dependent glucose-6-phosphate dehydrogenase, which catalyzes the oxidation of G6P to 6-phosphogluconolactone, is required for the intracellular reduction of the deazaflavin cofactor F420, which serves as the hydride donor to PA-824 in the *Rv3547* catalyzed reduction of this compound.

PA-824 shows potent bactericidal activity against both susceptible and multi drug resistant *Mtb* (Stover, Warrener et al. 2000), with aerobic MICs ranging from 0.015 to 0.5 μM, but is primarily distinguished from other anti-tubercular drugs by its activity against non-

replicating bacteria. Activated PA-824 inhibits the synthesis of proteins and cell wall lipids. PA-824 activity is limited to *Mtb* complex (Stover, Warrener et al. 2000; Manjunatha, Boshoff et al. 2006; Manjunatha, Lahiri et al. 2006).

Although the MIC of PA-824 for the Δ*glpX* strain is significantly higher (about 3-4 fold) than that observed for WT *Mtb* strain, it is within the normal MIC range of 0.015 to 0.5 μM. The observed higher inhibitory concentration of PA-824 for Δ*glpX* can possibly be correlated with the reduced intracellular fructose 6-phosphate and hence glucose 6-phosphate levels. FBPase activity and hence the intracellular F6P/G6P levels are significantly lower in the Δ*glpX* strain, which account for a reduced G6P dehydrogenase activity and hence reduced hydride donor activity of F420. It can be concluded that while Δ*glpX* strain is not resistant to PA-824, it offers resistance to PA-824 presumably due to the reduced intracellular G6P/F6P levels resulting from the loss of FBPase activity.

Mutations in the mycobacterial *Rv3547* gene are found in OPC-67683 resistant *Mtb* strains as well, suggesting that the *Rv3547*-encoded enzyme is required for activation of both PA-824 and OPC-67683 (Matsumoto, Hashizume et al. 2006). However, there is no conclusive evidence whether FGD1 and coenzyme F420 are also needed for activation of OPC-67683, since there were no OPC-67683 resistant strains with variations (mutations) in FGD1 or coenzyme F420. This indicates that possibly the activation mechanism of OPC-67683 is FGD1/F420 independent. This also explains our results wherein the MIC of OPC-67683 for Δ*glpX* strain is comparable to that of WT *Mtb* strain.

C. MIC of PA-824 against *ΔglpX* is dependent on the bacterial population in the assay: The effect of inoculum density on MIC of PA-824 was evaluated using two different bacterial

inoculum densities (i.e. 5 X 10^5 CFU/ml and 5 X 10^4 CFU/ml) (Table VI). While it is known that the antibacterial effects of certain drugs depend on the bacterial population in the assay, in case of PA-824 the MIC, varied with the bacterial population. The MICs for both *ΔglpX* and WT *Mtb* were on the lower side for lower bacterial population/CFU. Moreover this observation is consistent with our hypothesis that the higher inhibitory concentration of PA-824 for *ΔglpX* is possibly due to lower intracellular levels of F6P/G6P, which is in turn directly related to the total cellular G6P dehydrogenase activity and hence the bacterial population/CFU present in the assay.

Bacterial population (CFU/ml)	MIC (μM)*	
	ΔglpX	WT *Mtb* H37Rv
≈ 5 X 10^5 CFU/ml	0.417 ± 0.020	0.132 ± 0.013
≈ 5 X 10^4 CFU/ml	0.240 ± 0.009	0.090 ± 0.010

TABLE VI. EFFECT OF INOCULUM DENSITY ON PA-824 MIC AS DETERMINED BY MABA.

D. TMC207, an ATP synthase inhibitor: TMC207 (formerly known as R207910) is a diarylquinoline (Guillemont, Meyer et al. 2011) being developed in parallel for both MDR-TB (by Tibotec) and drug sensitive TB (by the TB Alliance). TMC207 is an inhibitor of the F0 subunit of the mycobacterial adenosine triphosphate (ATP) synthase proton pump (Matteelli, Carvalho et al. 2010; Guillemont, Meyer et al. 2011), a novel mechanism of action against *Mtb*, and demonstrates no cross-resistance against other anti-TB drugs (Koul, Dendouga et al. 2007). Importantly, it has also displayed more than 20,000-fold selectivity against the *Mtb* ATP synthase compared with a eukaryotic homologue (Haagsma, Abdillahi-Ibrahim et al. 2009). It is

highly potent *in vitro* against *Mtb*, having a minimum inhibitory concentration (MIC) of 0.03–0.12 μg/ml (0.05 – 0.21 μM) , and demonstrates marked bactericidal activity in both the acute and chronic mouse model of TB (Ibrahim, Truffot-Pernot et al. 2009).

E. The MIC of TMC207 for *ΔglpX* is about 3 fold lower compared to that for WT *Mtb* H37Rv: Our initial MABA results showed that the MIC of TMC207 is about 2-fold lower for the Δ*glpX* (< 0.031 μM) strain than that observed for corresponding WT *Mtb* (0.062 ± 0.005) control. However since our results did not give a definitive MIC endpoint (lowest concentrations used in assay demonstrate 95% inhibition), we used further lower concentrations to test the MIC. Repeated assay results are as shown in the table(s) below.

Bacterial population (CFU/ml)	MIC (μM)*	
	ΔglpX	WT *Mtb* H37Rv
≈ 5 X 10^5 CFU/ml	0.018 ± 0.001	0.073 ± 0.002
≈ 5 X 10^4 CFU/ml	0.018 ± 0.001	0.067 ± 0.002

TABLE VII. EFFECT OF INOCULUM DENSITY ON TMC207 MIC AS DETERMINED BY MABA.

Unlike PA-824, the MICs for TMC207 did not vary with varying bacterial population (Table VII). To cross verify the MABA assay results for TMC207, we measured the MIC with an additional, more specific method utilizing the ATP readout. Measuring the cellular ATP content, utilizing luminescence based detection for cell viability is highly specific and more suitable since TMC207 is an ATP synthase inhibitor (Haagsma, Abdillahi-Ibrahim et al. 2009). The cellular ATP assay MICs in Table VIII are comparable to those obtained with MABA. The MIC for TMC207 is consistently about 3-4 fold lower than that for WT *Mtb*. Interestingly, it's

not within the reported MIC range of 0.05–0.21 μM (rather lower than 0.05 μM), indicative of a possible hypersensitivity.

Bacterial population (CFU/ml)	MIC (μM)* by cellular ATP assay	
	ΔglpX	WT *Mtb* H37Rv
≈ 5 X 10^5 CFU/ml	0.019 ± 0.001	0.070 ± 0.002
≈ 5 X 10^4 CFU/ml	0.018 ± 0.001	0.066 ± 0.002

TABLE VIII. MICs FOR TMC207 AS DETERMINED BY THE CELLULAR ATP ASSAY, AT VARYING INOCULUM DENSITIES.

In summary, the *glpX* gene deletion renders *Mtb* more susceptible to TMC207, indicative of a possible synergistic effect (Δ*glpX* and TMC207 both together) on bacterial growth inhibition. TMC207 (earlier referred to as R207910) has also shown synergy with another known anti-TB drug pyrazinamide (Ibrahim, Andries et al. 2007). This finding offers an exciting possibility of a novel inhibition method (also synergistic with TMC207) provided the *glpX* gene product is a druggable target (i.e. the *glpX* gene product's FBPase activity can be inhibited by chemical means).

3.4 CONCLUSION

The Δ*glpX* strain showed similar MICs for all the drugs tested except for PA-824 and TMC207. While Δ*glpX* strain is not resistant to PA-824, the higher MIC for PA-824 can presumably be correlated with the reduced intracellular F6P and hence G6P levels. Δ*glpX* strains show significantly reduced FBPase activity, which could be correlated with lower intracellular F6P/G6P levels and hence reduced G6P dehydrogenase activity and the corresponding reduced hydride donor activity of F420. The MIC of TMC207 for *ΔglpX* is about 3 fold lower compared to that for WT *Mtb*, as verified by both the MABA and cellular ATP assays. The Δ*glpX* strain is

more susceptible to TMC207 compared to WT *Mtb* strain, indicative of a possible synergistic effect.

CHAPTER 4

BIOCHEMICAL CHARACTERIZATION OF *GLPX* ENCODED FRUCTOSE 1,6 BISPHOSPHATASE FROM *MYCOBACTERIUM TUBERCULOSIS*

(Manuscript 1: *glpX* gene of *Mycobacterium tuberculosis*: Heterologous Expression, Purification and Enzymatic Characterization of the encoded Fructose 1,6-bisphosphatase II)

4.1 RATIONALE AND HYPOTHESIS: Results from genetic and biochemical analyses revealed that the *Rv1099c* gene of *Mtb* encodes the missing mycobacterial FBPase (II) (Movahedzadeh, Rison et al. 2004). The protein encoded by the *Mtb glpX* (*Rv1099c*) gene is identical to other class II FBPase proteins from *Escherichia coli* (GlpX) (42% identity) (Brown, Singer et al. 2009) and *Corynebacterium glutamicum* FBPase II (65% identity) (Rittmann, Schaffer et al. 2003). Moreover, the *glpX* (*Rv1099c*) transposon mutant was predicted to be attenuated in Transposon Site Hybridization (TraSH) experiments (Sassetti and Rubin 2003), indicating a probable role of this enzyme in mycobacterial pathogenesis. Since *Mt*FBPase II constitutes the only known FBPase in *Mtb* and has no human homologue, biochemical and structural studies could reveal certain unique characteristics of this enzyme that can be exploited for species-specific drug design. An attempt to purify this enzyme was made in the past (Movahedzadeh, Rison et al. 2004), however the semi-purified protein was found to be enzymatically inactive. In this chapter, the successful expression, purification and enzymatic characterization of the functionally active recombinant *Mt*FBPase II is described. Further experiments were performed to define substrate specificity and to identify effectors (inhibitors/activators) of FBPase activity.

4.2 MATERIAL AND METHODS

A. Materials: Unless otherwise stated, all chemicals were of analytical grade and purchased from Fischer Scientific Company, USA. Restriction endonuclease and T4 DNA ligase were

obtained from New England Biolabs, USA. Coupling enzymes for activity assay (yeast glucose-6-phosphate dehydrogenase and yeast phosphoglucoisomerase), lysozyme and DNAse I were purchased form Sigma Aldrich, USA. The plasmid, pET15b, expression vector and Ni–NTA agarose beads were from Novagen, USA. QIA quick spin columns and plasmid purification kit were from Qiagen, Germany. Complete EDTA-free protease-inhibitor was obtained from Roche Molecular Biochemicals, USA. Hiload 16/60 and Hiload 26/60 FPLC columns packed with Superdex-200 were from GE Life Sciences, USA. Gel filtration standard used for molecular weight estimation was purchased from Bio-Rad, USA. SDS-PAGE profiles were performed using 4–12% NuPAGE Bis-Tris gradient gels from Invitrogen, USA. SeeBlue Plus2 and SimplyBlue SafeStain were also purchased from Invitrogen, USA. Pierce 660 nm Protein Assay kit was procured from Pierce Thermo Scientific, USA. *E. coli* DH5α and BL21 (DE3) strains were obtained from Novagen, USA.

B. Cloning of *Mt*FBPase: Cloning of the *glpX* gene of *M. tuberculosis* for overexpression of FBPase in *E.coli* was essentially the same as previously described (Movahedzadeh, Rison et al. 2004). PCR amplification of the *glpX* gene was carried out using specific primers that were designed based on the genome sequence of *Mtb* H37Rv from genolist.pasteur.fr/TubercuList (forward *Rv1099c*-NdeI 5-GGAATTCCATATGGAGCTGGTCCGGGT-3 and reverse *Rv1099c*-XhoI 3-TGACTCGAGGGCAATGGGTACACG-5). These primers introduced an NdeI site at the 5′ end and an XhoI at the 3′ end (NdeI and XhoI sites are underlined in the depicted sequence) to allow in-frame cloning of the gene into the expression vector, pET15b. Genomic DNA isolated from *Mtb* H37Rv was used as the template. The primers were each used at 300 nM final concentration. PCR was carried out using the Mastercycler pro PCR machine (Eppendorf) with *Mtb* DNA as the template and 2% DMSO. The temperature cycle used was: an initial 3 min

at 94 °C to denature high-GC DNA; 10 cycles of 45 s each at 94 °C, 1 min at 63 °C and 1 min at 72 °C; 25 cycles of 45 s at 94 °C, 1 min at 63 °C and 1 min at 72 °C (this last 1 min step increased by 20 s per cycle); and finally an extension step of 7 min at 72 °C to complete primer extension. The PCR product, purified using Qiagen mini columns (Qiagen, Germany) as per manufacturer's instructions, was digested with *Nde*I and *Xho*I and ligated using T4 DNA ligase into pET15b that was previously digested with the same restriction enzymes. This was electroporated into *E. coli* DH5α cells (Novagen, USA) and transformants selected on LB-Agar plates containing ampicillin (100 µg/ml). Recombinant colonies were analyzed by restriction digestion with *Xho*I and *Nde*I for the release of the insert. For recombinant expression of *Mt*FBPase, *E. coli* BL21 (DE3) (Novagen, USA) was transformed with pET15b-*glpX* and transformants selected on LB-Agar plates containing ampicillin (100 µg/ml).

C. Expression and purification of *Mt*FBPase: *E. coli* BL21 (DE3) harboring pET15b-*glpX* was grown in 100 ml of LB broth containing 100 µg/mL ampicillin with shaking (180 rpm) overnight at 37°C. The following day, 10 ml of the broth was transferred to 1 liter of LB broth containing 100 µg/mL ampicillin and shaking at 37°C was continued until the OD600 reached about 0.5, at that point isopropyl-D-1-thiogalactopyranoside (IPTG), at a final concentration of 0.5 mM, was added to induce the culture. The culture was further grown for 8 h at 25°C and then centrifuged at 4000g for 20 min at 4°C. The induction and growth conditions (IPTG concentration, induction temperature and time) were optimized in order to obtain high cell density and high protein expression. Cells were washed with phosphate-buffered saline and stored at −20°C until further use.

The cell pellet was resuspended in of buffer A (50 mM sodium phosphate buffer, pH 8.0, and 300 mM NaCl) with 0.50 mg/ml of lysozyme, 2 U/ml of DNAse I (Sigma-Aldrich, USA) and

complete EDTA-free protease-inhibitor (Roche Molecular Biochemicals, USA) followed by gentle stirring at 4°C for 45 min to obtain a homogeneous suspension. The lysed cell pellet was subjected to disruption by sonication on ice using a Branson 450 Sonifier for a total time of 20 min using a fixed protocol (1 pulse/s with the duty cycle dial at 50%). Insoluble cell debris was removed by centrifugation at 16,000 g for 30 min at 4 °C. The supernatant was clarified by passage through a 0.45 µm PVDF syringe filter (Millipore, USA) and loaded onto Ni-NTA bind resin (Novagen, USA) pre-equilibrated in buffer A (50 mM sodium phosphate buffer, pH 8.0, and 300 mM NaCl) with 10 mM imidazole. *Mt*FBPase bound to the Ni-NTA column, was then washed with 32 column volumes of buffer A with 20 mM imidazole followed by 40 column volumes of buffer A with 50 mM imidazole. *Mt*FBPase was eluted from the column with 6 column volumes of buffer A with 250 mM imidazole. The eluate was immediately subjected to buffer exchange and concentration in an Amicon-15 Ultracel 100 K (Millipore, USA) centrifuge concentrator against an exchange buffer (20 mM tricine (pH 7.7), 50 mM KCl, 1 mM $MgCl_2$, 0.1 mM DTT, 15% (v/v) glycerol) prior to running size exclusion chromatography (SEC) as described for FBPase II from *Corynebacterium glutamicum* (Rittmann, Schaffer et al. 2003). *Mt*FBPase was applied to a Superdex-200 Hiload 26/60 column (GE Healthcare Biosciences, USA) on an ÄKTA purifier FPLC system (GE Healthcare Biosciences, USA) pre-equilibrated with the same exchange buffer and eluted at a low flow rate of 1 ml/min. Fractions of 3 ml were collected and pooled according to the chromatogram and the observed purity profile on SDS-PAGE. The pooled fractions were concentrated and filtered through a 0.22-µm membrane filter. Equal volumes of different purification stage samples were analyzed by SDS–PAGE using 4–12% NuPAGE Bis-Tris gradient gels (Invitrogen, USA) and staining was performed using SimplyBlue SafeStain (Invitrogen, USA). All protein purification steps were performed at 4 °C.

The protein concentration was estimated by the Thermo Scientific's Pierce 660 nm protein assay kit using bovine serum albumin (BSA) as the standard following the manufacturer's instructions

D. Native Molecular Weight Estimation by Size Exclusion Chromatography: SEC was performed on an ÄKTA purifier FPLC system using a Hiload 16/60 packed with Superdex 200 (GE Life Sciences). The mobile phase was 20 mM tricine (pH 7.7), 50 mM KCl, 1 mM $MgCl_2$ and 0.1 mM DTT. Gel filtration standards from Bio-Rad was used for calibration of the column. Gel filtration standards included thyroglobulin (bovine) 670 kDa, gamma globulin (bovine) 158 kDa, ovalbumin (chicken) 44 kDa, myoglobulin (horse) 17 kDa and vitamin B12 1350 Da. Times of protein elution/retention were determined by monitoring the absorbance at 280 nm. *Mt*FBPase was injected at a concentration of 1.0 mg/ml. Flow rate was set at 1 ml/min. The relative elution was calculated as: $K_{av} = (V_e - V_o) / (V_t - V_o)$, where V_e is the elution volume, V_o is the void volume, and V_t is the total column volume. A standard curve was plotted as K_{av} vs. log molecular weight. The molecular weight for *Mt*FBPase was estimated from this standard curve obtained using the gel filtration standards.

E. Determination of Enzyme Activity: Fructose 1,6–bisphosphatase activity was determined using a coupled spectrophotometric assay as described previously (Rittmann, Schaffer et al. 2003) by monitoring the increase in absorbance due to the reduction of $NADP^+$ to NADPH at 340 nm. Standard assays were performed in a 96-well plate format in 100 µl reaction mixture aliquots containing 50 mM KCl, 20 mM tricine (pH 7.7), 8 mM $MgCl_2$, 1.0 mM $NADP^+$, yeast glucose-6-phosphate dehydrogenase (G6PDH)(1.0 U/ml), yeast phosphoglucoisomerase (PGI)(2.5 U/ml) and purified *Mt*FBPase at a concentration of 1 µg/ml (10 µl of 1:100 dilution of 1 mg/ml concentration of the protein added to a final volume of 100 µl reaction mixture). The reaction mixture was incubated at 30 °C for 10 min and the reaction was started by adding 250

μM of the substrate F1,6BP, followed by monitoring the increase in absorbance at 340 nm at a constant temperature of 30 °C. F6P formed by the reaction of FBPase is converted to G6P and subsequently to 6-phosphogluconate by the coupling enzymes phosphoglucoisomerase and G6P dehydrogenase and the concomitant formation of NADPH (ε340 nm=6.22 $mM^{-1}cm^{-1}$) from $NADP^+$ is followed at 340 nm. The activity was measured by monitoring the increase in the absorbance at 340 nm using a Spectramax384 high throughput microplate spectrophotometer (Molecular Devices, USA). The reaction was followed for 10 min at a constant temperature of 30 °C.

One unit of FBPase activity is defined as the amount of enzyme that catalyzes the conversion of 1 μM of fructose 1,6-bisphosphate to fructose 6-phosphate per minute under the above assay conditions.

F. Effect of metal ions on enzymatic activity: The effect of various metal ions on *Mt*FBPase activity was determined by addition of varying amounts of metal ions as their chloride or sulfate salts to the assay mixture followed by measuring the enzymatic activity. Most of the Li^+ sensitive phosphatases (FBPase and IMPase) are strongly inhibited by submillimolar concentrations of Li^+ (McAllister, Whiting et al. 1992; York, Ponder et al. 1995; Albert, Yenush et al. 2000; Patel, Martinez-Ripoll et al. 2002; Rittmann, Schaffer et al. 2003). We tested the effect of lithium chloride salt by incorporation into the assay mixture in a concentration range of 0–40 mM. The effect of bivalent metal ions Mg^{2+} and Mn^{2+} was determined by incorporation of various amounts (concentration range of 0.01 mM–40 mM) of these ions (as chloride salts) in the activity assay. Other bivalent metal ions such as Ca^{2+}, Zn^{2+}, Fe^{2+}, Cu^{2+}, Co^{2+} and Ni^{2+} as their chloride or sulfate salts were also tested in a concentration range of 0–40 mM in order to measure their ability to effectively substitute for magnesium in the assay mixture. While these bivalent metal ions were

tested, the existing Mg^{2+} in the protein solution was buffer exchanged (using an Amicon-15 Ultracel 100 K (Millipore, USA)) with each of the bivalent metal ions respectively. Performing this buffer exchange ensured that there was no background enzymatic activity arising from the presence of Mg^{2+} and the observed activity, if any, was due solely to the presence of the respective bivalent metal ion. The effect of monovalent metal cations, K^+ and Na^+, on enzymatic activity was measured by incubating these metals as their chloride salts in the assay mix. The effects of NH_4Cl and KH_2PO_4 (phosphate is also the catalytic product of FBPase reaction) were also determined by the addition of different amounts of these salts in the enzyme assay.

G. Determination of pH optima for *Mt*FBPase: For measurement of pH optima, enzyme activity was assayed in 100 mM buffers of various pH values in the range of 5–11 (acetate buffer, pH5.0–6.0; tris/tricine buffer, pH 7.0–9.0; bicarbonate buffer, pH 10.0–11.0).

H. Determination of thermal stability: Thermal stability of *Mt*FBPase was studied by incubating the purified protein at different temperatures for 30 min. Upon completion of incubation time, the samples were transferred immediately to ice until determination of residual enzymatic activity and were compared to the activity of unheated (stored at 4°C throughout) controls.

I. Malachite Green Assay for Free Phosphate Determination: The method described by Baykov (Baykov, Evtushenko et al. 1988) was suitably adapted with appropriate modifications.

Color reagent. Concentrated sulfuric acid (60 ml, d = 1.84 g/liter) was slowly added to 300 ml of water. The solution was then cooled to room temperature and supplemented with 0.44 g of malachite green (Sigma Aldrich). The resulting orange solution was stable for at least 1 year at room temperature. On the day of use, 2.5 ml of 7.5% ammonium molybdate (Sigma Aldrich)

was added to 10 ml of the dye solution followed by 0.2 ml of 11% Tween 20 (Fisher Scientific). This mixture is referred to as "color reagent".

Note: It is important to check that all enzyme preparations and assay buffers do not contain free phosphate. This was conveniently verified by adding 25 μL of the color reagent to 100 μL of sample solution. The blank OD values at 630 nm were found to be consistently less than 0.1.

Standard plot for phosphate estimation

Preparation of phosphate standards: A premix solution containing 40 μM phosphate was prepared by pipetting 40 μL of 1 mM phosphate standard to 960 μL distilled water or enzyme reaction buffer. Standard solutions (7 concentrations in the range of 40 - 4 μM) were subsequently prepared by serial dilutions.

Color formation and its measurement: 100 μL of standard solutions were transferred (in triplicate) into a clear-bottom 96-well plate. Appropriate blank controls containing water or reaction buffer only were also incorporated in the assay. 100 μL of test samples (appropriately diluted) were transferred into separate wells of the plate. 25 μL of color reagent was added to each well followed by gently mixing by tapping the plate. Plates were then incubated for 30 min at room temperature for color development. Absorbance was measured at (630 nm) on a plate reader. The green color complex has a λ_{max} in the range of 600 nm - 660nm.

Phosphate determination in samples (determination of substrate specificity): The enzyme reaction was set up in a similar manner as described under "Determination of Enzyme Activity", except for the addition of coupling enzymes (G6PDH and PGI) and $NADP^+$, which were replaced by equivalent amounts of deionized water. The enzyme reaction was initiated by the addition of the appropriate substrate, followed by incubation at 30 °C for 10 minutes. After 10 minutes of incubation, 10 μl of enzyme reaction mix was mixed with 25μl of the color reagent

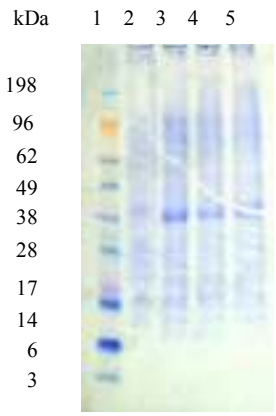
(thus terminating the enzyme reaction) followed by addition of 90µl deionized water. The mixture was left to stand for 10 min, before the absorbance at 630 nm was measured.

J. Incubation Assay (coupled spectrophotometric assay) for identification of effectors of FBPase activity: Effectors of *Mt*FBPase activity were identified by "Incubation assay" as described by Hines J.K. and co-workers (Hines, Fromm et al. 2007). The assay adopted is the same as that described under "Determination of Enzyme Activity", with an important modification that each of the effectors under consideration is incubated with *Mt*FBPase for 2 hours at 25°C. The coupling enzyme mixture, $NADP^+$, and the substrate F1,6BP were added after 2 hours incubation and the enzyme activity was assessed as described under "Determination of Enzyme Activity". Appropriate controls with F6P (catalytic product) were also run to rule out the possibility of coupling enzyme inhibition by the tested effectors. While it was expected that intermediates of TCA and glycolysis may have regulatory effects on the *Mt*FBPase enzyme, it had not been verified experimentally. Therefore, a wide array of glycolysis and TCA intermediates and adenosine phosphates (AMP, ADP, ATP) were chosen as effectors to identify all possible means of regulation of this mycobacterial enzyme. Moreover, adenosine phosphates have been known to be involved in regulating glycolysis/gluconeogenesis in the cell. In general, glycolysis is activated by ADP or AMP, and inhibited by ATP and PEP, whereas gluconeogenesis is activated by ATP and PEP, and inhibited by AMP and ADP (Hines, Kruesel et al. 2007). The *fbp*-encoded FBPase is inhibited by AMP (Hines, Kruesel et al. 2007) and activated by PEP and citrate (Hines, Fromm et al. 2007). It is very likely that *glpX* activity may also be regulated by such effectors which may, in turn, play an important role in preventing futile cycling and maintaining appropriate FBP/F6P levels in the cell.

4.3 RESULTS AND DISCUSSION:

A. Cloning, Expression and Purification of *Mt*FBPase: In order to produce recombinant His-tagged *Mt*FBPase protein in *E. coli*, *glpX* gene-encoding mycobacterial FBPase was cloned into pET15b, a T7 promoter based inducible bacterial expression vector. Thus the *Mt*FBPase was expressed as a fusion protein with an N-terminal histidine tag.

Significant expression of *Mt*FBPase was observed upon induction of *E. coli* BL21 (DE3) cell cultures harboring the *glpX* gene in pEt15b, with 0.5 mM IPTG (Figure 24a Lane 3). No prominent band was detected in the uninduced control (without IPTG) indicating that efficient expression was induced by IPTG (Figure 24a Lane 2). To verify whether the *Mt*FBPase is expressed in a soluble form or in inclusion bodies, the induced *E. coli* cells were disrupted by sonication, centrifuged and the resulting pellet and supernatant were analyzed. The mycobacterial FBPase (II) overexpressed in *E. coli* was present in the soluble fraction (Figure 24a Lane 4) as well as in the insoluble fraction (inclusion bodies). The Ni-NTA elute was relatively impure (Figure 24b Lane 8) with some high molecular weight impurities, therefore a SEC step was introduced to purify the protein to homogeneity. Fractions 18–20 from the size-exclusion column were the most pure and had high protein content (Fig. 1a); thus, they were pooled for further studies. Fractions 15–22 were pooled, concentrated, and used for setting up the crystallization of the protein.



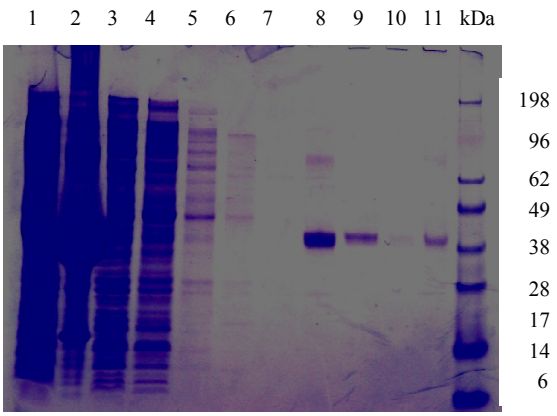


Figure 24a **Figure 24b**

Figure 24. Purification and analysis of *Mt*FBPase. 24a: SDS-PAGE analysis of induction and localization of *Mt*FBPase. an SDS-PAGE analysis of induction and localization of recombinant MtFBPase. Recombinant *Mt*FBPase was expressed in *E. coli* strain BL21 (DE3) and subjected to SDS-PAGE using 4–12% NuPAGE Bis-Tris gradient gels. The gel was stained with SimplyBlue Safe Stain. Lane 1: molecular weight standard SeeBlue Plus2; Lane 2: uninduced cell lysate; Lane 3: induced cell lysate with 0.5 mM IPTG; Lane 4: soluble fraction; Lane 5: insoluble fraction. **24b**: SDS-PAGE analysis of *Mt*FBPase. Lane 1: Lysis supernatant; Lane 2: pellet post lysis; Lane 3: Ni-NTA Flow through; Lane 4: Wash 1 concentration 25 mM imidazole; Lane 5: wash 2 concentration 25 mM imidazole; Lane 6: Wash 3 concentration 50 mM imidazole; Lane 7: Blank; Lane 8: Ni-NTA elute 250 mM imidazole (immediately dialyzed with Mg^{2+} containing buffer); Lane 9: *Mt*FBPase SEC main peak fractions pooled; Lane 10: Filtrate left after buffer exchange; Lane 11: *Mt*FBPase with other high molecular weight impurities (rejected fractions in SEC); Lane 12: molecular weight standard SeeBlue Plus2.

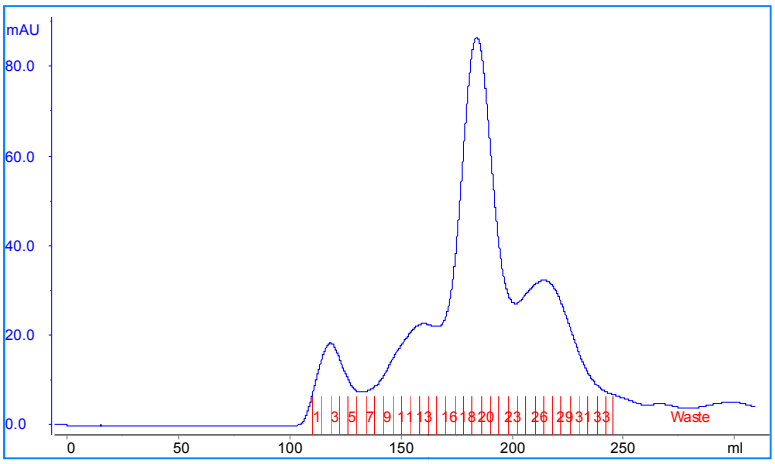


Figure 24c.

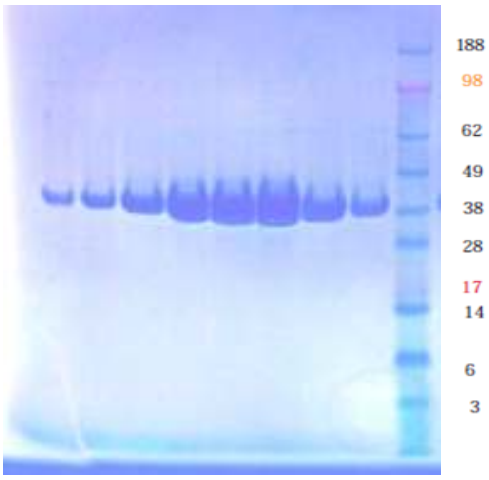


Figure 24d.

Figure 24c. Size Exclusion Profile of the affinity captured protein showing several low and high molecular weight impurities in addition to the main peak of *Mt*FBPase. The fractions (18-20) corresponding to the main peak, were used for the crystallization experiments. **Figure 24d.** Lanes 1 to 8 (from the left): SDS-PAGE profile of SEC fractions 15 to 22 (Main peak of chromatograph in Figure 24c.). Lane 9: Molecular weight standard SeeBlue Plus2.

B. Absolute requirement of Bivalent Metal Ions: The Ni-NTA elute was subjected to enzymatic assay (as described in methods) to measure the FBPase activity, however no activity was obtained initially. In fact, this loss of activity has been reported earlier (Movahedzadeh, Rison et al. 2004). In the case of *C. glutamicum* FBPase purification, only 16% of the FBPase activity present in the crude extract could be purified; that attributed to the absence of Mn^{2+} during elution (Rittmann, Schaffer et al. 2003). In the case of *Mt*FBPase we observed that the enzymatic activity was retained if the protein was immediately exchanged with an exchange buffer containing the bivalent metal ion magnesium. The fact that no FBPase activity was present in the initial Ni-NTA elute could be attributed to the absence of Mg^{2+} during elution. Thus for stabilization, the protein solution (post Ni-NTA elution) was immediately exchanged in an exchange buffer containing Mg^{2+}. In order to further purify the protein, the buffer exchanged protein solution was applied onto a Superdex-200 Hiload 26/60 FPLC column.

The recombinant *Mt*FBPase was purified to homogeneity using a two-step procedure comprising a Ni-NTA affinity capture followed by SEC. Approximately 1.90 mg of active enzyme with a specific activity of 1.308 U/mg was purified per liter of *E. coli* culture (Table IX). Introduction of the His-tag at the N-terminal proved effective for affinity capture of the recombinant protein. Moreover, the presence of the N-terminal His-tag did not interfere with catalytic activity as demonstrated for several other FBPases. GlpX protein from *E. coli* was also purified as N-terminal His-tagged protein and the crystal structure revealed that the N-terminal residues do not comprise the catalytic pocket of the enzyme (Brown, Singer et al. 2009).

Purification step	Protein (mg)	Total activity Units (U)	Sp. Act. (Units/mg) or μmol/min/mg[a]	Purification fold[b]	Yield (%)[c]
Cell lysate supernatant	347.1	6.42 U	0.019	1.0	100
Ni-NTA elute (buffer exchanged)	3.3	3.13 U	0.955	51.6	48.8
Final purified protein	1.9	2.49 U	1.300	70.7	39.0

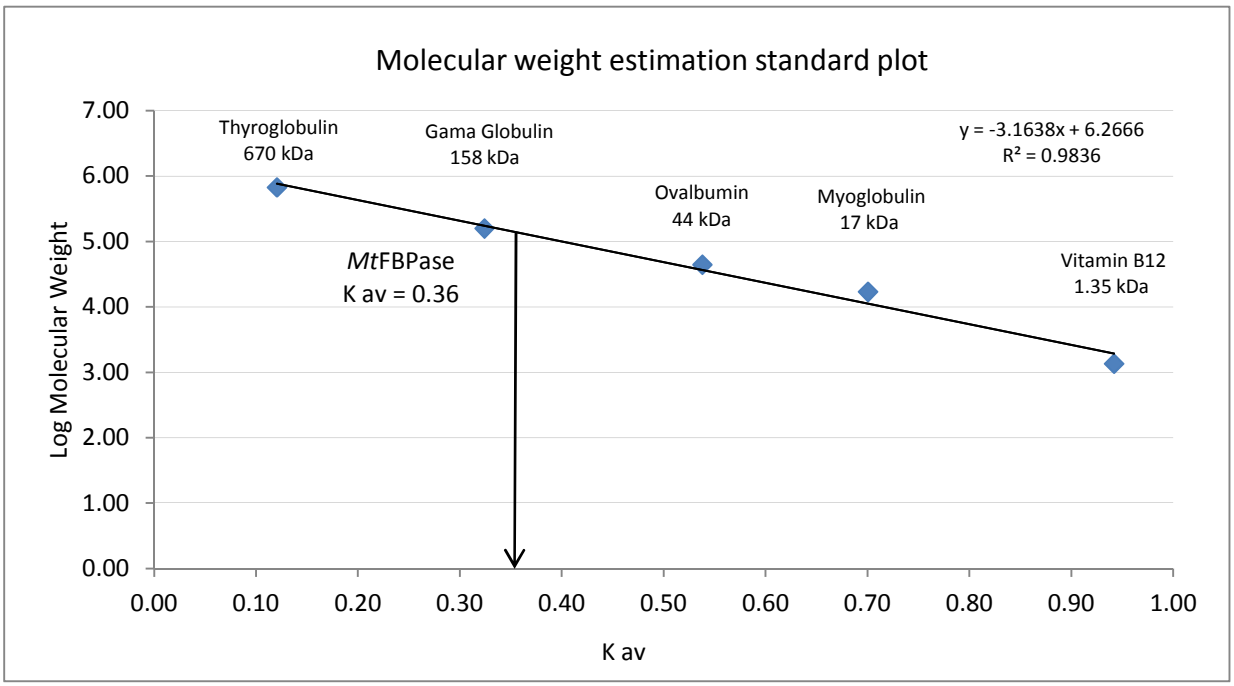
[a]: Specific activity (Units/mg): A unit of activity catalyzes formation of 1μmol of F6P per min.
[b]: Purification fold: ratio of specific activities during purification to that of the cell lysate supernatant.
[c]: % yield: comparison of total units (U) at each stage of purification to that of the cell lysate supernatant sample.

TABLE IX: PURIFICATION OF *MT*FBPASE FROM *E. COLI* BL21 (DE3) [PET15B-*GLPX*] CELLS IN A ONE LITER CULTURE.

C. *Mt*FBPase exists as a Tetramer (Size Exclusion Chromatography): The apparent protein subunit size as determined by SDS-PAGE analysis was about 39 kDa (Figure 24b) which is higher than the predicted theoretical mass of about 34.5 kDa. However, mass spectral analysis of the protein sample revealed the mass of *Mt*FBPase to be 36612.5 Da (accounting for the extra hexa-histidine tag). This abnormal electrophoretic mobility has been cited for some other FBPases (Rittmann, Schaffer et al. 2003), as well as IMPases such as *Mtb* SuhB (Nigou, Dover et al. 2002), and *E. coli* SuhB (Chen and Roberts 2000). Based on several citations, this observed abnormal electrophoretic mobility appears to be a striking feature of the super-family of lithium-sensitive phosphatases because it is also shared by the 3′(2′),5′-diphosphonucleoside 3′(2′)-phosphohydrolase (DNPPase, RHL protein) from rice (Peng and Verma 1995).

Molecular weight estimation by SEC, calibrated against the elution volumes of known protein standards, is an established procedure. SEC is also capable of providing an estimation of the oligomeric state of the protein in solution under selected buffer conditions. Void volume of the column as determined by dextran blue V_o =39.80 ml; Total column volume V_t (also referred to as geometric column volume) =120 ml. Using this method we estimated the molecular weight

for the purified *Mt*FBPase to be about 127 kDa (Figure 25), indicating that at a concentration of 1 mg/ml, it exists possibly as a tetramer (assuming monomer subunit of about 36 kDa). Moreover the predicted multimeric state of *Mt*FBPase based on crystalogrpahic symmetry and also PDBePISA (Protein Interfaces, Surfaces and Assemblies) is found to be tetrameric (dimer with crystallographic symmetry - tetramer) as discussed in detail in Chapter 5 (Krissinel and Henrick 2007). In the case of GlpX from *E. coli*, an anomalous high native molecular mass of 118 kDa was obtained in the gel filtration experiments (predicted monomer mass=36 kDa) indicating that it could be an elongated dimer, this data was further supported by the observed elongated shape of the GlpX dimer in the crystal structure (Brown, Singer et al. 2009). On the contrary, the purified His-tagged *C. glutamicum* FBPase (65% identity to *Mt*FBPase), eluted at about 140 kDa in size exclusion chromatography, indicating it to be a tetramer (Rittmann, Schaffer et al. 2003).



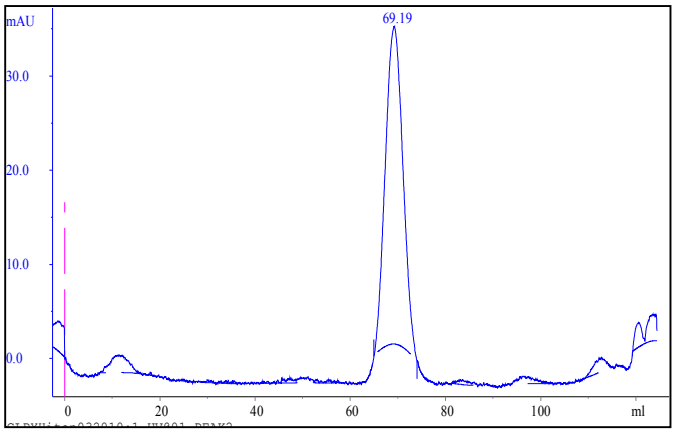


Figure 25. *Mt*FBPase molecular weight estimation by SEC. 25a. K_{av} vs. log (molecular weight) plot for molecular weight standard. The elution volumes of standard proteins were used to calculate the K_{av} values (K_{av}= ($V_e - V_o$) / ($V_t - V_o$)). The standard proteins of known molecular weight were thyroglobulin (bovine): 670 kDa, gama globulin (bovine): 158 kDa, ovalbumin (chicken): 44 kDa, myoglobulin (horse): 17 kDa, and vitamin B12: 1350 Da. K_{av} for *Mt*FBPase=0.36. **25b.** Size-exclusion chromatography (SEC) of *Mt*FBPase on Superdex-200.

Activity and stability: The enzymatic activity of the purified enzyme was determined by continuously monitoring the increase in absorbance due to the reduction of $NADP^+$ to NADPH at 340 nm for 10 min. Different amounts of the purified *Mt*FBPase were checked for activity. The activity of the *Mt*FBPase was linearly dependent on the amount of protein added to the reaction mix indicating that initial velocity is proportional to the total enzyme concentration. The minimum concentration of protein giving linear increase in absorbance at 340 nm for 10 min (0.1 μg/100 μl) was used for subsequent assays. The assay was standardized so that all reactants were in excess and only the recombinant *Mt*FBPase was the limiting factor in the reaction. The purified enzyme could easily be stored under refrigeration (2–8 °C) for a period of 1 month without any significant drop in enzymatic activity.

D. Effect of Bivalent Metal Ions: Like other lithium sensitive phosphatases (FBPase and IMPase) *Mt*FBPase also had an absolute requirement of bivalent metal ions Mg^{2+} and Mn^{2+} (Figure 26). The concentration required for maximal FBPase activity was at least 8 mM and the activity remained steady until a concentration of 40 mM (5-fold) (Figure 26b). This result was not consistent with those observed for *glpX* from *E. coli* (Donahue, Bownas et al. 2000; Brown, Singer et al. 2009) which require much lower concentrations Mg^{2+} for activity. Furthermore, mammalian IMPases (human, rat and bovine) (McAllister, Whiting et al. 1992; Peng and Verma 1995) and *E. coli* IMPase (Chen and Roberts 2000) are inhibited at much lower concentrations of Mg^{2+} (3–5 mM). In fact, this magnesium dependence of activity for this particular FBPase resembles the IMPase enzymes from *Mtb* SuhB (Nigou, Dover et al. 2002), hyperthermophilic bacteria such as *Thermotoga maritime* (Chen and Roberts 1999), archaea such as *Methanococcus jannaschii* MJ0109 (Chen and Roberts 1998), and the dual function FBPase and IMPase (Rv2131) from *Mtb* (Gu, Chen et al. 2006). Replacement of Mg^{2+} with 4 mM Mn^{2+} led to a 4-

fold increase in activity, but presence of Mn^{2+} at higher concentrations led to a decrease in this effect (Figure 26b). Omission of Mg^{2+} (in assay mixture) resulted in >90% loss of activity. Addition of EDTA at an equimolar concentration to the bivalent metal ions, gave similar results. This establishes the absolute requirement of the bivalent metal ions Mg^{2+} and Mn^{2+}. Among the other divalent metal ions tested, the only ion that could substitute for Mg^{2+} was Mn^{2+}. However, other divalent metal ions, including Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Fe^{2+} and Ca^{2+}, showed no significant activation of FBPase at any specific concentration (in the range of 0.01 mM–40 mM). In the absence of Mg^{2+}, Figure 26a shows the relative *Mt*FBPase activity in the presence of each of these bivalent metal ions at concentrations of 8 mM respectively.

