Structural and inhibition studies of thiamine monophosphate kinase from *Mycobacterium tuberculosis*

Dlamini Lenye Sebenzile

12140164

Submitted in partial fulfilment of the degree:

Master's in Biochemistry

Department of Biochemistry, Genetics and, Microbiology

Faculty of Natural and Agricultural Sciences

University of Pretoria

24 January 2020

KEYWORDS

Mycobacterium tuberculosis

Thiamine monophosphate kinase

X-ray crystallography

Structure-based inhibitor identification

ABSTRACT

Vitamin B1 is an indispensable co-factor for various enzymes *inter alia* in the Krebs cycle, pentose phosphate pathway, nucleotide and amino acid synthesis. Due to its importance in metabolism, proteins involved in the synthesis of vitamin B1 have been identified as potential drug targets. Thiamine monophosphate kinase (ThiL), catalyses the last reaction in the pathway, the ATP dependent phosphorylation of thiamine monophosphate (TMP) producing thiamine pyrophosphate (TPP) the active and co-factor form of vitamin B1.

In this study, thiamine monophosphate kinase from *Mycobacterium tuberculosis* (MtbThiL, ~36 kDa) was produced as an N-terminally His_6-tagged fusion protein, purified by affinity and size exclusion chromatography, and crystallised. Hexagonal MtbThiL crystals belonged to space group $P6_122$. Molecular replacement revealed a symmetric homodimer with a single monomer occupying the asymmetric unit. Analysis of the structure showed that each subunit of MtbThiL has an ATP and TMP binding site and is structurally related to other ThiL enzymes.

Ten lead compounds were identified from compound databases as potential ThiL inhibitors, and oxythiamine was chosen for further study. The binding affinities of oxythiamine and TMP to MtbThiL were determined by isothermal titration calorimetry and a pyruvate kinase-lactate dehydrogenase enzyme assay, which revealed that the binding affinity for oxythiamine by MtbThiL is lower than the substrate TMP.

DECLARATION OF ORIGINALITY

Full names: Lenye Sebenzile Dlamini

Student number: 12140164

Title: Structural and inhibition studies of thiamine monophosphate kinase from *Mycobacterium tuberculosis*

Declaration:

I, Lenye Sebenzile Dlamini declare that the thesis/dissertation, which I hereby submit for the degree Master's in Biochemistry at the University of Pretoria, is my own work and has not been previously submitted by me for a degree at this or any other tertiary institution.

1. I am aware of what plagiarism is and I am aware of the University's policy in this regard.

2. I declare that this dissertation is my own original work. Where other people's work has been used (either from a printed source, internet or any other source) this has been properly acknowledged and referenced in accordance with departmental requirements.

3. I have not used work previously produced by another student or any other person to hand in as my own.

4. I have not and will not allow anyone to copy my work with the intention of passing it off as his/her work.

Signature: Date: 24 January 2020

ACKNOWLEDGEMENTS

CONTENTS

LIST OF ABBREVIATIONS

ΔG	Change in Gibb's free energy
ΔH	Change in enthalpy
ΔS	Change in entropy
3D	Three dimensional
AaThiL	Thiamine monophosphate kinase from *Aquifex aeolicus*
AbThiL	Thiamine monophosphate kinase from *Acinetobacter baumanii*
ADMET	Absorption, distribution, metabolism, excretion, toxicology
ADP	Adenosine diphosphate
AEX	Anion exchange chromatography
AMP-PNP	Adenylyl imidodiphosphate
ANOVA	Analysis of variance
APC	Antigen presenting cells
APS	Ammonium per-sulphate
ATP	Adenosine triphosphate
BCG	Bacille Calmette-Guerin
C-terminus	Carboxy terminus
DLS	Dynamic light scattering
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylene diamine tetra-acetic acid
EI complex	Enzyme-inhibitor complex
ES complex	Enzyme-substrate complex
GlideSP	Glide standard precision
GlideXP	Glide extra precision
HIV	Human immunodeficiency virus
HTVS	High through-put virtual screening
His_6	Hexa-histidine (tag)
IEX	Ion exchange chromatography

IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K_a	Binding affinity constant
k_{cat}	Enzyme turn-over number
K_D	Equilibrium dissociation constant
kDa	Kilodalton
K_M	Michaelis constant
LB	Lysogeny broth
LDH	Lactate dehydrogenase
PK-LDH assay	Pyruvate kinase-lactate dehydrogenase coupled enzyme assay
MDR-TB	Multi-drug resistant TB
MM	Michaelis-Menten
Mtb	*Mycobacterium tuberculosis*
MTBC	*Mycobacterium tuberculosis* complex
MtbThiL	Thiamine monophosphate kinase from *Mycobacterium tuberculosis*
MW	Molecular weight
N-terminus	Amino terminus
NAD^+	Oxidised nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
Ni-NTA	Nickel nitrilotriacetic acid
NMP	Nucleoside monophosphate
OD_{600}	Optical density at 600 nm
OT	Oxythiamine
PDB	Protein data bank
PEG	Polyethylene glycol
PEP	Phosphoenolpyruvate
pI	Isoelectric point
PK	Pyruvate kinase
ROS	Reactive oxygen species
RTP	Room temperature and pressure
SBDD	Structure-based drug discovery
SDS	Sodium dodecyl sulphate

SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
TB	Tuberculosis
TDR-TB	Totally drug resistant tuberculosis
TEMED	Tetra-methyl-ethylene-diamine
TMP	Thiamine monophosphate
TPP	Thiamine pyrophosphate
UV	Ultra-violet
VLS	Virtual ligand screening
WHO	World Health Organisation
XDR-TB	Extensively drug resistant tuberculosis

LIST OF FIGURES

LIST OF TABLES

1. Introduction

1.1 Tuberculosis

Human tuberculosis (TB) is an airborne disease caused by the bacterium *Mycobacterium tuberculosis* (Mtb) [1]. It is the leading cause of death from a single infectious agent and the ninth most common cause of death worldwide. TB caused the deaths of an estimated 1.5 million people in 2018, 0.25 million of whom were co-infected with the human immune virus (HIV). The regions with the highest incidence rates in 2018 were southeast Asia, sub-Saharan Africa and the West-Pacific region [2] (Figure 1.1).
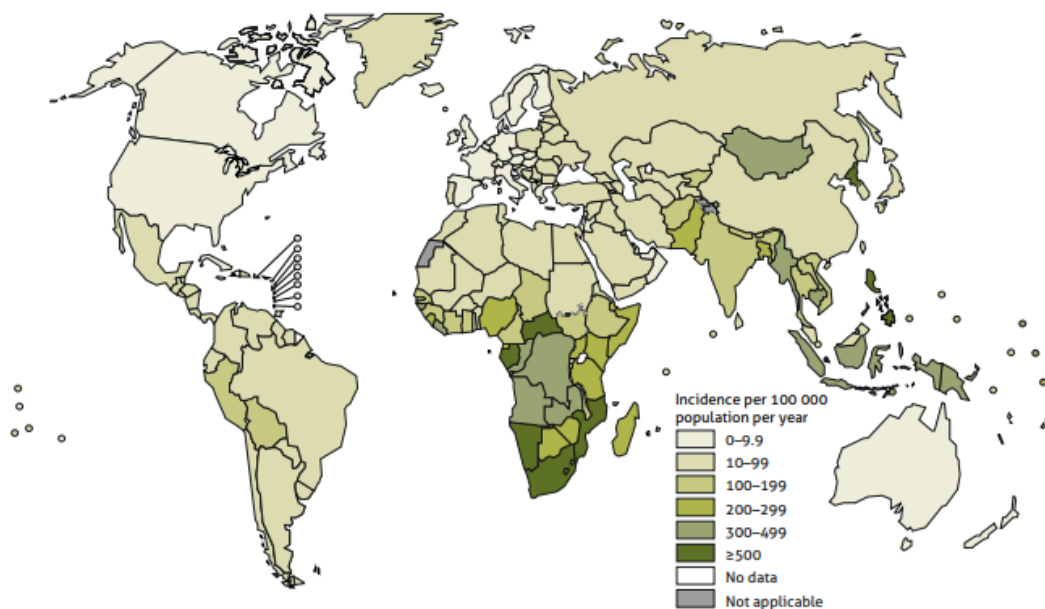
Figure 1.1: World tuberculosis incidence rates in 2018. Twenty countries carry the highest TB disease burden. Sub-Saharan Africa stands out as one of the most affected regions [2].