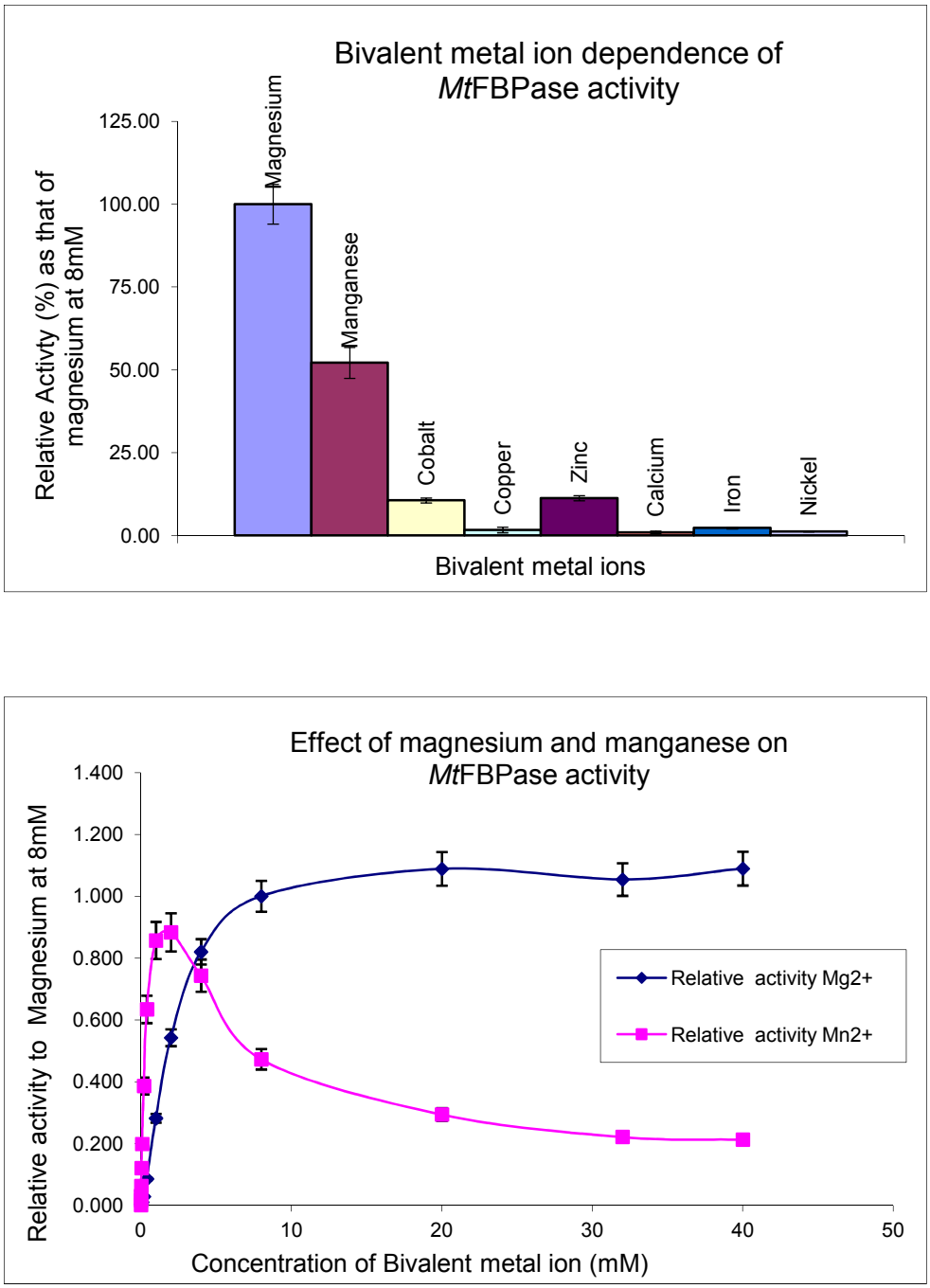


Figure 26. Bivalent metal ion dependence of *Mt*FBPase activity. **26a.** Metal ion dependence of *Mt*FBPase activity. The activities represented were determined as described in material and methods at 8 mM concentration of each cation and are normalized (percentage activity) to the corresponding value obtained with 8 mM $MgCl_2$. Metal ions Mg^{2+}, Mn^{2+}, Zn^{2+}, and Ca^{2+} were used as chloride salts, and Co^{2+}, Cu^{2+}, Fe^{2+}, and Ni^{2+} were used as sulfate salts. **26b.** Mg^{2+} and Mn^{2+} dependence of *Mt*FBPase activity; FBPase activity (relative activity to that at 8 mM Mg^{2+}) was assayed as described in material and methods, the concentrations of bivalent metal ions were varied in the range of 0.01–40 mM.

E. Effect of monovalent cations (Li^+ inhibition): The monovalent cation Li^+ is a potent inhibitor of the FBPase and IMPase family enzymes (York, Ponder et al. 1995; Albert, Yenush et al. 2000; Patel, Martinez-Ripoll et al. 2002; Rittmann, Schaffer et al. 2003). In this study, *Mt*FBPase activity was inhibited by Li^+, exhibiting 50% residual activity (IC_{50}) at 200 μM and <10% activity at (IC_{90}) 2.5 mM Li^+ (Figure 27). The inhibition constant for LiCl (IC_{50}=200 μM) is in close agreement with that observed for *C. glutamicum* FBPase II (Rittmann, Schaffer et al. 2003) (IC_{50}=140 μM) and also with that for mammalian FBPase (0.3 mM). These findings suggest that *Mt*FBPase belongs to the group of lithium sensitive phosphatases, unlike the *glpX*-encoded FBPase in *E. coli* which belongs to the more lithium-resistant sub group (Brown, Singer et al. 2009). The observed inhibition constant is significantly lower than reported previously (Movahedzadeh, Rison et al. 2004) for *Mt*FBPase cell free extracts (IC_{50}=1.1 mM, IC_{90}=10 mM), possibly due to interference by other components in the cell free extract and also moderately purified *Mt*FBPase. In order to determine inhibition specificity of Li^+, other monovalent cations including Na^+ and K^+ were also tested. The results showed that K^+ had no significant effect on FBPase activity even at 100 mM. Na^+ showed 27% inhibition in enzyme activity at a concentration of 80 mM. Among the other ions tested, NH_4^+ (added as NH_4Cl) did not inhibit activity in a concentration range of 0.1 mM–20 mM. Inorganic phosphate present at a concentration of 2 mM in the enzyme assay displayed 63% maximal activity as compared to control, indicating product (phosphate) dependent inhibition of enzyme activity, or possible allosteric effects via glycerol-3 phosphate.

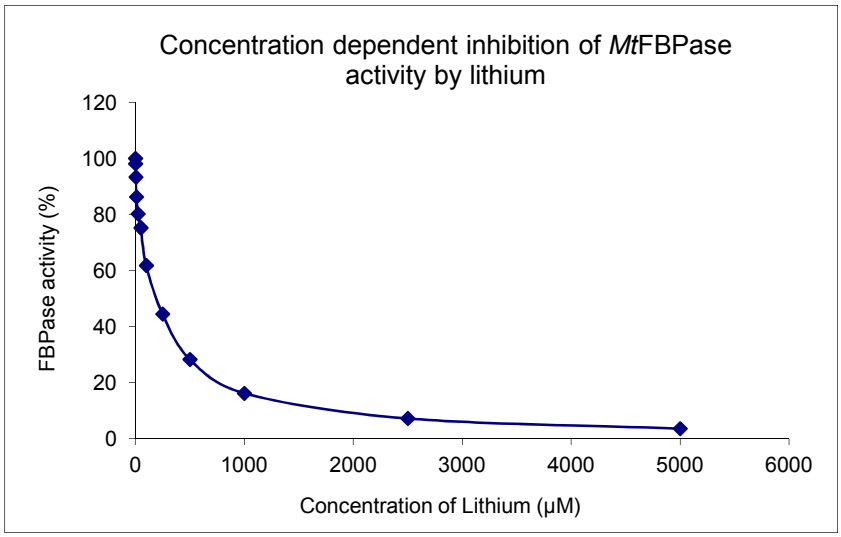
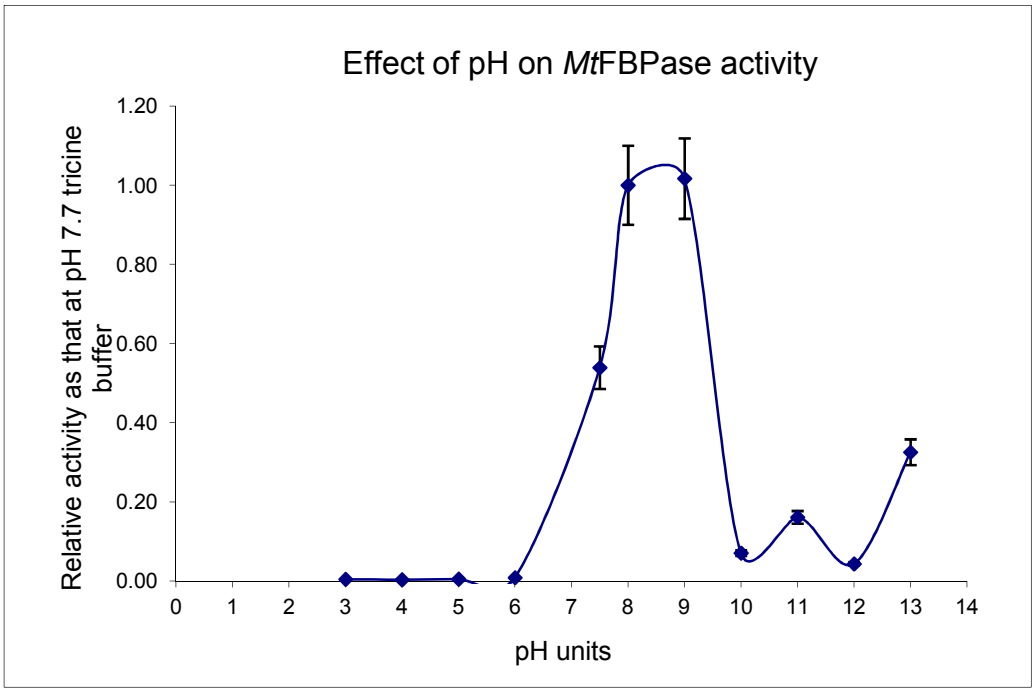


Figure 27: Lithium dependent inhibition of FBPase activity. 50% activity at 200 μM Li^+ and less than 10% activity at 2.5 mM Li^+; FBPase activity is reported as % relative activity where the activities are compared to enzyme control without any lithium added.

F. Effect of pH (activity dependence) and temperature (stability): The FBPase activity vs. pH plot exhibited a bell-shaped curve (Figure 28a). The relative activities (at respective pH values in comparison to that at pH 7.7) are plotted. FBPase was active over a pH range from 7.0 to 9.0, exhibiting maximum activity at pH 9.0. However, all assays were performed at pH 7.7 (physiologically relevant pH) since the enzyme was purified and stored, and also found to be stable, at pH 7.7. The thermal stability was tested from 10 to 80 °C by incubating the enzyme for 30 min. The enzyme retained its full activity until 30 °C. A significant loss in the activity (65% loss) was noted at 40 °C. Complete inactivation of the enzyme occurred at temperatures greater than 50 °C (Figure 28b).



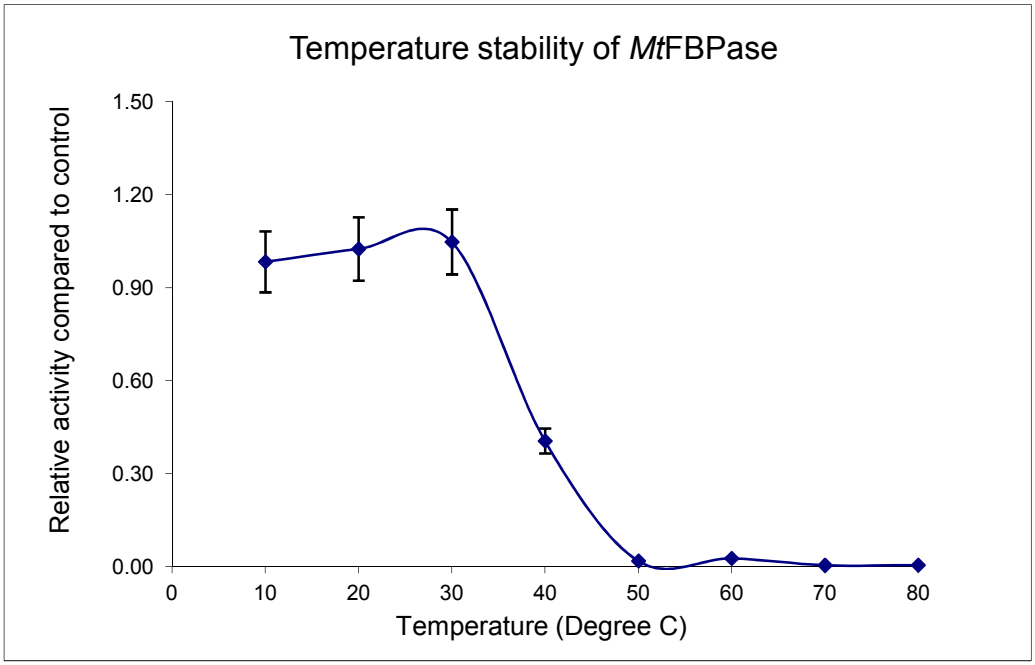


Figure 28. pH optima and thermostability of *Mt*FBPase. 28A. pH optima for *Mt*FBPase. For measurement of pH optima, enzyme activity was assayed in 100 mM of different buffers (refer to methods), relative activity compared to that of Tricine buffer pH 7.7 as indicated in the plot; **28B.** Thermostability of *Mt*FBPase was assayed after incubating aliquots of the enzyme in temperature ranging from 10–80 °C for 30 min followed by determination of residual activity. Relative activity compared to that of the control incubated at 4 °C is shown in the plot.

G. Catalytic properties: To determine the K_m and V_{max}, the assay contained fructose 1, 6-bisphosphate at concentrations of 3.9–1500 μM. Substrate dependent inhibition of enzymatic activity was observed at concentrations higher than 500 μM. Such substrate dependent inhibition has also been observed for *E. coli* GlpX protein. Such inhibition was also verified in using the malachite green assay for quantification of the phosphate released, as was done for *E. coli* GlpX protein. Cooperative behavior with regard to substrate concentration and Mg^{2+} concentration was also observed (Figure 26b shows that the activity does not change significantly after 8 mM concentration of Mg^{2+}). The double-reciprocal plots of the reaction velocity versus Mg^{2+} (data not shown) were found to be nonlinear at lower concentrations of Mg^{2+}. The plots were nonlinear because of the co-operative behavior of Mg^{2+} binding to the enzymes. Such a co-operative behavior has also been observed for FBPase from other sources. Kinetic parameters were determined by nonlinear curve fitting from the Lineweaver-Burk plot using GraphPad Prism software (version 5.00, Graph-Pad Software). A K_m of 44 μM, V_{max} of 1.6 U/mg and K_{cat} of 1.0 s−1 were determined. The observed K_m of 44 μM indicates high substrate affinity as evident in several other Class II FBPases (*E. coli* GlpX K_m =35 μM (Donahue, Bownas et al. 2000), *C. glutamicum* FBPase II K_m =14 μM (Rittmann, Schaffer et al. 2003) and *E. coli* Class I FBPase K_m =14 μM (Kelley-Loughnane, Biolsi et al. 2002). The specific activity and turnover number is lower than that for other reported Class II FBPases. Accordingly the calculated catalytic efficiency K_{cat}/ K_m was 1.0/44 μM=22.7 s^{-1} mM^{-1} which is significantly lower than that for the *C. glutamicum* FBPase (K_{cat}/ K_m =236 s^{-1} mM^{-1} determined at 30 °C). However, it is almost similar to that for GlpX *E. coli* (K_{cat}/ K_m =57 s^{-1} mM^{-1} determined at room temperature) and about 40 fold lower than that for *E. coli* FBPase I (K_{cat}/ K_m =948 s^{-1} mM^{-1}).

H. Substrate specificity of *Mt*FBPase II. Enzymatic assays to determine substrate specificity indicated that F1,6BP was the best substrate found for phosphatase activity, followed by glycerol 3-phosphate (G3P) (16.3 % activity). While 3-phosphoglycerate (3PG) (8.90 % activity), glucose 1,6–phosphate (7.18 % activity) and ribulose 1,5-bisphosphate (6.04 % activity) served as less efficient substrates (Table X).

Additionally, 1 mM fructose 1-phosphate, G6P, F6P, mannose 6-phosphate, glucose 1,6-bisphosphate, and inositol 1-phosphate were not significant substrates for FBPase II.

Compound[a]	Rate (%)[b]
F1,6BP	100
sn-glycerol 3-phosphate	16.3
3-phosphoglycerate	8.90
D-glucose 1,6-bisphosphate	7.18
D-ribulose 1,5-bisphosphate	6.04
F6P	2.60
D-mannose 6-phosphate	<2
G6P	<2
D-fructose 1-phosphate	<1
inositol 1-phosphate	<1

a Concentrations of potential substrates used in the assay were 1 mM.
b Enzyme activity was determined under the optimal reaction conditions as described in Materials and Methods, using the phosphate quantitation assay. Reaction rates were compared to the rate of FBPase activity with 1 mM F1,6BP. Rates were adjusted for autophosphohydrolysis.
TABLE X. SUBSTRATE SPECIFICITY OF *MT*FBPASE II.

Tetrameric FBPase I in *E. coli* and its regulatory mechanism: FBPase I is required for growth of *E. coli* on gluconeogenic carbon sources (Donahue et al. 2001). The overall structure and the regulation of FBPase I activity in *E. coli* has been extensively studied and characterized

(Hines, Fromm et al. 2006; Hines, Chen et al. 2007; Hines, Fromm et al. 2007; Hines, Kruesel et al. 2007). Phosphoenolpyruvate (PEP) and citrate (intermediates of glycolysis and the tricarboxylic acid cycle), are identified as activators of *E. coli* FBPase I. At physiologically relevant concentrations, PEP and citrate stabilize a tetramer in the active R-state (Hines, Fromm et al. 2007). This allosteric activation (of FBPase) by PEP and citrate implies a rapid feed-forward mechanism of activation of gluconeogenesis in heterotrophic bacteria. Interestingly these bacteria can also rapidly down-regulate the activated FBPase and thus avoid futile cycling. This down-regulation occurs via allosteric inhibition of *E. coli* FBPase by G6P or AMP. FBPase undergoes a transition from the active R-state to a T-state in response to G6P and AMP ligation. Moreover G6P and AMP are synergistic inhibitors of *E. coli* FBPase (Hines, Kruesel et al. 2007). Allosteric regulation of FBPase I constitutes an important part of the metabolic switch that allows rapid changeover between gluconeogenesis and glycolysis in response to nutrient availability and ATP synthesis. A similar experimental approach, as described by Hines et al. was taken to identify possible effectors of *Mt*FBPase.

I. Effectors of *Mt*FBPase II: An extensive set of effectors was tested for their ability to inhibit or activate the *Mt*FBPase activity. Effectors comprising intermediates of the tricarboxylic acid (TCA) cycle, glycolytic intermediates and adenosine phosphates (ATP, ADP, AMP) were rationally chosen based on previous literature reports on effectors of FBPase I activity (Table XI). Three different concentrations (50, 1000, and 5000 μM) of these effectors were chosen to assess their concentration dependent effect on FBPase activity.

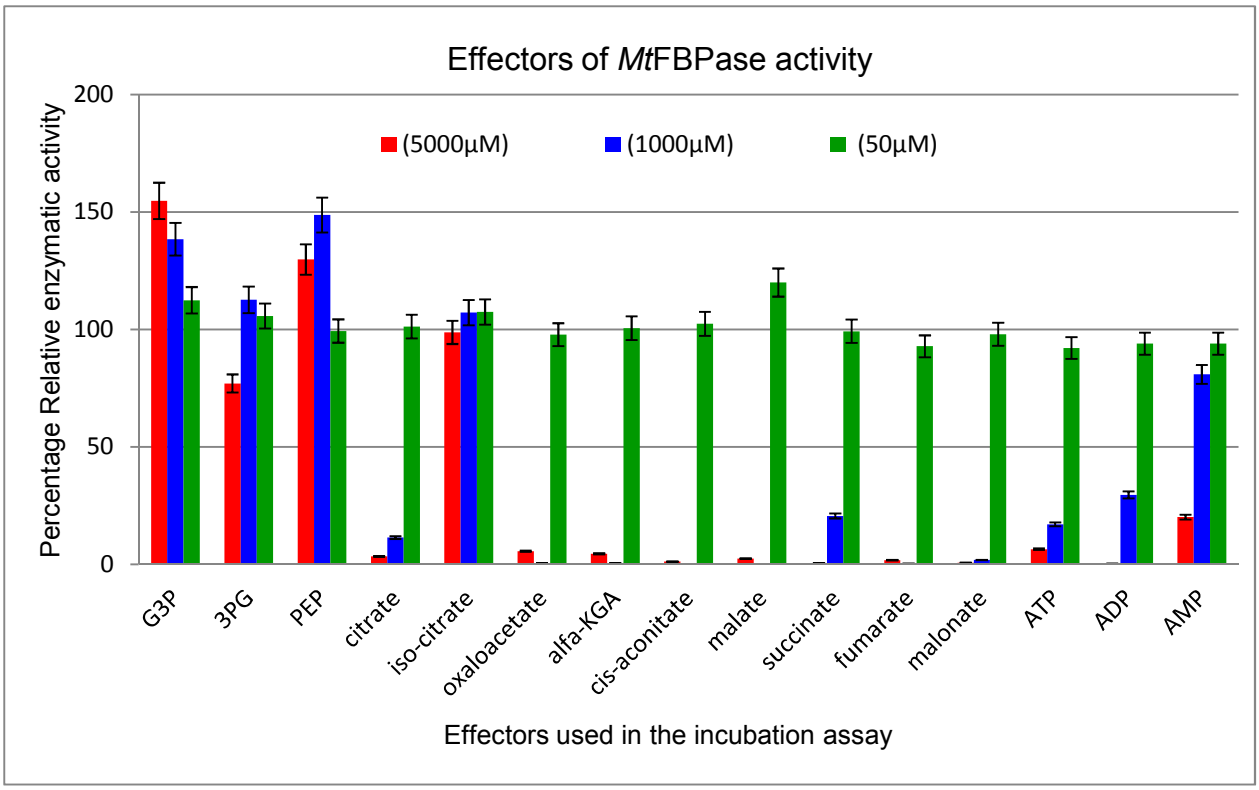


Figure 29. Activation of *Mt*FBPase by intermediates of glycolysis and the TCA cycle. Assay performed as described in methods, bar heights are the means of three independent assays, indicating percent activity relative to a control with no effector. Error bars represent 5% SD. Enzyme activity was determined by coupled spectrophotometric assay as described under "Incubation assay for identification of effectors". Reaction rates were compared to the control rate with no effector added. Appropriate controls with the catalytic product fructose 6-phosphate were run to rule out the possibility of inhibition of coupling enzymes by the tested effectors. None of the effectors had inhibitory effect on the coupling enzymes.

Phosphorylated three-carbon derivatives G3P, 3PG and PEP had marginal activation effects on FBPase II activity at 1 mM concentration. Such a marginal activation effect was also observed for G3P and PEP at the higher concentration of 5mM, but not for 3PG (Figure 29). 3PG showed rather marginal inhibitory effects on FBPase II activity at the higher concentration of 5mM (Figure 29).

ADP and ATP inhibited the FBPase II reaction with Fru 1,6-BP substrate at sub millimolar concentrations (Figure 29), however, sub millimolar concentrations of AMP did not

significantly affect *Mt*FBPase activity. At concentrations as high as 5mM all the adenosine phosphates tested inhibited FBPase activity. This high concentration of 5 mM is rather non-physiological and, hence, not relevant to regulation of enzyme activity.

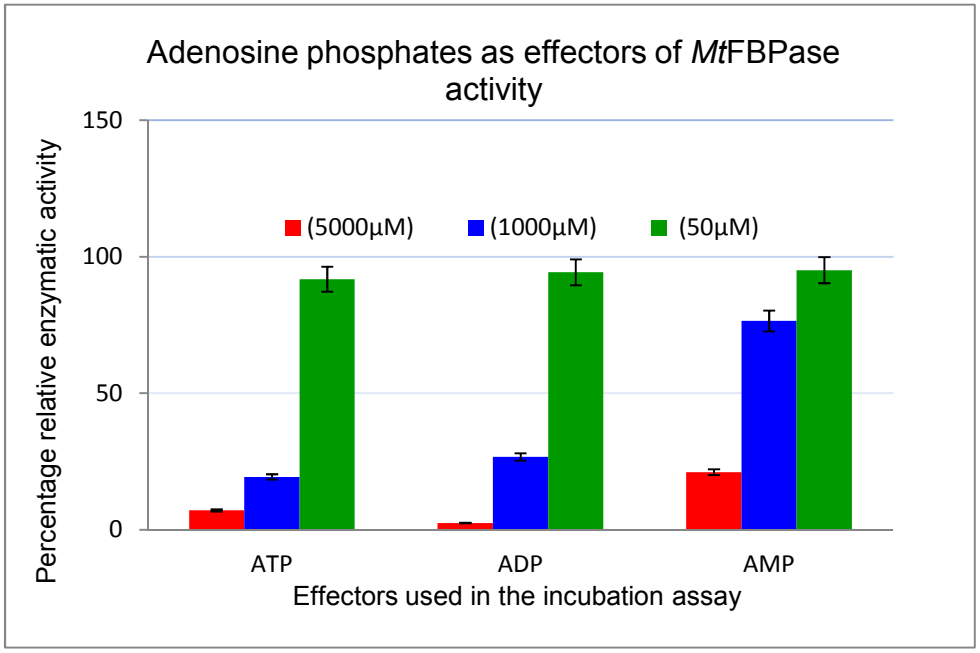


Figure 30. Inhibition of *Mt*FBPase by adenosine phosphates. Assay performed with higher amounts of $MgCl_2$ (1mM higher than concentration of the effector). Enzyme activity was determined by coupled spectrophotometric assay as described under "Incubation assay for identification of effectors". Reaction rates were compared to the control rate with no effector added.

Assays involving evaluation of adenosine phosphates as effectors, were repeated with slightly higher amounts of the bivalent metal ion Mg^{2+} (> 8mM $MgCl_2$). Based on earlier published reports, it is possible that Mg^{2+} required for FBPase activity is bound by ADP or ATP, thus decreasing the FBPase II activity. The inhibitory effects of ATP or ADP are probably due to low concentration of free Mg^{2+} available for FBPase activity. Therefore, 1 mM Mg^{2+} in excess of the nucleotide concentration was added to the reaction involving AMP, ATP, and ADP as

effectors (Figure 30). With excess Mg^{2+} however, both ADP and ATP had similar inhibitory effects on the activity of *Mt*FBPase as was observed in the initial screening of effectors.

Incubation of *Mt*FBPase individually with intermediates of the tricarboxylic acid cycle (at 1 mM concentrations) abolishes the enzymatic activity (Figure 29). The TCA cycle intermediates citrate, oxaloacetate, alfa–ketoglutarate, *cis*-aconitate, malate and malonate all inhibit enzyme activity (50% inhibition) at submillimolar concentrations (Table XI). Precise IC_{50} values of these effectors were determined by repeating the incubation assay with serial dilutions of the effectors (Table XI). Interestingly, iso-citrate did not have as pronounced inhibitory effect on enzyme activity as its isomer citrate which becomes a subject matter of further investigation. Although these observed inhibition concentrations are very high and not indicative of a possible allosteric regulation for *Mt*FBPase activity (as is the case with FBPase I), a careful evaluation of these TCA intermediates is needed to characterize the inhibitor leads for structure based drug discovery. Furthermore an important consideration is that all the activity assays are performed in the presence of glycerol (required for stability of *Mt*FBPase), future experiments need to be performed in the absence of glycerol/use alternate stabilizers for *Mt*FBPase.

Effector	IC_{50} (50% inhibition of FBPase activity under ideal assay conditions) [a, b]
citrate	445.26 ± 11.33 μM
iso-citrate	ND[c]
oxaloacetate	549.35 ± 16.79 μM
alfa-keto glutaric acid	499.91 ± 09.52μM
cis-aconitate	459.51 ± 07.49 μM
malate	616.25 ± 25.64 μM
malonate	430.11 ± 8.93 μM

[a] Enzyme activity was determined by coupled spectrophotometric assay as described in the "Incubation assay for identification of effectors" section. Reaction rates were compared to the rate with no effector added.
[b] Appropriate controls were run using the catalytic product fructose 6-phosphate to rule out the possibility of inhibition of coupling enzymes in the assay.
[c]Iso-citrate did not show any inhibition in the tested range of 50 μM to 5000μM.

TABLE XI. IC_{50} VALUES OF EFFECTORS OF *MT*FBPASE (50% INHIBITION OF FBPASE ACTIVITY WITH F1,6BP SUBSTRATE).

From a structural regulation standpoint, the inhibition of *Mt*FBPase by the above effectors is important since there is a possibility of identifying allosteric regulation sites in the enzyme tertiary structure (for details refer to the structural biology section chapter 5) if any. Any information on the possible allosteric regulation, presents an exciting opportunity to begin structure-guided drug discovery towards both the allosteric sites in addition to targeting the catalytic site of this enzyme.

4.4 CONCLUSION

The chapter describes successful purification and biochemical characterization of *Mt*FBPase. The SEC results suggest that *Mt*FBPase exists as a tetramer in solution (estimated molecular weight =127 kDa). Like other FBPases, *Mt*FBPases requires bivalent metal ions for enzymatic activity. Based on the observed IC_{50} for Li^+, *Mt*FBPase is classified as a lithium sensitive phosphatase unlike the *E. coli* GlpX protein which is more resistant to Li^+. The thermostability and pH optima for activity have been determined. Substrate specificity findings demonstrate that *Mt*FBPase is a true FBPase with no IMPase activity, ruling out the possibility of it being a dual activity FBPase/IMPase. Also the activity towards other substrates tested (less than 15% of that for F1,6BP) is rather non-specific indicating that *Mt*FBPase has no other significant activity. Of the several TCA intermediates tested, PEP, 3PG and G3P showed marginal enzyme activation at millimolar concentration. TCA intermediates citrate, oxaloacetate, αKG, *cis*-aconitate, malate and malonate all inhibit enzyme activity (50% inhibition) at submillimolar concentrations. These findings suggest that unlike FBPase I, *Mt*FBPase does not have similar allosteric regulation (activation by citrate and PEP and inhibition by AMP and G6P). Also the inhibition by TCA intermediates is an important finding which needs to be validated by additional co-crystallization experiments in order to rule out the possibility of non-specific inhibition. This finding has an evolutionary relevance since presence of an FBPase II which appears to have different (compared to the FBPase I) or no allosteric regulation might be a survival strategy adopted by *Mtb* to prevent a possible metabolic “choke-point” during intracellular growth and survival.

CHAPTER 5

STRUCTURAL CHARACTERIZATION OF FRUCTOSE 1,6-BISPHOSPHATASE II

5.1 RATIONAL AND HYPOTHESIS

Since the biological significance of *glpX* encoded FBPase has been verified and the enzyme has been successfully purified and characterized biochemically, the next important step in the drug discovery/target validation process would be to determine its three-dimensional structure. X-ray crystallography has been the most reliable technique for structural biology since the 1960's. Macromolecular crystallography started in the late 1950s, beginning with the structures of haemoglobin and myoglobin by Max Perutz and John Kendrew respectively, for which they were awarded the Nobel Prize in Chemistry in 1962 (Kendrew, Dickerson et al. 1960; Perutz, Rossmann et al. 1960). Currently, over 80,000 X-ray crystal structures of proteins, nucleic acids and other biological molecules have been determined as described in the current pdb holdings (http://www.rcsb.org/pdb/statistics/holdings.do). Obviously, two important determinants of successful utilization of macromolecular crystallography would be the crystallization of the said target in a form suitable to achieve the structure solution and the diffraction of crystals to high resolution. The structure of a homologue from *E. coli* (GlpX protein) has been solved (**PDB ID:** 3big, 3bih, 3d1r, 1ni9) (Brown, Singer et al. 2009), creating the possibility of a structure solution for *Mt*FBPase by a relatively simple molecular replacement (MR) method thus avoiding rather time-consuming and resource intensive methods such as Multi-wavelength anomalous diffraction (MAD), Single-wavelength anomalous dispersion (SAD) or the more classic Multiple Isomorphous Replacement (MIR).

Like any other protein, the crystallization of the target protein (*Mt*FBPase), is the major obstacle for success. This chapter describes the successful crystallization of *Mt*FBPase in several

forms, further it also presents the structure solution and refinement of *Mt*FBPase in apo- and F6P-bound forms. The chapter discusses the similarities and differences between the *E. coli* GlpX and *Mt*FBPase proteins, provides conclusive evidence that *Mt*FBPase exists as a homotetramer (quaternary structure) and the results provide fundamental structural data and methodology for structure-based drug discovery focused on this essential *Mtb* protein.

5.2 MATERIAL AND METHODS

A. Crystallization:

The process of crystallization can be well described by understanding the crystallization phase diagram (Figure 31). One of the several adjustable parameters (precipitant or additive concentration, pH and temperature) can be varied along with varying protein concentration so as to generate a two dimensional phase diagram. The crystallization phase diagram can be further distinguished into four distinct zones:

i. Precipitation zone - very high supersaturation hence protein precipitates;
ii. Nucleation zone - moderate supersaturation wherein spontaneous nucleation takes place;
iii. metastable zone) - lower supersaturation, best possible condition for growth of large, stable, well ordered crystals;
iv. undersaturation zone - protein is completely dissolved and does not crystallize.

Nucleation being a prerequisite for crystallization, each method reaches the nucleation zone and then the metastable zone via a different route as depicted in the phase diagram (dashed lines) (Figure 31).

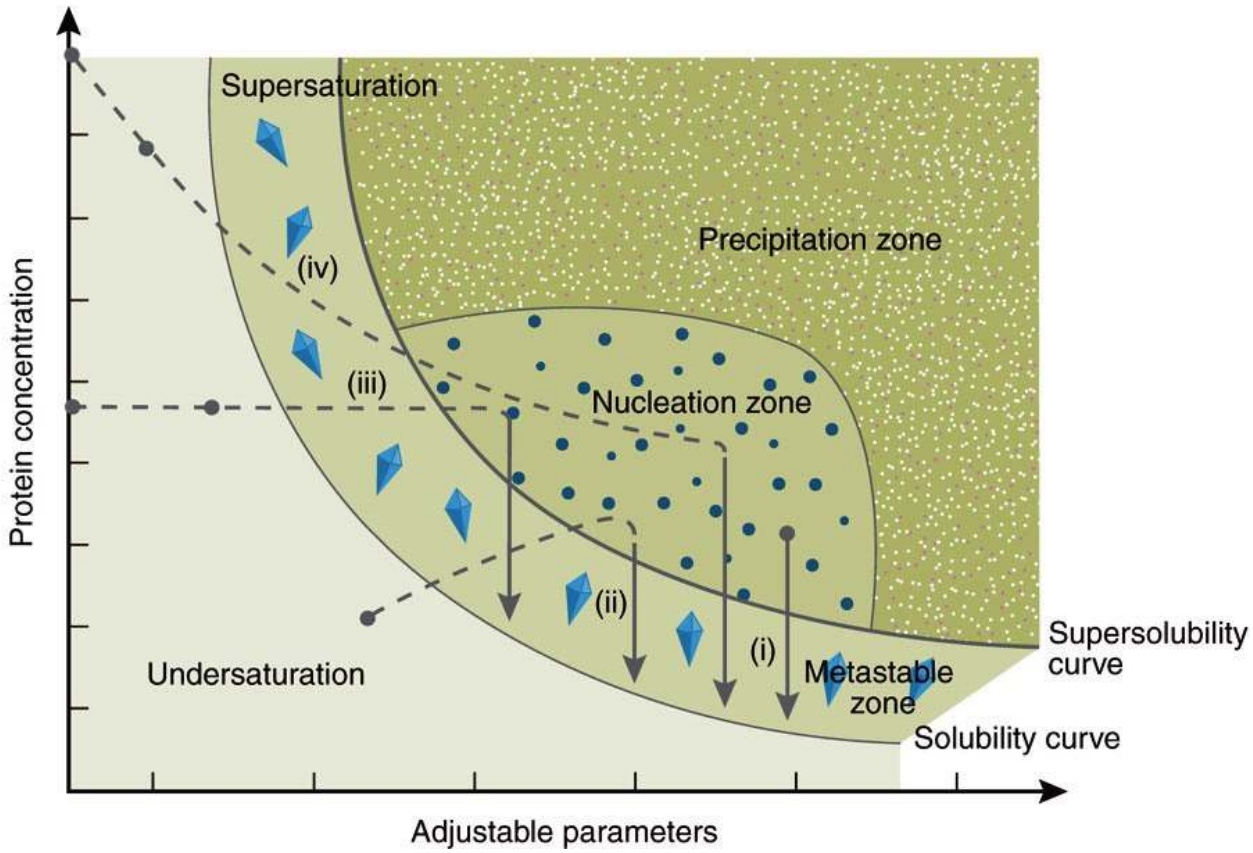


Figure 31. Schematic illustration of a protein crystallization phase diagram. Adjustable parameters include precipitant or additive concentration, pH and temperature. The four major crystallization methods are represented: (i) microbatch, (ii) vapor diffusion, (iii) dialysis and (iv) free interface diffusion (FID). Each involves a different route to reach the nucleation and metastable zones, assuming the adjustable parameter is precipitant concentration. The filled black circles represent the starting conditions. Two alternative starting points are shown for FID and dialysis because the undersaturated protein solution can contain either protein alone or protein mixed with a low concentration of the precipitating agents. The solubility is defined as the concentration of protein in the solute that is in equilibrium with crystals. The supersolubility curve is defined as the line separating conditions under which spontaneous nucleation (or phase separation or precipitation) occurs from those under which the crystallization solution remains clear if left undisturbed. **Adapted with permissions from Chayen et al.** (Chayen 2004; Chayen and Saridakis 2008)

For crystallization of *Mt*FBPase, purified His-tagged recombinant *Mt*FBPase was concentrated to 10 mg/ml in the purification exchange buffer (refer the procedure for purification). Crystallization trials were performed by the hanging-drop vapor-diffusion method

and suitable crystallization conditions were found from several pre-formulated crystallization screens from Hampton Research (Hampton Research, Aliso Viejo, California, USA) (Gutka, Franzblau et al. 2011).

1. Crystal screening and crystal optimization: Several crystallization screens from Hampton Research were used for crystal screening. Further optimization of the crystals was performed using stock precipitant solutions from Hampton Research. Index, Crystal screen, PEGRx, SaltRx and PEG/Ion were the crystallization screens used for identifying possible conditions for *Mt*FBPase. After successful crystallization conditions were found, the first test undertaken was to reproduce these crystals in the same conditions and further optimize them in size and quality. Volumes of the precipitant and protein solution were varied so as to evaluate their respective effect on the crystal formation for *Mt*FBPase. The first optimization screens undertaken were designed by varying the successful precipitant concentration (from ~0.5 to 2 times the original concentration). This was often followed by a narrower screen around the resulting successful precipitant concentrations. In cases where it was applicable, a pH screen (varying the pH in increments of 0.2-0.4 units over a range of ±1 unit either side of the original pH) using several of the most successful precipitant concentrations from the refined precipitant screen, was also undertaken. If crystals were grown from these screens the size of the drop and the ratio of protein to precipitant were varied to optimize the crystals further (generally for obtaining larger crystals). All crystallization experiments were undertaken using the hanging drop vapour diffusion method. In this method a small volume of protein (typically 1-3 μL) mixed with a similar volume of precipitant solution, is placed on a coverslip and sealed over a well containing a much larger volume of the same precipitant solution. Water and other volatile components equilibrate by vapour diffusion causing the hanging drop to shrink and the

precipitant and protein concentration within it to increase. In some cases, depending upon the conditions, this results in amorphous precipitation or crystallization of the protein.

2. Hanging drop Vapor Diffusion method: Protein crystallization by the hanging drop vapor diffusion technique is based on the equilibration of salt/buffer concentrations in the protein solution and the buffer/reservoir. In this method, a drop consisting of a mixture of protein and precipitating solution (buffer) is placed on an inverted pre-siliconized glass cover slip which equilibrates with the buffer reservoir. The air-gap between the reservoir and the cover slip is sealed (crystallization trays have sealant) such that a salt concentration difference is generated between the protein drop and the buffer reservoir. Over a period of few weeks, equilibration occurs, leading to favorable conditions for protein crystallization. Crystallization experiments were carried out using 48-well VDX48 crystallization trays (with sealant) and 12 mm circular (0.22mm or 0.96 mm thickness) pre-siliconized glass cover slips supplied by Hampton research. All experiments were carried at 25ºC. A 1-2 µL drop of protein (usually centrifuged at 12000 rpm for 5 minutes prior to use) was mixed with 1-2 µL of precipitant solution on a siliconized cover slip, which was then inverted and sealed over pre-greased well containing 500-600 µL of the same precipitant solution, further details can be found in study by Gutka et al. (Gutka, Franzblau et al. 2011).

3. Crystal Scoring System: All drops were scored after being laid down to monitor their progress towards crystallization. Drops were examined at least once a week over the first month, after which they were scored on a less regular basis. The scoring system adopted was as described in the crystallization kits from Hampton research as indicated in table (**Table XII**).

Score	Crystal description	Score	Crystal description
1	Clear Drop	5	Posettes or Spherulites
2	Phase Separation	6	Needles (1D Growth)
3	Regular Granular Precipitate	7	Plates (2D Growth)
4	Birefringent Precipitate or Microcrystals	8	Single Crystals (3D Growth < 0.2 mm)
		9	Single Crystals (3D Growth > 0.2 mm)

TABLE XII. CRYSTAL SCORING SYSTEM AS ADOPTED FROM THE CRYSTALLIZATION SCREENS FROM HAMPTON RESEARCH.

4. Co-crystallization experiments with substrate F1,6BP and catalytic product F6P: In some cases protein was directly mixed with compounds of interest and then set up in crystallization trials. The crystallization trials were set up with the compound of interest included in the crystallization buffer. The *Mt*FBPase complex structures with substrate F1,6BP, and product F6P were obtained by subtle variations in the crystallization conditions as follows. *Mt*FBPase-F1,6 BP crystals were obtained by mixing 100 µl protein solution + 10 µl 10mM substrate solution (F1,6BP), incubating for 30 min at 30°C and then mixing 1 µl of the resultant mix with 1 µl reservoir solution consisting of 2.9M Sodium Malonate pH 4.0 (Malonate grid from Hampton Research).

*Mt*FBPase-F6BP crystals were obtained by mixing 100 µl protein solution + 10 µl 10mM substrate solution (F1,6BP), incubating for 30 min at 30°C and then mixing 1 µl of the resultant mix with 1 µl reservoir solution consisting of 2.9M Sodium Malonate pH 7.0 (Malonate grid

from Hampton Research). The protein solution contained magnesium chloride at a concentration of 1mM, no extra magnesium solution was added to the protein precipitant mix.

The pH of the precipitant mixture plays an important role in trapping the substrate and product co-crystal structures. The enzyme being virtually inactive at pH 4.0 (Gutka H.J. et. al. 2011), we could obtain the co-crystal structure of *Mt*FBPase-F1,6BP. The co-crystal structure of *Mt*FBPase-F6P was obtained at pH 7.0 since the enzyme is active in the pH range of 7-9 and capable of catalyzing the breakdown of F1,6BP to F6P (Gutka H.J. et. al. 2011).

B. Data collection and processing:

1. Cryo-mounting (Cryoprotectant Testing): Mother liquor (and also the final exchange buffer for *Mt*FBPase) already contained 10% v/v glycerol, this mother liquor was modified to include larger amounts of glycerol (typically 20% v/v) or PEG 4000 in varying percentages. These solutions were used for cryoprotection of the crystals prior to freezing and subsequent X-ray exposure.

2. Flash-freezing of crystals: Crystals for the synchrotron data collection were frozen into liquid nitrogen after equilibration with cryoprotectant solution. Crystals were transferred (using a cryo mount loop) to about 10 µl of cryoprotectant solution and equilibrated for about 10-20 seconds before being transferred quickly into a vacuum flask containing liquid nitrogen. Further the frozen crystals were held onto the liquid nitrogen cryostream for data collection.

3. Synchrotron Data Collection: X-ray diffraction data for all the *Mt*FBPase crystal forms (APO and complexes) were collected from a single frozen crystal using a MAR300 CCD detector at the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory, Illinois, USA.

4. General Data Collection and Processing Strategy: The crystal was mounted, centered in the beam, and several frames of X-ray diffraction data were collected. In most cases two frames

90º apart were collected followed by three to five consecutive frames starting from one orientation. These were then processed using HKL2000 to get estimates of unit cell parameters and some indication of the crystal lattice and space space group. In general for data collections, the aim was to collect as redundant a data set as time would allow. For each crystal form, several frames (typically 180 degrees, minimum) were collected with 1° scan width and 1 s exposure time. The crystal-to-detector distance was set to about 280 mm. The data were indexed and scaled with the HKL-2000 package (Otwinowski and Minor 1997).

5. **Integration and Scaling with HKL2000:** The indexing and scaling of diffraction data was undertaken using the HKL2000 suite of programs, XdisplayF, DENZO and SCALEPACK. DENZO was used to index the raw diffraction images, estimate the cell constants, and refine the crystal and detector parameters and integrate the diffraction maxima (Otwinowski and Minor 1997). XdisplayF was used to visualize the diffraction images and the integration carried out by DENZO. After each refinement cycle, DENZO updates the display with new values for the refined parameters and shifts in their values. After this SCALEPACK was used to find the relative scale factors between the images, refine the crystal parameters using the entire data set (global refinement), and then merge multiple measurements and measurements related by space group symmetry. SCALEPACK estimates scale and B-factors for each processed frame and after averaging intensities of symmetry related reflections, it writes out final intensities and errors and provides data quality statistics.

6. **Cell content (Matthews coefficient) analysis:** Matthews analysis is used to determine the number of molecules in the asymmetric unit, post scaling of the datasets. Matthews coefficient can be calculated in the CCP4 suite. The Matthews coefficient (V_M) is determined by dividing the volume of the asymmetric unit by the molecular weight of the protein in the asymmetric unit.

The usual range for V_M is 1.6-3.5 $Å^3$/Da for macromolecular protein crystals, corresponding to a solvent content of 27-65% (Matthews 1968). Number of molecules in the asymmetric unit (N) is an integer number that gives a protein molecular weight per asymmetric unit wherein the V_M is in an observed range of 1.6-3.5 $Å^3$/Da.

C. Structure solution by Molecular Replacement (MR) and Model Refinement

1. Molecular Replacement with MOLREP: Molecular replacement (MR) is the method of choice to obtain the phase information (position and orientation of the molecules in the unit cell) when the structure of a homologous protein is already available. The known protein should have at least 30% or greater sequence identity to the unknown protein. This known protein structure can be used to obtain initial phase information for the unknown dataset, by using it as a model of the unknown structure (Rossmann 1990; Rossmann 2001; Rhodes 2006). The process of molecular replacement positions the known structure model (search model) within the unit cell (orientation and position, a total of six parameters per chain) of the unknown structure. Initial phase information, which is improved by subsequent model building and refinement, is then obtained from the calculated structure factors of this approximation to the unknown crystal structure. Successful positioning of the known structure model into the unit cell of the unknown structure requires both angular reorientation (rotation) and relocation (translation) of the model in three dimensions. The search which is six dimensional, since both the rotation (3D) and translation (3D) require three values to be defined is generally broken down into two searches. A rotation search to find the orientation of the model is followed by a translation search (using the reoriented model) to find the location of the model in the unit cell of the unknown structure. Successful solution by MR is dependent upon several factors such as data quality, completeness of data, homology between the models and the number of molecules in the asymmetric unit.

Since originally proposed in the 1960's (Rossmann and Blow 1962), several computer packages have been developed to apply the method with increasing probability of success. Common MR packages include AMoRe, MOLREP, PHASER: 1. AMoRe, 2. MOLREP and 3. PHASER. They are all included in the CCP4 program suite.

AMoRe is a significant improvement on conventional methods, based on more powerful algorithms and a new conception that enables automation and rapid solution (Navaza 1994; Navaza 2001). AMoRe has the options for searching long lists of possible solutions; it is very fast and can search for model of more than one type to assemble multidomain solutions. The current CCP4 suite has all three of the above mentioned molecular replacement tools included. MOLREP and PHASER are the current tools for routine molecular replacement efforts. MOLREP is an automated molecular replacement program that was initially developed (Vagin and Teplyakov 1997) as a fast and efficient alternative to few existing program packages and has subsequently been expanded with new algorithms and adapted to various operating systems. MOLREP was originally part of the program suite BLANC (Vagin, Murshudov et al. 1998) but was also included in the CCP4 suite (Collaborative Computational Project 1994). There have been several recent reviews describing the utility of MOLREP as an indispensible molecular replacement program in macromolecular crystallography (Lebedev, Vagin et al. 2008; Vagin and Teplyakov 2010). PHASER is a program for phasing macromolecular crystal structures by both molecular replacement and experimental phasing methods (McCoy 2007; McCoy, Grosse-Kunstleve et al. 2007). The phasing algorithms implemented in PHASER have been developed using maximum likelihood and multivariate statistics (McCoy, Grosse-Kunstleve et al. 2005). PHASER in particular has a convenient and compact format and can make selections of rotation grid size etc that relieve the user from knowing all the details of and also selecting molecular

replacement parameters. PHASER performs anisotropy correction, cell content analysis, rotation and translation functions, phasing and refinement and provides a solution without user defined parameters. PHASER is more effective in structure solutions when the model and the unknown structure differ considerably.

2. Model building and Refinement (COOT, REFMAC5): COOT is a model building, density fitting and refinement molecular-graphics application package for crystallography of biological macromolecules. The program displays electron-density maps and atomic models and allows manipulations of the model such as real-space refinement, manual rotation/translation, rigid-body fitting, ligand search, water addition, mutations, rotamers and validation tools like Ramachandran idealization as well as interfaces to external programs for refinement, validation and graphics (Emsley and Cowtan 2004; Emsley, Lohkamp et al. 2010). The software is extremely user friendly for novice users, providing tools for common tasks which are accessible through menus and toolbars or by mouse controls. COOT was used for reiterative model refinement and building of missing chains/residues. Also its interface with the CCP4 provided with the possibility of running REFMAC5 iteratively and presented a better refined model for *Mt*FBPase. In case of *Mt*FBPase structure models, COOT was used for model refinement, building the short distorted helix residues corresponding to Arg235- Glu258 of the *E. coli* GlpX (significantly shorter in *Mt*FBPase because of a deletion of 13 residues). Also ligand fitting, finding water molecules and other refinement related functions were performed using COOT.

REFMAC5 is distributed as a part of the CCP4 suite and in conjunction with COOT was used for reiterative refinement of the model (Murshudov, Vagin et al. 1997). Reliable models at resolutions at least as low as 4 Å can be achieved with the low-resolution refinement tools such as secondary-structure restraints, restraints to known homologous structures, automatic global

and local NCS restraints (Nicholls, Long et al. 2012) . REFMAC5 is a flexible and highly optimized refinement package that is ideally suited for refinement across the entire resolution spectrum encountered in macromolecular crystallography (Murshudov, Skubak et al. 2011);(Vagin, Steiner et al. 2004).

3. Model Validation (*MolProbity*)

MolProbity was used for model validation and determination of the quality parameters for the refined structure. *MolProbity* is a general-purpose structure-validation web tool for 3D structures of proteins, nucleic acids and complexes. *MolProbity* provides broad-spectrum solidly based evaluation of model quality at both the global (overall structural fold) and local (Cβ outliers of individual residues, Ramachandran outliers, flipped branched protein side chains etc.) levels. It provides detailed all-atom contact analysis of any steric problems within the molecules as well as updated dihedral-angle diagnostics, and it can calculate and display the H-bond and van der Waals contacts in the interfaces between components. *MolProbity* results are displayed in multiple forms: as overall numeric scores, as lists or charts of local problems, as downloadable PDB and graphics files. *MolProbity* has an additional feature wherein results can be seen as 3D kinemage graphic images in the KiNG viewer. *MolProbity* is available free to all users at http://molprobity.biochem.duke.edu. (Davis, Leaver-Fay et al. 2007) *MolProbity* is one such tool addressing the model validation problems, the general capabilities, enhanced features and usage have been recently reviewed (Chen, Arendall et al. 2010).

4. Sequence and Structure analysis (ClustalW, PyMOL, PDBePISA)

The Clustal series of programs find wide application in molecular biology for multiple alignment of both nucleic acid and protein sequences and for preparing phylogenetic trees(Higgins and Sharp 1988; Thompson, Higgins et al. 1994). Clustal programs are popular because of their robustness, portability, the accuracy of the results and user-friendliness. Clustal

is available as a web based resource at the EBI (European Bioinformatics Institute) (http://www.ebi.ac.uk/clustalw/) (Chenna, Sugawara et al. 2003). ClustalW was rewritten in C++ and can be ported to the latest versions of Linux, Macintosh and Windows operating systems. The program can be run on-line from the EBI web server: http://www.ebi.ac.uk/tools/clustalw2 (Larkin, Blackshields et al. 2007). ClustalW was used for all primary sequence alignments described in this chapter.

PyMOL is a molecular graphics system with an embedded Python interpreter designed for real-time visualization and rapid generation of high-quality molecular graphics images and animations. It can also perform many other valuable tasks (such as editing PDB files and displaying maps) (DeLano 2002). PyMOL was employed for generating high-quality graphics images of MtFBPase structures as described in the results and discussion section. Recently there have been applications discussing the use of PyMOL as a structure based drug discovery tool for ligand docking and binding site analysis in conjunction with Autodock/Vina (Seeliger and de Groot 2010; Lill and Danielson 2011). PyMOL was used as a structure visualization graphics application software, all images described in the results section being taken using PyMOL.

PDBePISA (Protein, Interfaces, Surfaces and Assemblies; PISA for short) is a web based interactive tool offered by the PDBe to investigate stability of formation of macromolecular complexes. PDBePISA provides detailed analysis of surfaces, interfaces and assemblies for all entries in the Protein Data Bank (PDB), the service also allows upload of one's own PDB format coordinate files (Krissinel and Henrick 2007; Krissinel 2010; Krissinel 2011).

The stability of a macromolecular complex depends upon several physicochemical properties including the free energy of formation, interfacial surface area, chemical interactions such as hydrogen bonds, salt bridges and hydrophobic interactions between the interfacing

components. Such interactions are also likely to be present in crystal packings. Therefore a careful analysis of the crystal interfaces and the prediction of higher order structures may actually predict the actual composition (correct stoichiometry/quaternary structure) of the macromolecular complexes. PISA takes into account all the above mentioned physicochemical properties listed above during the analysis of a given structure and makes prediction of the possible stable complexes. It is important to realize that PISA results are predictions based on the calculations of various physico-chemical properties of the molecular structure and any conclusions drawn from them must be backed with additional biological and experimental data. In our case we describe how PISA along with other experimental data (Self-rotation function and size exclusion chromatography) provided additional support to the notion that the quaternary structure of MtFBPase is a tetramer.

5.3 RESULTS AND DISCUSSION

A. Manuscript 2: Crystallization and preliminary X-ray characterization of the *glpX* encoded Class II Fructose 1,6-bisphosphatase from *Mycobacterium tuberculosis*.

1. **Crystallization of *Mt*FBPase II in the native form and as a complex with substrate (F1,6BP) and product (F6P) respectively:** The native *Mt*FBPase crystals were obtained by mixing 1 µl protein solution with 1 µl reservoir solution consisting of 1.8 M ammonium citrate tribasic pH 7.0 (Index screen from Hampton Research). As depicted in Figure 32, the crystallization was reproducible and easily scalable with increasing drop volumes (2 µl, 4 µl, 6 µl).

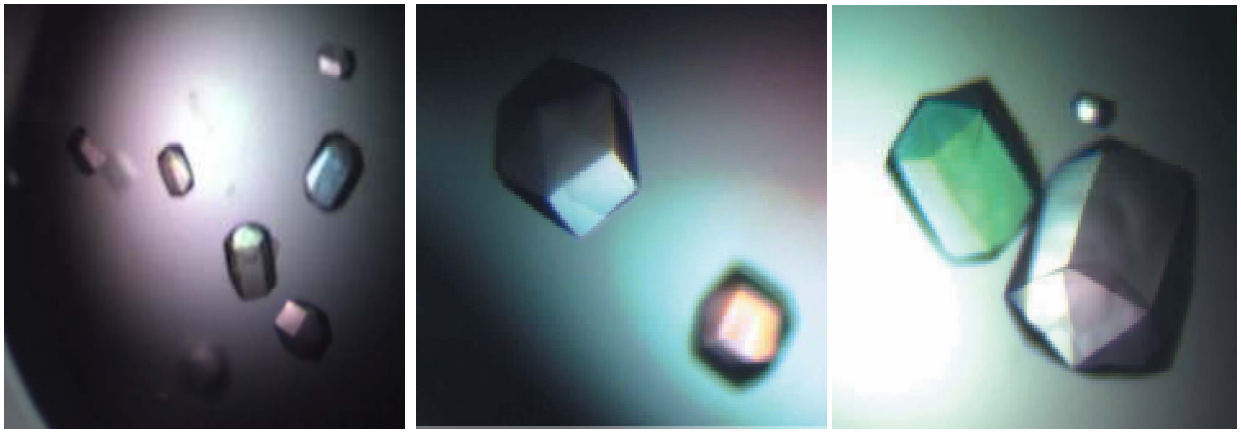


Figure 32. Native *Mt*FBPase crystals: (a) Hexagonal crystals of *Mtb* FBPase obtained at pH 7.0. (b, c) Larger crystals of Mtb FBPase obtained under similar conditions with increased drop volumes (4 µl and 6 µl). Approximate dimensions of the crystals (2a) 45 x 45 x 50 um, 2b. 110 x 110 x 130 um, 2c. 210 x 210 x 230 um.

The *Mt*FBPase complex structures with substrate F1,6BP and product F6P were obtained by subtle variations in the crystallization conditions as follows. *Mt*FBPase- F1,6BP crystals (Figure 33) were obtained by mixing 100 µl protein solution + 10 µl 10mM substrate solution (F1,6BP), incubating for 30 min at 30°C and then mixing 1 µl of the resultant combination with 1 µl reservoir solution consisting of 2.9 M Sodium Malonate pH 4.0 (Malonate grid from Hampton Research).

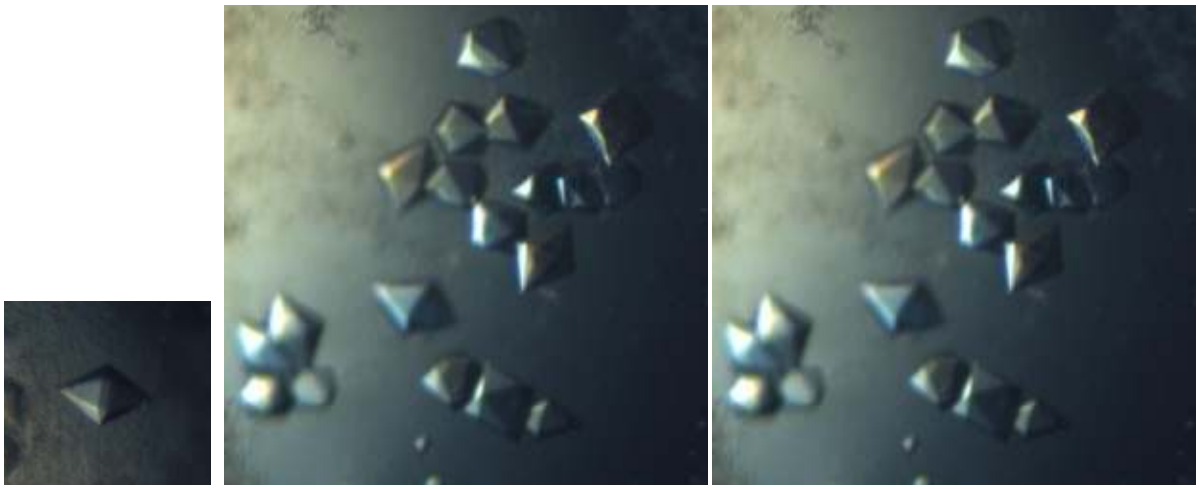


Figure 33. *Mt*FBPase – F1,6BP complex crystals. Hexagonal bipyramidal crystals of *Mt*FBPase complex with F1,6BP. Inset, single hexagonal bipyramidal crystal. Approximate dimensions of the crystals 25 x 27 x 52 μm

*Mt*FBPase- F6P crystals (Figure 34) were obtained by mixing 100 μl protein solution + 10 μl 10mM substrate solution (F1,6BP), incubating for 30 min at 30°C and then mixing 1 μl of the resultant solution with 1 μl reservoir solution consisting of 2.9M Sodium Malonate pH 7.0 (Malonate grid from Hampton Research). The protein solution contained magnesium chloride at a concentration of 1 mM, no extra magnesium solution was added to the protein precipitant mix. The pH of the precipitant solution plays an important role in trapping the substrate and product co-crystal structures. The enzyme being virtually inactive at pH 4.0 (Gutka, Rukseree et al. 2011), we could obtain the co-crystal structure of *Mt*FBPase-F1,6BP. The co-crystal structure of *Mt*FBPase-F6P was obtained at pH 7.0 since the enzyme is active in the pH range of 7- 9 and capable of catalyzing the breakdown of F1,6BP to F6P (Gutka, Franzblau et al. 2011).

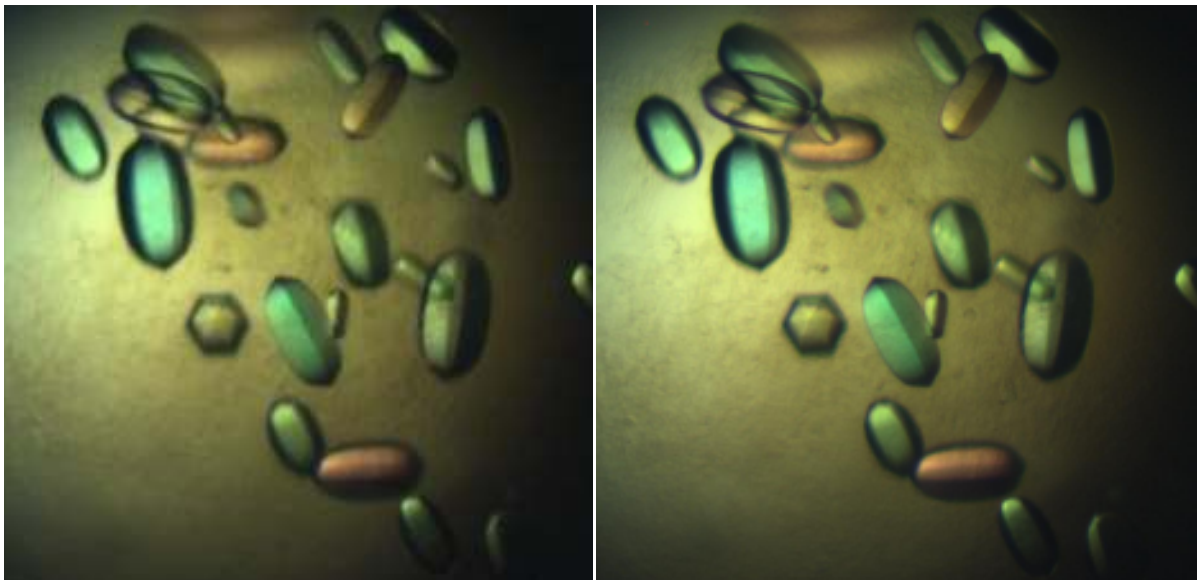


Figure 34. *Mt*FBPase–F6P complex: Soft edge prismatic crystals of *Mt*FBPase-F6P complex. Approximate dimensions of the crystals 22 x 19 x 57 µm.

2. Data Collection, Processing and Structural Refinement Statistics for different crystal forms:

2.1 Data Collection and Processing: The native *Mt*FBPase crystals were soaked in a cryoprotectant consisting of the reservoir solution (1.8 M ammonium citrate tribasic pH 7.0) plus 20% (v/v) glycerol. *Mt*FBPase-F6P crystals were soaked in a cryoprotectant consisting of the reservoir solution (2.9 M sodium malonate pH 7.0) plus 20% (v/v) glycerol. *Mt*FBPase-F1,6BP crystals were soaked in a cryoprotectant consisting of the reservoir solution (2.9 M sodium malonate pH 4.0) plus 20% (v/v) glycerol.

All the crystals were flash-frozen in liquid nitrogen at 100 K. X-ray diffraction data for all crystal forms were collected from a single frozen crystal using a MAR 300 CCD detector on the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory, Illinois, USA. 180 frames were collected with 1° scan width and 1 s exposure time for all crystal forms. The diffraction data were integrated, reduced, processed and scaled with HKL2000 (Otwinowski and Minor 1997).

Symmetry and systematically absent reflections suggested that the native *Mt*FBPase crystals belonged to space group $P6_122$ (or its enantiomer $P6_522$), with unit-cell parameters a = b = 131.3, c = 143.2 Å, even though the weak reflections in the higher resolution shells had a high R_{merge} (Table 2) which could indicate a lower symmetry space group. Initial cross-rotation function ($P6_1$ or $P6_5$) was calculated using the program Phaser from the CCP4 suite (Winn, Ballard et al. 2011).

Determination of the Matthews coefficient suggested the presence of 50.3% solvent content in the unit cell (V_M = 2.47 $Å^3$ /Da) with a dimer in the asymmetric unit of the higher symmetry space groups ($P6_122$ or $P6_522$) (Matthews 1968). The data-collection statistics for native *Mt*FBPase crystals are presented in Table XIII. Figure 35 depicts a representative diffraction pattern for native *Mt*FBPase crystals.

Synchrotron-radiation source	22-ID, SER-CAT, APS
Wavelength (Å)	1.0000
Space group	$P6_122$
Unit-cell parameters (Å)	a= b=131.3, c = 143.2.
Total No. of reflections	212726
No. of unique reflections	20561
Redundancy (overall)	10.4 (7.2)
Resolution (Å)	(40-2.7) (2.75-2.70)[a]
Completeness (%)	98.5 (88.1)
R_{merge}[b]	0.092 (0.686)
I/σ(I) overall	5.3 (2.0)
Mosaicity	0.43
Wilson plot (B-overall, $Å^2$)	71.5

[a] In parenthesis the resolution and values for the last shell as indicated.
[b] $R_{merge} = \Sigma hkl\ \Sigma i\ |Ii(hkl) - \langle I(hkl)\rangle| / \Sigma hkl\ \Sigma i\ Ii(hkl)$

TABLE XIII. DATA COLLECTION AND PROCESSING STATISTICS.

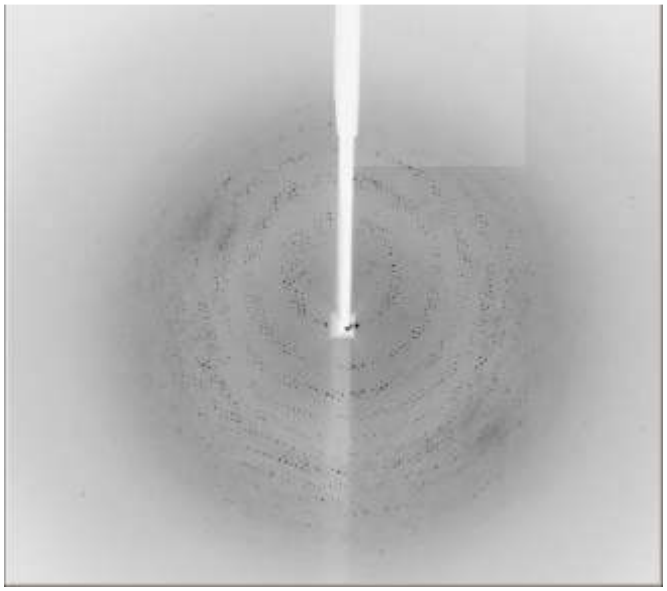


Figure 35: Representative diffraction pattern for native *Mt*FBPase crystal.

2.2 Structure Solution and Refinement

As described in materials and methods, the structure solution for all crystal forms was obtained using the molecular-replacement method using the program MOLREP (Vagin and Teplyakov 1997; Vagin and Teplyakov 2010). Table XIV provides a summary of the different crystal forms and their crystallographic parameters.

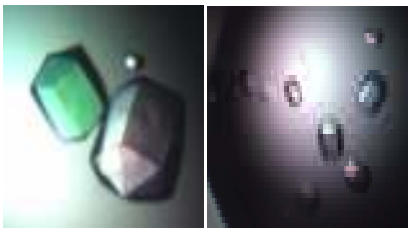
Crystal Identity	Image(s), Form, morphology and crystallization conditions	Resolution	Space group	Cell constants
Native *Mt*FBPase	 Hexagonal 1.8 M ammonium citrate pH 7.0	2.6 Å	$P6_122$	a= b = 131.3, c = 143.2 Å
*Mt*FBPase-F6P complex	 Hexagonal 2.9 M Sodium malonate pH 7.0	2.3 Å	$P6_122$	a= b = 130.2, c = 139.9 Å
*Mt*FBPase-F 1,6BP complex	 Hexagonal bipyramidal 2.9 M Sodium malonate pH 4.0	3.7 Å	$P6_122$	a=b= 114.2, c=327.1 Å

TABLE XIV. COMPILATION OF THE CRYSTALLIZATION AND CRYSTALLOGRAPHIC PARAMETERS FOR THE DIFFERENT CRYSTAL FORMS OF *MT*FBPASE.

For native *Mt*FBPase structure, the initial possibility of the space group being $P6_1$ could not be excluded. Thus, an initial search was undertaken using the Phaser program (McCoy 2007; McCoy, Grosse-Kunstleve et al. 2007) from the CCP4 suite (Winn, Ballard et al. 2011) using the *E. coli* GlpX monomer as a probe (PDB entry 3d1r) (Brown, Singer et al. 2009) in space group $P6_1$ and looking for four monomers in the asymmetric unit. No complete solution was found; however, a reproducible dimer was recognized among the partial solutions. This dimer differed from that suggested previously as a functional unit for the *E. coli* enzyme (Brown, Singer et al. 2009) and was used as a search model in space group $P6_122$. A solution was found with MOLREP as implemented in the CCP4 suite (Winn et al., 2011). The crystallographic statistics for $P6_122$ with one dimer in the asymmetric unit were an R factor of 0.504 and a correlation coefficient of 0.574; those for $P6_522$ were an R factor of 0.632 and a correlation coefficient of 0.308, indicating $P6_122$ to be the correct solution.

This preliminary model allowed the construction of a revised native *Mt*FBPase model with the correct sequence and initial refinement (using REFMAC as implemented in the Coot software) proceeded normally (R_{free} = 0.390, $R_{\text{value (working + test)}}$ = 0.332 and $R_{\text{value (working)}}$ = 0.329). COOT (Emsley and Cowtan 2004; Emsley, Lohkamp et al. 2010), was used for manually rebuilding the model utilizing the correct amino-acid sequence, particularly the sequence between Arg^{230}-Leu^{249}. Water molecules were added during successive rounds of refinement using the COOT find water options. *Mt*FBPase complex structures (with F6P and F1,6P respectively) were solved using the refined native structure coordinates as the search model. The solution clearly showed two protein molecules in the asymmetric unit. Difference electron-density maps for the *Mt*FBPase F1,6BP and F6P complex(s) clearly indicated the presence of the respective ligand in each subunit.

In case of the *Mt*FBPase-F6P complex, F6P was fit in the corresponding electron density map using find ligand option in Coot. The corresponding coordinate for F6P was obtained from the SMILES library in Coot. Iterative refinement and manual rebuilding cycles finally lead to low fraction of Ramachandran outliers as identified in COOT. Final refinement paramaters are presented in Table XV. The models were evaluated for quality using the *MolProbity*, the corresponding Ramachandran plots and statistics are described in Figures 36 (A. B). The structural coordinates and structure factors for both structures shall be deposited in PDB in due course.

Parameters	Native *Mt*FBPase	*Mt*FBPase-F6P complex
Protein details	Met1Ala, Residues 1- 304 chain length, covering all the catalytic residues	Met1Ala, Residues 1- 304 chain length, covering all the catalytic residues
Synchrotron-radiation source	22-ID, SER-CAT, APS	
Wavelength (Å)	1.0000	
Space group	$P6_122$	$P6_122$
Resolution limits (Å)	40 – 2.6 Å	40 – 2.3 Å
Unit-cell parameters (Å)	a = b = 131.3, c = 143.2.	a=b= 130.2, c= 139.9
Number of reflections used	19378	26588
R value (work + test)	0.211	0.211
R value (work)	0.207	0.208
Free R value	0.286	0.266
RMS Bond length (Å)	0.012	0.012
RMS Bond angle (degrees)	1.91	1.89
RMS ChirVol (cubic Ang)	0.127	0.151
Ramachandran statistics: Preferred regions: Allowed regions: Outliers:	 567 (93.3 %) 41 (6.7 %) 0 (0 %)	 574 (94.4 %) 34 (4.13%) 0 (0 %)
Model composition: Amino acids	 608	 608
Water/solvent molecules	87 (HOH)	152 (HOH)
Number of ligands	0	2 (F6P), 4 (Glycerol)
Metal ions	6 (Magnesium)	6 (Magnesium)
Number of protein chains in asymmetric unit	2	2

TABLE XV. STRUCTURAL REFINEMENT STATISTICS FOR THE NATIVE *MT*FBPASE and *MT*FBPASE-F6P COMPLEX CRYSTALS.

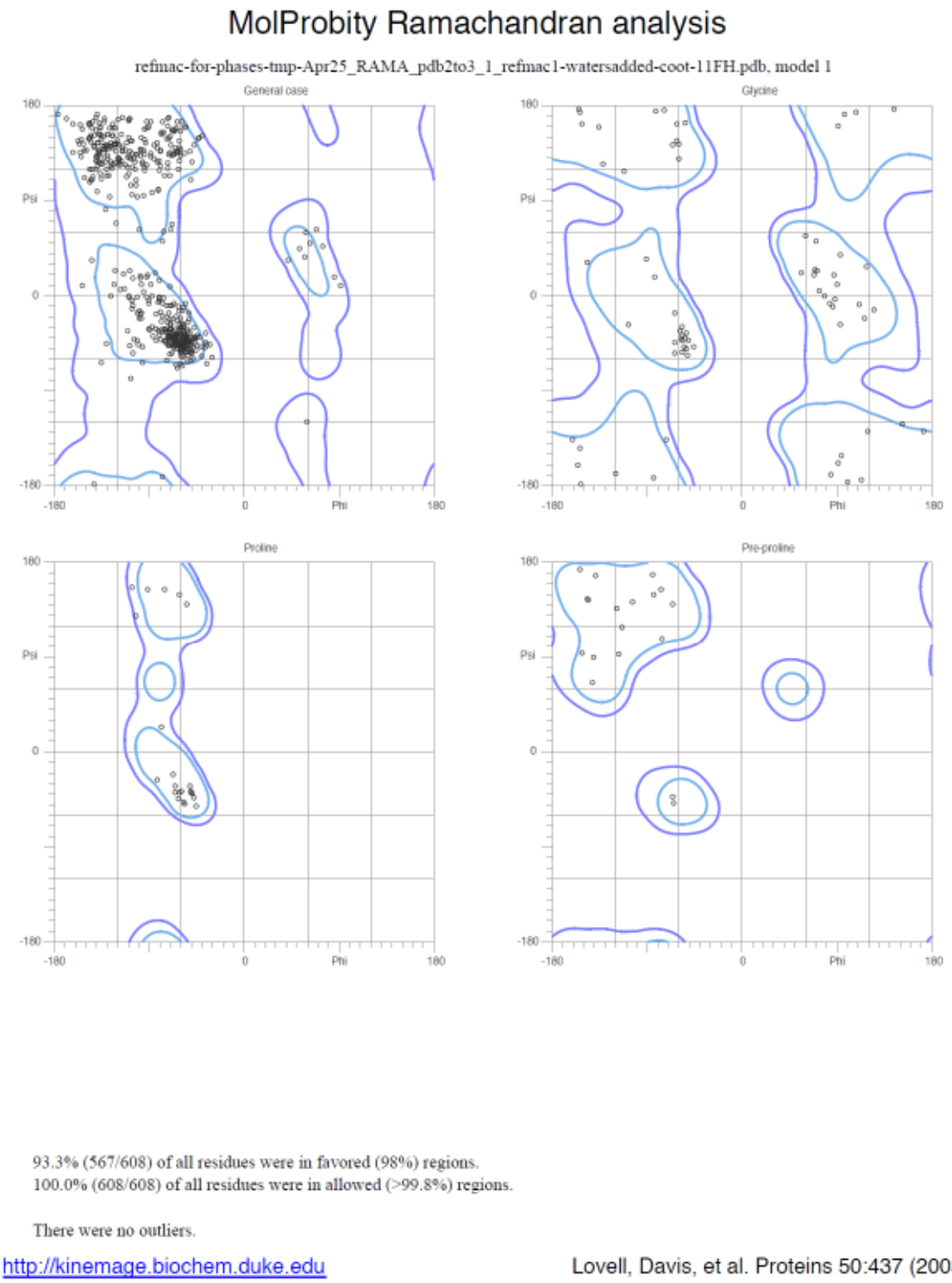


Figure 36. Ramachandran plot analysis as generated by *MolProbity* A. Native *Mt*FBPase structure

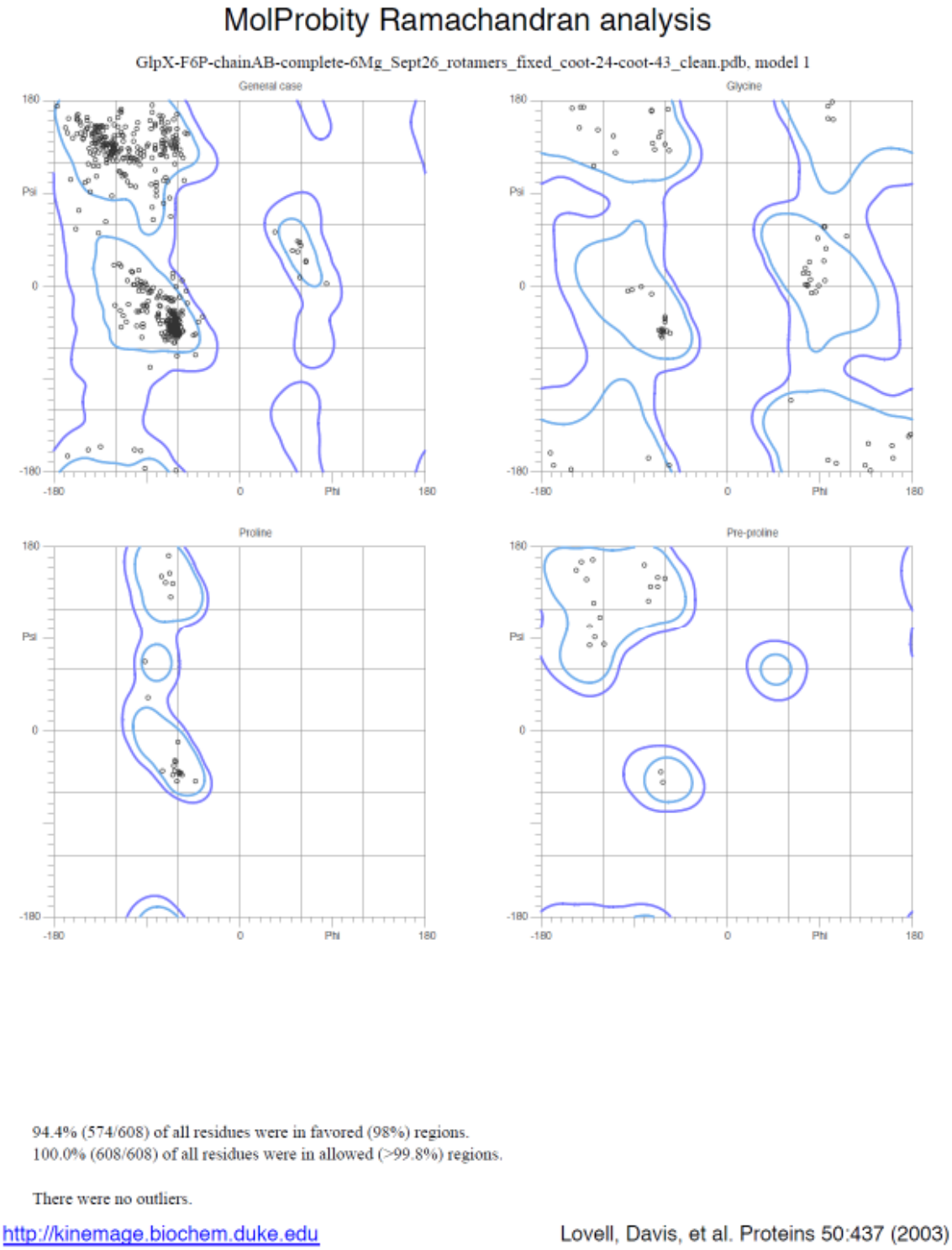


Figure 36. Ramachandran plot analysis as generated by *MolProbity*. B. *Mt*FBPase-F6P complex structure.

B. Structural Analysis:

The structure solution and refinement of the two crystal forms of *Mt*FBPase (native and F6P complex) characterized here provide a three-dimensional framework to analyze the quaternary structure, the monomeric fold and the active site interactions. Our finding that the quaternary structure of *Mt*FBPase is a tetramer is novel and never proposed before in case of *E. coli* GlpX. Furthermore the *Mt*FBPase-F6P complex structure provides a unique snapshot into the catalytic mechanism and adds on to the information available for previously characterized complex structures of *E. coli* GlpX with F1,6BP and phosphate respectively (PDB id: 3big, 3bih).

1. Quaternary Structure of *Mt*FBPase: *Mt*FBPase exists as a tetramer

The crystal structures of native *Mt*FBPase and *Mt*FBPase-F6P complex were solved by molecular replacement (MR), to 2.6 and 2.3 Å resolution respectively in the space group $P6_122$ (Table XV). The native *Mt*FBPase crystal structure (Figure 37) had two FBPase in the asymmetric unit (a.u) forming a dimer. This dimer (Figure 37) is clearly different from the one described in the *E.coli* GlpX structure (PDB id: 3d1r). The *E. coli* FPBase II dimer contacts are very different. The dimer is elongated but does not explain the anomalous MW by SEC misconstrued to be an elongated dimer since it had an anomalously high native molecular weight (118 kDa as estimated by SEC, monomer molecular weight about 36kDa). The dimer in the asymmetric unit (au.) and the crystallographic two fold create a tetramer of *Mt*FBPase, which also includes the dimer analogous to the one described for *E.coli* GlpX. This finding supports our initial finding that *Mt*FBPase exists as an oligomer (most likely tetramer) based on the molecular weight (127 kDa) estimation using size exclusion chromatography. The Self Rotation Function (SRF) (Rossmann and Blow 1962) provides additional evidence of what is present in the crystal. It is the crystallographic symmetry that creates a tetramer with molecular weight in the range of

127kDa as suggested by SEC. Additionally the PISA analysis suggests the quaternary structure of *Mt*FBPase to be tetrameric based on the contact surfaces among the different monomers.

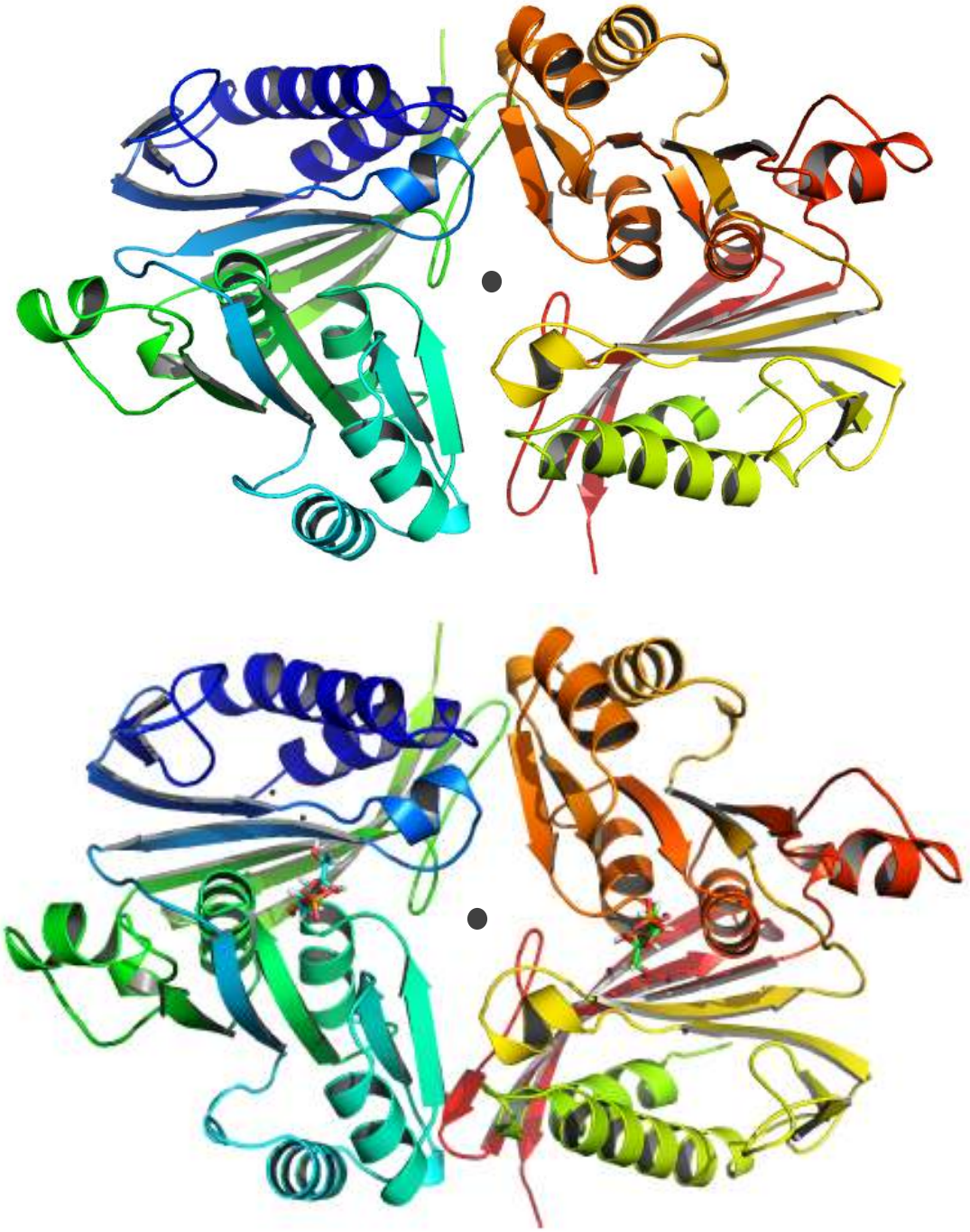


Figure 37. Dimer structures of MtFBPase as found in the asymmetric unit. A. Native *Mt*FBPase dimer; B. *Mt*FBPase –F6P complex dimer. The symmetry axis is perpendicular to the plane of paper in between the two subunits (grey dots).