The epidemiology of TB hinges on two main factors, the competency of the host immune system and the efficiency of health care systems in diagnosing and treating all forms of TB. The highest TB incidence rates are observed in the most impoverished regions in the world with poor healthcare systems such that TB may be classified as a poverty related disease [3]. Individuals co-infected with HIV and TB often die from the latter because the HIV infection impairs the

immune response to Mtb infection. Regions with high HIV disease burden, such as sub-Saharan Africa, therefore also show high TB mortality rates [4].

1.1.2 Tuberculosis diagnosis and treatment

The diagnosis of TB remains as challenging as its treatment. Many early detection techniques are still used today such as sputum smear microscopy, as well as skin-, blood-, and culture-based tests. More recent diagnostic tools include nucleic acid amplification tests and the Xpert MTB/RIF diagnostic tool that detects early infection and rifampicin resistance [5, 6].

The first major breakthrough in anti-Mtb treatment was the development of the bacilli Calmette- Guerin vaccine (BCG). It is still in use today – mainly for infants [7-9].

Drugs used to treat active Mtb infection are classified as first- or second-line drugs based on their potency. First-line drugs are used to treat drug susceptible TB and early infection immediately after diagnosis. These drugs include isoniazid, pyrazinamide and ethambutol. Isoniazid and ethambutol inhibit mycolic acid cell wall synthesis by primarily targeting InhA, a nicotinamide dinucleotide dependent enoyl-acyl carrier protein reductase [10] and arabinosyltransferases EmbA, EmbB and EmbC respectively [11]. Pyrazinamide only acts against dormant mycobacteria, targeting ribosomal protein S1 involved in trans-translation and aspartate decarboxylase which is required in the synthesis of vitamin B5, a derivative of coenzyme A and vital in energy metabolism [12, 13]. The emergence of drug resistant Mtb has, however, necessitated combination treatments, where first- and second-line drugs are taken in specific, predefined combinations. Second-line drugs include streptomycin and amikamycin. Treatment regimens can last from six to twenty months depending on the type of resistance and patient response to treatment [14, 15].

1.2 *Mycobacterium tuberculosis*

In 1882 Robert Koch identified Mtb as the agent responsible for human TB [16]. The bacterium is part of the so-called "*Mycobacterium tuberculosis* complex" (MTBC), a collection of closely related Mtb species found across the world targeting different animals. The MTBC includes *M. africanum*, a strain with lower human pathogenicity common to West Africa [17] as well as *M. bovis*, *M. caprae*, and *M. pinnipedi* that infect animals [18, 19].

Mtb are gram-positive, obligate intracellular bacteria characterised by a hydrophobic mycolic acid cell wall [20]. They are rod-shaped but adopt a spore-like morphology under stressful conditions linked to its virulence [21, 22]. Its wide distribution and persistence over millennia have earned Mtb the title of most successful intracellular pathogen [23, 24].

1.2.1 *Mycobacterium tuberculosis* pathogenesis and immunological response to infection

Mtb is primarily transmitted through aerosols, meaning that airways and lungs are primary sites of infection. In the airways, Mtb infects non-immune microfold (M) cells that allow bacteria to cross the mucosa. In alveolar tissue, by contrast, Mtb infects innate immune cells particularly macrophages [25-27]. Rarer sites of infection include bones, lymph nodes, the central nervous system and the genitourinary tract [28].

Mtb infection results in either active disease, clearing of the infection or persistence of Mtb in a metabolically dormant state. Infection initially mostly triggers an innate immune response through macrophages and dendritic cells [29]. These cells phagocytose the bacteria and present Mtb antigens to effector and regulatory T-cells. Other processes of the host immune system involve phagosome-lysosome fusion, the production of reactive oxygen (and nitrogen) species (ROS), programmed cell death and inflammation. While individual bacteria are killed by these mechanisms, the overall infection is often not cleared [30].