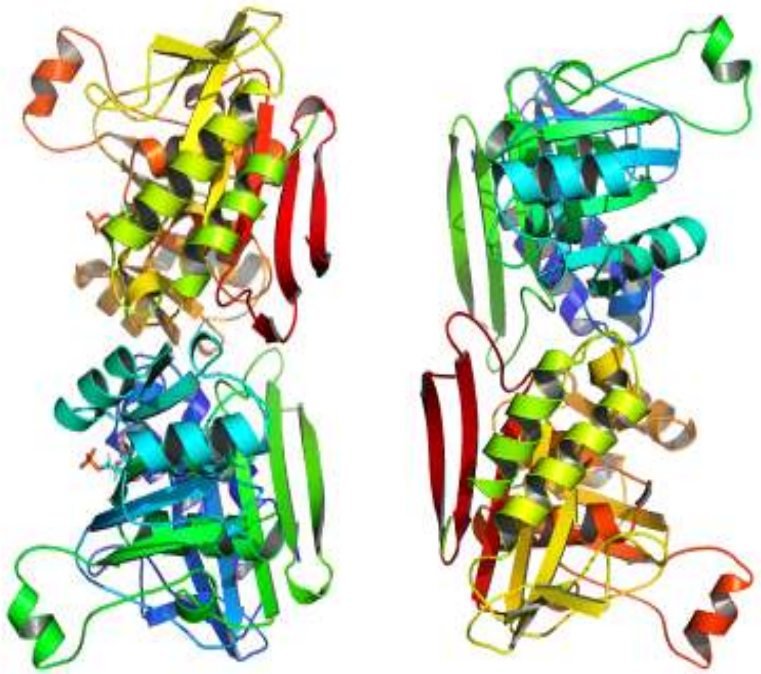


Figure 37 (Continued) Dimer structures of MtFBPase as found in the asymmetric unit. C. *Mt*FBPase –F6P complex (left) and native *Mt*FBPase (right) dimer structure compassion**.** F6P in the MtFBPase –F6P complex is represented as a stick model.

SRF for both native *Mt*FBPase and *Mt*FBPase-F6P complex models was calculated. The F6P complex SRF plots show sharp maxima at 35° (P-direction) and 55°(R-direction) in addition to the 2 fold crystallographic symmetry (Q-direction) in the x-y plane. The PQR directions mark the three orthogonal 2-fold axes (note 35°+55°=90°) (Figure 38). This observation substantiates the hypothesis (based on SEC and crystallographic symmetry) of *Mt*FBPase existing as a tetrameric unit with 222 symmetry. Figure 39 (A,B), depict the SRF for both structures at 4 standard chi values (60°, 90°, 120°, 180°). Figure 40 depicts the PDBe PISA (Protein Interfaces, Surfaces and Assemblies) output for *Mt*FBPase quaternary structure. PDBePISA is an interactive tool for the exploration of macromolecular interfaces, prediction of probable quaternary structures (assemblies). The highlighted red box includes the composition and the dissociation pattern and 4 F6P for A_2B_2 unit (tetramer). This data further supports the hypothesis of *Mt*FBPase existing as a tetramer as found in the crystal lattice, consistent with the molecular

weight determination by SEC. Additionally *C. glutamicum* FBPase has an estimated molecular weight of 140 kDa (tetramer with a monomeric mass of 35kDa) indicating its tetrameric nature. Also based on the ClustalW alignment of primary sequences for *M. smegmatis, F. tularensis* its highly likely that they too would exist as tetramer.

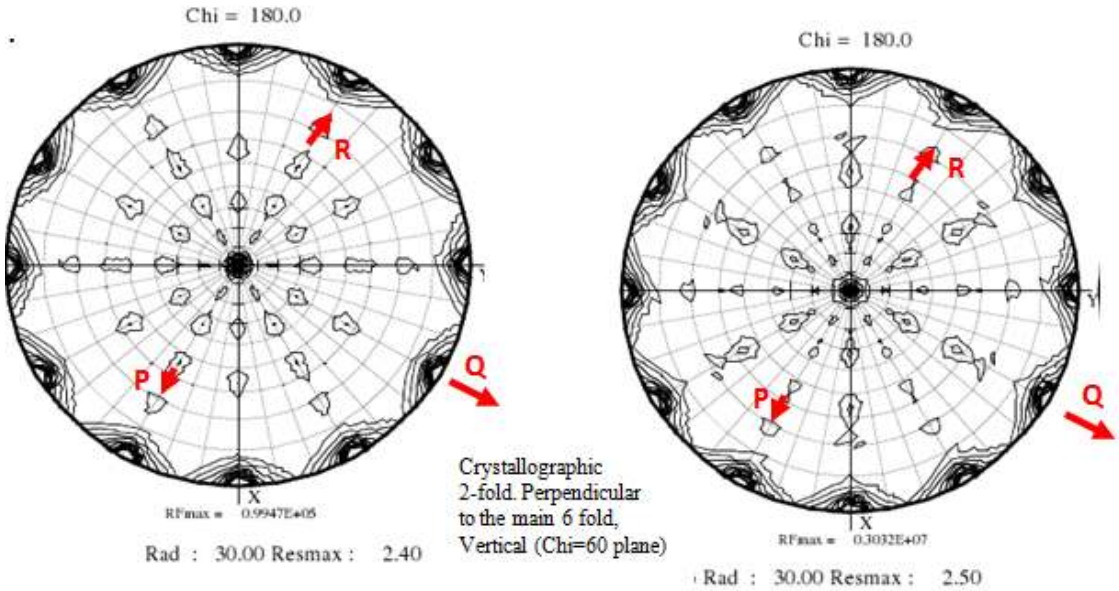


Figure 38. Self rotation function (SRF) plots for native *Mt*FBPase and F6P complex. 38A. SRF for native model with Resmax = 2.40 Å; **38B.** SRF for F6P bound complex with Resmax = 2.50Å. The F6P complex SRF plots for Chi = 180°, show sharp peaks/intensities at 35°(P) and 55°(R) in addition to the 2 fold crystallographic symmetry (Q-direction) as indicated by the red arrow heads. The angular distance around the P-Q and Q-R directions can be measured to be 90°.

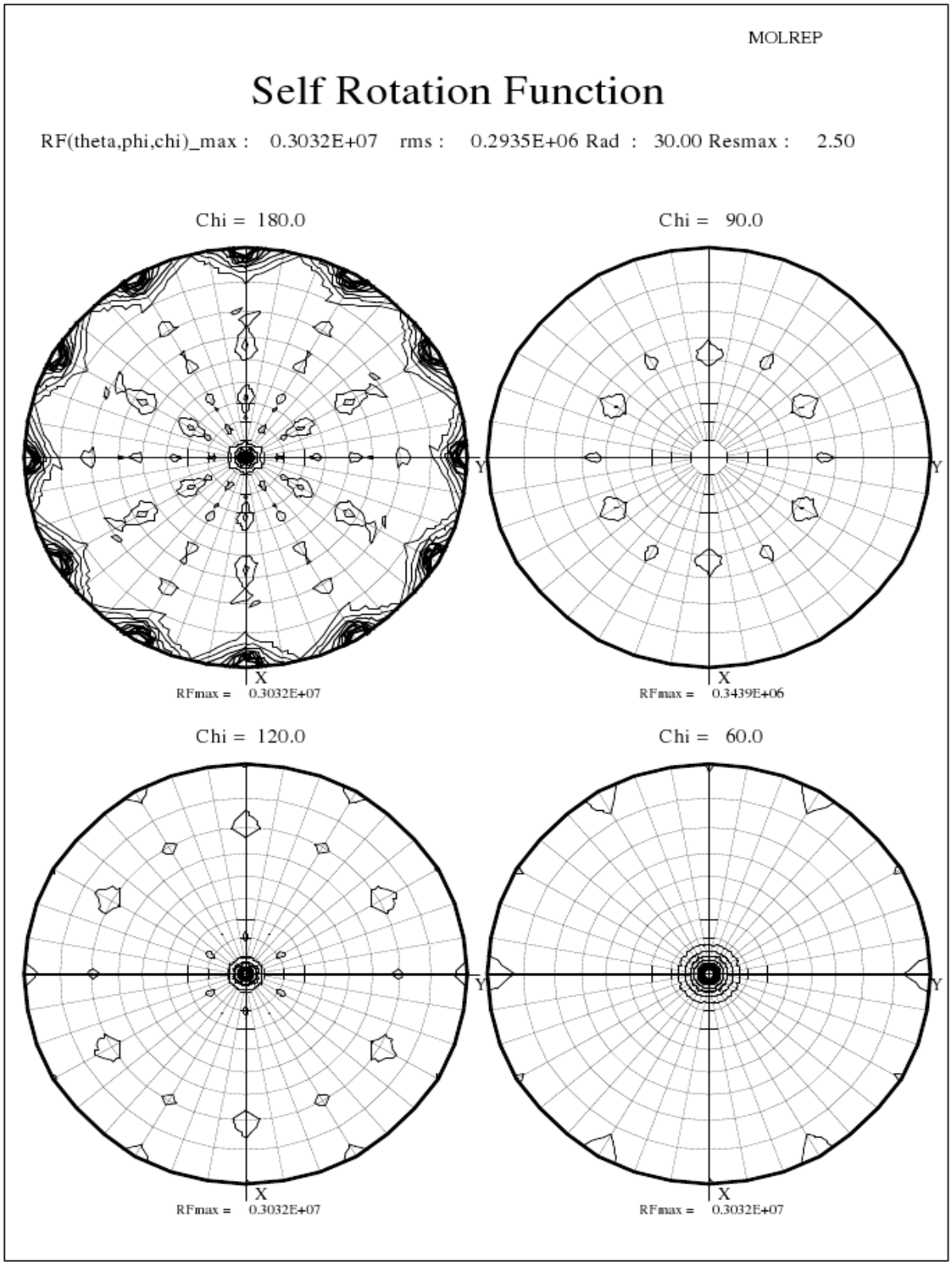


Figure 39. Self Rotation Function. A. *Mt*FBPase–F6P complex, Rad: 30.00 Å, Resmax: 2.50 Å

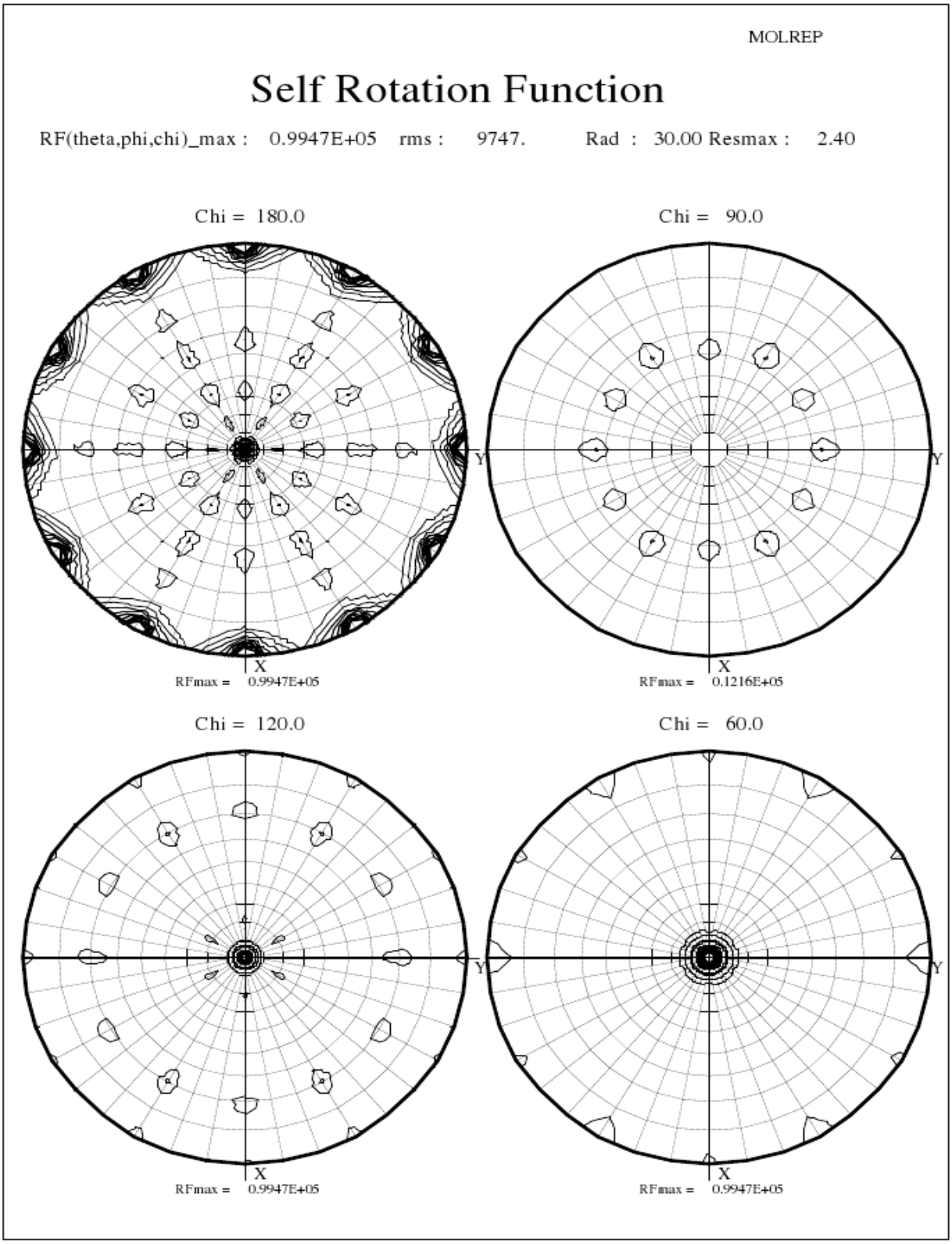


Figure 39. Self Rotation Function (continued). B. Native *Mt*FBPase; Rad: 30.00 Å, Resmax: 2.40 Å

Assembly Summary XML

Multimeric state	4	Surface area, Å²	46212.2	ΔG^int, kcal/mol	0.9	TΔS^diss, kcal/mol	41.9
Copies in unit cell	6	Buried area, Å²	10487.2	ΔG^diss, kcal/mol	-0.5	Symmetry number	4
Formula	A4a4						
Composition	A2B2[FBP]4						
Dissociation pattern	2(A[FBP]) + 2(B[FBP])						

View | View Dissociated | Download Assembly | Remark 350 | Viewer: Jmol

Engaged interfaces XML

Id	##		Interfacing structures	Nocc	Diss.	Sym.ID	Buried area, Å²	ΔiG, kcal/mol	NHB	NSB	NDS	CSS
1	1	◉	B + A	2	×	1_555	1348.4 (26%)	-0.2 (52%)	24 (43%)	2 (22%)	0	0.043
2	2	○	B + A	2	×	10_664	814.3 (16%)	-1.6 (344%)	15 (27%)	7 (78%)	0	0.036
7	7	○	[FBP]A:3499 + A	2		1_555	223.9 (4%)	1.4 (295%)	9 (16%)	0 (0%)	0	0.100
	8	○	[FBP]B:3499 + B	2		1_555	223.4 (4%)	1.2 (251%)	8 (14%)	0 (0%)	0	0.100
						Average:	223.7 (4%)	1.3 (273%)	9 (16%)	0 (0%)	0	0.100
11	12	○	B + B	1	×	10_664	16.8 (0%)	-0.2 (26%)	0 (0%)	0 (0%)	0	0.001
	13	○	A + A	1	×	10_664	6.8 (0%)	-0.2 (24%)	0 (0%)	0 (0%)	0	0.001
						Average:	11.8 (0%)	-0.2 (25%)	0 (0%)	0 (0%)	0	0.001

View | Details

Assembled monomers XML

##		Range	Surface area, Å²	Buried area, Å²	ΔG^int kcal/mol	NHB	NSB	Symmetry operation	Sym.ID	Transformation matrix Rotation	Translation
1	◉	A	13715.1	2393.4	0.67	48	9	x,y,z	1_555	1.000 0.000 -0.000 0.000 1.000 0.000 0.000 0.000 1.000	0.000 0.000 0.000
2	○	[FBP]A:3499	385.4	223.9	0.24	9	0	x,y,z	1_555	1.000 0.000 -0.000 0.000 1.000 0.000 0.000 0.000 1.000	0.000 0.000 0.000
3	○	B	13864.1	2402.9	-0.59	47	9	x,y,z	1_555	1.000 0.000 -0.000 0.000 1.000 0.000 0.000 0.000 1.000	0.000 0.000 0.000
4	○	[FBP]B:3499	385.1	223.4	0.14	8	0	x,y,z	1_555	1.000 0.000 -0.000 0.000 1.000 0.000 0.000 0.000 1.000	0.000 0.000 0.000
5	○	A	13715.1	2393.4	0.67	48	9	-y+1,-x+1,-z-1/6	10_664	0.500 -0.866 0.000 -0.866 -0.500 -0.000 0.000 0.000 -1.000	65.075 112.712 -23.308
6	○	[FBP]A:3499	385.4	223.9	0.24	9	0	-y+1,-x+1,-z-1/6	10_664	0.500 -0.866 0.000 -0.866 -0.500 -0.000 0.000 0.000 -1.000	65.075 112.712 -23.308
7	○	B	13864.1	2402.9	-0.59	47	9	-y+1,-x+1,-z-1/6	10_664	0.500 -0.866 0.000 -0.866 -0.500 -0.000 0.000 0.000 -1.000	65.075 112.712 -23.308
8	○	[FBP]B:3499	385.1	223.4	0.14	8	0	-y+1,-x+1,-z-1/6	10_664	0.500 -0.866 0.000 -0.866 -0.500 -0.000 0.000 0.000 -1.000	65.075 112.712 -23.308

Figure 40: PDBe PISA (Protein Interfaces, Surfaces and Assemblies) output for the tetrameric assembly of *Mt*FBPase. The highlighted red box includes the composition and the dissociation pattern and 4 F6P for A_2B_2 unit (tetramer) (Krissinel and Henrick 2007; Krissinel 2010; Krissinel 2011).

Analysis of the packing and subunit interactions (PDBe-PISA) (Figure 40), and re-evaluation of the SEC data for molecular weight estimation, provides strong evidence that *Mt*FBPase is a functional tetramer (Figures 41. 42 a-c). The structures of the PDB entries for *E.coli* GlpX (1ni9, 3bih, 3big, 3d1r), all contain only a monomer of 36 KDa in the asymmetric unit. The suggestion of an elongated crystallographic dimer (upper part of the tetramer depicted

in panel B of Figure 41) was proposed as a functional unit (Figure 42d.) to explain the anomalously large molecular weight (117 kDa) observed in *E. coli*. Our findings suggest an alternative explanation for this observation and rather demonstrate that like *Mt*FBPase, *E. coli* GlpX as well exists as a functional tetramer. A corresponding model was also generated using the crystallographic symmetry and the results show that like in the *Mt*FBPase crystal, *E. coli* GlpX also forms a 222 oligomer and exists as a tetramer (Figure 41B.).

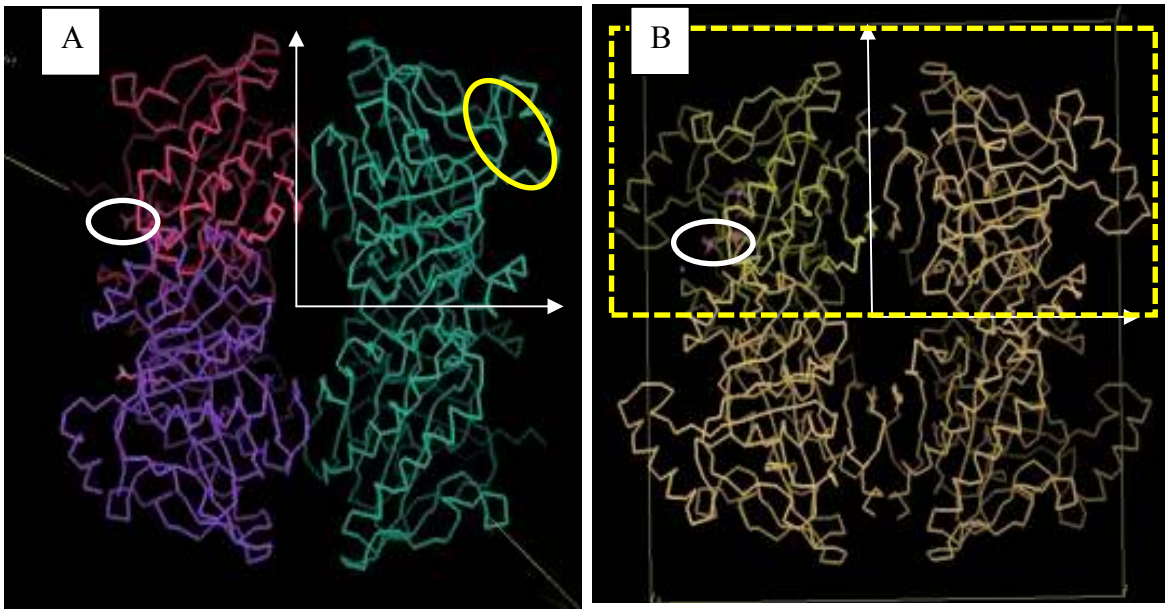
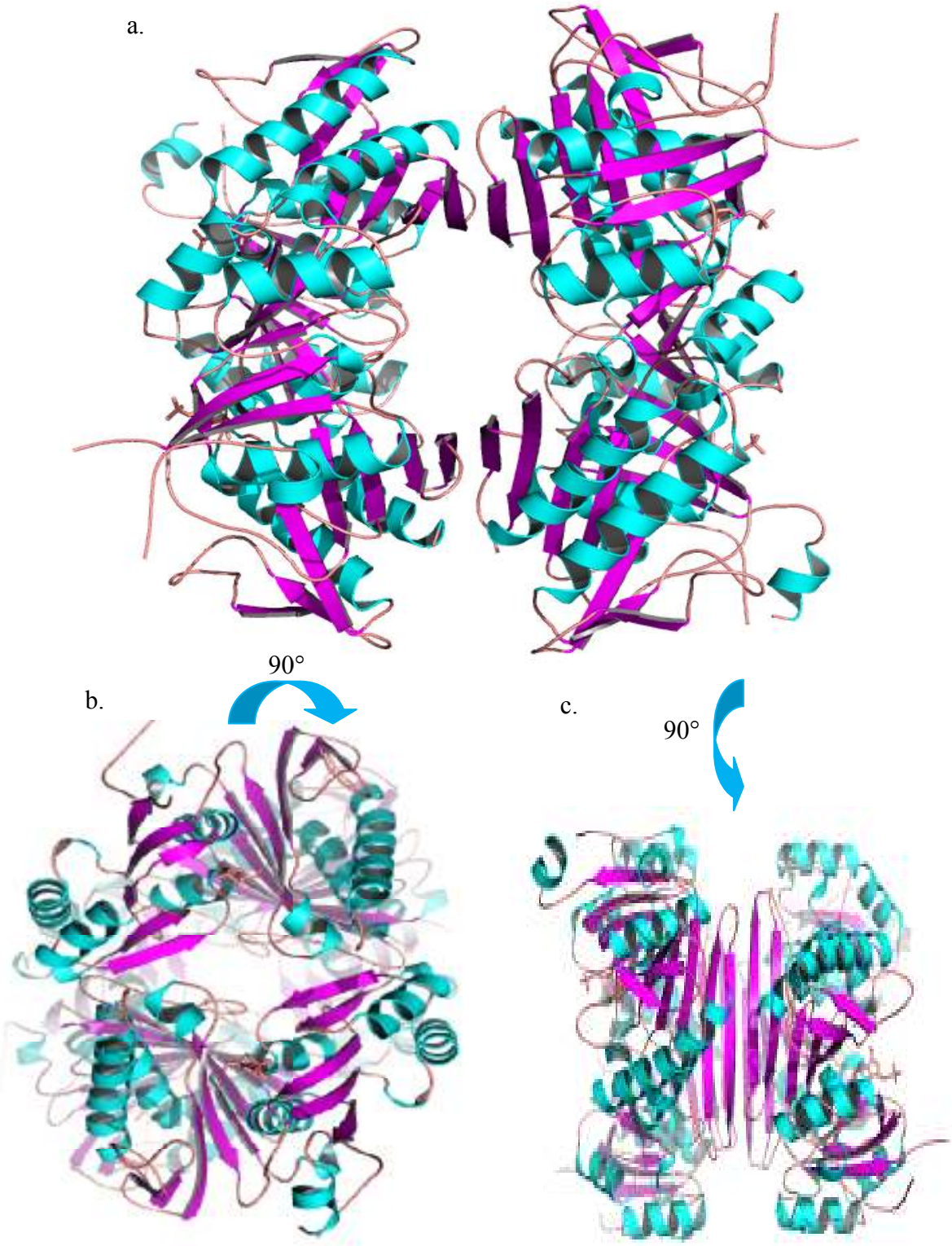


Figure 41. Comparison of tetrameric assemblies of *Mt*FBPase and *E. coli* GlpX. 41A. *Mt*FBPase-F6P in P6$_1$22 cell: White Oval marks the position of F6P. The outermost helix (yellow) on the symmetry equivalent B units is not very well ordered in the four chains of *Mtb* FBPase when compared with *E.coli.* The yellow lines are from axes of the unit cell (COOT program). **41B. *E. coli* GlpX tetramer in P422 cell:** PDB id: 3d1r. Oval marks the position of F1,6DP. The outermost helix is very well ordered. The dashed rectangle marks the elongated dimer described and proposed for the functional unit of *E. coli* GlpX. Arrows mark the direction of two ortogonal dyads forming a 222 tetramer (third is perpendicular to image towards the viewer).


a.
90°
b.
c.
90°

d.

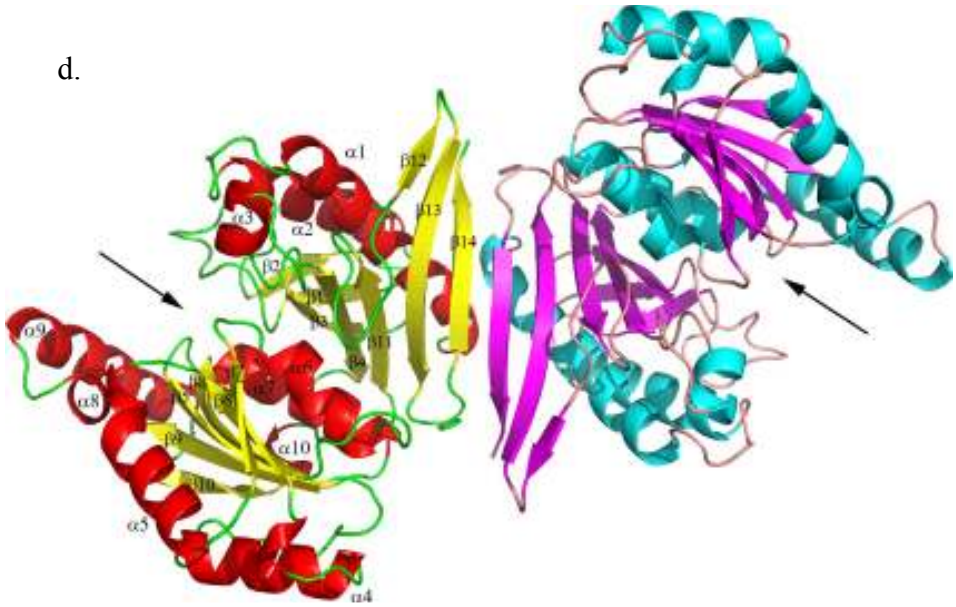
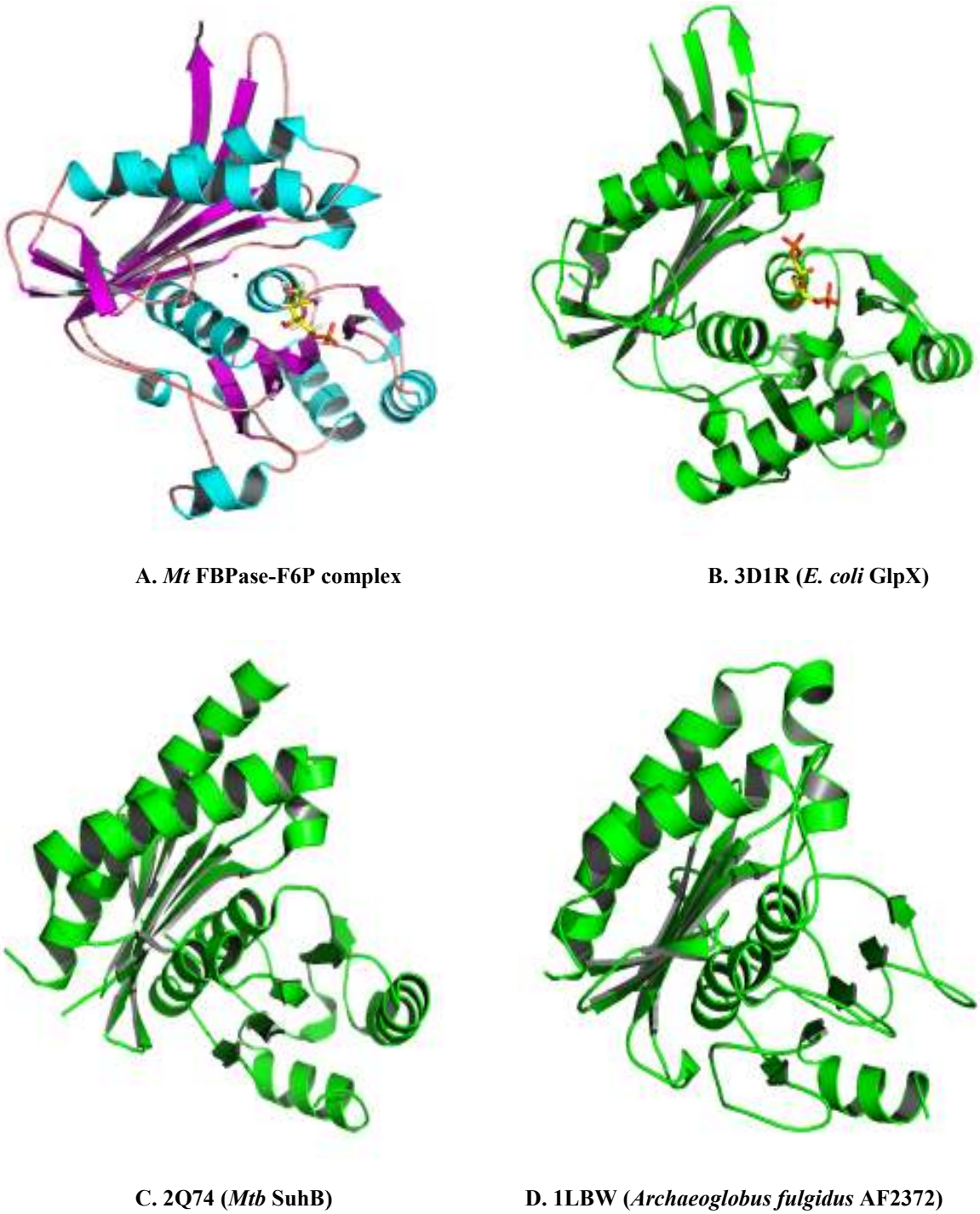


Figure 42. *Mtb*FBPase-F6P complex tetramer assembly. A. B. and C. Different views of the *Mtb*FBPase-F6P complex tetramer assembly: The model is built before each of the residues in the 242 – 260 α9- helix region could be modeled into the structure. The secondary structures of *Mt*FBPase structures have been represented in cyan for α helix, magenta for β strand and pink for loops. F6-P in the active site is represented as a stick model; **D. Dimer for the *E. coli* GlpX:** Adapted with permission from Brown *et al.* 2009. The crystallographic dimer contact depicted in Figure 39B corresponds to the one described for *E. coli* GlpX (Figure 40D.) The A-B non-crystallographic dimer observed in the present crystal structure (Refer PISA output in Figure 40) is different. The presence of both the contacts in the crystal structure implies 222 symmetry for the tetrameric A_2B_2 aggregate (see Figure 41B. for the corresponding tetramer in the *E. coli* GlpX structure).

2. Overview of the monomer structural fold (native *Mt*FBPase and *Mt*FBPase-F6P complex): conservation in the lithium sensitive phosphatase superfamily

The *Mt*FBPase monomer is a compact, globular potein with two distinct α/β-type domains arranged in alternating multilayer fashion of α/β/α/β/α fold characteristic for the superfamily of lithium sensitive phosphatases (Figure 43). Secondary structure analysis of the *Mt*FBPase monomer reveals that there are 12 α helices and 13 β-strands connected by loops of varying lengths. These secondary structures can be grouped together into 2 distinct domains (domains A and B respectively). The two-layer domain A (also referred to as the hot dog-like configuration in the *E. coli* GlpX structure) (top half of Figure 43a.) includes the sequences from the N-terminal (Ala^{1}–Arg^{104}) and C-terminal parts (Leu^{272}–Arg^{304}) of *Mt*FBPase and is formed by the antiparallel β-sheet with seven β-strands, which is partially wrapped around two long α-helices. The three-layered domain B (also referred to as α/β/α-sandwich configuration in the *E. coli* GlpX structure) is formed by the central part of the *Mt*FBPase sequence (Gly^{105}–Asp^{271}). Both these domains are joined by two loops between β4-β5 (12 residues) and β11-β12 (10 residues) regions respectively. The active site of *Mt*FBPase is accommodated between the α/β region of the domain B (α/β/α-sandwich) and comprises of residues highly conserved among GlpX encoded FBPases from different sources (*C. glutamicum, E. coli* and *F. tularensis*).

The two FBPase subunits in the a.u. have closely interacting residues at the interface with the residues of αH3 interacting with those of β7. Also the αH6 residues interact with loop residues between strands β11 and β12. These dimer interface interactions are different from the β-β strand interactions between the two subunits as observed for *E coli* GlpX (Brown, Singer et al. 2009). Such β-β strand interactions are also observed in the overall *Mt*FBPase tetramer (Figure 42c.).



A. *Mt* FBPase-F6P complex

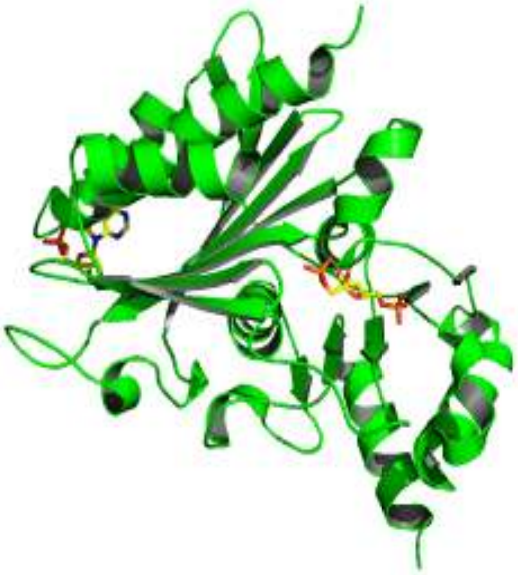
B. 3D1R (*E. coli* GlpX)

C. 2Q74 (*Mtb* SuhB)

D. 1LBW (*Archaeoglobus fulgidus* AF2372)

Figure 43: *Mt*FBPase monomer α-β-α-β-α structural fold similar to that observed with other known phosphatases belonging to the superfamily of lithium sensitive phosphatases. A. *Mt*FBPase-F6P complex; B. *E. coli* GlpX (PDB id 3D1R); C. *Mtb* SuhB (PDB id 2Q74); D. *Archaeoglobus fulgidus* AF2372 (PDB id 1LBW).

E. 1DK4 (*Methanoccocus jannaschii* MJ0109)



F. 2Q8ME (*E coli* FBPase)

G. 2QFL (*E. coli* SuhB)

H. 3B8B (*Bacteroides thetaiotaomicron* CysQ)

Figure 43 (Continued): E. *Methanoccocus jannaschii* MJ0109 (PDB id 1DK4); F. *E coli* FBPase (PDB id 2Q8ME); G. *E. coli* SuhB (PDB id 2QFL); H. *Bacteroides thetaiotaomicron* CysQ (PDB id 3B8B). All the proteins are shown in their respective monomeric forms in orientations showing the characteristic α-β-α-β-α structure. The α helices and β strands for *Mt*FBPase-F6P complex structure are shown in cyan and magenta color respectively. The ligands present if any in all structures have been modeled as stick.

The monomeric structure is very similar to the *E coli* GlpX structure used as a search model for molecular replacement except for a marked difference in the αH9 region of the parent structure. *Mt*FBPase lacks about 13 residues which constitute a part of the long αH9 and the preceding 5 residue long loop in the *E coli* GlpX structure. As a result, the corresponding αH9 in the *Mt*FBPase structure is much shorter and the overall loop-H9-loop region (Arg^{230} – Tyr^{243}) does not protrude as far away from the manifold structure (Figures 44 A. B.). Such an analogous loop region (also referred to as the mobile catalytic loop) has been implicated in determining lithium sensitivity of dual activity IMPase/FBPase. MJ0109 and AF2372 also referred to as lithium insensitive IMPase (refer Table XVI), have a distinct mobile loop structure different from several other IMPases/FBPase. This structural differentiation appears to be useful in classifying IMPase-like proteins as lithium sensitive or insensitive respectively. The size and orientation of the mobile catalytic loop is a key determinant of lithium sensitivity for the said proteins (Stieglitz, Johnson et al. 2002; Li, Stieglitz et al. 2010).

Enzyme	IC_{50} *mM*	Reference
*Mt*FBPase	0.2	(Gutka, Rukseree et al. 2011)
C. Glutamicum FBPase	0.14	(Rittmann, Schaffer et al. 2003)
E. coli GlpX	70	(Brown, Singer et al. 2009)
E. coli YggF	15.8	(Brown, Singer et al. 2009)
AF2372	290	(Stieglitz, Johnson et al. 2002)
MJ0109	250	(Chen and Roberts 1998)
Pig kidney FBPase	1–3	(Zhang, Villeret et al. 1996)
Human IMPase	0.3	(McAllister, Whiting et al. 1992)

TABLE XVI. SUMMARY OF FBPASE/IMPASE FAMILY ENZYMES INHIBITION (IC_{50}) BY LITHIUM.

Based on the primary sequence comparison with other analogues (*C. glutamicum*, *Mt*FBPase, *F. tularensis glpX*), *E. coli* GlpX has a longer loop-αH9-loop region, it remains to be

verified experimentally (in a manner similar to that verified for MJ109) whether this structural difference is a possible reason behind lithium resistance of *E. coli* GlpX (IC_{50} approximately 70 mM). In case of MJ109, a series of mobile catalytic loop mutations were performed (Asp38Ala, Asp26Ala, Glu39Asp, and Glu41Ala) and the corresponding effects on Mg^{2+} affinity and lithium sensitivity evaluated. While lithium ion inhibition of the MJ109 is very poor (IC_{50} approximately 250 mM), the Asp38Ala enzyme has a dramatically enhanced sensitivity to lithium with an IC_{50} of 12 mM (Li, Stieglitz et al. 2010). Like MJ109, *E. coli* GlpX too has several negatively charged residues (Asp^{214}, Glu^{243}, Glu^{244}, Glu^{250} and Glu^{252}), reiterating the need for a systematic mutational study on the α9H and evaluate the corresponding effect on Mg^{2+} affinity and lithium sensitivity. A general question arises as to whether why Li^+ but not other monovalent metal ions like Na^+ or K^+ are inhibitors of IMPase/FBPase. In order to perform the catalytic function and promote the nucleophile formation, the metal ion needs to insert into the active site (indicated by arrow in Figure 44C). The radius for Li^+ ion (0.68 Å) is close enough to that for Mg^{2+} ion (0.66 Å) and hence Li^+ ion can substitute structurally for the Mg^{2+} ion. However, because of its monovalent charge Li^+ ion cannot create a nucleophile and exerts inhibitory effects on the enzyme's catalytic function. Other monovalent ions such as Na^+ or K^+ ions (radius 0.97 Å and 1.33 Å, respectively) would be unable to substitute structurally for Mg^{2+} due to their larger size. Based on the findings, we can hypothesize that a longer αH9 in *E. coli* GlpX, facilitates closer interaction with the catalytic active site and also probably the substrate/product (regulate the entry and binding of bivalent metal ions in the active site) (Figure 44B) and based on the above discussion/analogy can be implicated as a probable reason for the resistance to lithium.

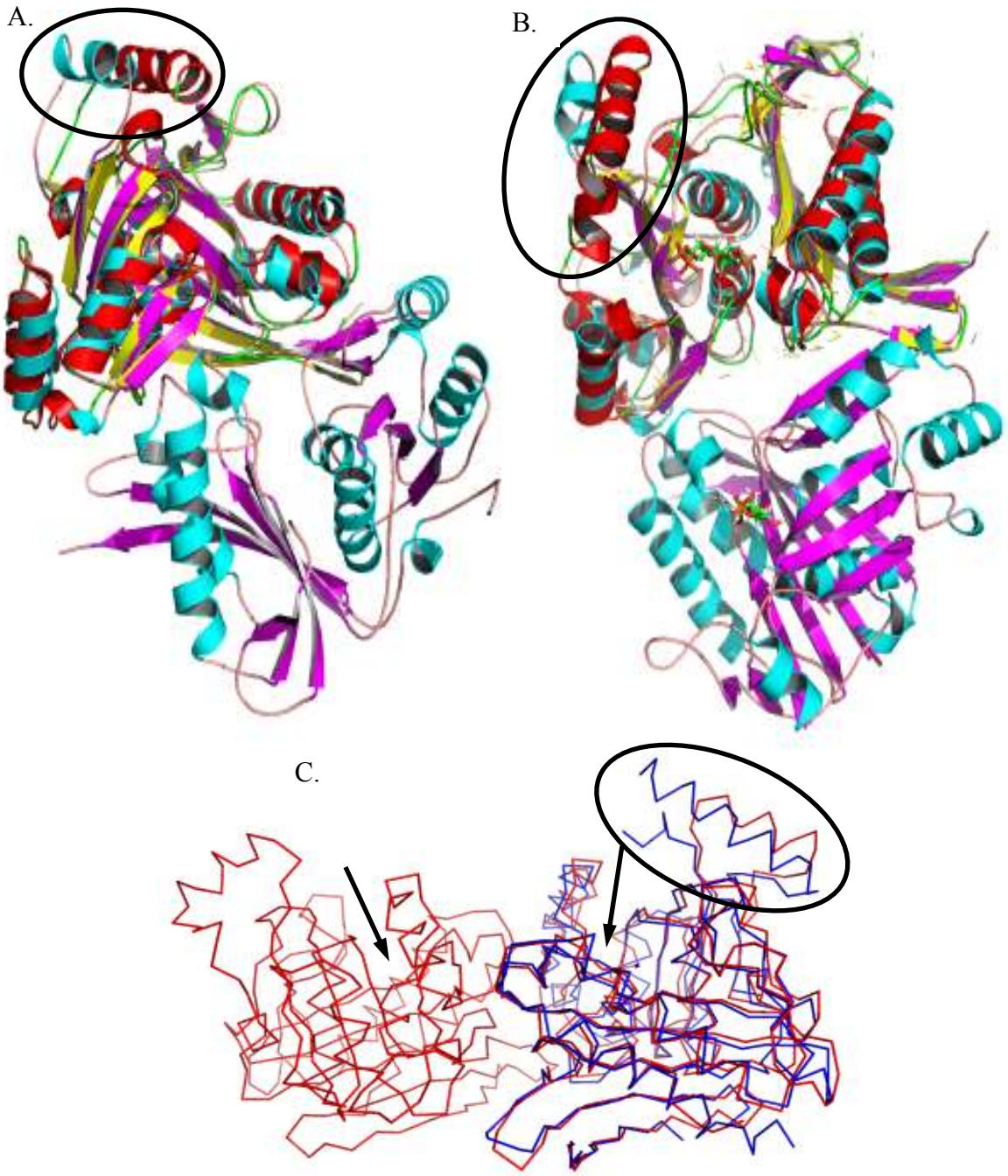


Figure 44. *Mt*FBPase structural overlay with monomer of *E. coli* GlpX. A. Native *Mt*FBPase with *E. coli* GlpX (PDB id 3bih); B. *Mt*FBPase-F6P complex *E. coli* GlpX-F1,6BP(PDB id 3d1r); The secondary structures of *Mt*FBPase structures have been represented in cyan for α helix, magenta for β strand and pink for loops and that for *E coli* GlpX have been represented in red for α helix, yellow for β strand and green for loops. F1,6BP and F6P have been represented as stick model;. **C. Native *Mt*FBPase structural overlay with monomer of *E. coli* GlpX native structure (PDB id 3bih);** both *Mt*FBPase and GlpX are represented as C-α models. Arrow points towards the active site. Area of marked difference between the two structures is encircled (α9 helix region). The dimer proposed for *E. coli* GlpX has a different conformation than that observed for *Mt*FBPase.

Sequence annotated by Structure (SAS), predicted secondary structural features for *Mt*FBPase.

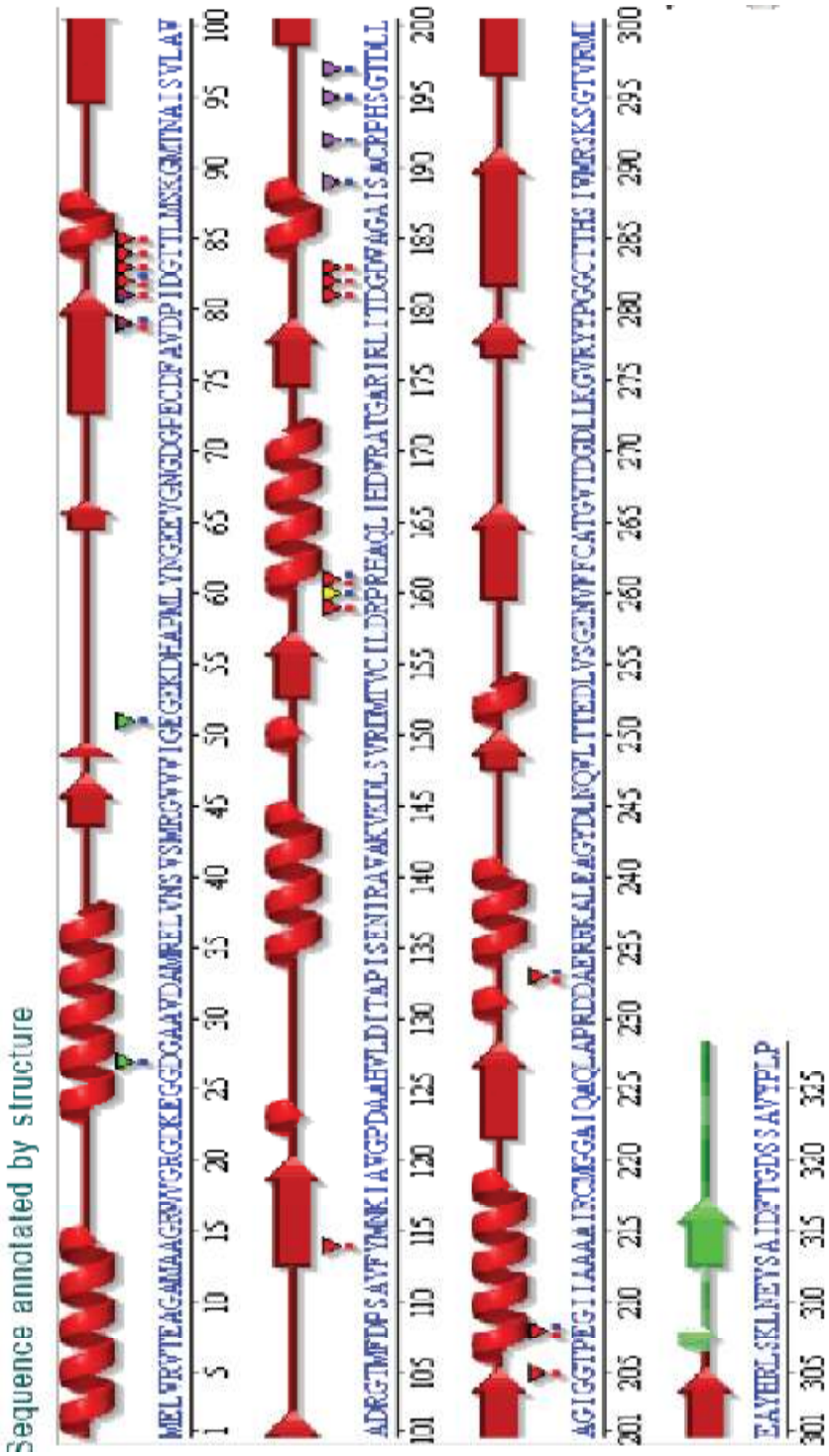


Figure 45. Sequence comparison of *Mt*FBPase protein with other class II FBPases A. Sequence annotated by Structure (SAS), predicted secondary structural features for *Mt*FBPase. Sequence features with homology to *E. coli* GlpX are in red, the sequence portion in green is the predicted structure. Standard annotations have been made for α helix and β strands.

```
                 1         2         3         4         5         6
        123456789012345678901234567890123456789012345678901234567890123456789012345
        ---------+---------+---------+---------+---------+---------+-----
        ------MELVRVTEAGAMAAGRWVGRGDKEGGDGAAVDAMRELVNSVSMRGVVVIGEGEKDHAPM
3big:A  MRRELAIEFSRVTESAALAGYKWLGRGDKNTADGAAVNAMRIMLNQVNIDGTIVIGEGEIAEAPM
3bih:A  MRRELAIEFSRVTESAALAGYKWLGRGDKNTADGAAVNAMRIMLNQVNIDGTIVIGEGEIAEAPM
3d1r:A  MRRELAIEFSRVTESAALAGYKWLGRGDKNTADGAAVNAMRIMLNQVNIDGTIVIGEGEIAEAPM

                                          1         1         1         1
            7         8         9         0         1         2         3
        678901234567890123456789012345678901234567890123456789012345678901234567890
        ----+---------+---------+---------+---------+---------+---------+
        LYNGEEVGNGDGPECDFAVDPIDGTTLMSKGMTNAISVLAVADRGTMFDPSAVFYMNKIAVGPDA
3big:A  LYIGEKVGTGRGDAVDIAVDPIEGTRMTAMGQANALAVLAVGDKGCFLNAPDM-YMEKLIVGPGA
3bih:A  LYIGEKVGTGRGDAVDIAVDPIEGTRMTAMGQANALAVLAVGDKGCFLNAPDM-YMEKLIVGPGA
3d1r:A  LYIGEKVGTGRGDAVDIAVDPIEGTRMTAMGQANALAVLAVGDKGCFLNAPDM-YMEKLIVGPGA

                 1         1         1         1         1         1
                 4         5         6         7         8         9
        123456789012345678901234567890123456789012345678901234567890123456789012345
        AHVLDITAPISENIRAVAKVKDLSVRDMTVCILDRPRHAQLIHDVRATGARIRLITDGDVAGAIS
3big:A  KGTIDLNLPLADNLRNVAAALGKPLSELTVTILAKPRHDAVIAEMQQLGVRVFAIPDGDVAASIL
3bih:A  KGTIDLNLPLADNLRNVAAALGKPLSELTVTILAKPRHDAVIAEMQQLGVRVFAIPDGDVAASIL
3d1r:A  KGTIDLNLPLADNLRNVAAALGKPLSELTVTILAKPRHDAVIAEMQQLGVRVFAIPDGDVAASIL

            2         2         2         2         2         2         2
            0         1         2         3         4         5         6
        678901234567890123456789012345678901234567890123456789012345678901234567890
        ----+---------+---------+---------+-----------+---------+---------+
        ACRPHSGTDLLAGIGGTPEGIIAAAAIRCMGGAIQAQLAPRDDA----ERRKALEA--------G
3big:A  TCMPDSEVDVLYGIGGAPEGVVSAAVIRALDGDMNGRLLARHDVKNE-ENRRIGEQELARCKAMG
3bih:A  TCMPDSEVDVLYGIGGAPEGVVSAAVIRALDGDMNGRLLARHDVKNE-ENRRIGEQELARCKAMG
3d1r:A  TCMPDSEVDVLYGIGGAPEGVVSAAVIRALDGDMNGRLLARHDVKGDNENRRIGEQELARCKAMG
                                                  Short α-helix

                 2         2         2         3         3         3
                 7         8         9         0         1         2
        123456789012345678901234567890123456789012345678901234567890123456789012345
        ---------+---------+---------+---------+---------+---------+-----
        YDLNQVLTTEDLVSGENVFFCATGVTDGDLLKGVRYYPGGCTTHSIVMRSKSGTVRMIEAYHRLS
3big:A  IEAGKVLRLGDMARSDNVIFSATGITKGDLLEGISRKGNIATTETLLIRGKS-TIRRIQSIHYL-
3bih:A  IEAGKVLRLGDMARSDNVIFSATGITKGDLLEGISRKGNIATTETLLIRGKSRTIRRIQSIHYL-
3d1r:A  IEAGKVLRLGDMARSDNVIFSATGITKGDLLEGISRKGNIATTETLLIRGKS-TIRRIQSIHY--

            3         3
            3         4
        678901234567890123456
        ----+---------+------
        KLNEYSAIDFTGDSSAVYPLP
3big:A  ----------------------
3bih:A  ----------------------
3d1r:A  ----------------------
```

Figure 45. Sequence comparison of *Mt*FBPase protein with other class II FBPases B. Sequence Annotated by Structure (SAS) multimple sequence overlay (*Mt*FBPase with the closest homologues in pdb (3big, 3bih, 3d1r; structures of *E. coli* GlpX)). Residues forming α helix are colored in red, those forming β strands in blue and loops are in black.

CLUSTAL 2.0.12 multiple sequence alignment

```
Mtube           --------------MELVRVTEAGAMAAGRWVGRGDKEGGDGAAVDAMRELVNSVSMRGV 46
Msmeg           MTPSRGEAPDRNLALELVRVTEAGAMAAGRWVGRGDKEGGDGAAVDAMRELVNSVSMRGV 60
Cglut           MNLKNPETPDRNLAMELVRVTEAAALASGRWVGRGMKNEGDGAAVDAMRQLINSVTMKGV 60
Ftula           --------MNRKVALEAVRVTELAALASWSQMGRGDKIAADQAAVDAMRKALNEVDIDGT 52
Ecoli           --------MRRELAIEFSRVTESAALAGYKWLGRGDKNTADGAAVNAMRIMLNQVNIDGT 52
                              :*  **** .*:*.   :*** *  .* ***:***  :*.* : *.

Mtube           VVIGEGEKDHAPMLYNGEEVGNGDGPECDFAVDPIDGTTLMSKGMTNAISVLAVADRGTM 106
Msmeg           VVIGEGEKDNAPMLYNGEEVGNGDGPECDFAVDPIDGTTLMSKGMPNAISVLAVAERGAM 120
Cglut           VVIGEGEKDEAPMLYNGEEVGTGFGPEVDIAVDPVDGTTLMAEGRPNAISILAAAERGTM 120
Ftula           VVIGEGELDEAPMLYIGEKVGAG-GCEVDIALDPLEGTTITSKGGANALTVLAMADKGGF 111
Ecoli           IVIGEGEIDEAPMLYIGEKVGTGRGDAVDIAVDPIEGTRMTAMGQANALAVLAVGDKGCF 112
                :****** *.***** **:** * *   *:*:**::** : : * .**:::** .::* :

Mtube           FDPSAVFYMNKIAVG--PDAAHVLDITAPISENIRAVAKVKDLSVRDMTVCILDRPRHAQ 164
Msmeg           FDPSAVFYMNKIAAG--PDVADFIDITSPIAANIQRIAKMRKASVSDITVCILDRPRHAK 178
Cglut           YDPSSVFYMKKIAVG--PEAAGKIDIEAPVAHNINAVAKSKGINPSDVTVVVLDRPRHIE 178
Ftula           LN-APDVYMQKIAVGGINAPKGIVDLDDSVTNNLKRIAEFKGVHMSALVVCTMDRPRHEH 170
Ecoli           LN-APDMYMEKLIVG--PGAKGTIDLNLPLADNLRNVAAALGKPLSELTVTILAKPRHDA 169
                : :. .**:*: .*        :*:  .:: *:. :*         :.*  : :***

Mtube           LIHDVRATGARIRLITDGDVAGAISACRPHSGTDLLAGIGGTPEGIIAAAAIRCMGGAIQ 224
Msmeg           LIADVRAAGARIRLISDGDVAGAISACRPDSGTDLLAGIGGTPEGIITAAAIRCMGGEIQ 238
Cglut           LIADIRRAGAKVRLISDGDVAGAVAAAQDSNSVDIMMGTGGTPEGIITACAMKCMGGEIQ 238
Ftula           IIKEARECGARVILINDGDVSGVIATATENSGIDVYIGTGGAPEGVLAAAALKCLGGQMQ 230
Ecoli           VIAEMQQLGVRVFAIPDGDVAASILTCMPDSEVDVLYGIGGAPEGVVSAAVIRALDGDMN 229
                :* : :  *.::  * ****:. : :.   .  *:  * **:***:::*..::.:.* ::

Mtube           AQLAPRD-------------DAERRKALEAG-YDLNQVLTTEDLVSGENVFFCATGVTDG 270
Msmeg           ATLAPTD-------------DEERQKAIDRG-HDLDRVLTTKDLVSGENVFFCATGVTDG 284
Cglut           GILAPMN-------------DFERQKAHDAG-LVLDQVLHTNDLVSSDNCYFVATGVTNG 284
Ftula           ARLIFND-------------EEEIKRAHRLGITDLNKKYDIDDLASGD-IVFAATGVTDG 276
Ecoli           GRLLARHDVKGDNEENRRIGEQELARCKAMGIEAG-KVLRLGDMARSDNVIFSATGITKG 288
                . *    .  Short α-helix  : *  :.   *     :     *:. .:   * ***:*.*

Mtube           DLLKGVR-----YYPGGCTTHSIVMRSKSGTVRMIEAYHRLSKLNEYSAIDFTGDSSAVY 325
Msmeg           DLLKGVR-----FFGGGCTTQSIVMRSKSGTVRMIEAYHRLSKLNEYSAINFTGDSSAAY 339
Cglut           DMLRGVS-----YRANGATTRSLVMRAKSGTIRHIESVHQLSKLQEYSVVDYTTAT---- 335
Ftula           NMLQGVKRVNSTRRGSYAVTHSVVMRSTTKTVRHITAEHSFDFKEG-IEKFMS------- 328
Ecoli           DLLEGIS-----RKGNIATTETLLIRGKSRTIRRIQSIHYLDRKDPEMQVHIL------- 336
                ::*.*:Insertion  . ..*.::::*..: *:* * : * :.  :

Mtube           PLP 328
Msmeg           PLP 342
Cglut           ---
Ftula           ---
Ecoli           ---
```

Figure 45. Sequence comparison of *Mt*FBPase protein with other class II FBPases. C. CLUSTALW alignment of *Mt*FBPase *(Rv1099c)* as compared to other class II FBPases from *M. smegmatis*, *C. glutamicum* , *F. tularensis* GlpX and *E. coli.* (41–65 % sequence identity) identified over 50 conserved residues. Invariant positions are indicated by asterisks (*) under the alignment, while highly conserved and weakly conserved positions are indicated by colons and periods. Residues in the active catalytic site of *E. coli* GlpX are indicated in red. Importantly, all residues of the active catalytic site are highly conserved in all the FBPases considered for sequence comparison. Characteristic conserved blocks have been boxed. Also, conserved residues among the catalytically important residues have been boxed. The short α-helix (Asp[231] residue onwards) has been indicated in a black box. The *E. coli* sequence was obtained from SWISSPROT (accession no. P28860), the *C. glutamicum* sequence was obtained from GenBank

(accession no. 19552240), the *Mtb* sequence was obtained from the Tuberculist server (http://genolist.pasteur.fr/TubercuList/; gene name Rv1099c), the *glpX* sequence of *F. tularensis* was obtained from KEGG database (FTF1631c) and the *M. smegmatis* sequence was derived from genomic sequences obtained from The Institute for Genomic Research website (http://www.tigr.org/). There is an extra insertion between residues 284 – 288 in *F. tularensis* GlpX protein.

Sequence Annotated by Structure (SAS) overlay of the *Mt*FBPase model and primary sequence with those of the *E. coli* GlpX PDB models, suggest that the helix corresponding to αH9 in *E. coli* GlpX is about 10 residues shorter in *Mt*FBPase (Figure 45 A.B.). A CLUSTAL primary sequence overlay of several known FBPase II (Figure 45C.), shows that the helix corresponding to αH9 in *E. coli* GlpX is about 10 residues shorter in all of the FBPase II described herein (*Mt*FBPase, *M. smegmatis, C. glutamicum, F. tularensis*).

3. Active site/catalytic pocket (conservation of catalytic reaction mechanism)

The electron density map for *Mt*FBPase-F6P complex structure shows strong density corresponding to F6P in the active site. This finding adds to the already available information on structural forms of *E. coli* GlpX.

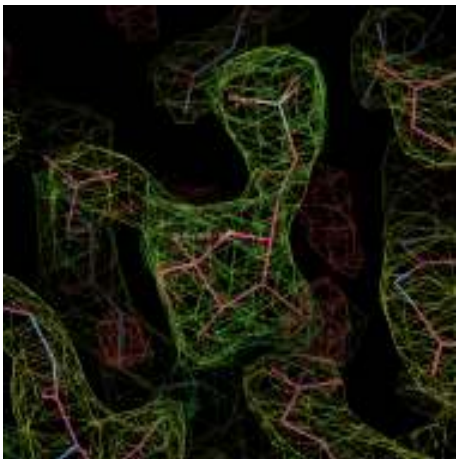


Figure 46. Electron density map corresponding to F6P, as seen in the *Mt*FBPase-F6P complex refined structure model . Image taken using snapshot option in COOT.

Comparing the catalytic residues for *E. coli* GlpX and *Mt*FBPase, we find 100% functional conservation among the two proteins. Asp^{27}, Glu^{51}, Asp^{79}, Asp^{82}, Thr^{84}, Tyr^{113}, Arg^{159}, Arg^{161}, Asp^{183} and Glu^{208} of *Mt*FBPase constitute the catalytic site, these residues correspond to Asp^{33}, Glu^{57}, Asp^{85}, Glu^{88}, Thr^{90}, Tyr^{119}, Lys^{164}, Arg^{166}, Asp^{188} and Glu^{213} in *E coli* GlpX active site. Asp^{82} and Arg^{159} in *Mt*FBPase are not identical to their corresponding counterparts in *E. coli* GlpX (Glu^{88} and Lys^{164}) but are functionally equivalent (Asp is functionally equivalent to Glu and likewise Arg is functionally equivalent to Lys). Overall, based on the high level of sequence identity and functional conservation, it can be expected that the catalytic mechanism for *Mt*FBPase would be similar to that described for GlpX (Brown, Singer et al. 2009) (Figure 47, 48 A,B). Like the parent *E. coli* GlpX structure *Mt*FBPase structures (native and F6P complex) show convincing electron densities for the three metal ions (Mg^{2+}) present in the active site. The atomic distance for metal ion at site 2 (Ca^{2+} site in the GlpX structure) is significantly different from that observed in the parent GlpX structure, with distances between Asp^{27} and Mg^{2+} being about 4.9 Å and Glu^{51} and Mg^{2+} being about 5.6 Å (Figure 48C.). The Mg^{2+} ion is rather close to the Thr^{84} (2.9 Å). This is probably due to the coordination by this metal site with the leaving phosphate group during catalysis of F1,6BP.

Arg^{159}, Arg^{161}, Tyr^{113} form a characteristic triad that interacts with the 6-phosphate group in F6P. This is a characteristic feature among FBPases, not observed in IMPase, suggesting that such an architecture determines substrate specificity to F1,6BP. Additionally a Mg^{2+} ion also coordinates with the 6-phosphate group at this site. The 4-OH group in F6P is recognized by side chain of Asp^{183}, the 3-OH group by side chain of Asp^{82} and the 2-OH group by side chain of Asp^{181}. In case of the *E. coli* GlpX protein, the interaction of the side chains of Thr^{90} (Thr^{84} for *Mt*FBPase) and Asp^{33} (Asp^{27} for *Mt*FBPase) is required for catalysis (Figures 47, 48). This

interaction increases the nucleophilicity of the Thr^{90} hydroxyl oxygen, thereby resulting in the formation of hydroxyl nucleophile. The hydroxyl ion subsequently attacks the phosphorus at 1-phosphate position in F1,6BP, thereby generating a transition structure which is stabilized by the two Mg^{2+} ions in the active site. Eventually the 1-phosphate leaves resulting in the catalytic product F6P. Since the catalytic residues involved in this mechanism and the metal ions at the active site are highly conserved among the two proteins (*Mt*FBPase and GlpX), the *Mt*FBPase can be assumed to follow a similar catalytic mechanism as that already described for *E. coli* GlpX. Our F6P complex structure adds valuable information to the already available structural information of GlpX protein (enzyme in action with its all possible complexes i.e. F6P, F1,6BP and Phosphate).

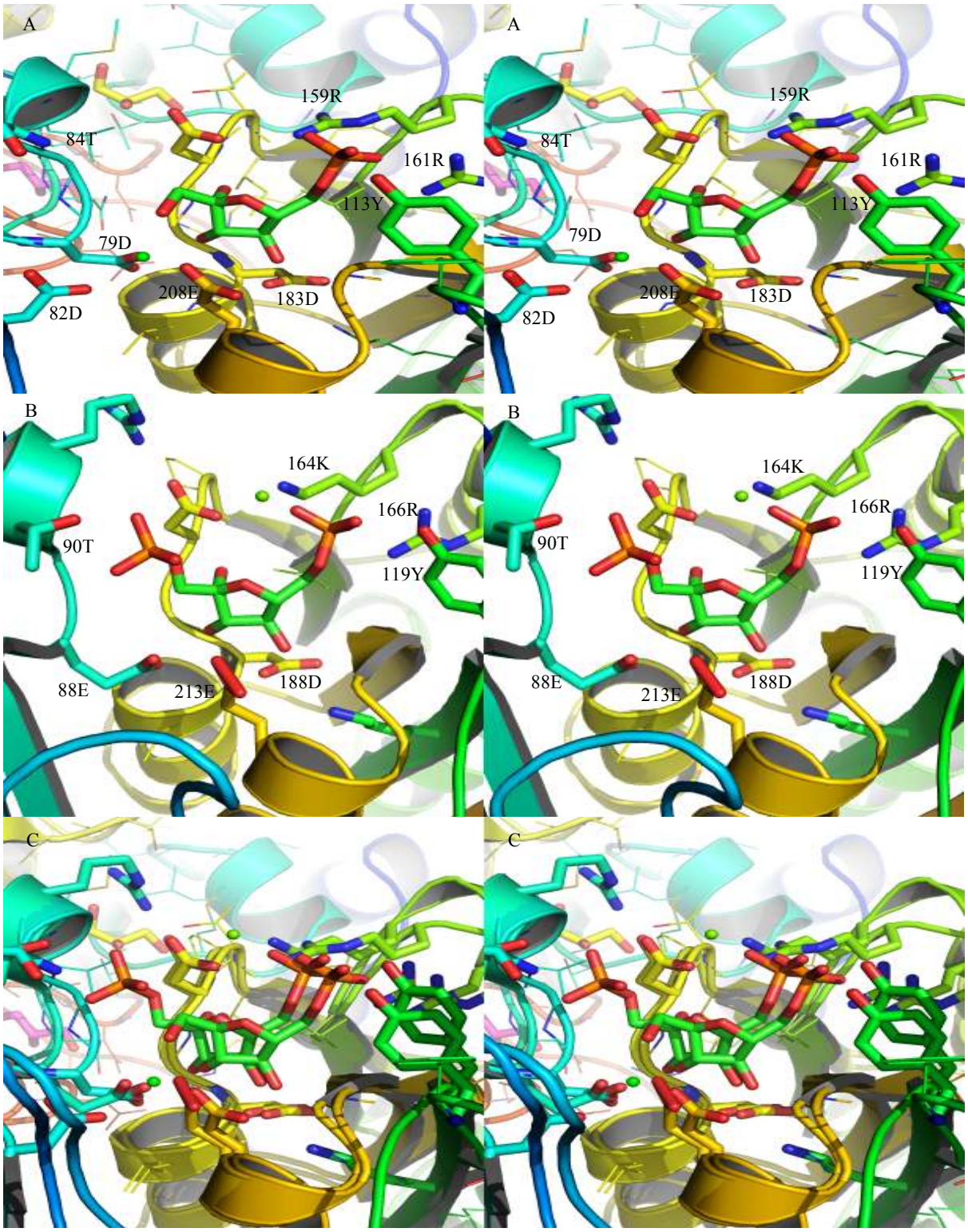


Figure 47. Stereo view of *Mt*FBPase active site. A. F6P bound in active site of *Mt*FBPase; B. F1,6BP bound *E. coli* GlpX active site (PDB id: 3d1r); C. overlay of *Mt*FBPase and GlpX active sites as shown in A and B. The amino acid residues have been annotated in standard 1 letter format.

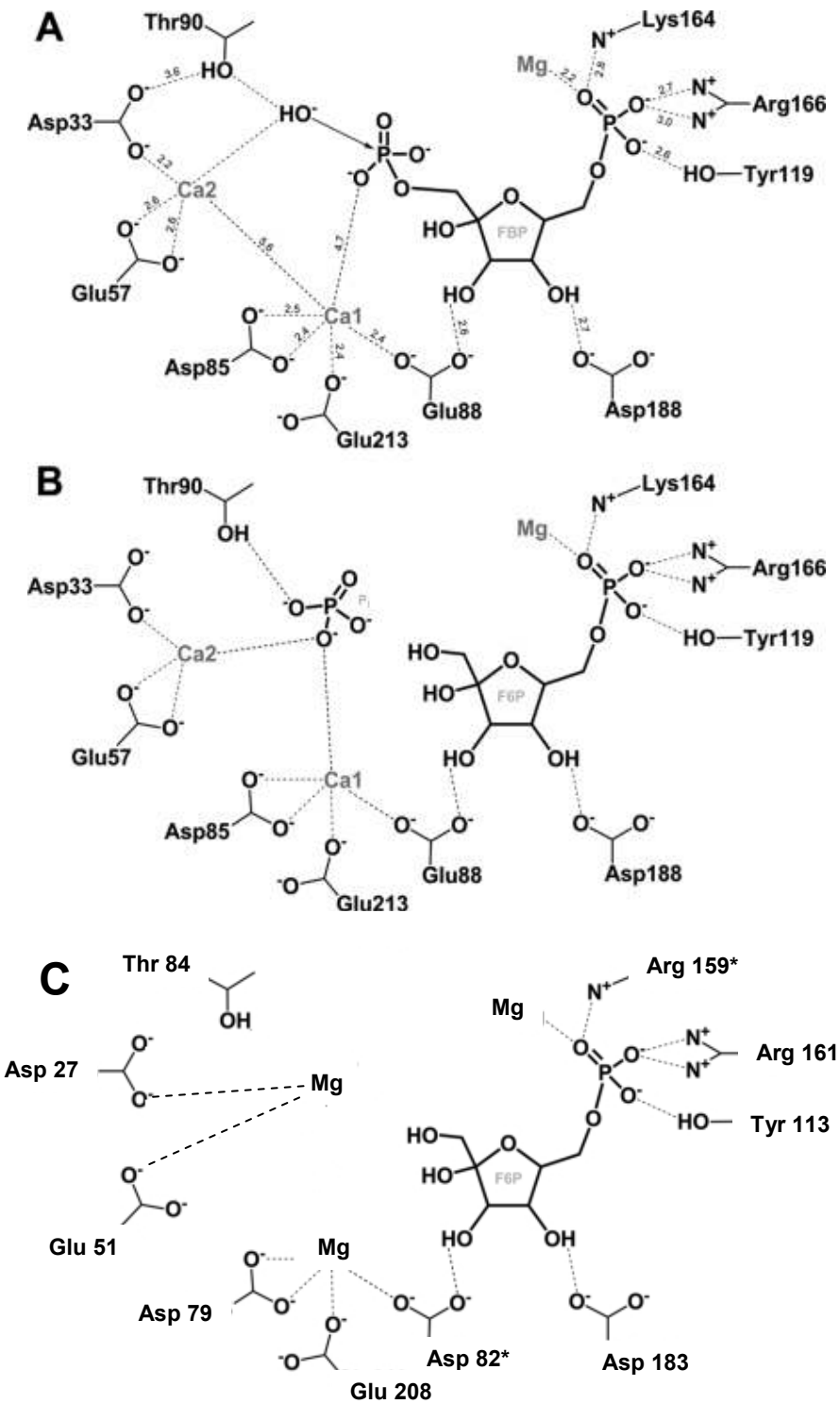


Figure 48. Reaction mechanism of F1,6BP hydrolysis by GlpX. A. and B. diagrams showing the coordination of FBP in the active site of *E. coli* GlpX (Reproduced with permission from Brown et al. 2009). C. Schematic representation of the F6P bound *Mt*FBPase structure; the corresponding active site residues of *Mt*FBPase are indicated in red. Residues Asp^{82} and Arg^{159} are different from the corresponding *E. coli* GlpX protein.

4. Structural Characterization as a foundation for Target Based Drug Discovery – Future perspective

The structural findings described in this chapter strongly indicate that the quaternary structure of *Mt*FBPase (and *E. coli* GlpX too) is a tetramer. The classical FBPase I (present in mammals and several prokaryotes) also exist as a tetramer. The initial findings of biochemical characterization suggest that, though there are several biochemical effectors that inhibit *Mt*FBPase activity (at submillimolar concentrations) (Table XI), it is still uncertain whether or not there is any allosteric regulation in *Mt*FBPase as widely observed for FBPase I (micromolar inhibitory concentrations for AMP, activation by citrate and PEP). Approaches towards design of allosteric inhibitors of FBPase I as potential anti-diabetic drug candidates has been recently reviewed (van Poelje, Potter et al. 2011). CS-917 is one such potent FBPase I inhibitor which has shown promising results in experimental animal model (Erion, van Poelje et al. 2005; van Poelje, Potter et al. 2006; Yoshida, Okuno et al. 2008). Our findings suggest of three possible alternative strategies for target based drug discovery based on the essentiality information already available about this target and presented earlier.

i. High throughput screening (HTS) against the said target (*Mt*FBPase) – Since we already have a purified target (established purification protocols for *Mt*FBPase), easily scalable screening assay (coupled spectrophotometric assay and the malachite green assay for phosphate release estimation) and information on basic biochemical properties of the enzyme, HTS against the *Mt*FBPase target can be performed.

ii. Virtual screening against *Mt*FBPase using computational methods – Since the native *Mt*FBPase and *Mt*FBPase-F6P crystal structures are available, virtual screening of chemical libraries (either physical or electronic) against *Mt*FBPase could identify a subset of promising compounds that could be later tested for activity in the enzymatic

assays to identify robust hit compounds. Validation of the initial hits will be done with further secondary assays and the confirmed, most efficienty hits, will be further analyzed by co-crystallization experiments with *Mt*FPBase. The ability to obtain routinely crystallographic structures of *Mt*FBPase-ligand complexes will be the cornerstone of hit and lead optimization using Structure Based Drug Design strategies.

iii. Identifying inhibitors specifically for the active site (F1,6BP mimetics) – F1,6BP mimetics by computational methods, as described in the literature (Sarker, Talcott et al. 2012), would be an alternative strategy. However, first it will be necessary to confirm that the proposed compounds are indeed inhibitors of *Mt*FBPase.

While all of these strategies have their advantages and limitations, a multidimensional approach using the combination of the above mentioned strategies seems to be more rational (optimal utilization of available resources). The available high resolution structure of *Mt*FBPase (F6P complex) would play an important role in the drug design effort. Some of the areas wherein a high resolution structure of *Mt*FBPase as available herein would have a tremendous impact are:

i. Virtual screening of inhibitors using computational methods.

ii. Validating the putative hits obtained by any discovery method method (Virtual screening, high throughput screening, or even the proposed F1,6BP mimetic (Figure 49).

iii. Lead optimization by Structural Based Drug Design methods.

Designing F1,6BP mimetics is a rather challenging task. Sarker et al. describe a cheminformatics; pathway analysis coupled with whole-cell assays based approach to search for active compounds against *Mtb* (Sarker, Talcott et al. 2012). There are several such recent studies

describing the combined application of cheminformatics and whole-cell assays in tuberculosis drug discovery (Ekins, Kaneko et al. 2010; Ekins and Freundlich 2011; Ekins, Freundlich et al. 2011). Studies described by Sarker et al. provide additional evidence proposing *glpX* (Rv1099c) as an essential gene and have identified two suggested D-fructose 1,6 bisphosphate mimics (DFP000133SC and DFP000134S) with MIC activities of 40 and 20 ug/ml, respectively in whole-cell assays. These compounds are available from Maybridge (Cat. Nos. DFP00134 (a), DFP00133 (b)) and can be described as 5-nitro-3-thienyl-thioureas (Figure 48).

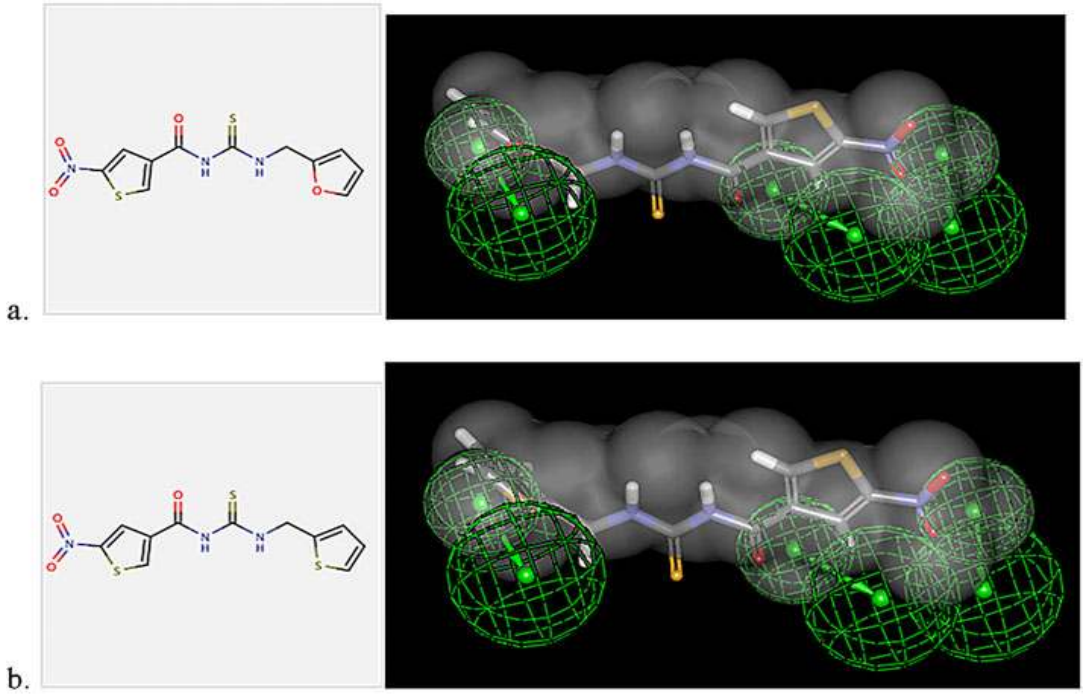


Figure 49. The proposed F1,6BP mimics identified using a combination of chemiformatics and pathway analysis for the metabolic pathways of *Mtb*. The green contours represent hydrogen bond acceptors in their pharmacophore description. Adapted with permission from Sarker et al. 2012.

Tanimoto similarity scores of these two compounds compared to F1,6BP are low (0.28 and 0.24, respectively) and it is likely that their mode of action is not as inhibitors of *Mt*FBPase but rather they could act because of their NO_2 group, in a manner similar to PA-824

(Manjunatha, Boshoff et al. 2006; Sarker, Talcott et al. 2012) . As a solid basis for further work, a systematic experimental evaluation of these compounds as *Mt*FBPase inhibitors is warranted.

5.4 CONCLUSIONS

Structural biology efforts centered on *glpX*, have yielded extremely valuable information on *Mt*FBPase. *Mt*FBPase has been successfully crystallized in several forms (native, F1,6BP and F6P bound complexes), and the structures of the native form and F6P complex have been successfully determined.

While there are many similarities between the overall monomeric structures of *Mt*FBPase and *E. coli* GlpX, the following novel findings from this work are of interest in understanding the structural properties of FBPase II.

a. The longer αH9 and the possible correlation of loop-αH9-loop structure in *E. coli* GlpX to lithium resistance. The possible correlation needs to be investigated in *E. coli* GlpX by designing rational point mutations and evaluating its lithium sensitivity in a manner similar to that described for inositol monophosphatase (IMPase) enzyme from the hyperthermophilic archaeon *Methanocaldococcus jannaschii*. The size and orientation of this catalytic loop has been implicated in determining lithium sensitivity of dual activity IMPases/FBPases AF2372 and MJ0109 (Stieglitz, Johnson et al. 2002; Li, Stieglitz et al. 2010).

b. The functional unit of *Mt*FBPase appears to be a tetramer in solution with a novel protein-protein interface that could be functional in allosteric regulation. While the known biochemical regulators of FBPase I activity (PEP and citrate as activators and

AMP as inhibitor), do not show significant effect on *Mt*FBPase, the possibility of allosteric regulation for *Mt*FBPase cannot be completely ruled out. The inhibition of *Mt*FBPase by intermediates of TCA cycle at submillimolar concentrations is an indicator for this possible alternate regulatory mechanism which needs to be further verified by co-crystallization and mechanistic studies.

c. The refined structures of native *Mt*FBPase and F6P-*Mt*FBPase will be essential pieces of a Structure Based Drug Design strategy/program to identify promising leads, and lead-optimization efforts to achieve proof-of-concept inhibitors that can be tested in cell (and possibly) animal models as therapeutic agents against *Mtb*.

CHAPTER 6
GLPX GENE OF FRANCISELLA TULARENSIS ENCODES A FUNCTIONAL FRUCTOSE 1,6-BISPHOSPHATASEII

6.1 RATIONAL AND HYPOTHESIS

As described previously, the *glpX* gene is required for the virulence of *F. tularensis in vivo* (Kadzhaev, Zingmark et al. 2009). Like *Mtb*, there is no Class I FBPase in *F. tularensis* and it is expected that the *glpX* gene product is responsible for catalysis of F1,6BP. However, the exact biochemical activity, or function of the *glpX* gene product has not been experimentally verified. In this chapter the biochemical activity of the *glpX* gene product is discussed. This chapter also describes purification and crystallization of the *glpX* gene product which serves as a foundation for a Structure Based Drug Design strategy targeting the *glpX* gene of *F. tularensis*.

6.2 MATERIALS AND METHODS

A. Cloning of *F. tularensis glpX* gene

The *F. tularensis glpX* gene was codon optimized, synthesized, and cloned into a pET-15b vector suitable for genetic complementation studies and recombinant expression by CelTek Inc. (Refer APPENDIX A for details). The construct was sequenced for its appropriateness (Refer APPENDIX B and C for details), additionally, the construct was subjected to restriction digestion with BamHI/XbaI and the insertion of the *FtglpX* gene was confirmed.

B. Complementation of *E. coli* Δ*glpX*/Δ*fbp* strains

The pET-15b-*FtglpX* construct was transformed into *E. coli* strains JLD2402, which lacks both *fbp* and *glpX*, and JLD2404, which lacks only *fbp.* This has been described elsewhere (Movahedzadeh, Rison et al. 2004). Antibiotic resistant transformants were grown on minimal media plates containing glucose or glycerol as the sole carbon source, and IPTG to induce expression of the *glpX* gene. Additionally, the pET-15b-*FtglpX* construct was transformed into

E. coli strain JB108 (BL21[λDE3 (lacUV5-T7gene1)]Δ*fbp*287) which was grown in minimal medium containing glycerol as previously described (Donahue, Bownas et al. 2000). Appropriate control strains of *E. coli* BL21 (DE3), untransformed JB108, and JB108 strain (transformed with pET-15b) were also run with appropriate concentrations of the inducer, IPTG. The IPTG concentrations were increased from 0 mM to 1 mM to identify the effect of the inducer on the growth of JB108 harboring pET-15b-*FtglpX*.

C. Expression and purification of *F. tularensis* GlpX

The pET-15b-*FtglpX* construct expressing N-terminal histidine-tagged *glpX* was transformed into *E. coli* strain BL21 (DE3). The positive transformants were selected on LB agar containing ampicillin (100µg/ml). Then transformants were grown overnight in LB broth (with ampicillin at 100µg/ml) at 37°C in an incubator shaker set at 180 rpm. The following morning, 1 ml of this culture was transferred to 100 ml of LB broth (with ampicillin at 100µg/ml). The culture was grown at 37°C and the OD_{600} was monitored frequently until it reached to 0.6. Subsequently, the culture was induced with IPTG at a final concentration of 1mM. The growth was continued at 37°C for 5 hours post induction. The cell pellet was harvested and frozen at -80°C until further use.

Purification of the *F. tularensis* GlpX gene product was performed in a similar manner as described for *Mt*FBPase (Chapter 4), except that the final buffer exchange and concentration of the purified protein was performed using an Amicon-15 Ultracel 30 K (Millipore, USA) centrifuge concentrator. This was to ensure that the protein was completely retained during centrifugation irrespective of its oligomeric state in solution (molecular weight of the monomer is about 36 kDa).

D. Molecular weight estimation of *F. tularensis* GlpX by Size Exclusion Chromatography

SEC was performed on an ÄKTA purifier FPLC system in a similar manner as that described for *Mt*FBPase (Chapter 4). Elution/retention times were determined by monitoring the absorbance at 280 nm. *F. tularensis* GlpX was injected at a concentration of 1.0 mg/ ml. The relative elution (K_{av}) and hence the molecular weight for *F. tularensis* GlpX was estimated as described for *Mt*FBPase (Chapter 4).

E. Crystallization of *F. tularensis* GlpX

Crystallization trials were performed by the hanging-drop vapor-diffusion method (2µl drop volume, 300µl reservoir volume), and suitable crystallization conditions were found from selected crystallization screens from Hampton Research (Hampton Research, Aliso Viejo, California, USA). Crystal development was scored using the criteria as provided by Hampton. All the crystals grew at temperatures between 22–27 °C. While the protein crystallized under several different conditions, only a few conditions (as described in results) were considered for further optimization of crystal diffraction and data collection, based on the preliminary diffraction observed. For conditions where protein co-crystals with substrate were set, F 1,6-BP was mixed with the protein solution at a concentration of 10mM, the resultant mixture was incubated at 30°C for 30 min followed by storage on ice. This mixture was used for crystallization in a similar manner as that used for APO protein crystallization.

F. Data Collection and Processing

The crystals were soaked in a cryoprotectant consisting of the respective reservoir solution plus 15% PEG 10000 (v/v). Subsequently, the crystals were flash-frozen in liquid nitrogen at 100 K. X-ray diffraction data for the crystals were collected from a single frozen crystal using a MAR 300 CCD detector on the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beam line at the Advanced Photon Source, Argonne National Laboratory,

Illinois, USA. A total of 240 frames were collected for crystal ID I65 with 1° scan width and 1 s exposure time. The crystal-to-detector distance was set to 280 mm. The data were indexed and scaled with the HKL- 2000 package.

6.3 RESULTS AND DISCUSSION

A. *F. tularensis glpX* encodes a Class II FBPase which functionally complements an *E. coli* Δ*glpX*/Δ*fbp* strain

The pET15b-*FtglpX* construct was introduced into *E. coli* strains JLD2402 and JLD2404, and the ampicillin resistant colonies streaked on minimal media plates containing either glucose or glycerol as the sole carbon source, and IPTG to induce the expression of *glpX* gene. While the control strains (untransformed JLD2402 and JLD2404) only grew on glucose as a sole carbon source, the pET15b-*FtglpX* construct transformed into JLD2402 and JLD2404, grew on either glucose or glycerol as the sole carbon source (Figure 50). This genetic complementation is indicative of FBPase activity of the *F. tularensis glpX* gene product. However, it is noteworthy that such complementation occurs at 27°C over the incubation period of 36 hours (as opposed to the normal growth condition of 37°C overnight) which does not rule out the possibility of non-specific complementation. Although the complementation of JLD2402 and JLD2404 was not expected from the pET15b-*FtglpX* vector in the true sense (because of the lack of a host T7 polymerase gene), it was still detected, indicating a possibility of read-through from other promoters. Since this complementation result is not confirmatory and rather non-specific, an additional (more specific) genetic complementation was performed.