To isolate the Mtb infection, the immune system amasses immune cells such as macrophages, neutrophils, dendritic cells, natural killer cells and Langhan's giant cells around the bacteria. The resulting physiological structures are known as granulomas. Conditions in the granuloma are often acidic, nutrient deficient and high in ROS, retarding Mtb replication and growth. While the granuloma contains the infection, it also shields bacteria from host immune responses [29, 31-34], allowing bacteria to remain metabolically dormant. The dormant bacteria are reactivated by constant exposure to Mtb and co-infection with other pathogens [35-38], reinitiating active TB (Fig. 1.2) [39].
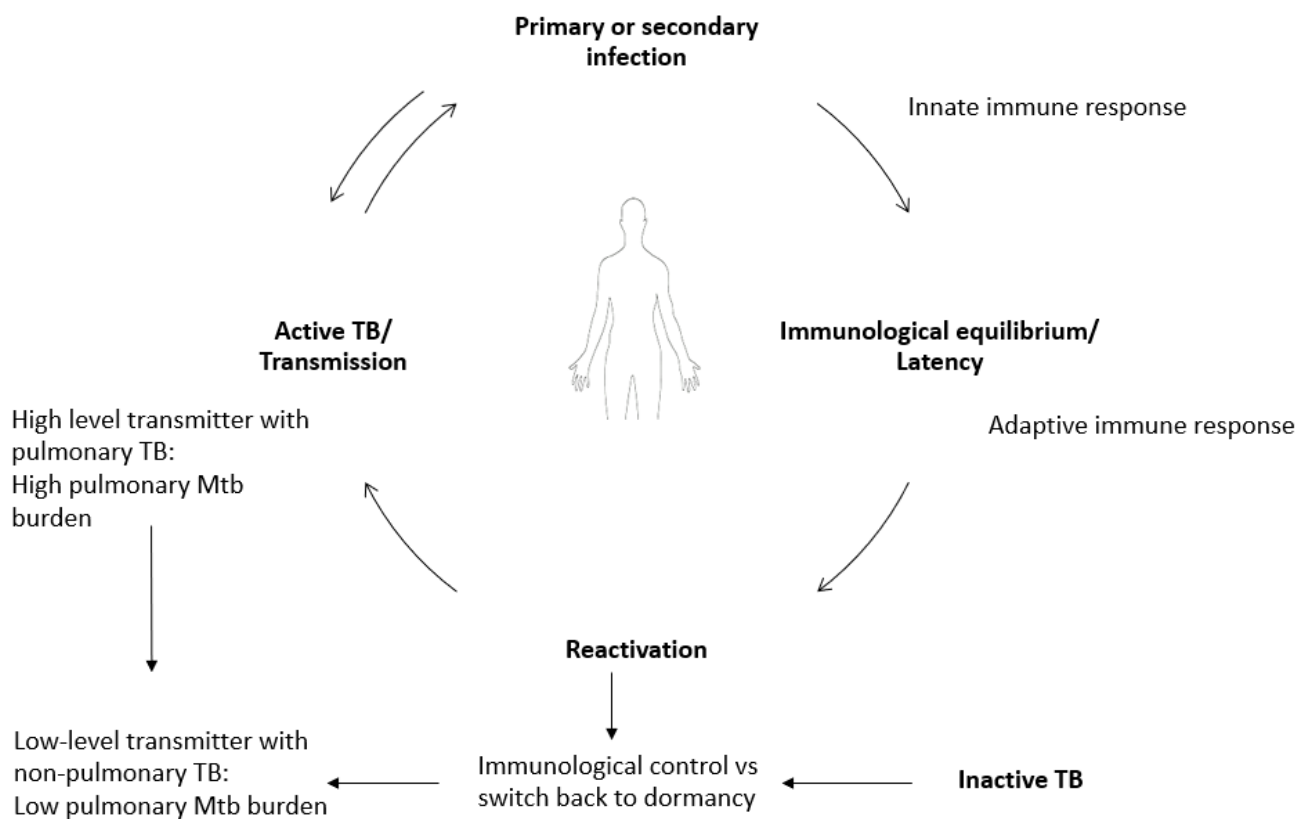
Figure 1.2: _Mycobacterium tuberculosis_ immunological life cycle. Mtb infection results in two distinct metabolic outcomes, active or latent Mtb. Initial immune responses clear the infection through adaptive immune responses. However, Mtb will often become dormant in granulomas. These latent Mtb cells can be reactivated resulting in transmittable, active TB. After reactivation, the immune system may once again contain Mtb infection reverting it back to latency, at this point, the host is referred to as a low-level transmitter with low pulmonary disease burden. However, depending on the host immune system, reactivation may result in active and transmittable TB [40].

Active TB is accompanied by symptoms such as chronic coughing with blood- stained sputum, severe weight loss and night sweats [40, 41]. Active Mtb infection leads to alveolar tissue necrosis and excessive inflammation damaging the lung tissue. This often results in respiratory failure and pneumonia culminating in death.

1.2.2 Drug resistance

The development and spread of drug resistant bacterial strains is a global health problem and poses a risk to the treatment of many infectious diseases such as pneumonia, influenza and TB [42].