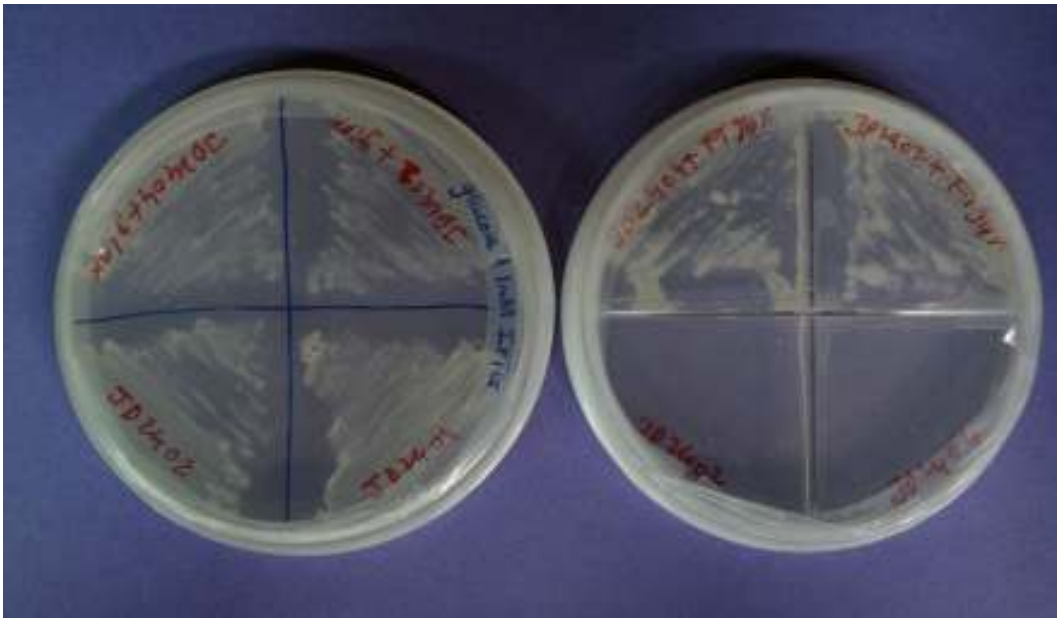


Figure 50. Growth pattern of the glucose dependent *E. coli* strains transformed with *FtglpX*. JLD2402 and JLD2404 transformed with the pET15b-*FtglpX* construct were streaked onto minimal agar containing glucose (plate on the left) or glycerol (plate on right) as the carbon source. The plates were incubated at 27°C for 36 hours so as to visually confirm bacterial growth. As observed above, the transformed strains grew on minimal agar with glycerol but not the control (untransformed) strains.

F. tularensis GlpX protein expression from the T7 promoter complements growth of JB108 Δ*fbp E. coli* strain on glycerol minimal media in an IPTG concentration dependent manner (Figure 51). Strains JB108 (Δ*fbp glpX*$^{+}$), JB108 (pET15b) and JB108 (pET15b-*FtglpX*), in which *glpX* is expressed from the T7 promoter, and control strain BL21 (DE3) (*fbp*$^{+}$ *glpX*$^{+}$) were grown in minimal glycerol media. Variable concentrations of IPTG (0 mM to 1 mM) were used. The growth of JB108 (pET15b-*FtglpX*) and the corresponding FBPase activity was directly proportional to the effective IPTG concentration used for growth. The coupled spectrophotometric assay (as described in chapter 4) was used to measure the FBPase activity in cell-free extracts. The *E. coli* JB108 (Δ*fbp*) host gave low background levels (Table XVII), while

the mutant complemented with pET15b-*FtglpX* showed elevated levels of FBPase activity in an IPTG concentration dependent manner.

Growth profile of *E. coli* strains expressing *F. tularensis* GlpX protein

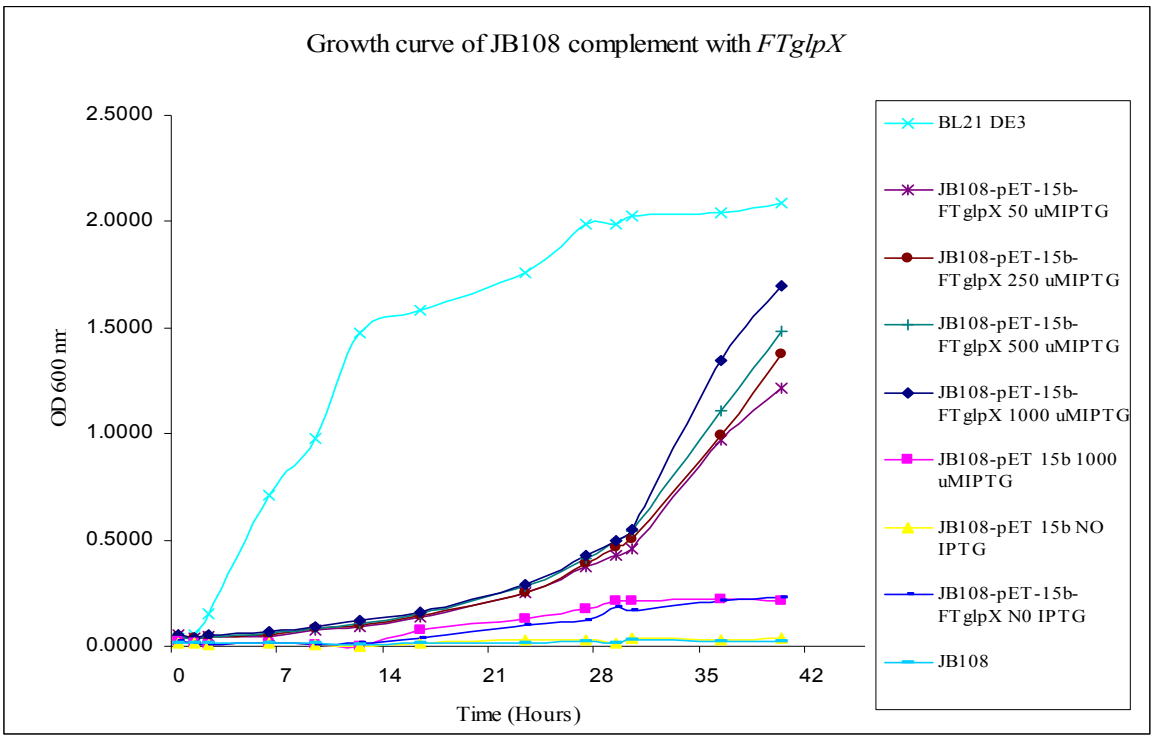


Figure 51. Growth profiles of *E. coli* strains expressing *Ft*GlpX protein. Effect of increased *Ft*GlpX on growth of Δ*fbp E. coli* strains. Strains used were JB108 (Δ*fbp glpX*$^{+}$), JB108 (pET15b) and JB108 (pET15b-*FtglpX*) in which *glpX* is expressed from the T7 promoter, BL21 (DE3) (*fbp*$^{+}$ *glpX*$^{+}$) was a control strain. All strains were grown in minimal medium containing glycerol. The growth of JB108 (pET15b-*FtglpX*) increased in an IPTG concentration (0 mM to 1 mM) dependent manner. The starter culture was grown in minimal medium containing glucose overnight. The IPTG levels utilized for strain JB108 (pET15b-*FtglpX*) during growth were 0 mM, 0.05 mM, 0.25 mM , 0.5 mM, and 1 mM. No IPTG was used for each of BL21 (DE3) (*fbp*$^{+}$ *glpX*$^{+}$), JB108 (Δ*fbp glpX*$^{+}$), JB108 (pET15b) and JB108 (pET15b-*FtglpX*) strains which served as controls. This experiment was repeated twice. Samples for measuring FBPase activity were withdrawn at 24 hours post IPTG induction; the cultures were also supplemented with ampicillin (100µg/ml).

Strain/construct	IPTG conc. (mM)	n[a]	FBPase specific activity [nmol min^{-1} (mg protein)$^{-1}$][b]
BL21(DE3)	0	3	16.3±1.22
JB108 [*Δfbp*][c]	0	3	0.33±0.08
JB108(pET15-b)	0	3	0.09±0.03
JB108(pET15-b)	1000	3	0.16±0.07
JB108(pET15-b-*FtglpX*)	0	6	0.33±0.09
JB108(pET15-b-*FtglpX*)	50	6	5.53±1.13
JB108(pET15-b-*FtglpX*)	250	6	9.89±1.75
JB108(pET15-b-*FtglpX*)	500	6	16.72±1.25
JB108(pET15-b-*FtglpX*)	1000	6	21.46±1.87

[a], n=number of determinations.
[b],FBPase activity is measured using the coupled spectrometric assay as developed for the purified *Mt*FBPase (Refer Chapter 4), reported values are Mean ± SD.
[c]Square bracket indicates relevant genotype of strain.

TABLE XVII. FBPASE ACTIVITY IN *E. COLI* STRAINS EXPRESSING *FT*GLPX PROTEIN. Samples taken at 24 hrs after IPTG induction.

Primary sequence comparison of *Ft*GlpX, with other known class II FBPases (41- 47% sequence identity), suggests that several characteristic sequence blocks are conserved, also all the important catalytic residues are conserved (Figure 52).

Sequence comparison of *F. tularensis* GlpX protein with other class II FBPases:

CLUSTAL 2.0.12 multiple sequence alignment

```
Ftula          --------MNRKVALEAVRVTELAALASWSQMGRGDKIAADQAAVDAMRKALNEVDIDGT 52
Mtube          --------------MELVRVTEAGAMAAGRWVGRGDKEGGDGAAVDAMRELVNSVSMRGV 46
Msmeg          MTPSRGEAPDRNLALELVRVTEAGAMAAGRWVGRGDKEGGDGAAVDAMRELVNSVSMRGV 60
Cglut          MNLKNPETPDRNLAMELVRVTEAAALASGRWVGRGMKNEGDGAAVDAMRQLINSVTMKGV 60
Ecoli          --------MRRELAIEFSRVTESAALAGYKWLGRGDKNTADGAAVNAMRIMLNQVNIDGT 52
                              :*  **** .*:*.   :*** *  .* ***:***  :*.* : *.

Ftula          VVIGEGELDEAPMLYIGEKVGAG-GCEVDIALDPLEGTTITSKGGANALTVLAMADKGGF 111
Mtube          VVIGEGEKDHAPMLYNGEEVGNGDGPECDFAVDPIDGTTLMSKGMTNAISVLAVADRGTM 106
Msmeg          VVIGEGEKDNAPMLYNGEEVGNGDGPECDFAVDPIDGTTLMSKGMPNAISVLAVAERGAM 120
Cglut          VVIGEGEKDEAPMLYNGEEVGTGFGPEVDIAVDPVDGTTLMAEGRPNAISILAAAERGTM 120
Ecoli          IVIGEGEIDEAPMLYIGEKVGTGRGDAVDIAVDPIEGTRMTAMGQANALAVLAVGDKGCF 112
               :****** *.**** **:** * *   *:*:**::** : : * .**:::** .::* :

Ftula          LN-APDVYMQKIAVGGINAPKGIVDLDDSVTNNLKRIAEFKGVHMSALVVCTMDRPRHEH 170
Mtube          FDPSAVFYMNKIAVG--PDAAHVLDITAPISENIRAVAKVKDLSVRDMTVCILDRPRHAQ 164
Msmeg          FDPSAVFYMNKIAAG--PDVADFIDITSPIAANIQRIAKMRKASVSDITVCILDRPRHAK 178
Cglut          YDPSSVFYMKKIAVG--PEAAGKIDIEAPVAHNINAVAKSKGINPSDVTVVVLDRPRHIE 178
Ecoli          LN-APDMYMEKLIVG--PGAKGTIDLNLPLADNLRNVAAALGKPLSELTVTILAKPRHDA 169
               : :. .**:*: .*        :*:  .:: *:. :*         :.*  : :***

Ftula          IIKEARECGARVILINDGDVSGVIATATENSGIDVYIGTGGAPEGVLAAAALKCLGGQMQ 230
Mtube          LIHDVRATGARIRLITDGDVAGAISACRPHSGTDLLAGIGGTPEGIIAAAAIRCMGGAIQ 224
Msmeg          LIADVRAAGARIRLISDGDVAGAISACRPDSGTDLLAGIGGTPEGIITAAAIRCMGGEIQ 238
Cglut          LIADIRRAGAKVRLISDGDVAGAVAAAQDSNSVDIMMGTGGTPEGIITACAMKCMGGEIQ 238
Ecoli          VIAEMQQLGVRVFAIPDGDVAASILTCMPDSEVDVLYGIGGAPEGVVSAAVIRALDGDMN 229
               :* : :  *.::  * ****:. : :.   .  *:  * **:***:::*..::.:.* ::

Ftula          ARLIFND-------------EEEIKRAHRLGITDLNKKYDIDDLASGD-IVFAATGVTDG 276
Mtube          AQLAPRD-------------DAERRKALEAG-YDLNQVLTTEDLVSGENVFFCATGVTDG 270
Msmeg          ATLAPTD-------------DEERQKAIDRG-HDLDRVLTTKDLVSGENVFFCATGVTDG 284
Cglut          GILAPMN-------------DFERQKAHDAG-LVLDQVLHTNDLVSSDNCYFVATGVTNG 284
Ecoli          GRLLARHDVKGDNEENRRIGEQELARCKAMGIEAG-KVLRLGDMARSDNVIFSATGITKG 288
               . *   .             : *  :.   *     :     *:. .:   * ***:*.*

Ftula          NMLQGVKRVNSTRRGSYAVTHSVVMRSTTKTVRHITAEHSFDFKEG-IEKFMS------- 328
Mtube          DLLKGVR-----YYPGGCTTHSIVMRSKSGTVRMIEAYHRLSKLNEYSAIDFTGDSSAVY 325
Msmeg          DLLKGVR-----FFGGGCTTQSIVMRSKSGTVRMIEAYHRLSKLNEYSAINFTGDSSAAY 339
Cglut          DMLRGVS-----YRANGATTRSLVMRAKSGTIRHIESVHQLSKLQEYSVVDYTTAT---- 335
Ecoli          DLLEGIS-----RKGNIATTETLLIRGKSRTIRRIQSIHYLDRKDPEMQVHIL------- 336
               ::*.*:         . ..*.::::*..: *:* * : * :.  :

Ftula          ---
Mtube          PLP 328
Msmeg          PLP 342
Cglut          ---
Ecoli          ---
```

Figure 52. Sequence comparison of *Ft*GlpX protein with other class II FBPases. CLUSTALW alignment of *F. tularensis* GlpX with *Mtb (Rv1099c), M. smegmatis*, *C. glutamicum* , and *E. coli.* (41–47% sequence identity) identified over 50 conserved residues. Invariant positions are indicated by asterisks (*) under the alignment, while highly conserved and weakly conserved positions are indicated by colons and periods. Residues in the active catalytic site of *E. coli* GlpX are indicated in red. Importantly, all residues of the active catalytic site are highly conserved in all the FBPases considered for sequence comparison. Characteristic conserved blocks have been boxed. Also, conserved residues among the catalytically important residues have been boxed. The *glpX* sequence of *F. tularensis* was obtained from the KEGG database (FTF1631c), the *E. coli* sequence was obtained from SWISSPROT (accession no.

P28860), the *C. glutamicum* sequence was obtained from GenBank (accession no. 19552240), the *Mtb* sequence was obtained from the Tuberculist server (http://genolist.pasteur.fr/TubercuList/; gene name Rv1099c) and the *M. smegmatis* sequence was derived from genomic sequences obtained from The Institute for Genomic Research website (http://www.tigr.org/). The exact start residue of the proteins is known only for *C. glutamicum* (Rittmann et al., 2003).

B. Expression, Purification, Crystallization and Preliminary X-ray Characterization of *F. tularensis* GlpX

1. Purification of *F. tularensis* GlpX:

Significant expression of *F. tularensis* GlpX was observed upon induction with 1 mM IPTG in the *E. coli* cell cultures transformed with the pET15b-*FtglpX* construct (Figure 53. Lane 1). The *F. tularensis* GlpX, which was overexpressed in *E. coli*, was present in the soluble fraction (Figure 53. Lane 2), and the insoluble fraction (inclusion bodies) (Figure 53. Lane 3). The Ni-NTA elute was relatively impure (Figure 53. Lanes 5, 6) containing some high molecular weight impurities. *F. tularensis* GlpX was then purified to homogeneity by a subsequent SEC step (Figure 54).

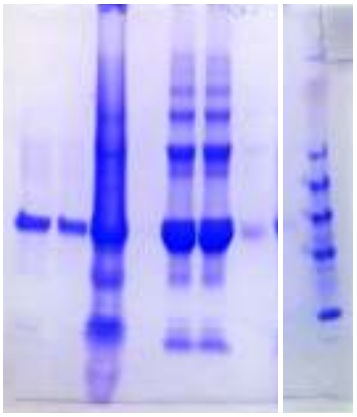


Figure 53. SDS-PAGE analysis of *Ft*GlpX expression, cell lysis and affinity capture. From left to right Lane 1: lysed cells; Lane 2: Supernatant of lysed cells obtained by high speed centrifugation; Lane 3: Cell pellet-insoluble cell debris; Lane 4: Flow through; Lane 5: Elution 1 (2ml volume); Lane 6: Elution 2 (2ml volume); Lane 7: 50 mM imidazole washing; Lane 8: SeeBlue Plus2 marker.

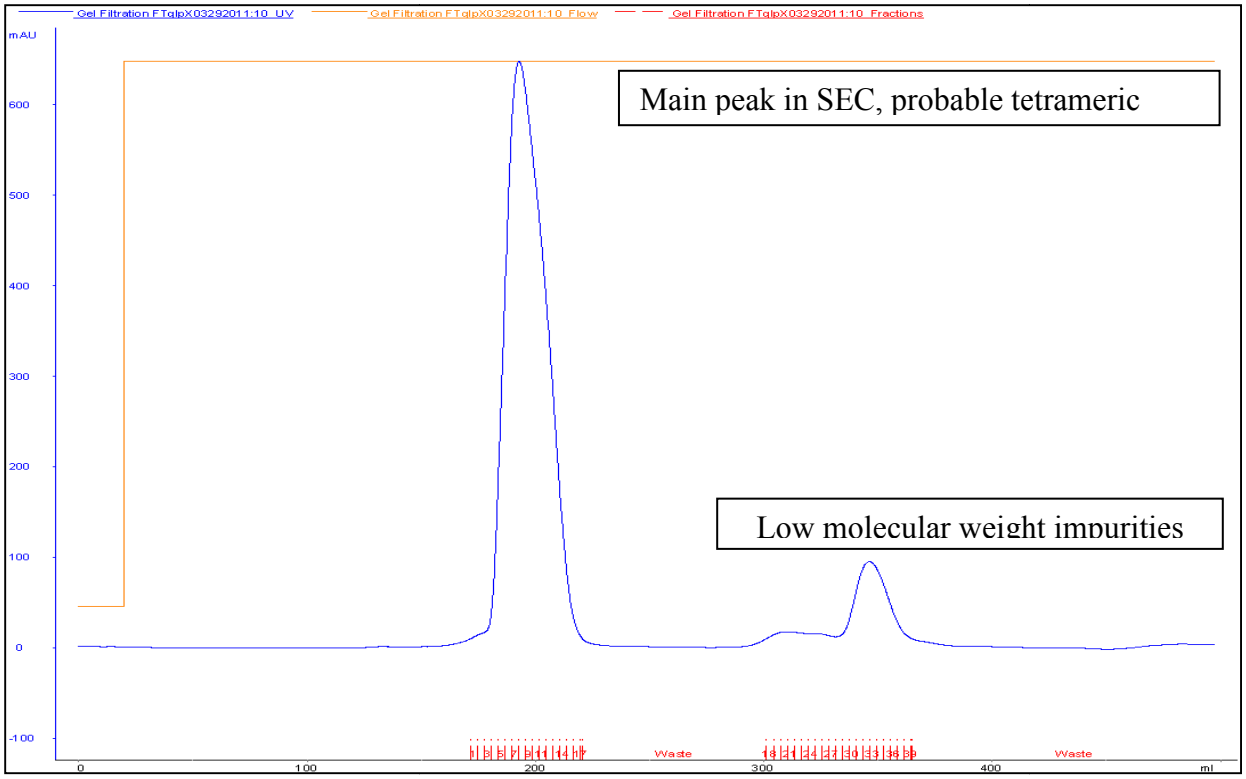
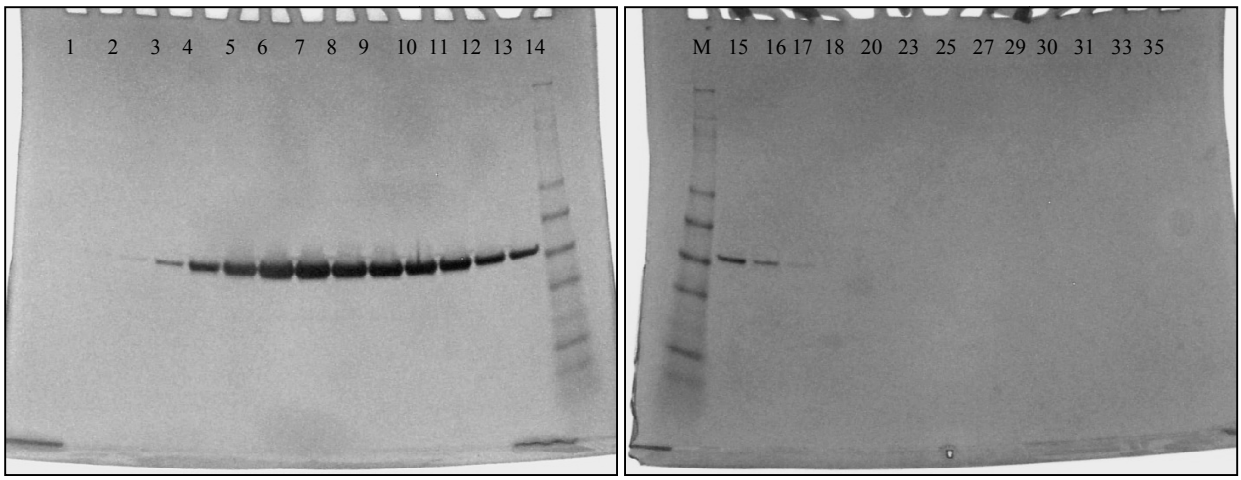


Figure 54. SEC chromatogram for *Ft*GlpX. Main peak at about 190 ml corresponds to the *Ft*GlpX tetramer (Fractions 1 to 17). Small peak at 350 ml (fractions 18 to 37) might be the monomeric form of GlpX protein or other buffer ingredients.



Gel 1 Gel 2

Figure 55. SDS-PAGE profile for SEC fractions of *Ft*GlpX. SDS-PAGE Gels 1 and 2 showing the profile of SEC fractions. The respective fractions are indicated on the gels. M molecular weight marker. Fractions 1 – 17 were diluted 1:10 with the elution buffer and fractions 18 – 37 were diluted 1:4 using the elution buffer. (Fractions as identified in the SEC chromatogram Figure 54).

Fractions 4 – 16 appeared to have high purity (Figure 55) and, hence, were pooled. Fractions 18 – 37 were confirmed to contain non-protein impurities, possibly buffer ingredients. Total protein quantification was performed using the Bradford protein assay. Bovine serum albumin (BSA) was used as the standard (Bradford 1976). Ultimately the protein solution was concentrated to 10 ± 0.5 mg/ml and crystallization performed.

2. Molecular weight of *F. tularensis* GlpX determined by Size Exclusion Chromatography: The exact oligomeric state of the *F. tularensis* GlpX in solution state was determined by SEC. As described in chapter 4, the void volume of the column, as determined by dextran blue, is V_o=39.80 ml; Total column volume, V_t (also referred to as geometric column volume) =120 ml. Using this method, two protein peaks corresponding to molecular weights of 63.8 kDa and 124.9 kDa were observed (Figure 56). This observation indicates that at a concentration of 1 mg/ml, *F. tularensis* GlpX exists as a mixture of dimer and tetramer (assuming the monomer subunit is about 35 kDa).

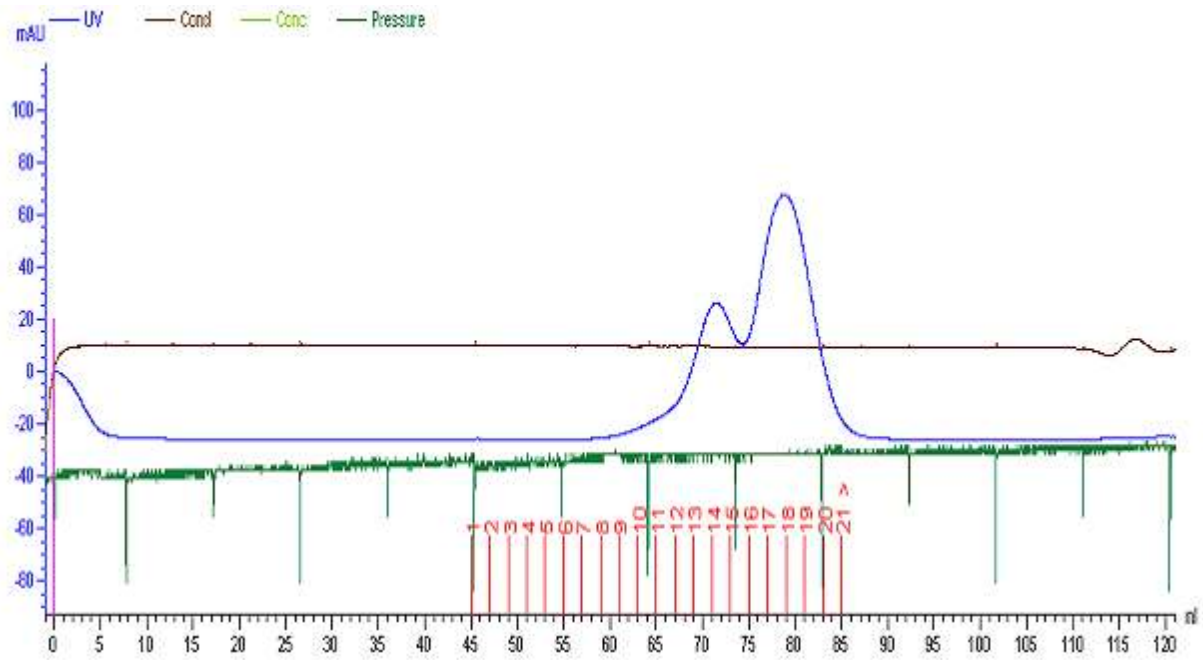


Figure 56. SEC chromatogram for *Ft*GlpX protein for molecular weight estimation. The chromatogram indicating the presence of dimeric and tetrameric species (63.8 kDa and 124.9 kDa, respectively).

3. Crystallization and Preliminary X-ray Characterization of *F. tularensis* GlpX

The following crystallization screens from Hampton Scientific were tested extensively for crystallization of *F. tularensis* GlpX protein:

1. Index solution set (1 -96)
2. Crystal screen (1-50)

Details of crystallization condition, and diffraction details have been summarized for all preliminary hits in Table XVIII. Data collection and processing statistics of one of the crystal forms (I65) have been described in Table XIX.

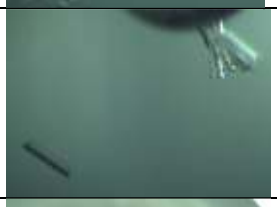
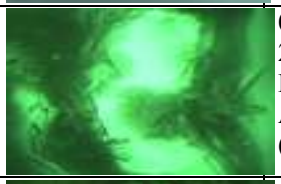
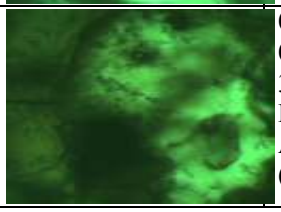
ID	Image/photograph of crystals	Crystallization condition (chemical composition of crystallization solutions, protein concentration, additional components)	Crystal diffraction details Resolution (Å)
D5		0.1 M HEPES sodium pH 7.5, 2% v/v Polyethylene glycol 400 2.0 M Ammonium sulfate Protein Conc.: 10mg/ml Substrate F 1,6-BP (10mM) (protein solution + precipitant in 1:1 ratio)	Diffraction >3 Å
F5		0.1 M HEPES sodium pH 7.5, 2% v/v Polyethylene glycol 400 2.0 M Ammonium sulfate Protein Conc.: 5 mg/ml Substrate F 1,6-BP (10mM) (protein solution + precipitant in 1:1 ratio)	Diffraction >3 Å
C4		0.2 M Sodium acetate trihydrate 0.1 M Tris HCl pH 8.5, 30% w/v Polyethylene glycol 4000 Protein Conc.: 5 mg/ml APO condition (protein solution + precipitant in 1:1 ratio)	Diffraction >3 Å
I65		Index 65 0.1 M Ammonium sulfate 0.1M Bis-Tris pH 5.5, 17% Polyethylene glycol 10000 Protein Conc.: 10mg/ml Substrate F 1,6BP (10mM) (protein solution + precipitant in 1:1 ratio)	Diffraction 240 frames collected, 200 frames selected for data processing, resolution ~ 2.8 Å.
D4		0.1 M Tris HCl pH 8.5 2.0 M Ammonium sulfate Protein Conc.: 10mg/ml Substrate F 1,6BP (10mM) (protein solution + precipitant in 1:1 ratio)	Poor diffraction
F4		0.1 M Tris HCl pH 8.5 2.0 M Ammonium sulfate Protein Conc.: 5 mg/ml Substrate F 1,6BP (10mM) (protein solution + precipitant in 1:1 ratio)	Poor diffraction
CS18		0.1 M Tris HCl pH 8.5 2.0 M Ammonium sulfate Protein Conc.: 10mg/ml APO condition (protein solution + precipitant in 1:1 ratio)	Poor diffraction
CS22		0.2 M Sodium acetate trihydrate 0.1 M Tris HCl pH 8.5 30% w/v Polyethylene glycol 4,000 Protein Conc.: 10 mg/ml APO condition (protein solution + precipitant in 1:1 ratio)	Poor diffraction

TABLE XVIII. PRELIMINARY CRYSTALLIZATION HITS FOR *FT*GLPX.

Synchrotron-radiation source	22ID-SER-CAT, APS
Wavelength (Å)	1.0000
Space group	Space group P21, No. 4, with 4 chains in the asymmetric unit.
Unit-cell parameters (Å)	a =76.02, b = 101.26, c = 92.59, α = 90, β = 90.061, γ = 90
Total No. of reflections	142758
No. of unique reflections	33841
Redundancy	3.8 (3.1)
Resolution (Å)	20 - 2.77 Å (2.77 Å)
Completeness (%)	91.6 (82.5) %
R_{merge}[a]	14.9 (70.5) %
I/σ(I) overall	10.3 (2.0)
Wilson plot (B-overall, $Å^2$)	53.3 $Å^2$

TABLE XIX. CRYSTAL DATA-COLLECTION AND PROCESSING STATISTICS FOR *FT*GLPX CRYSTAL. In parenthesis is the resolution and values for the last shell as indicated.
[a]$R_{merge} = \Sigma hkl\ \Sigma i\ |Ii(hkl) - \langle I(hkl)\rangle| / \Sigma hkl\ \Sigma i\ Ii(hkl)$

6.4 CONCLUSION

The genetic complementation results prove that the *F. tularensis* GlpX protein possesses FBPase activity and can complement the *E. coli* Δ*fbp* strain. Bioinformatics results indicate that, like *Mt*FBPase, it is a class II FBPase. *F. tularensis* GlpX was easily purified following the standardized protocols in a manner similar to that used for *Mt*FBPase. Hence, biochemical properties of this essential enzyme could be studied in a similar manner to that for *Mt*FBPase. Biochemical characterization would provide valuable and novel information (not just reiterative information) for drug discovery and should be pursued as an ongoing research activity.

Moreover, preliminary crystallization of this important enzyme has been successful. An initial structural model of *F. tularensis* GlpX (dataset collected at resolution of 2.7 Å) can be easily built (to satisfactory statistical parameters) by iterative refinement.

In addition to the initial structural model, this work has generated an exhaustive list of possible crystallization conditions (Table XVIII) for this target protein. This can ultimately lead to better quality diffraction data sets and thus resolve the structure of *F. Tularensis* GlpX.

The major bottlenecks in the process of structure-based drug discovery are the availability of purified protein target and the ability to crystallize the protein, both of which have been adequately addressed in this chapter. The biochemical and crystallographic understanding of this validated enzyme target can serve as a starting point for a structure-based drug-discovery expedition.

FUTURE PERSPECTIVE

Our findings suggest that the *glpX* gene is required for eugonic growth on selective gluconeogenic substrates (acetate, glycerol and oleic acid) and for survival *in vivo*. The *in vivo* survival profile suggests that *glpX* is required to achieve an initial instillation dose and also maintain high bacterial loads at later time points (acute and chronic phases). A 2-3 log lower bacterial count is considered to be a significant indication of the *glpX* gene essentiality *in vivo*. The lower bacterial counts at later time points could be argued due to an initially low instillation dose (<100 CFU at 1 day post infection). The significance or essentiality of genes can be efficiently studied at varying stages of disease progression (acute vs. chronic infection) by generating conditional mutants to understand the function of such essential genes. The most common way to construct a conditional mutant is by controlling expression of the gene with an inducible promoter. The development of tetracycline-inducible systems (P_{tet}) in many microorganisms, including mycobacteria, has provided an unprecedented opportunity to study the essential genes (Ehrt and Schnappinger 2006). This system could be used to controllably express the *glpX* gene and evaluate the effects of turning-off *glpX* expression at specific stages of infection (acute or chronic) in the mice model. The P_{tet} system has been used to regulate gene expression both *in vitro* and *in vivo* in *mycobacteria* (Blokpoel, Murphy et al. 2005; Ehrt, Guo et al. 2005; Ehrt and Schnappinger 2006; Guo, Monteleone et al. 2007; Klotzsche, Ehrt et al. 2009). This system functions efficiently *in vivo* as well, since the inducer (tetracycline) can enter eukaryotic cells and be administered to animals. This allows control of mycobacterial gene expression *in vivo* as well as *in vitro*. Thus, the effect of withdrawal of gene function can be monitored by controlling tetracycline dosing at early, mid, and late stages of infection. Since it is known that there is a significant difference in the *in vivo* survival profile of the Δ*glpX* strain in

BALB/c mice, the therapeutic potential of the Δ*glpX* strain as a vaccine candidate should be evaluated.

The Δ*glpX* strain showed similar MIC for all the drugs tested except for PA-824 and TMC207. Identifying a chemical means to inhibit *glpX* gene function would provide a unique opportunity to find a possible explanation for the higher MIC of PA-824 and also evaluate the possible synergistic effect with TMC207.

*Mt*FBPase has been successfully crystallized in several forms (native, F1,6BP and F6P bound complexes), and the structures of the native form and F6P complex have been determined.

While there are many similarities between the overall monomeric structures of *Mt*FBPase and *E. coli* GlpX, two novel findings from this work are of interest in understanding the structural properties of FBPase II. The longer αH9 and the possible correlation of loop-αH9-loop structure in *E. coli* GlpX to lithium resistance should be investigated by designing rational point mutations and evaluating its lithium sensitivity in a manner similar to that described for inositol monophosphatase (IMPase) enzyme from the hyperthermophilic archaeon *Methanocaldococcus jannaschii*. Point mutations on acidic residues Asp^{241}, Glu^{243}, Glu^{244} and Glu^{250} similar to the acidic residues found in mobile catalytic loop of archeal IMPase might provide an insight into the reason behind lithium resistance in case of *E. coli* GlpX. An alternative strategy would be to introduce additional acidic residues between the Ala^{233}-Glu^{234} regions of *Mt*FBPase and evaluate the lithium sensitivity of such mutants.

While the known biochemical regulators of FBPase I activity (PEP and citrate as activators and AMP as inhibitor), do not show a significant effect on *Mt*FBPase, the possibility of allosteric regulation for *Mt*FBPase cannot be completely ruled out. The inhibition of *Mt*FBPase by intermediates of TCA cycle at submillimolar concentrations is an indicator for this

possible alternate regulatory mechanism which needs to be further verified by co- crystallization and mechanistic studies in light of understanding the novel protein-protein interface in the tetrameric *Mt*FBPase structure that could be involved in allosteric regulation.

The refined structures of native *Mt*FBPase and F6P-*Mt*FBPase are critical pieces information for Structure Based Drug Design strategy/program.

The coupled spectrophotometric assay (used for biochemical characterization of the enzyme) and the malachite green assay (used to determine substrate specificity) in their current forms can be easily scaled to medium and high throughput formats for screening of potential inhibitors. The potential leads (identified by virtual screen or HTS) can be independently validated by co-crystalization with *Mt*FBPase and elucidation of their complex structures. The crystallization of *Mt*FBPase is highly robust and easily reproducible with similar crystals being obtained on different occasions (*Mt*FBPase purified on different occasions and crystals set using standard established protocols). The crystal packing in both structures (same space group: $P6_122$) is quite open with ample solvent channels that would most likely permit soaking of ligands directly into pre-grown crystals. Retaining the diffraction quality of soaked crystals would certainly be a challenge. Novel methods based on the control of the humidity of the crystals as applied to PurE, a potential bacterial drug target (Abad-Zapatero, Oliete et al. 2011), could be used to improve the diffraction quality.

F. tularensis GlpX protein too can be pursued along the similar path as that adopted for *Mt*FBPase (biochemical characterization of the purified protein, structural biology, and drug discovery) to explore its potential as a validated drug target in pathogenic bacteria.

CITED LITERATURE

Abad-Zapatero, C., R. Oliete, et al. (2011). "Humidity control can compensate for the damage induced in protein crystals by alien solutions." Acta Crystallogr Sect F Struct Biol Cryst Commun 67(Pt 10): 1300-8.

Albert, A., L. Yenush, et al. (2000). "X-ray structure of yeast Hal2p, a major target of lithium and sodium toxicity, and identification of framework interactions determining cation sensitivity." J Mol Biol 295(4): 927-38.

Arcus, V. L., J. S. Lott, et al. (2006). "The potential impact of structural genomics on tuberculosis drug discovery." Drug Discov Today 11(1-2): 28-34.

Arora, A., N. R. Chandra, et al. (2011). "Structural biology of Mycobacterium tuberculosis proteins: the Indian efforts." Tuberculosis (Edinb) 91(5): 456-68.

Awasthy, D., A. Ambady, et al. (2010). "Essentiality and functional analysis of type I and type III pantothenate kinases of Mycobacterium tuberculosis." Microbiology 156(Pt 9): 2691-701.

Bai, H., X. Xue, et al. (2010). "Antisense antibiotics: a brief review of novel target discovery and delivery." Curr Drug Discov Technol 7(2): 76-85.

Baker, E. N. (2007). "Structural genomics as an approach towards understanding the biology of tuberculosis." J Struct Funct Genomics 8(2-3): 57-65.

Balasubramanian, V., M. S. Pavelka, Jr., et al. (1996). "Allelic exchange in Mycobacterium tuberculosis with long linear recombination substrates." J Bacteriol 178(1): 273-9.

Bashiri, G., A. M. Rehan, et al. (2010). "Metabolic engineering of cofactor F420 production in Mycobacterium smegmatis." PLoS One 5(12): e15803.

Bashiri, G., C. J. Squire, et al. (2007). "Expression, purification and crystallization of native and selenomethionine labeled Mycobacterium tuberculosis FGD1 (Rv0407) using a Mycobacterium smegmatis expression system." Protein Expr Purif 54(1): 38-44.

Bashiri, G., C. J. Squire, et al. (2008). "Crystal structures of F420-dependent glucose-6-phosphate dehydrogenase FGD1 involved in the activation of the anti-tuberculosis drug candidate PA-824 reveal the basis of coenzyme and substrate binding." J Biol Chem 283(25): 17531-41.

Baykov, A. A., O. A. Evtushenko, et al. (1988). "A malachite green procedure for orthophosphate determination and its use in alkaline phosphatase-based enzyme immunoassay." Anal Biochem 171(2): 266-70.

Belisle, J. T., S. B. Mahaffey, et al. (2009). Isolation of Mycobacterium Species Genomic DNA. Mycobacteria Protocols. T. Parish and A. C. C. Brown, Humana Press. 465: 1-12.

Bellinzoni, M. and G. Riccardi (2003). "Techniques and applications: The heterologous expression of Mycobacterium tuberculosis genes is an uphill road." Trends Microbiol 11(8): 351-8.

Blokpoel, M. C., H. N. Murphy, et al. (2005). "Tetracycline-inducible gene regulation in mycobacteria." Nucleic Acids Res 33(2): e22.

Boshoff, H. I. and C. E. Barry, 3rd (2005). "Tuberculosis - metabolism and respiration in the absence of growth." Nat Rev Microbiol 3(1): 70-80.

Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." Anal Biochem 72: 248-54.

Brown, A. C. and T. Parish (2006). "Instability of the acetamide-inducible expression vector pJAM2 in Mycobacterium tuberculosis." Plasmid 55(1): 81-6.

Brown, G., A. Singer, et al. (2009). "Structural and biochemical characterization of the type II fructose-1,6-bisphosphatase GlpX from Escherichia coli." J Biol Chem 284(6): 3784-92.

Carroll, P., D. G. Muttucumaru, et al. (2005). "Use of a tetracycline-inducible system for conditional expression in Mycobacterium tuberculosis and Mycobacterium smegmatis." Appl Environ Microbiol 71(6): 3077-84.

Chayen, N. E. (2004). "Turning protein crystallisation from an art into a science." Curr Opin Struct Biol 14(5): 577-83.

Chayen, N. E. and E. Saridakis (2008). "Protein crystallization: from purified protein to diffraction-quality crystal." Nat Methods 5(2): 147-53.

Chen, L. and M. F. Roberts (1998). "Cloning and expression of the inositol monophosphatase gene from Methanococcus jannaschii and characterization of the enzyme." Appl Environ Microbiol 64(7): 2609-15.

Chen, L. and M. F. Roberts (1999). "Characterization of a tetrameric inositol monophosphatase from the hyperthermophilic bacterium Thermotoga maritima." Appl Environ Microbiol 65(10): 4559-67.

Chen, L. and M. F. Roberts (2000). "Overexpression, purification, and analysis of complementation behavior of E. coli SuhB protein: comparison with bacterial and archaeal inositol monophosphatases." Biochemistry 39(14): 4145-53.

Chen, V. B., W. B. Arendall, III, et al. (2010). "MolProbity: all-atom structure validation for macromolecular crystallography." Acta Crystallographica Section D 66(1): 12-21.

Chenna, R., H. Sugawara, et al. (2003). "Multiple sequence alignment with the Clustal series of programs." Nucleic Acids Res 31(13): 3497-500.

Chim, N., J. E. Habel, et al. (2011). "The TB Structural Genomics Consortium: a decade of progress." Tuberculosis (Edinb) 91(2): 155-72.

Cho, S. H., S. Warit, et al. (2007). "Low-oxygen-recovery assay for high-throughput screening of compounds against nonreplicating Mycobacterium tuberculosis." Antimicrob Agents Chemother 51(4): 1380-5.

Choi, K. P., T. B. Bair, et al. (2001). "Use of transposon Tn5367 mutagenesis and a nitroimidazopyran-based selection system to demonstrate a requirement for fbiA and fbiB in coenzyme F(420) biosynthesis by Mycobacterium bovis BCG." J Bacteriol 183(24): 7058-66.

Choi, K. P., N. Kendrick, et al. (2002). "Demonstration that fbiC is required by Mycobacterium bovis BCG for coenzyme F(420) and FO biosynthesis." J Bacteriol 184(9): 2420-8.

Cole, S. T., R. Brosch, et al. (1998). "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence." Nature 393(6685): 537-44.

Collaborative Computational Project, n. (1994). "The CCP4 suite: programs for protein crystallography." Acta Crystallogr D Biol Crystallogr 50(Pt 5): 760-3.

Das, A. R., N. Dattagupta, et al. (2003). "A novel thiocationic liposomal formulation of antisense oligonucleotides with activity against Mycobacterium tuberculosis." Scand J Infect Dis 35(3): 168-74.

Davis, I. W., A. Leaver-Fay, et al. (2007). "MolProbity: all-atom contacts and structure validation for proteins and nucleic acids." Nucleic Acids Research 35(suppl 2): W375-W383.

de Carvalho, L. P., S. M. Fischer, et al. (2010). "Metabolomics of Mycobacterium tuberculosis reveals compartmentalized co-catabolism of carbon substrates." Chem Biol 17(10): 1122-31.

DeLano, W. L. (2002). "The PyMOL Molecular Graphics System ".