Pathogens develop drug resistance by various mechanisms such as physical barriers that limit drug access or efflux pumps that actively transport the drug out of the cell. Alternatively, the antibiotic or its target are modified to prevent interaction [43, 44]. Drug resistance has long been hypothesized to result from improper and over use of antimicrobial drugs [1, 45, 46]. Resistance mechanisms encoded on chromosomes are classified as innate, while acquired resistance comes about through the acquisition of foreign DNA encoding resistance genes in a process known as horizontal gene transfer [43, 47]. Errors during DNA replication further cause spontaneous mutations in chromosomal DNA also resulting in drug resistant strains of bacteria.

1.2.3 Drug resistance in *Mycobacterium tuberculosis*

Over the years, Mtb has developed resistance to many Mtb drugs. First "multi-drug resistant" (MDR) Mtb strains were identified, followed by "extensively drug resistant" (XDR) and finally "totally drug resistant" (TDR) strains [46]. MDR strains are defined as being resistant to isoniazid and rifampicin, the two most potent Mtb antimicrobial drugs. Such strains may also be resistant to other first-line drugs. XDR Mtb is defined as resistance to isoniazid and rifampicin, any fluoroquinolone and any of the three second-line drugs (amikacin, capreomycin, and kanamycin)

[48]. TDR Mtb strains are resistant to all available drugs. They have been reported in Iran, Italy, India and South Africa [16, 49-51]. The World Health Organisation (WHO) currently does not recognise TDR Mtb or the corresponding form of TB.

Constant exposure to antibiotics results in sustained selection pressure on bacteria promoting the survival of drug resistant bacteria [52]. During treatment, anti-TB drug combinations are prescribed for up to twenty months creating such selection pressures [1, 14]. In addition, the pro- drugs isoniazid and ethionamide are activated by redox enzymes in the granuloma releasing ROS that can damage DNA increasing the rate of mutation and potentially contributing to development of drug-resistance [53].

Properties of Mtb that contribute to its innate resistance include its characteristic mycolic acid cell wall and its ability for growth arrest. The hydrophobic cell wall dramatically reduces drug permeability especially of hydrophilic molecules [54]. Growth arrest, metabolic latency and non- replicating persistence occurs in response to hypoxia and acidity as found in granuloma [8, 55] and enables Mtb to survive with minimal energy sources for extended periods. When conditions for growth return, the bacteria revert to their normal metabolic state causing post primary infections (Fig. 1.2) [56, 57].

Mtb adapts to different stressors, including drugs and the conditions in the granuloma, through mutations on chromosomal genes and altered gene expression [58, 59]. Mutations on genes encoding proteins essential for activity of the drug result in defective proteins and enzymes decreasing or cancelling the effects of the drug. For example, Mtb resistance to pro-drug isoniazid results from defective catalase-peroxidase (KatG) encoded by a mutated *katG* gene. This defective enzyme cannot activate isoniazid, nullifying the effect of the drug on Mtb [60]. Additionally, isoniazid and rifampicin resistant Mtb had increased expression of genes that encode efflux pumps in the presence of these drugs demonstrating the differential expression of

genes in the presence of drugs [61]. This also occurs in proteins of metabolic pathways to compensate for deficient nutrients, co-factors and reaction intermediates. Mtb is dependent on proteins and enzymes to adapt to changes in its environment. As such, proteins and enzymes involved in lipid, carbohydrate and amino acid metabolism and vitamin and cofactor biosynthesis, especially those unique to Mtb, are frequent targets for drug development against drug resistant bacteria.

1.3 Enzymology

1.3.1 Enzyme catalysis

Enzymes are substrate-specific protein catalysts that speed up the rate of chemical reactions in cells by lowering their activation energy [62]. The reaction occurs in an active site, a conserved three-dimensional (3D) pocket within the protein that is structurally and chemically complementary to the substrate or reaction intermediate [63]. Functional groups mostly from amino acid side chains interact with and orient the substrate(s) and/or co-factors in the most energetically favourable positions allowing catalysis, during which bonds are formed, broken or functional groups transferred [64, 65].

To extend the range of chemical reactions beyond those based on the functional groups of the twenty standard amino acid side chains, enzymes bind non-protein molecules known as co-factors and co-enzymes [66, 67]. Whereas co-factors are molecules or ions covalently bound to the enzyme, co-enzymes are non-covalently bound organic compounds or vitamins. Enzymes are assigned to six categories according to the main reaction catalysed including oxidoreductases, transferases, hydrolases, lyases, isomerases, and ligases [68]. Further grouping into families is based on similarities in their amino acid sequences and structures.