Donahue, J. L., J. L. Bownas, et al. (2000). "Purification and characterization of glpX-encoded fructose 1, 6-bisphosphatase, a new enzyme of the glycerol 3-phosphate regulon of Escherichia coli." J Bacteriol 182(19): 5624-7.

Dubos, R. J., B. D. Davis, et al. (1946). "The effect of water soluble lipids on the growth and biological properties of tubercle bacilli." Am Rev Tuberc 54(3): 204-12.

Dunn, M. F., J. A. Ramirez-Trujillo, et al. (2009). "Major roles of isocitrate lyase and malate synthase in bacterial and fungal pathogenesis." Microbiology 155(Pt 10): 3166-75.

Ehrt, S., X. V. Guo, et al. (2005). "Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor." Nucleic Acids Res 33(2): e21.

Ehrt, S. and D. Schnappinger (2006). "Controlling gene expression in mycobacteria." Future Microbiol 1(2): 177-84.

Ekins, S. and J. S. Freundlich (2011). "Validating new tuberculosis computational models with public whole cell screening aerobic activity datasets." Pharm Res 28(8): 1859-69.

Ekins, S., J. S. Freundlich, et al. (2011). "Computational databases, pathway and cheminformatics tools for tuberculosis drug discovery." Trends Microbiol 19(2): 65-74.

Ekins, S., T. Kaneko, et al. (2010). "Analysis and hit filtering of a very large library of compounds screened against Mycobacterium tuberculosis." Mol Biosyst 6(11): 2316-24.

Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallographica Section D 60(12 Part 1): 2126-2132.

Emsley, P., B. Lohkamp, et al. (2010). "Features and development of Coot." Acta Crystallographica Section D 66(4): 486-501.

Erion, M. D., P. D. van Poelje, et al. (2005). "MB06322 (CS-917): A potent and selective inhibitor of fructose 1,6-bisphosphatase for controlling gluconeogenesis in type 2 diabetes." Proc Natl Acad Sci U S A 102(22): 7970-5.

Goldstone, R. M., N. J. Moreland, et al. (2008). "A new Gateway vector and expression protocol for fast and efficient recombinant protein expression in Mycobacterium smegmatis." Protein Expr Purif 57(1): 81-7.

Goulding, C. W., L. J. Perry, et al. (2003). "Structural genomics of Mycobacterium tuberculosis: a preliminary report of progress at UCLA." Biophys Chem 105(2-3): 361-70.

Gu, X., M. Chen, et al. (2006). "Rv2131c gene product: an unconventional enzyme that is both inositol monophosphatase and fructose-1,6-bisphosphatase." Biochem Biophys Res Commun 339(3): 897-904.

Guillemont, J., C. Meyer, et al. (2011). "Diarylquinolines, synthesis pathways and quantitative structure--activity relationship studies leading to the discovery of TMC207." Future Med Chem 3(11): 1345-60.

Guo, X. V., M. Monteleone, et al. (2007). "Silencing Mycobacterium smegmatis by using tetracycline repressors." J Bacteriol 189(13): 4614-23.

Gutierrez-Lugo, M. T., G. L. Newton, et al. (2006). "Cloning, expression and rapid purification of active recombinant mycothiol ligase as B1 immunoglobulin binding domain of streptococcal protein G, glutathione-S-transferase and maltose binding protein fusion proteins in Mycobacterium smegmatis." Protein Expr Purif 50(1): 128-36.

Gutka, H. J., S. G. Franzblau, et al. (2011). "Crystallization and preliminary X-ray characterization of the glpX-encoded class II fructose-1,6-bisphosphatase from

Mycobacterium tuberculosis." Acta Crystallogr Sect F Struct Biol Cryst Commun 67(Pt 6): 710-3.

Gutka, H. J., K. Rukseree, et al. (2011). "glpX gene of Mycobacterium tuberculosis: heterologous expression, purification, and enzymatic characterization of the encoded fructose 1,6-bisphosphatase II." Appl Biochem Biotechnol 164(8): 1376-89.

Guy, E. S. and A. Mallampalli (2008). "Managing TB in the 21st century: existing and novel drug therapies." Ther Adv Respir Dis 2(6): 401-8.

Haagsma, A. C., R. Abdillahi-Ibrahim, et al. (2009). "Selectivity of TMC207 towards mycobacterial ATP synthase compared with that towards the eukaryotic homologue." Antimicrob Agents Chemother 53(3): 1290-2.

Harth, G., M. A. Horwitz, et al. (2002). "Targeting the Mycobacterium tuberculosis 30/32-kDa mycolyl transferase complex as a therapeutic strategy against tuberculosis: Proof of principle by using antisense technology." Proc Natl Acad Sci U S A 99(24): 15614-9.

Harth, G., P. C. Zamecnik, et al. (2007). "Hairpin extensions enhance the efficacy of mycolyl transferase-specific antisense oligonucleotides targeting Mycobacterium tuberculosis." Proc Natl Acad Sci U S A 104(17): 7199-204.

Harth, G., P. C. Zamecnik, et al. (2000). "Treatment of Mycobacterium tuberculosis with antisense oligonucleotides to glutamine synthetase mRNA inhibits glutamine synthetase activity, formation of the poly-L-glutamate/glutamine cell wall structure, and bacterial replication." Proc Natl Acad Sci U S A 97(1): 418-23.

Hatzios, S. K., A. T. Iavarone, et al. (2008). "Rv2131c from Mycobacterium tuberculosis is a CysQ 3'-phosphoadenosine-5'-phosphatase." Biochemistry 47(21): 5823-31.

Hatzios, S. K., M. W. Schelle, et al. (2011). "The Mycobacterium tuberculosis CysQ phosphatase modulates the biosynthesis of sulfated glycolipids and bacterial growth." Bioorg Med Chem Lett 21(17): 4956-9.

Higgins, D. G. and P. M. Sharp (1988). "CLUSTAL: a package for performing multiple sequence alignment on a microcomputer." Gene 73(1): 237-44.

Hinds, J., E. Mahenthiralingam, et al. (1999). "Enhanced gene replacement in mycobacteria." Microbiology 145 (Pt 3): 519-27.

Hines, J. K., X. Chen, et al. (2007). "Structures of mammalian and bacterial fructose-1,6-bisphosphatase reveal the basis for synergism in AMP/fructose 2,6-bisphosphate inhibition." J Biol Chem 282(49): 36121-31.

Hines, J. K., H. J. Fromm, et al. (2006). "Novel allosteric activation site in Escherichia coli fructose-1,6-bisphosphatase." J Biol Chem 281(27): 18386-93.

Hines, J. K., H. J. Fromm, et al. (2007). "Structures of activated fructose-1,6-bisphosphatase from Escherichia coli. Coordinate regulation of bacterial metabolism and the conservation of the R-state." J Biol Chem 282(16): 11696-704.

Hines, J. K., C. E. Kruesel, et al. (2007). "Structure of inhibited fructose-1,6-bisphosphatase from Escherichia coli: distinct allosteric inhibition sites for AMP and glucose 6-phosphate and the characterization of a gluconeogenic switch." J Biol Chem 282(34): 24697-706.

Holton, S. J., M. S. Weiss, et al. (2007). "Structure-based approaches to drug discovery against tuberculosis." Curr Protein Pept Sci 8(4): 365-75.

Hugonnet, J. E., L. W. Tremblay, et al. (2009). "Meropenem-clavulanate is effective against extensively drug-resistant Mycobacterium tuberculosis." Science 323(5918): 1215-8.

Ibrahim, M., K. Andries, et al. (2007). "Synergistic activity of R207910 combined with pyrazinamide against murine tuberculosis." Antimicrob Agents Chemother 51(3): 1011-5.

Ibrahim, M., C. Truffot-Pernot, et al. (2009). "Sterilizing activity of R207910 (TMC207)-containing regimens in the murine model of tuberculosis." Am J Respir Crit Care Med 180(6): 553-7.

Ioerger, T. R. and J. C. Sacchettini (2009). "Structural genomics approach to drug discovery for Mycobacterium tuberculosis." Curr Opin Microbiol 12(3): 318-25.

Kadzhaev, K., C. Zingmark, et al. (2009). "Identification of genes contributing to the virulence of Francisella tularensis SCHU S4 in a mouse intradermal infection model." PLoS One 4(5): e5463.

Kelley-Loughnane, N., S. A. Biolsi, et al. (2002). "Purification, kinetic studies, and homology model of Escherichia coli fructose-1,6-bisphosphatase." Biochim Biophys Acta 1594(1): 6-16.

Kendrew, J. C., R. E. Dickerson, et al. (1960). "Structure of myoglobin: A three-dimensional Fourier synthesis at 2 A. resolution." Nature 185(4711): 422-7.

Klotzsche, M., S. Ehrt, et al. (2009). "Improved tetracycline repressors for gene silencing in mycobacteria." Nucleic Acids Res 37(6): 1778-88.

Korepanova, A., F. P. Gao, et al. (2005). "Cloning and expression of multiple integral membrane proteins from Mycobacterium tuberculosis in Escherichia coli." Protein Sci 14(1): 148-58.

Korepanova, A., J. D. Moore, et al. (2007). "Expression of membrane proteins from Mycobacterium tuberculosis in Escherichia coli as fusions with maltose binding protein." Protein Expr Purif 53(1): 24-30.

Koul, A., N. Dendouga, et al. (2007). "Diarylquinolines target subunit c of mycobacterial ATP synthase." Nat Chem Biol 3(6): 323-4.

Koul, A., L. Vranckx, et al. (2008). "Diarylquinolines are bactericidal for dormant mycobacteria as a result of disturbed ATP homeostasis." J Biol Chem 283(37): 25273-80.

Krissinel, E. (2010). "Crystal contacts as nature's docking solutions." J Comput Chem 31(1): 133-43.

Krissinel, E. (2011). "Macromolecular complexes in crystals and solutions." Acta Crystallogr D Biol Crystallogr 67(Pt 4): 376-85.

Krissinel, E. and K. Henrick (2007). "Inference of macromolecular assemblies from crystalline state." J Mol Biol 372(3): 774-97.

Larkin, M. A., G. Blackshields, et al. (2007). "Clustal W and Clustal X version 2.0." Bioinformatics 23(21): 2947-8.

Lebedev, A. A., A. A. Vagin, et al. (2008). "Model preparation in MOLREP and examples of model improvement using X-ray data." Acta Crystallogr D Biol Crystallogr 64(Pt 1): 33-9.

Li, Y., Z. Chen, et al. (2007). "Inositol-1-phosphate synthetase mRNA as a new target for antisense inhibition of Mycobacterium tuberculosis." J Biotechnol 128(4): 726-34.

Li, Y., Z. Chen, et al. (2009). "Inhibition of inositol-1-phosphate synthetase in <i>Mycobacterium tuberculosis</i> by chitosan-antisense nanoparticles." Journal of Wuhan University of Technology--Materials Science Edition 24(1): 87-90.

Li, Z., K. A. Stieglitz, et al. (2010). "Mobile loop mutations in an archaeal inositol monophosphatase: modulating three-metal ion assisted catalysis and lithium inhibition." Protein Sci 19(2): 309-18.

Lill, M. A. and M. L. Danielson (2011). "Computer-aided drug design platform using PyMOL." J Comput Aided Mol Des 25(1): 13-9.

Liu, K., J. Yu, et al. (2003). "pckA-deficient Mycobacterium bovis BCG shows attenuated virulence in mice and in macrophages." Microbiology 149(Pt 7): 1829-35.

Manjunatha, U. H., H. Boshoff, et al. (2006). "Identification of a nitroimidazo-oxazine-specific protein involved in PA-824 resistance in Mycobacterium tuberculosis." Proc Natl Acad Sci U S A 103(2): 431-6.

Manjunatha, U. H., R. Lahiri, et al. (2006). "Mycobacterium leprae is naturally resistant to PA-824." Antimicrob Agents Chemother 50(10): 3350-4.

Marrero, J., K. Y. Rhee, et al. (2010). "Gluconeogenic carbon flow of tricarboxylic acid cycle intermediates is critical for Mycobacterium tuberculosis to establish and maintain infection." Proc Natl Acad Sci U S A 107(21): 9819-24.

Matsumoto, M., H. Hashizume, et al. (2006). "OPC-67683, a nitro-dihydro-imidazooxazole derivative with promising action against tuberculosis in vitro and in mice." PLoS Med 3(11): e466.

Matsuyama, A. and M. Yoshida (2009). "Systematic cloning of an ORFeome using the Gateway system." Methods Mol Biol 577: 11-24.

Matteelli, A., A. C. Carvalho, et al. (2010). "TMC207: the first compound of a new class of potent anti-tuberculosis drugs." Future Microbiol 5(6): 849-58.

Matthews, B. W. (1968). "Solvent content of protein crystals." J Mol Biol 33(2): 491-7.

McAllister, G., P. Whiting, et al. (1992). "cDNA cloning of human and rat brain myo-inositol monophosphatase. Expression and characterization of the human recombinant enzyme." Biochem J 284 (Pt 3): 749-54.

McCoy, A. (2007). "Solving structures of protein complexes by molecular replacement with Phaser." Acta Crystallographica Section D 63(1): 32-41.

McCoy, A. J., R. W. Grosse-Kunstleve, et al. (2007). "Phaser crystallographic software." Journal of Applied Crystallography 40(4): 658-674.

McCoy, A. J., R. W. Grosse-Kunstleve, et al. (2005). "Likelihood-enhanced fast translation functions." Acta Crystallogr D Biol Crystallogr 61(Pt 4): 458-64.

McKinney, J. D., K. Honer zu Bentrup, et al. (2000). "Persistence of Mycobacterium tuberculosis in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase." Nature 406(6797): 735-8.

Movahedzadeh, F., S. C. Rison, et al. (2004). "The Mycobacterium tuberculosis Rv1099c gene encodes a GlpX-like class II fructose 1,6-bisphosphatase." Microbiology 150(Pt 10): 3499-505.

Movahedzadeh, F., D. A. Smith, et al. (2004). "The Mycobacterium tuberculosis ino1 gene is essential for growth and virulence." Mol Microbiol 51(4): 1003-14.

Movahedzadeh, F., P. R. Wheeler, et al. (2010). "Inositol monophosphate phosphatase genes of Mycobacterium tuberculosis." BMC Microbiol 10: 50.

Munoz-Elias, E. J. and J. D. McKinney (2005). "Mycobacterium tuberculosis isocitrate lyases 1 and 2 are jointly required for in vivo growth and virulence." Nat Med 11(6): 638-44.

Murry, J. P., C. M. Sassetti, et al. (2008). "Transposon site hybridization in Mycobacterium tuberculosis." Methods Mol Biol 416: 45-59.

Murshudov, G. N., P. Skubak, et al. (2011). "REFMAC5 for the refinement of macromolecular crystal structures." Acta Crystallographica Section D 67(4): 355-367.

Murshudov, G. N., A. A. Vagin, et al. (1997). "Refinement of macromolecular structures by the maximum-likelihood method." Acta Crystallogr D Biol Crystallogr 53(Pt 3): 240-55.

Musa, T. L., T. R. Ioerger, et al. (2009). "The tuberculosis structural genomics consortium: a structural genomics approach to drug discovery." Adv Protein Chem Struct Biol 77: 41-76.

Navaza, J. (1994). "AMoRe: an automated package for molecular replacement." Acta Crystallographica Section A 50(2): 157-163.

Navaza, J. (2001). "Implementation of molecular replacement in AMoRe." Acta Crystallogr D Biol Crystallogr 57(Pt 10): 1367-72.

Nicholls, R. A., F. Long, et al. (2012). "Low-resolution refinement tools in REFMAC5." Acta Crystallographica Section D 68(4): 404-417.

Nigou, J., L. G. Dover, et al. (2002). "Purification and biochemical characterization of Mycobacterium tuberculosis SuhB, an inositol monophosphatase involved in inositol biosynthesis." Biochemistry 41(13): 4392-8.

Nishimasu, H., S. Fushinobu, et al. (2004). "The first crystal structure of the novel class of fructose-1,6-bisphosphatase present in thermophilic archaea." Structure 12(6): 949-59.

Noens, E. E., C. Williams, et al. (2011). "Improved mycobacterial protein production using a Mycobacterium smegmatis groEL1DeltaC expression strain." BMC Biotechnol 11: 27.

Otwinowski, Z. and W. Minor (1997). Methods Enzymol. 276: 307–326.

Parish, T., B. G. Gordhan, et al. (1999). "Production of mutants in amino acid biosynthesis genes of Mycobacterium tuberculosis by homologous recombination." Microbiology 145 (Pt 12): 3497-503.

Parish, T. and N. G. Stoker (2000). "Use of a flexible cassette method to generate a double unmarked Mycobacterium tuberculosis tlyA plcABC mutant by gene replacement." Microbiology 146 (Pt 8): 1969-75.

Parish, T. and N. G. Stoker (2002). "The common aromatic amino acid biosynthesis pathway is essential in Mycobacterium tuberculosis." Microbiology 148(Pt 10): 3069-77.

Pashley, C. A. and T. Parish (2003). "Efficient switching of mycobacteriophage L5-based integrating plasmids in Mycobacterium tuberculosis." FEMS Microbiol Lett 229(2): 211-5.

Patel, S., M. Martinez-Ripoll, et al. (2002). "Structural enzymology of Li(+)-sensitive/Mg(2+)-dependent phosphatases." J Mol Biol 320(5): 1087-94.

Pavelka, M. S., Jr. and W. R. Jacobs, Jr. (1999). "Comparison of the construction of unmarked deletion mutations in Mycobacterium smegmatis, Mycobacterium bovis bacillus Calmette-Guerin, and Mycobacterium tuberculosis H37Rv by allelic exchange." J Bacteriol 181(16): 4780-9.

Peng, Z. and D. P. Verma (1995). "A rice HAL2-like gene encodes a Ca(2+)-sensitive 3'(2'),5'-diphosphonucleoside 3'(2')-phosphohydrolase and complements yeast met22 and Escherichia coli cysQ mutations." J Biol Chem 270(49): 29105-10.

Perutz, M. F., M. G. Rossmann, et al. (1960). "Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-A. resolution, obtained by X-ray analysis." Nature 185(4711): 416-22.

Qin, H., J. Hu, et al. (2008). "Construction of a series of vectors for high throughput cloning and expression screening of membrane proteins from Mycobacterium tuberculosis." BMC Biotechnol 8: 51.

Rhodes, G. (2006). Chapter 6 - Obtaining Phases. Crystallography Made Crystal Clear (Third Edition). Burlington, Academic Press: 109-143.

Rittmann, D., S. Schaffer, et al. (2003). "Fructose-1,6-bisphosphatase from Corynebacterium glutamicum: expression and deletion of the fbp gene and biochemical characterization of the enzyme." Arch Microbiol 180(4): 285-92.

Rossmann, M. G. (1990). "The molecular replacement method." Acta Crystallogr A 46 (Pt 2): 73-82.

Rossmann, M. G. (2001). "Molecular replacement--historical background." Acta Crystallogr D Biol Crystallogr 57(Pt 10): 1360-6.

Rossmann, M. G. and D. M. Blow (1962). "The Detection of Subunits within the Crystallographic Asymmetric Unit." Acta Crystallographica 15: 24-31.

Rupp, B. (2003). "High-throughput crystallography at an affordable cost: the TB Structural Genomics Consortium Crystallization Facility." Acc Chem Res 36(3): 173-81.

Rupp, B., B. W. Segelke, et al. (2002). "The TB structural genomics consortium crystallization facility: towards automation from protein to electron density." Acta Crystallogr D Biol Crystallogr 58(Pt 10 Pt 1): 1514-8.

Russell, D. G., P. J. Cardona, et al. (2009). "Foamy macrophages and the progression of the human tuberculosis granuloma." Nat Immunol 10(9): 943-8.

Sareen, D., G. L. Newton, et al. (2003). "Mycothiol is essential for growth of Mycobacterium tuberculosis Erdman." J Bacteriol 185(22): 6736-40.

Sarker, M., C. Talcott, et al. (2012). "Combining cheminformatics methods and pathway analysis to identify molecules with whole-cell activity against Mycobacterium tuberculosis." Pharm Res 29(8): 2115-27.

Sassetti, C. M., D. H. Boyd, et al. (2001). "Comprehensive identification of conditionally essential genes in mycobacteria." Proc Natl Acad Sci U S A 98(22): 12712-7.

Sassetti, C. M., D. H. Boyd, et al. (2003). "Genes required for mycobacterial growth defined by high density mutagenesis." Mol Microbiol 48(1): 77-84.

Sassetti, C. M. and E. J. Rubin (2003). "Genetic requirements for mycobacterial survival during infection." Proc Natl Acad Sci U S A 100(22): 12989-94.

Seeliger, D. and B. L. de Groot (2010). "Ligand docking and binding site analysis with PyMOL and Autodock/Vina." J Comput Aided Mol Des 24(5): 417-22.

Sharma, V., S. Sharma, et al. (2000). "Structure of isocitrate lyase, a persistence factor of Mycobacterium tuberculosis." Nat Struct Biol 7(8): 663-8.

Singh, R., U. Manjunatha, et al. (2008). "PA-824 kills nonreplicating Mycobacterium tuberculosis by intracellular NO release." Science 322(5906): 1392-5.

Smith, C. V., C. C. Huang, et al. (2003). "Biochemical and structural studies of malate synthase from Mycobacterium tuberculosis." J Biol Chem 278(3): 1735-43.

Smith, C. V. and J. C. Sacchettini (2003). "Mycobacterium tuberculosis: a model system for structural genomics." Curr Opin Struct Biol 13(6): 658-64.

Stieglitz, K. A., K. A. Johnson, et al. (2002). "Crystal structure of a dual activity IMPase/FBPase (AF2372) from Archaeoglobus fulgidus. The story of a mobile loop." J Biol Chem 277(25): 22863-74.

Stover, C. K., P. Warrener, et al. (2000). "A small-molecule nitroimidazopyran drug candidate for the treatment of tuberculosis." Nature 405(6789): 962-6.

Sun, R., P. J. Converse, et al. (2004). "Mycobacterium tuberculosis ECF sigma factor sigC is required for lethality in mice and for the conditional expression of a defined gene set." Mol Microbiol 52(1): 25-38.

Terwilliger, T. C., M. S. Park, et al. (2003). "The TB structural genomics consortium: a resource for Mycobacterium tuberculosis biology." Tuberculosis (Edinb) 83(4): 223-49.

Thompson, J. D., D. G. Higgins, et al. (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucleic Acids Res 22(22): 4673-80.

Underwood, B. A., R. Vanderhaeghen, et al. (2006). "Simultaneous high-throughput recombinational cloning of open reading frames in closed and open configurations." Plant Biotechnol J 4(3): 317-24.

Vagin, A. and A. Teplyakov (1997). "MOLREP: an Automated Program for Molecular Replacement." Journal of Applied Crystallography 30(6): 1022-1025.

Vagin, A. and A. Teplyakov (2010). "Molecular replacement with MOLREP." Acta Crystallogr D Biol Crystallogr 66(Pt 1): 22-5.

Vagin, A. A., G. N. Murshudov, et al. (1998). "BLANC: the program suite for protein crystallography." Journal of Applied Crystallography 31(1): 98-102.

Vagin, A. A., R. A. Steiner, et al. (2004). "REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use." Acta Crystallogr D Biol Crystallogr 60(Pt 12 Pt 1): 2184-95.

van Poelje, P. D., S. C. Potter, et al. (2006). "Inhibition of fructose 1,6-bisphosphatase reduces excessive endogenous glucose production and attenuates hyperglycemia in Zucker diabetic fatty rats." Diabetes 55(6): 1747-54.

van Poelje, P. D., S. C. Potter, et al. (2011). "Fructose-1, 6-bisphosphatase inhibitors for reducing excessive endogenous glucose production in type 2 diabetes." Handb Exp Pharmacol(203): 279-301.

Venugopal, A., R. Bryk, et al. (2011). "Virulence of Mycobacterium tuberculosis depends on lipoamide dehydrogenase, a member of three multienzyme complexes." Cell Host Microbe 9(1): 21-31.

Walhout, A. J., G. F. Temple, et al. (2000). "GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes." Methods Enzymol 328: 575-92.

Watkins, H. A. and E. N. Baker (2008). "Cloning, expression, purification and preliminary crystallographic analysis of the RNase HI domain of the Mycobacterium tuberculosis protein Rv2228c as a maltose-binding protein fusion." Acta Crystallogr Sect F Struct Biol Cryst Commun 64(Pt 8): 746-9.

WHO (2008). World Health Organization. Global tuberculosis control: surveillance, planning, financing. WHO Report.

Winn, M. D., C. C. Ballard, et al. (2011). "Overview of the CCP4 suite and current developments." Acta Crystallogr D Biol Crystallogr 67(Pt 4): 235-42.

York, J. D., J. W. Ponder, et al. (1995). "Definition of a metal-dependent/Li(+)-inhibited phosphomonoesterase protein family based upon a conserved three-dimensional core structure." Proc Natl Acad Sci U S A 92(11): 5149-53.

Yoshida, T., A. Okuno, et al. (2008). "CS-917, a fructose 1,6-bisphosphatase inhibitor, improves postprandial hyperglycemia after meal loading in non-obese type 2 diabetic Goto-Kakizaki rats." Eur J Pharmacol 601(1-3): 192-7.

Youmans, A. S. and G. P. Youmans (1953). "Studies on the metabolism of Mycobacterium tuberculosis. II. The effect of compounds related to the Kreb's tricarboxylic acid cycle on the growth of Mycobacterium tuberculosis var. hominis." J Bacteriol 65(1): 96-9.

Youmans, A. S. and G. P. Youmans (1953). "Studies on the metabolism of Mycobacterium tuberculosis. III. The growth of Mycobacterium tuberculosis var. hominis in the presence of various intermediates of the dissimilation of glucose to pyruvic acid." J Bacteriol 65(1): 100-2.

Youmans, A. S. and G. P. Youmans (1954). "Studies on the metabolism of Mycobacterium tuberculosis. IV. The effect of fatty acids on the growth of M. tuberculosis var. hominis." J Bacteriol 67(6): 731-3.

Youmans, A. S. and G. P. Youmans (1954). "Studies on the metabolism of Mycobacterium tuberculosis. V. The effect of amino acids on the growth of M. tuberculosis var. hominis." J Bacteriol 67(6): 734-7.

Youmans, G. P. and A. S. Youmans (1953). "Studies on the metabolism of Mycobacterium tuberculosis. I. The effect of carbohydrates and alcohols on the growth of Mycobacterium tuberculosis var. hominis." J Bacteriol 65(1): 92-5.

Zahrt, T. C. and V. Deretic (2000). "An essential two-component signal transduction system in Mycobacterium tuberculosis." J Bacteriol 182(13): 3832-8.

Zhang, R., V. Villeret, et al. (1996). "Kinetics and Mechanisms of Activation and Inhibition of Porcine Liver Fructose-1,6-bisphosphatase by Monovalent Cationsâ€ " Biochemistry 35(9): 3038-3043.

APPENDICES

1. APPENDIX A: Certificate of Analysis for pET15b-*F. tularensis* GlpX construct.

Celtek Genes Celtek Bioscience, LLC, 1515 Elm Hill Pike, Suite 104, Nashville, TN 37210
Tel: 615-399-7010 Fax: 615-399-7090 E-mail: info@celtek-genes.com website: www.celtek-genes.com

PRODUCT INFORMATION
Quality Assurance Certificate
Order ID: CG-1252 Clone ID: N0376-1

QC Items	Specifications		Results
Sequencing Alignment	Sequencing results are consistent with the target		☑ Pass ☐ Fail
Vector Sequence	The flanking sequences of the cloning site are correct		☑ Pass ☐ Fail
Reading Frame	☑ Correct and consisted with client's requirement		☑ Pass ☐ Fail
Restriction Digests	The size of inserted fragment is right and no any contaminated bands		☑ Pass ☐ Fail
PCR Amplification	Correct and no any contaminated bands		☑ Pass ☐ Fail
DNA Quality	☑ Miniprep: 10 ug/tube ☐ Maxiprep: 50 ug/tube $OD_{260/280}$=1.8~2.0, No any contamination		☑ Pass ☐ Fail
Appearance	Clear, no foreign particles		☑ Pass ☐ Fail
orientation	T7-TER - gene - T7		
Gene Name:	NOT DISCLOSED		
Gene ID:	X3890L		
Order Date:	June 3, 2010		
Shipping Date:	July 12, 2010		
Gene length:	996 bp		
Sequencing No.:	2		
Vector:	pET15b	Cloning Site:	NdeI/BamHI

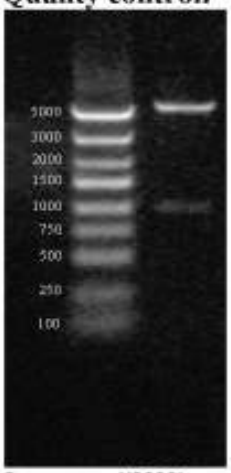
Note:
Plasmid-DNA was isolated from an E.coli strain containing methylases therefore methylation-sensitive restriction enzymes are blocked.

APPENDIX A (Continued)

Celtek Genes Celtek Bioscience, LLC, 1515 Elm Hill Pike, Suite 104, Nashville, TN 37210
Tel: 615-399-7010 Fax: 615-399-7090 E-mail: info@celtek-genes.com website: www.celtek-genes.com

Product of synthesis according to sequence chromatograms:
Sequencing showed correct sequences of the synthesis product.

Quality control:



Gene name: X3890L
Clone ID#:N0376-1
RES:BamHI/XbaI

Construct map: (Reverse complementary)

APPENDIX B

2. APPENDIX B: Sequence of the pET15b *F. tularensis* GlpX gene construct.

Celtek Genes Celtek Bioscience, LLC, 1515 Elm Hill Pike, Suite 104, Nashville, TN 37210
Tel: 615-399-7010 Fax: 615-399-7090 E-mail: info@celtek-genes.com website: www.celtek-genes.com

Detailed sequence of the whole construct: (Reverse complementary)

```
1     TTCTCATGTT TGACAGCTTA TCATCGATAA GCTTTAATGC GGTAGTTTAT CACAGTTAAA
61    TTGCTAACGC AGTCAGGCAC CGTGTATGAA ATCTAACAAT GCGCTCATCG TCATCCTCGG
121   CACCGTCACC CTGGATGCTG TAGGCATAGG CTTGGTTATG CCGGTACTGC CGGGCCTCTT
181   GCGGGATATC CGGATATAGT TCCTCCTTTC AGCAAAAAAC CCCTCAAGAC CCGTTTAGAG
241   GCCCCAAGGG GTTATGCTAG TTATTGCTCA GCGGTGGCAG CAGCCAACTC AGCTTCCTTT
301   CGGGCTTTGT TAGCAGCCGG ATCTTTAGCT CATGAATTTT TCAATAGCCT CCTTAAAGTC
361   AAAGCTATGT TCCGCGGTGA TGTGGCGAAC GGTTTTGGTG GTGCTACGCA TAACAACGCT
421   ATGGGTCACC GCATAGCTAC CACGAGCGCT GCTGTTTCAGA CGCTTAAGAC CCTGCAGCAT
481   ATTGCCGTCG GTCACGCCCG TCGCGGCAAA AACGATATCG CCGCTCGCCA GGTCGTCGAT
541   GTCATACTTT TTGTTCAGGT CCGTGATGCC CAGAGGGTGT GCGCGTTTAA TCTCCTCCTC
601   ATCGTTGAAG ATCAGACGCG CCTGCATCTG ACCGCCCAGG CATTTCAGCG CCGCTGCTGC
661   CAGAACACCT TCTGGCGCAC CGCCCGTACC GATGTAAACG TCAATGCCGC TGTTTTCGGT
721   TGCGGTCGCA ATCACACCGC TCACATCGCG ATCATTAATC AGAATCACGC GGGCACCACA
781   TTCACGCGCT TCTTTGATAA TATGTTGATG ACGCGGACGG TCCATGGTGC AAACAACCAG
841   GCCGCTCATG TGCAGGCCCT TAAATTGCGC AATACGCTTC AGATTGTTCG TCACGCTGTC
901   GTCCAGATCC ACGATGCCTT TCGGGCCGTT GATGCCACCA ACGGCGATTT TCTGCATATA
961   CACGTCCGGT GCGTTCAGGA AGCCGCCTTT ATCCGCCATT GCCAGCACGG TCAGTGCATT
1021  TGCGCCACCC TTGCTGGTAA TGTTCGTACC TTCCAGCGGG TCCAGTGCAA TGTCAACTTC
1081  GCAACCACCC GCACCACCTT TTTCGCTAAT ATACAGCATT GGGGCGTCAT CCAGTTCGCC
1141  TTCGCCGATC ACAACGGTGT CATCAATGTC CACCTCGTTC AGTGCTTTAC GCATCGCATC
1201  CACTGCCGCT TGATCTGCGG CGATCTTGTC GCCACGACCC ATCTGGCTGC AGCTTGCCAG
1261  TGCCGCCAGC TCGGTAACAC GCACGGCCTC CAGCGCCACT TTACGATTCA TATGGCTGCC
1321  GCGCGGCACC AGGCCGCTGC TGTGATGATG ATGATGATGG CTGCTGCCCA TGGTATATCT
1381  CCTTCTTAAA GTTAAACAAA ATTATTTCTA GAGGGGAATT GTTATCCGCT CACAATTCCC
1441  CTATAGTGAG TCGTATTAAT TTCGCGGGAT CGAGATCTCG ATCCTCTACG CCGGACGCAT
1501  CGTGGCCGGC ATCACCGGCG CCACAGGTGC GGTTGCTGGC GCCTATATCG CCGACATCAC
1561  CGATGGGGAA GATCGGGCTC GCCACTTCGG GCTCATGAGC GCTTGTTTCG GCGTGGGTAT
1621  GGTGGCAGGC CCCGTGGCCG GGGGACTGTT GGGCGCCATC TCCTTGCATG CACCATTCCT
1681  TGCGGCGGCG GTGCTCAACG GCCTCAACCT ACTACTGGGC TGCTTCCTAA TGCAGGAGTC
1741  GCATAAGGGA GAGCGTCGAG ATCCCGGACA CCATCGAATG GCGCAAAACC TTTCGCGGTA
1801  TGGCATGATA GCGCCCGGAA GAGAGTCAAT TCAGGGTGGT GAATGTGAAA CCAGTAACGT
1861  TATACGATGT CGCAGAGTAT GCCGGTGTCT CTTATCAGAC CGTTTCCCGC GTGGTGAACC
1921  AGGCCAGCCA CGTTTCTGCG AAAACGCGGG AAAAAGTGGA AGCGGCGATG GCGGAGCTGA
1981  ATTACATTCC CAACCGCGTG GCACAACAAC TGGCGGGCAA ACAGTCGTTG CTGATTGGCG
2041  TTGCCACCTC CAGTCTGGCC CTGCACGCGC CGTCGCAAAT TGTCGCGGCG ATTAAATCTC
2101  GCGCCGATCA ACTGGGTGCC AGCGTGGTGG TGTCGATGGT AGAACGAAGC GGCGTCGAAG
2161  CCTGTAAAGC GGCGGTGCAC AATCTTCTCG CGCAACGCGT CAGTGGGCTG ATCATTAACT
2221  ATCCGCTGGA TGACCAGGAT GCCATTGCTG TGGAAGCTGC CTGCACTAAT GTTCCGGCGT
2281  TATTTCTTGA TGTCTCTGAC CAGACACCCA TCAACAGTAT TATTTTCTCC CATGAAGACG
2341  GTACGCGACT GGGCGTGGAG CATCTGGTCG CATTGGGTCA CCAGCAAATC GCGCTGTTAG
2401  CGGGCCCATT AAGTTCTGTC TCGGCGCGTC TGCGTCTGGC TGGCTGGCAT AAATATCTCA
2461  CTCGCAATCA AATTCAGCCG ATAGCGGAAC GGGAAGGCGA CTGGAGTGCC ATGTCCGGTT
2521  TTCAACAAAC CATGCAAATG CTGAATGAGG GCATCGTTCC CACTGCGATG CTGGTTGCCA
2581  ACGATCAGAT GGCGCTGGGC GCAATGCGCG CCATTACCGA GTCCGGGCTG CGCGTTGGTG
2641  CGGATATCTC GGTAGTGGGA TACGACGATA CCGAAGACAG CTCATGTTAT ATCCCGCCGT
2701  TAACCACCAT CAAACAGGAT TTTCGCCTGC TGGGGCAAAC CAGCGTGGAC CGCTTGCTGC
2761  AACTCTCTCA GGGCCAGGCG GTGAAGGGCA ATCAGCTGTT GCCCGTCTCA CTGGTGAAAA
2821  GAAAAACCAC CCTGGCGCCC AATACGCAAA CCGCCTCTCC CCGCGCGTTG GCCGATTCAT
2881  TAATGCAGCT GGCACGACAG GTTTCCCGAC TGGAAAGCGG GCAGTGAGCG CAACGCAATT
2941  AATGTAAGTT AGCTCACTCA TTAGGCACCG GGATCTGAC CGATGCCCTT GAGAGCCTTC
3001  AACCCAGTCA GCTCCTTCCG GTGGGCGCGG GGCATGACTA TCGTCGCCGC ACTTATGACT
3061  GTCTTCTTTA TCATGCAACT CGTAGGACAG GTGCCGGCAG CGCTCTGGGT CATTTTCGGC
```

APPENDIX B (Continued)

Celtek Genes Celtek Bioscience, LLC, 1515 Elm Hill Pike, Suite 104, Nashville, TN 37210
Tel: 615-399-7010 Fax: 615-399-7090 E-mail: info@celtek-genes.com website: www.celtek-genes.com

```
3121 GAGGACCGCT TTCGCTGGAG CGCGACGATG ATCGGCCTGT CGCTTGCGGT ATTCGGAATC
3181 TTGCACGCCC TCGCTCAAGC CTTCGTCACT GGTCCCGCCA CCAAACGTTT CGGCGAGAAG
3241 CAGGCCATTA TCGCCGGCAT GGCGGCCGAC GCGCTGGGCT ACGTCTTGCT GGCGTTCGCG
3301 ACGCGAGGCT GGATGGCCTT CCCCATTATG ATTCTTCTCG CTTCCGGCGG CATCGGGATG
3361 CCCGCGTTGC AGGCCATGCT GTCCAGGCAG GTAGATGACG ACCATCAGGG ACAGCTTCAA
3421 GGATCGCTCG CGGCTCTTAC CAGCCTAACT TCGATCACTG GACCGCTGAT CGTCACGGCG
3481 ATTTATGCCG CCTCGGCGAG CACATGGAAC GGGTTGGCAT GGATTGTAGG CGCCGCCCTA
3541 TACCTTGTCT GCCTCCCCGC GTTGCGTCGC GGTGCATGGA GCCGGGCCAC CTCGACCTGA
3601 ATGGAAGCCG GCGGCACCTC GCTAACGGAT TCACCACTCC AAGAATTGGA GCCAATCAAT
3661 TCTTGCGGAG AACTGTGAAT GCGCAAACCA ACCCTTGGCA GAACATATCC ATCGCGTCCG
3721 CCATCTCCAG CAGCCGCACG CGGCGCATCT CGGGCAGCGT TGGGTCCTGG CCACGGGTGC
3781 GCATGATCGT GCTCCTGTCG TTGAGGACCC GGCTAGGCTG GCGGGGTTGC CTTACTGGTT
3841 AGCAGAATGA ATCACCGATA CGCGAGCGAA CGTGAAGCGA CTGCTGCTGC AAAACGTCTG
3901 CGACCTGAGC AACAACATGA ATGGTCTTCG GTTTCCGTGT TTCGTAAAGT CTGGAAACGC
3961 GGAAGTCAGC GCCCTGCACC ATTATGTTCC GGATCTGCAT CGCAGGATGC TGCTGGCTAC
4021 CCTGTGGAAC ACCTACATCT GTATTAACGA AGCGCTGGCA TTGACCCTGA GTGATTTTTC
4081 TCTGGTCCCG CCGCATCCAT ACCGCCAGTT GTTTACCCTC ACAACGTTCC AGTAACCGGG
4141 CATGTTCATC ATCAGTAACC CGTATCGTGA GCATCCTCTC TCGTTTCATC GGTATCATTA
4201 CCCCCATGAA CAGAAATCCC CCTTACACGG AGGCATCAGT GACCAAACAG GAAAAAACCG
4261 CCCTTAACAT GGCCCGCTTT ATCAGAAGCC AGACATTAAC GCTTCTGGAG AAACTCAACG
4321 AGCTGGACGC GGATGAACAG GCAGACATCT GTGAATCGCT TCACGACCAC GCTGATGAGC
4381 TTTACCGCAG CTGCCTCGCG CGTTTCGGTG ATGACGGTGA AAACCTCTGA CACATGCAGC
4441 TCCCGGAGAC GGTCACAGCT TGTCTGTAAG CGGATGCCGG GAGCAGACAA GCCCGTCAGG
4501 GCGCGTCAGC GGGTGTTGGC GGGTGTCGGG GCGCAGCCAT GACCCAGTCA CGTAGCGATA
4561 GCGGAGTGTA TACTGGCTTA ACTATGCGGC ATCAGAGCAG ATTGTACTGA GAGTGCACCA
4621 TATATGCGGT GTGAAATACC GCACAGATGC GTAAGGAGAA AATACCGCAT CAGGCGCTCT
4681 TCCGCTTCCT CGCTCACTGA CTCGCTGCGC TCGGTCGTTC GGCTGCGGCG AGCGGTATCA
4741 GCTCACTCAA AGGCGGTAAT ACGGTTATCC ACAGAATCAG GGGATAACGC AGGAAAGAAC
4801 ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA AGGCCGCGTT GCTGGCGTTT
4861 TTCCATAGGC TCCGCCCCCC TGACGAGCAT CACAAAAATC GACGCTCAAG TCAGAGGTGG
4921 CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC
4981 TCTCCTGTTC CGACCCTGCC GCTTACCGGA TACCTGTCCG CCTTTCTCCC TTCGGGAAGC
5041 GTGGCGCTTT CTCATAGCTC ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC
5101 AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC GCTGCGCCTT ATCCGGTAAC
5161 TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC CACTGGCAGC AGCCACTGGT
5221 AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG AGTTCTTGAA GTGGTGGCCT
5281 AACTACGGCT ACACTAGAAG GACAGTATTT GGTATCTGCG CTCTGCTGAA GCCAGTTACC
5341 TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA CCACCGCTGG TAGCGGTGGT
5401 TTTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAAG GATCTCAAGA AGATCCTTTG
5461 ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT CACGTTAAGG GATTTTGGTC
5521 ATGAGATTAT CAAAAAGGAT CTTCACCTAG ATCCTTTTAA ATTAAAAATG AAGTTTTAAA
5581 TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT ACCAATGCTT AATCAGTGAG
5641 GCACCTATCT CAGCGATCTG TCTATTTCGT TCATCCATAG TTGCCTGACT CCCCGTCGTG
5701 TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA GTGCTGCAAT GATACCGCGA
5761 GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC AGCCAGCCGG AAGGGCCGAG
5821 CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT CTATTAATTG TTGCCGGGAA
5881 GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG TTGTTGCCAT TGCTGCAGGC
5941 ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA GCTCCGGTTC CCAACGATCA
6001 AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG TTAGCTCCTT CGGTCCTCCG
6061 ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA TGGTTATGGC AGCACTGCAT
6121 AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG TGACTGGTGA GTACTCAACC
6181 AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT CTTGCCCGGC GTCAACACGG
6241 GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA TCATTGGAAA ACGTTCTTCG
6301 GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA GTTCGATGTA ACCCACTCGT
6361 GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG TTTCTGGGTG AGCAAAAACA
```

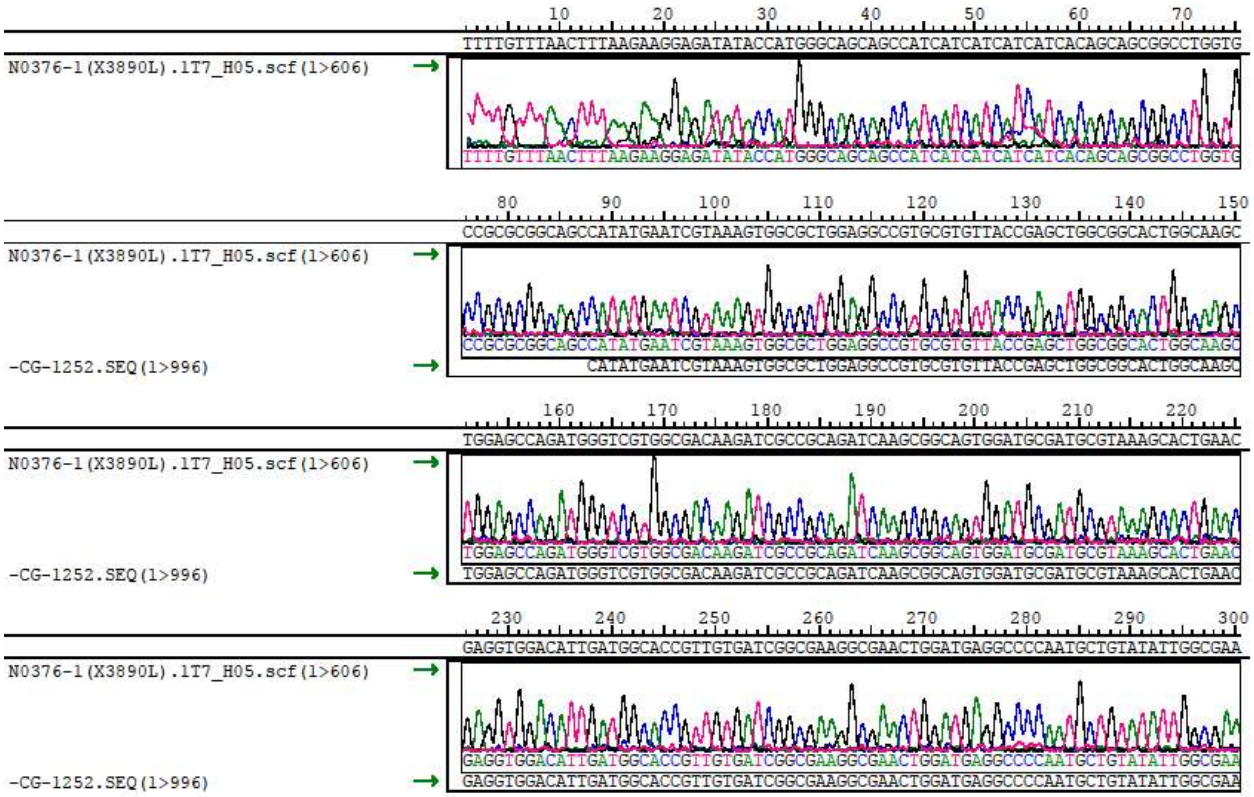
APPENDIX B (Continued)

Celtek Genes Celtek Bioscience, LLC, 1515 Elm Hill Pike, Suite 104, Nashville, TN 37210
Tel: 615-399-7010 Fax: 615-399-7090 E-mail: info@celtek-genes.com website: www.celtek-genes.com

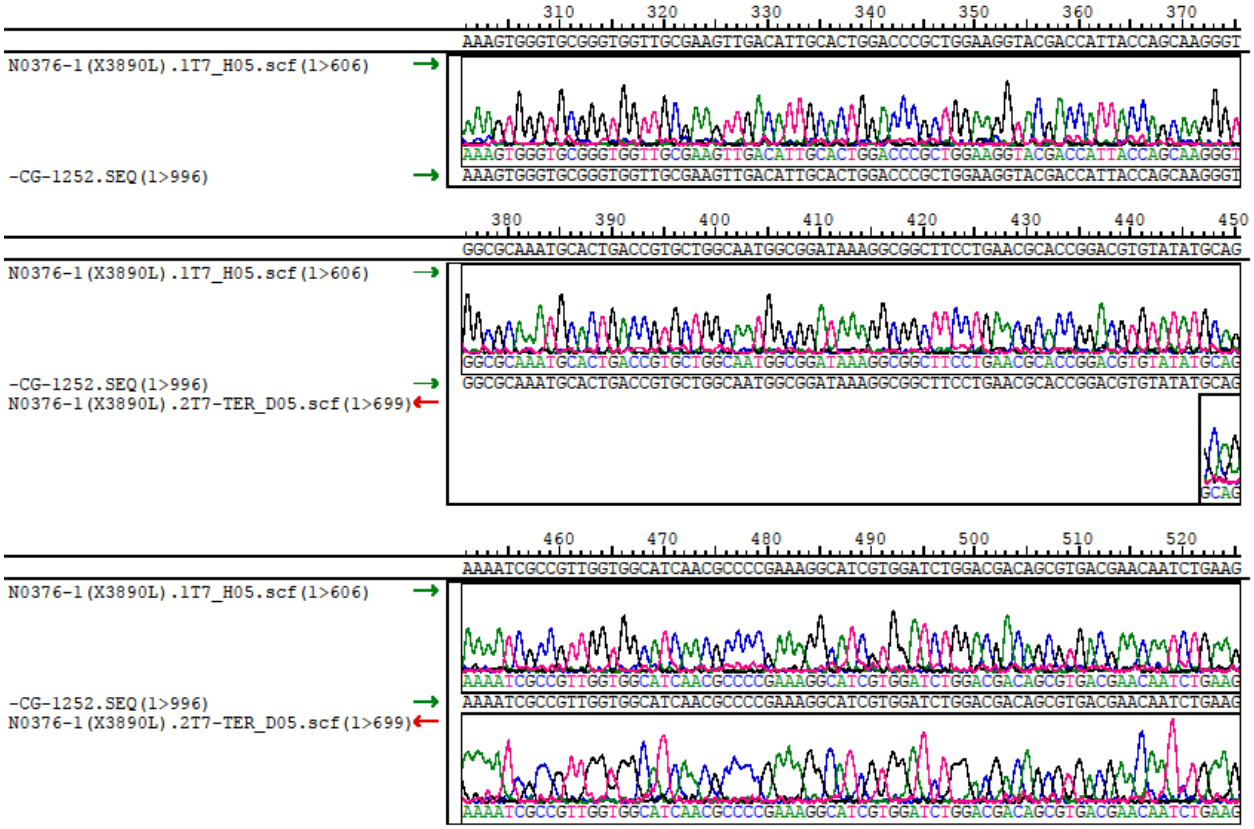
```
6421  GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC GGAAATGTTG AATACTCATA
6481  CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT ATTGTCTCAT GAGCGGATAC
6541  ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTC CGCGCACATT TCCCCGAAAA
6601  GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT TAACCTATAA AAATAGGCGT
6661  ATCACGAGGC CCTTTCGTCT TCAAGAA
```

APPENDIX C

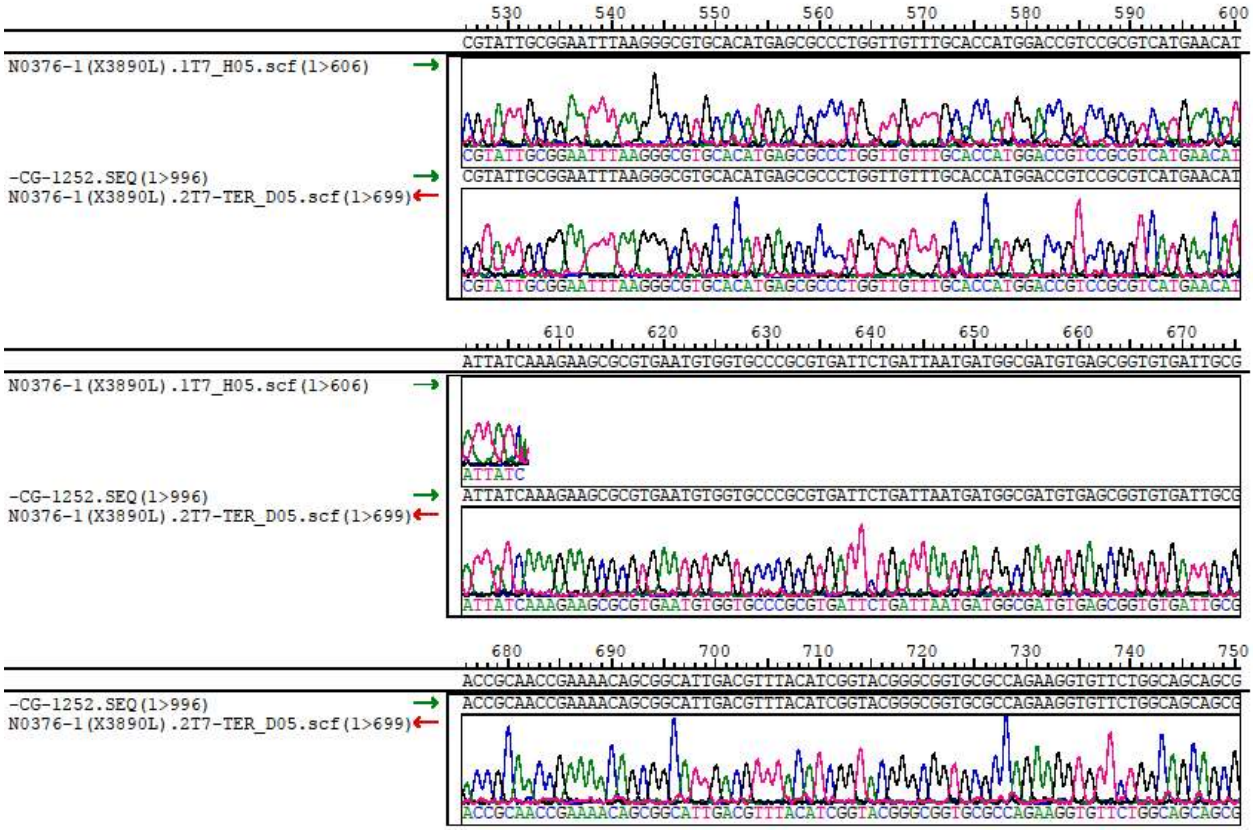
3. APPENDIX C: Raw data for sequencing of the pET15b *F. tularensis* GlpX construct.



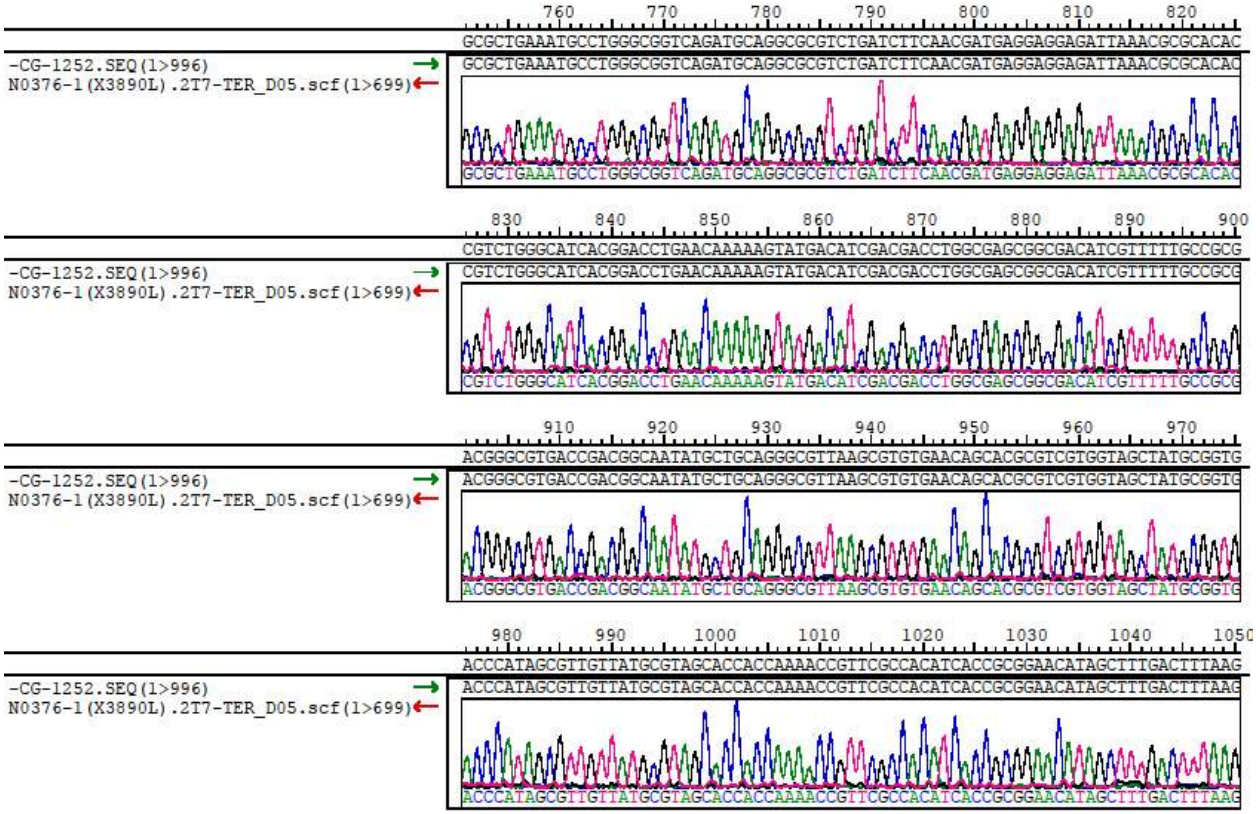
APPENDIX C (Continued)



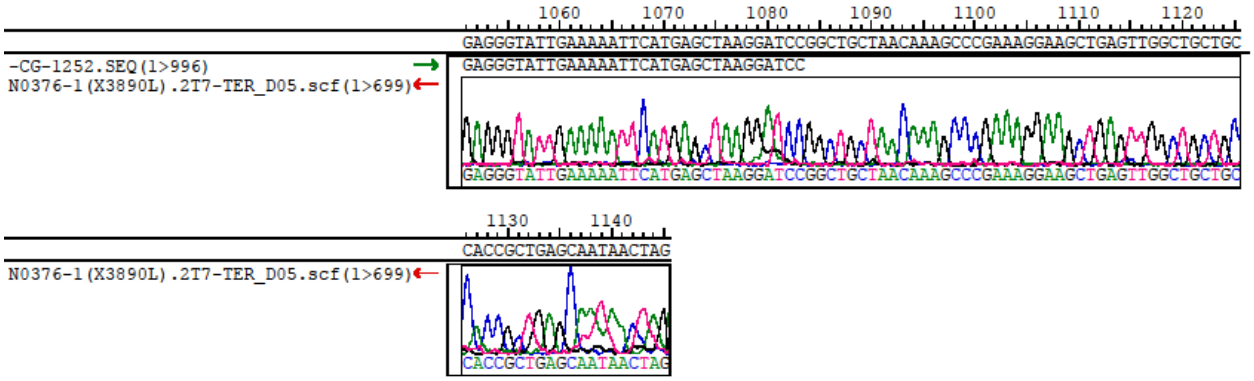
APPENDIX C (Continued)



APPENDIX C (Continued)



APPENDIX C (Continued)



APPENDIX D

4. APPENDIX D: Permissions for reproduction of text, figures and schemes in thesis.

Gmail by Google

Hiten Gutka <hiten1980@gmail.com>

Permission for reprint of article in my dissertation/thesis

2 messages

Hiten Gutka <hiten1980@gmail.com> **Tue, Dec 27, 2011 at 5:28 PM**
To: med@iucr.org

Dear Managing Editor,

I wished to request permission for reprint of the below mentioned article in my dissertation/thesis

- The title, author(s) and page extent of the article you wish to republish, or, in the case where you do not wish to reproduce the whole article, details of the section(s) to be reproduced.

- **Crystallization and preliminary X-ray characterization of the *glpX*-encoded class II fructose-1,6-bisphosphatase from *Mycobacterium tuberculosis***

  H. J. Gutka, S. G. Franzblau, F. Movahedzadeh and C. Abad-Zapatero

  I wish to reproduce the whole article.

- The journal title, volume/issue number and year of publication in which the article appeared, and the page numbers of the article.
- *Acta Cryst.* (2011). F**67**, 710-713 [doi:10.1107/S1744309111014722]
-
- Details of the publication in which you wish to reprint the Journal material.
- My thesis/dissertation entitled "Studies on fructose 1,6-bisphosphatase from pathogenic bacteria".
-
-
- Contact details for yourself, including a postal address.
- Contact details: Hiten J Gutka, Graduate Research Assistant, Room 412, Institute for Tuberculosis Research (M/C 964), 833 S Wood Street, Chicago IL 60612 USA.

I am a contributing author to the article and wished to get permission for reproducing it in my thesis/dissertation.

–
Thanks and Regards

Hiten J. Gutka
Graduate Student,
Centre for Pharmaceutical Biotechnology
& Institute for Tuberculosis Research,
College of Pharmacy,
University of Illinois at Chicago IL 60612
email: hgutka2@uic.edu
www.tuberculosisdrugresearch.org
An eye for an eye makes the whole world blind - Mahatma Gandhi

Peter Strickland <ps@iucr.org> **Wed, Dec 28, 2011 at 3:01 AM**

APPENDIX D (Continued)

Reply-To: ps@iucr.org
To: Hiten Gutka <hiten1980@gmail.com>
Cc: Peter Strickland <ps@iucr.org>

Dear Dr Gutka

Thank you for your email.

Permission is hereby granted, on behalf of the IUCr, for you to reproduce the material specified below, subject to the following conditions:

1. Reproduction is intended in a primary journal, secondary journal, CD-ROM, book or thesis.

2. The original article in which the material appeared is cited.

Best wishes
Peter Strickland

Managing Editor
IUCr Journals

IUCr Editorial Office, 5 Abbey Square, Chester CH1 2HU, England
Phone: 44 1244 342878 Fax: 44 1244 314888 Email: ps@iucr.org
Ftp: ftp.iucr.org WWW: http://journals.iucr.org/

> Dear Managing Editor,
>
> I wished to request permission for reprint of the below mentioned article
> in my dissertation/thesis
>
> - The title, author(s) and page extent of the article you wish to
> republish, or, in the case where you do not wish to reproduce the whole
> article, details of the section(s) to be reproduced.
> - Crystallization and preliminary X-ray characterization of the
> *glpX*-encoded
> class II fructose-1,6-bisphosphatase from *Mycobacterium tuberculosis*H.
> J.
Gutka<http://scripts.iucr.org/cgi-bin/citedin?search_on=name&author_name=Gutka,
%20H.J.>
> , S. G.
Franzblau<http://scripts.iucr.org/cgi-bin/citedin?search_on=name&author_name=Franzblau,
%20S.G.>
> , F.
Movahedzadeh<http://scripts.iucr.org/cgi-bin/citedin?search_on=name&author_name=Movahedzadeh,
%20F.>
> and C.
Abad-Zapatero<http://scripts.iucr.org/cgi-bin/citedin?search_on=name&author_name=Abad-Zapatero,
%20C.>
>
> I wish to reproduce the whole article.
>

APPENDIX D (Continued)

> - The journal title, volume/issue number and year of publication in
> which the article appeared, and the page numbers of the article.
> - *Acta Cryst.* (2011). F*67*, 710-713 [
> doi:10.1107/S1744309111014722<http://dx.doi.org/10.1107/S1744309111014722>
>]
> -
> - Details of the publication in which you wish to reprint the Journal
> material.
> - My thesis/dissertation entitled "Studies on fructose
> 1,6-bisphosphatase from pathogenic bacteria".
> -
> -
> - Contact details for yourself, including a postal address.
> - Contact details: Hiten J Gutka, Graduate Research Assistant, Room 412,
> Institute for Tuberculosis Research (M/C 964), 833 S Wood Street, Chicago
> IL 60612 USA.
>
> *I am a contributing author to the article and wished to get permission for
> reproducing it in my thesis/dissertation.*
> –
> Thanks and Regards
>
> Hiten J. Gutka
> Graduate Student,
> Centre for Pharmaceutical Biotechnology
> & Institute for Tuberculosis Research,
> College of Pharmacy,
> University of Illinois at Chicago IL 60612
> email: hgutka2@uic.edu
> www.tuberculosisdrugresearch.org
> **
> *An eye for an eye makes the whole world blind* - Mahatma Gandhi
>

APPENDIX D (Continued)

Copyright Clearance Center RightsLink®

Home | Account Info | Help

jbc

Title: Structural and Biochemical Characterization of the Type II Fructose-1,6-bisphosphatase GlpX from Escherichia coli

Author: Greg Brown, Alexander Singer, Vladimir V. Lunin, Michael Proudfoot, Tatiana Skarina, Robert Flick, Samvel Kochinyan, Ruslan Sanishvili, Andrzej Joachimiak, Aled M. Edwards, Alexei Savchenko, Alexander F. Yakunin

Publication: Journal of Biological Chemistry

Publisher: The American Society for Biochemistry and Molecular Biology

Date: Feb 6, 2009

Logged in as:
Hiten Gutka
Account #:
3000480188

LOGOUT

Download Letter of Permission for Reuse in Thesis/Dissertation

The publisher allows article reuse in thesis or dissertations at no charge. Click here to receive a letter of permission for this purpose.

BACK | CLOSE WINDOW

APPENDIX D (Continued)

ASBMB

American Society for Biochemistry and Molecular Biology

11200 Rockville Pike
Suite 302
Rockville, Maryland 20852

August 19, 2011

To whom it may concern,

It is the policy of the American Society for Biochemistry and Molecular Biology to allow reuse of any material published in its journals (the Journal of Biological Chemistry, Molecular & Cellular Proteomics and the Journal of Lipid Research) in a thesis or dissertation at no cost and with no explicit permission needed. Please see our copyright permissions page on the journal site for more information.

Best wishes,

Sarah Crespi

American Society for Biochemistry and Molecular Biology
11200 Rockville Pike, Rockville, MD
Suite 302
240-283-6616
JBC | MCP | JLR

Tel: 240-283-6600 • Fax: 240-881-2080 • E-mail: asbmb@asbmb.org

APPENDIX D (Continued)

Copyright Clearance Center

1 PAYMENT 2 REVIEW 3 CONFIRMATION

Step 3: Order Confirmation

Thank you for your order! A confirmation for your order will be sent to your account email address. If you have questions about your order, you can call us at 978-646-2600, M-F between 8:00 AM and 6:00 PM (Eastern), or write to us at info@copyright.com.

Confirmation Number: 11017714
Order Date: 08/04/2012

If you pay by credit card, your order will be finalized and your card will be charged within 24 hours. If you pay by invoice, you can change or cancel your order until the invoice is generated.

Payment Information

Hiten Gutka
hgutka2@uic.edu
+1 (312)2085883
Payment Method: CC ending in 5859

Order Details

Microbiology

Order detail ID:	62685846
ISSN:	1465-2080
Publication year:	2012
Publication Type:	e-Journal
Publisher:	SOCIETY FOR GENERAL MICROBIOLOGY
Rightsholder:	SOCIETY FOR GENERAL MICROBIOLOGY
Author/Editor:	Parish T, Stoker NG.
Your reference:	Hiten Gutka's thesis chapter 2

Permission Status:	**Granted**
Permission type:	Republish or display content
Type of use:	Republish in a dissertation
Republication title:	Studies on Fructose 1,6-bisphosphatases from pathogenic bacteria
Republishing organization:	University of Illinois at Chicago
Organization status:	Non-profit 501(c)(3)
Republication date:	03/31/2013
Circulation/ Distribution:	5
Type of content:	Figure/ diagram/ table
Description of requested content:	Use of a flexible cassette method to generate a double unmarked Mycobacterium tuberculosis tlyA plcABC mutant by gene replacement
Page range(s):	1971
Translating to:	No Translation
Requested content's publication date:	08/03/2012

Rightsholder terms apply (see terms and conditions)

$ 3.50

Total order items: 1 **Order Total: $ 3.50**

APPENDIX D (Continued)

APPENDIX D (Continued)

Copyright Clearance Center

Confirmation Number: 11017714
Order Date: 08/04/2012

Customer Information

Customer: Hiten Gutka
Account Number: 3000480188
Organization: Hiten Gutka
Email: hgutka2@uic.edu
Phone: +1 (312)2085883
Payment Method: Credit Card ending in 5859

Order Details

Microbiology

Billing Status:
Charged to Credit Card

Order detail ID:	62685846
ISSN:	1465-2080
Publication year:	2012
Publication Type:	e-Journal
Publisher:	SOCIETY FOR GENERAL MICROBIOLOGY
Rightsholder:	SOCIETY FOR GENERAL MICROBIOLOGY
Author/Editor:	Parish T, Stoker NG.
Your reference:	Hiten Gutka's thesis chapter 2

Permission Status:	**Granted**
Permission type:	Republish or display content
Type of use:	Republish in a dissertation
Requested use:	Dissertation
Republication title:	Studies on Fructose 1,6-bisphosphatases from pathogenic bacteria
Republishing organization:	University of Illinois at Chicago
Organization status:	Non-profit 501(c)(3)
Republication date:	03/31/2013
Circulation/ Distribution:	5
Type of content:	Figure/ diagram/ table
Description of requested content:	Use of a flexible cassette method to generate a double unmarked Mycobacterium tuberculosis thyA plcABC mutant by gene replacement
Page range(s):	1971
Translating to:	No Translation
Requested content's publication date:	08/03/2012
Payment Method:	CC ending in 5859

Rightsholder terms apply (see terms and conditions)

$ 3.50

Total order items: 1

Order Total: $3.50

APPENDIX D (Continued)

SPRINGER LICENSE
TERMS AND CONDITIONS

Jan 03, 2012

This is a License Agreement between Hiten J Gutka ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	2817280978811
License date	Dec 27, 2011
Licensed content publisher	Springer
Licensed content publication	Applied Biochemistry and Biotechnology
Licensed content title	*glpX* Gene of *Mycobacterium tuberculosis*: Heterologous Expression, Purification, and Enzymatic Characterization of the Encoded Fructose 1,6-bisphosphatase II
Licensed content author	Hiten J. Gutka
Licensed content date	Jan 1, 2011
Volume number	164
Issue number	8
Type of Use	Thesis/Dissertation
Portion	Figures
Author of this Springer article	Yes and you are a contributor of the new work
Order reference number	None
Title of your thesis / dissertation	studies on Fructose bisphosphatase from pathogenic bacteria
Expected completion date	Feb 2013
Estimated size(pages)	150
Total	**0.00 USD**

Terms and Conditions

Introduction

The publisher for this copyrighted material is Springer Science + Business Media. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

Limited License

With reference to your request to reprint in your thesis material on which Springer Science and Business Media control the copyright, permission is granted, free of charge, for the use indicated in your enquiry. Licenses are for one-time use only with a maximum distribution equal to the number that you identified in the licensing process.

This License includes use in an electronic form, provided it is password protected or on the university's intranet, destined to microfilming by UMI and University repository. For any other electronic use, please contact Springer at (permissions.dordrecht@springer.com or permissions.heidelberg@springer.com)

The material can only be used for the purpose of defending your thesis, and with a maximum of 100 extra copies in paper.

APPENDIX D (Continued)

Although Springer holds copyright to the material and is entitled to negotiate on rights, this license is only valid, provided permission is also obtained from the (co) author (address is given with the article/chapter) and provided it concerns original material which does not carry references to other sources (if material in question appears with credit to another source, authorization from that source is required as well). Permission free of charge on this occasion does not prejudice any rights we might have to charge for reproduction of our copyrighted material in the future.

Altering/Modifying Material: Not Permitted

However figures and illustrations may be altered minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of the author(s) and/or Springer Science + Business Media. (Please contact Springer at permissions.dordrecht@springer.com or permissions.heidelberg@springer.com)

Reservation of Rights

Springer Science + Business Media reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

Copyright Notice:

Please include the following copyright citation referencing the publication in which the material was originally published. Where wording is within brackets, please include verbatim.

"With kind permission from Springer Science+Business Media: <book/journal title, chapter/article title, volume, year of publication, page, name(s) of author(s), figure number(s), and any original (first) copyright notice displayed with material>."

Warranties: Springer Science + Business Media makes no representations or warranties with respect to the licensed material.

Indemnity

You hereby indemnify and agree to hold harmless Springer Science + Business Media and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

No Transfer of License

This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without Springer Science + Business Media's written permission.

No Amendment Except in Writing

This license may not be amended except in a writing signed by both parties (or, in the case of Springer Science + Business Media, by CCC on Springer Science + Business Media's behalf).

Objection to Contrary Terms

Springer Science + Business Media hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and Springer Science + Business Media (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

Jurisdiction

All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by the country's law in which the work was originally published.

Other terms and conditions:

v1.2

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK500689818. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

APPENDIX D (Continued)

For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

APPENDIX D (Continued)

SPRINGER LICENSE
TERMS AND CONDITIONS

Aug 12, 2012

This is a License Agreement between Hiten J Gutka ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	2966691312044
License date	Aug 12, 2012
Licensed content publisher	Springer
Licensed content publication	Pharmaceutical Research
Licensed content title	Combining Cheminformatics Methods and Pathway Analysis to Identify Molecules with Whole-Cell Activity Against *Mycobacterium Tuberculosis*
Licensed content author	Malabika Sarker
Licensed content date	Jan 1, 2012
Volume number	29
Issue number	8
Type of Use	Thesis/Dissertation
Portion	Figures
Author of this Springer article	No
Order reference number	
Title of your thesis / dissertation	studies on Fructose bisphosphatase from pathogenic bacteria
Expected completion date	Feb 2013
Estimated size(pages)	150
Total	0.00 USD

Terms and Conditions

Introduction

The publisher for this copyrighted material is Springer Science + Business Media. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

Limited License

With reference to your request to reprint in your thesis material on which Springer Science and

APPENDIX D (Continued)

Business Media control the copyright, permission is granted, free of charge, for the use indicated in your enquiry.

Licenses are for one-time use only with a maximum distribution equal to the number that you identified in the licensing process.

This License includes use in an electronic form, provided its password protected or on the university's intranet or repository, including UMI (according to the definition at the Sherpa website: http://www.sherpa.ac.uk/romeo/). For any other electronic use, please contact Springer at (permissions.dordrecht@springer.com or permissions.heidelberg@springer.com).

The material can only be used for the purpose of defending your thesis, and with a maximum of 100 extra copies in paper.

Although Springer holds copyright to the material and is entitled to negotiate on rights, this license is only valid, provided permission is also obtained from the (co) author (address is given with the article/chapter) and provided it concerns original material which does not carry references to other sources (if material in question appears with credit to another source, authorization from that source is required as well).

Permission free of charge on this occasion does not prejudice any rights we might have to charge for reproduction of our copyrighted material in the future.

Altering/Modifying Material: Not Permitted
You may not alter or modify the material in any manner. Abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of the author(s) and/or Springer Science + Business Media. (Please contact Springer at (permissions.dordrecht@springer.com or permissions.heidelberg@springer.com)

Reservation of Rights
Springer Science + Business Media reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

Copyright Notice:Disclaimer
You must include the following copyright and permission notice in connection with any reproduction of the licensed material: "Springer and the original publisher /journal title, volume, year of publication, page, chapter/article title, name(s) of author(s), figure number(s), original copyright notice) is given to the publication in which the material was originally published, by adding; with kind permission from Springer Science and Business Media"

Warranties: None

Example 1: Springer Science + Business Media makes no representations or warranties with respect to the licensed material.

Example 2: Springer Science + Business Media makes no representations or warranties with respect to the licensed material and adopts on its own behalf the limitations and disclaimers

APPENDIX D (Continued)

established by CCC on its behalf in its Billing and Payment terms and conditions for this licensing transaction.

Indemnity
You hereby indemnify and agree to hold harmless Springer Science + Business Media and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

No Transfer of License
This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without Springer Science + Business Media's written permission.

No Amendment Except in Writing
This license may not be amended except in a writing signed by both parties (or, in the case of Springer Science + Business Media, by CCC on Springer Science + Business Media's behalf).

Objection to Contrary Terms
Springer Science + Business Media hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and Springer Science + Business Media (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

Jurisdiction
All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by arbitration, to be held in The Netherlands, in accordance with Dutch law, and to be conducted under the Rules of the 'Netherlands Arbitrage Instituut' (Netherlands Institute of Arbitration).***OR:***

All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by arbitration, to be held in the Federal Republic of Germany, in accordance with German law.

Other terms and conditions:

v1.3

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK500837083. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center

APPENDIX D (Continued)

ELSEVIER LICENSE TERMS AND CONDITIONS

Dec 27, 2011

This is a License Agreement between Hiten J Gutka ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Hiten J Gutka
Customer address	#1109, 1926 W Harrison Street Chicago, IL 60612
License number	2813371419474
License date	Dec 20, 2011
Licensed content publisher	Elsevier
Licensed content publication	Chemistry & Biology
Licensed content title	Metabolomics of *Mycobacterium tuberculosis* Reveals Compartmentalized Co-Catabolism of Carbon Substrates
Licensed content author	Luiz Pedro S. de Carvalho,Steven M. Fischer,Joeli Marrero,Carl Nathan,Sabine Ehrt,Kyu Y. Rhee
Licensed content date	29 October 2010
Licensed content volume number	17
Licensed content issue number	10
Number of pages	10
Start Page	1122
End Page	1131
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Order reference number	None
Title of your thesis/dissertation	studies on Fructose bisphosphatase from pathogenic bacteria

APPENDIX D (Continued)

Expected completion date	Feb 2013
Estimated size (number of pages)	150
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
VAT/Local Sales Tax	0.0 USD / 0.0 GBP
Total	**0.00 USD**

Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.
3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:
"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."
4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.
5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)
6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.
7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.
9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.
10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.
11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.
12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).
13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you

APPENDIX D (Continued)

and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages.

16. **Website**: The following terms and conditions apply to electronic reserve and author websites:

Electronic reserve: If licensed material is to be posted to website, the web site is to be password-protected and made available only to bona fide students registered on a relevant course if:

This license was made in connection with a course,

This permission is granted for 1 year only. You may obtain a license for future website posting,

All content posted to the web site must maintain the copyright information line on the bottom of each image,

A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxx or the Elsevier homepage for books at http://www.elsevier.com , and

Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

17. **Author website** for journals with the following additional clauses:

All content posted to the web site must maintain the copyright information line on the bottom of each image, and

he permission granted is limited to the personal version of your paper. You are not allowed to download and post the published electronic version of your article (whether PDF or HTML, proof or final version), nor may you scan the printed edition to create an electronic version,

A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxx . As part of our normal production process, you will receive an e-mail notice when your article appears on Elsevier's online service ScienceDirect (www.sciencedirect.com). That e-mail will include the article's Digital Object Identifier (DOI). This number provides the electronic link to the published article and should be included in the posting of your personal version. We ask that you wait until you receive this e-mail and have the DOI to do any posting.

Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

18. **Author website** for books with the following additional clauses:

Authors are permitted to place a brief summary of their work online only.

A hyper-text must be included to the Elsevier homepage at http://www.elsevier.com

All content posted to the web site must maintain the copyright information line on the bottom of each image

You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version.

Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

19. **Website** (regular and for author): A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxx. or for books to the Elsevier homepage at http://www.elsevier.com

20. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

21. **Other Conditions**:

APPENDIX D (Continued)

NATURE PUBLISHING GROUP LICENSE
TERMS AND CONDITIONS

Oct 08, 2012

This is a License Agreement between Hiten J Gutka ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3004470503755
License date	Oct 08, 2012
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature Methods
Licensed content title	Protein crystallization: from purified protein to diffraction-quality crystal
Licensed content author	Naomi E Chayen and Emmanuel Saridakis
Licensed content date	Feb 1, 2008
Volume number	5
Issue number	2
Type of Use	reuse in a thesis/dissertation
Requestor type	non-commercial (non-profit)
Format	electronic
Portion	Full article
Circulation/distribution	<501
Author of this NPG article	no
Your reference number	crystallization image
Title of your thesis / dissertation	studies on Fructose bisphosphatase from pathogenic bacteria
Expected completion date	Feb 2013
Estimated size (number of pages)	150
Total	0.00 USD
Terms and Conditions	

Terms and Conditions for Permissions

Nature Publishing Group hereby grants you a non-exclusive license to reproduce this material for this purpose, and for no other use, subject to the conditions below:

APPENDIX D (Continued)

1. NPG warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to Nature Publishing Group and does not carry the copyright of another entity (as credited in the published version). If the credit line on any part of the material you have requested indicates that it was reprinted or adapted by NPG with permission from another source, then you should also seek permission from that source to reuse the material.

2. Permission granted free of charge for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to the work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version. Where print permission has been granted for a fee, separate permission must be obtained for any additional, electronic re-use (unless, as in the case of a full paper, this has already been accounted for during your initial request in the calculation of a print run). NB: In all cases, web-based use of full-text articles must be authorized separately through the 'Use on a Web Site' option when requesting permission.

3. Permission granted for a first edition does not apply to second and subsequent editions and for editions in other languages (except for signatories to the STM Permissions Guidelines, or where the first edition permission was granted for free).

4. Nature Publishing Group's permission must be acknowledged next to the figure, table or abstract in print. In electronic form, this acknowledgement must be visible at the same time as the figure/table/abstract, and must be hyperlinked to the journal's homepage.

5. The credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)
   For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

   Note: For republication from the *British Journal of Cancer*, the following credit lines apply.
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication) For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

6. Adaptations of single figures do not require NPG approval. However, the adaptation should be credited as follows:

APPENDIX D (Continued)

Adapted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)

Note: For adaptation from the *British Journal of Cancer*, the following credit line applies.
Adapted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

7. Translations of 401 words up to a whole article require NPG approval. Please visit http://www.macmillanmedicalcommunications.com for more information. Translations of up to a 400 words do not require NPG approval. The translation should be credited as follows:

   Translated by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication).

   Note: For translation from the *British Journal of Cancer*, the following credit line applies.
   Translated by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

We are certain that all parties will benefit from this agreement and wish you the best in the use of this material. Thank you.

Special Terms:

v1.1

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK500872912. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

APPENDIX D (Continued)

ELSEVIER LICENSE
TERMS AND CONDITIONS

Oct 08, 2012

This is a License Agreement between Hiten J Gutka ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Hiten J Gutka
Customer address	#1109, 1926 W Harrison Street
	Chicago, IL 60612
License number	3004480253952
License date	Oct 08, 2012
Licensed content publisher	Elsevier
Licensed content publication	Current Opinion in Structural Biology
Licensed content title	Turning protein crystallisation from an art into a science
Licensed content author	Naomi E Chayen
Licensed content date	October 2004
Licensed content volume number	14
Licensed content issue number	5
Number of pages	7
Start Page	577
End Page	583
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this	No

APPENDIX D (Continued)

advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

APPENDIX D (Continued)

advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

APPENDIX D (Continued)

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages.

16. **Website**: The following terms and conditions apply to electronic reserve and author websites:
Electronic reserve: If licensed material is to be posted to website, the web site is to be password-protected and made available only to bona fide students registered on a relevant course if:
This license was made in connection with a course,
This permission is granted for 1 year only. You may obtain a license for future website posting,
All content posted to the web site must maintain the copyright information line on the bottom of each image,
A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com , and
Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

17. **Author website** for journals with the following additional clauses:

All content posted to the web site must maintain the copyright information line on the bottom of each image, and the permission granted is limited to the personal version of your paper. You are not allowed to download and post the published electronic version of your article (whether PDF or HTML, proof or final version), nor may you scan the printed edition to create an electronic version. A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx . As part of our normal production process, you will receive an e-mail notice when your article appears on Elsevier's online service ScienceDirect (www.sciencedirect.com). That e-mail will include the article's Digital Object Identifier (DOI). This number provides the electronic link to the published article and should be included in the posting of your personal version. We ask that you wait until you receive this e-mail and have the DOI to do any posting.

Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

18. **Author website** for books with the following additional clauses:
Authors are permitted to place a brief summary of their work online only.
A hyper-text must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.
You are not allowed to download and post the published electronic version of your chapter, nor

APPENDIX D (Continued)

may you scan the printed edition to create an electronic version.

Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

19. **Website** (regular and for author): A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx. or for books to the Elsevier homepage at http://www.elsevier.com

20. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

21. **Other Conditions**:

v1.6

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK500872917. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

VITA

NAME: Hiten Jayanti Gutka

EDUCATION:

- B. Pharm. Sci. University of Mumbai, Mumbai, 2002
- M. Pharm. Sci. (Medicinal Natural Products) University of Mumbai, Mumbai, 2004
- Ph.D., Pharmacognosy (Pharmaceutical Biotechnology) University of Illinois at Chicago, Chicago, 2012

WORK EXPERIENCE:

- June 2009- December 2012 Graduate Research Assistant, University of Illinois at Chicago, IL
- June 2012 – Aug 2012 Summer Intern Protein Formulations, Formulation Sciences, Medimmune LLC, Gaithersburg, MD
- Aug 2008- May 2009 Graduate Teaching Assistant, University of Illinois at Chicago, IL
- May 2007-Jul 2008 Sr. Executive Biopharmaceuticals (Protein Analytics & Quality), Reliance Life Sciences, Navi Mumbai, INDIA
- Feb 2006 to May 2007 Sr. Executive Biologics QA, USV Limited, Mumbai ,INDIA
- Sep 2004 to Jan 2006 Executive QA Biopharmaceuticals, Biocon Limited, Bangalore, INDIA

AWARDS AND HONOURS:

- AMGEN's Travelship, National Biotechnology Conference, San Diego: May 2012
- UIC Graduate College W.C. and May Deiss Award for Biomedical Research: Fall 2010
- UIC Graduate College Image of Research 2010 competition Finalist: Spring 2010
- UIC College of Pharmacy Image of Research 2011 competition, Third place: Fall 2011
- Bill and Melinda Gates foundation global health travel award, Keystone symposia (J4) Mycobacteria: Physiology, Metabolism and Pathogenesis - Back to the Basics. Vancouver Canada Jan 2011.
- University Grants Commission INDIA, University Institute of Chemical Technology, Mumbai INDIA, Junior Research Fellowship: Fall 2002 – Summers 2004

PROFESSIONAL MEMBERSHIP:

- Member, American Association of Pharmaceutical Scientist (AAPS)
- Member, American Crystallographic Association (ACA)
- Life Member, Indian Pharmaceutical Association (IPA)
- Registered Pharmacist The Pharmacy Council of India Maharashtra State Branch.

PUBLICATIONS:

Gutka H. J., Rukseree, K., Wheeler, P. R., Franzblau, S. G., Movahedzadeh. F. *glpX* Gene of *Mycobacterium tuberculosis*: Heterologous Expression, Purification, and Enzymatic Characterization of the Encoded Fructose 1,6-bisphosphatase II. Appl Biochem Biotechnol (2011), Aug; 164(8):1376-89.

Gutka, H.J., Franzblau, S.G., Movahedzadeh, F., Abad-Zapatero, C. Crystallization and preliminary X-ray characterization of the glpX-encoded class II fructose-1,6-bisphosphatase from *Mycobacterium tuberculosis.* Acta Cryst Section F Structural Biology and Crystallization Communications (2011), 67 (Pt 6):710-3.

Movahedzadeh, F. Frita, R. **Gutka, H.J.** A two-step strategy for the complementation of *M. tuberculosis* mutants. Genetics and Molecular Biology (2011), 34(2), 286-289.

SCIENTIFIC COMMUNICATIONS:

Poster Presentations

Structure of the Fructose 1,6-bisphosphatase (II) from M. Tuberculosis: Insights into target based drug discovery. National Biotechnology Conference, San Diego. May 2012.

Structural Characterization of *glpX* encoded Fructose 1,6 - bisphosphatase. IUCr International congress XXII, Madrid, Spain August 2011.

Enzymatic and Preliminary Structural Characterization of *glpX* encoded Fructose 1, 6 – bisphosphatase from *M. tuberculosis*. Keystone symposia (J4) Mycobacteria: Physiology, Metabolism and Pathogenesis - Back to the Basics. Vancouver Canada Jan 2011.

Biochemical characterization of *glpX* (Rv1099c) encoded Fructose 1,6 – bisphosphatase from *Mycobacterium tuberculosis*, 48th Annual MIKI meeting,University of Illinois at Chicago IL.

Formulation and HPTLC based standardization of whole leaf Aloe vera juice at International Conference on Botanicals, Jadhavpur University School of Natural Products, Calcutta Jan 2005.

Oral presentations

Tuberculosis research - a graduate student's perspective, at Harold Washington College (City College of Chicago) May 2010.

Simultaneous Quantitative estimation of Benzoic acid and Sorbic acid in Aloe vera gel by HPLC, at 56th Indian Pharmaceutical Congress Calcutta Dec 2004.