1.3.2 Enzyme inhibition

Enzyme inhibitors are small molecules that impede or stop the activity of enzymes. They either bind to the enzyme or the enzyme-substrate (ES) complex forming an enzyme-inhibitor (EI) complex with reduced or abrogated activity. Physiological inhibitors often regulate cellular processes to control the flow of metabolites in a cell [69]. Activators are the opposite of inhibitors and stimulate catalysis when binding to enzymes [70].

Inhibitors are categorised as reversible or irreversible depending on how they interact with the enzyme. Irreversible inhibitors modify key catalytic residues through covalent interactions preventing normal substrate recognition and positioning [71]. Reversible inhibitors, by comparison, bind through transient, non-covalent interactions either directly to the active site to compete with substrate binding or elsewhere in the enzyme deforming the active site. High substrate concentrations can displace reversible inhibitors allowing the enzyme to regain some activity [64, 72].

Reversible inhibitors are subdivided into competitive, uncompetitive and non-competitive categories. Competitive inhibitors bind to and compete for the active site with the substrate through structural similarity to form similar interactions with the enzyme [70, 73]. Uncompetitive and non-competitive inhibitors bind to sites on enzymes distinct from the active site but change the conformation of the active site preventing normal substrate binding resulting in allosteric inhibition. Whereas uncompetitive inhibitors bind to the ES complex, non-competitive inhibitors bind to both free enzyme and the ES complex.

Detailed structural information particularly high resolution crystal structures of substrate/ enzyme complexes may be used in designing competitive inhibitors [64]. This approach was used in this study.

1.3.3 Kinases

Kinases catalyse the transfer of a phosphate group from a high energy molecule such as ATP, GTP or phosphoenolpyruvate to a substrate [74, 75] using a hydroxyl, amino, carboxyl, or phosphate group as an acceptor [76]. Kinases are either protein or nonprotein kinases. Protein kinases phosphorylate amino acid side chains, typically serine, threonine and tyrosine on target proteins [74] to alter their activity or change their cellular localisation. Cellular processes regulated by protein kinases include metabolism, transcription, cell cycle regulation, apoptosis, cell motility and signal transduction [77].

The nonprotein kinases phosphorylate targets such as sugars, nucleosides or lipids [78]. Nucleoside monophosphate kinases, for example, phosphorylate and activate vitamin precursors, which function as enzyme co-factors in purine biosynthesis [79].

Because kinases are directly and indirectly involved in regulating metabolic pathways in organisms, they are a major target for drug development and design. Examples of these include multiple tyrosine kinases and the cyclin-dependent kinase family [80].

1.3.4 Structure-based drug discovery and virtual ligand screening

The principle of molecular recognition forms the basis of all biochemical reactions and is critical for the rational design of novel drugs [81]. Molecular recognition encapsulates the ability of macromolecules to recognize each other and small molecules through highly specific interactions based on structural and chemical complementarity. The specificity of interactions determine the affinity of targets for ligands and ensure that low affinity ligands do not replace high affinity ligands regardless of their concentration [82]. This principle forms the basis of multiple drug development strategies including fragment- and structure-based drug discovery.

Structure-based drug discovery (SBDD) attempts to complement the chemical functionality of an existing binding pocket in the form of a ligand. SBDD may employ virtual ligand screening (VLS), a computer-based mechanism of selecting known compounds by how well they complement the binding site and hence estimate their binding affinities [83]. In practice this is implemented through *in silico* docking and scoring functions. Docking programs sample ligands from a library and orient each to achieve favourable energies. Each resulting position is referred to as a pose. Scoring functions compare poses by weighting interactions such as hydrophobicity, hydrogen bonds, van der Waals interactions and solvation electrostatics. They calculate a final score and a predicted binding affinity [83-85].

Experimental binding affinities describe the equilibrium between ligand-free and -bound protein at constant temperature and pressure. They are related to the change in Gibb's free energy change (ΔG) achieved during binding and ΔG must be negative for ligand-binding to occur spontaneously [82, 86]. A compound found to have a similar or higher affinity than the natural substrate for its intended target is a good starting point in developing a selective and efficacious inhibitor or drug. All atomic interactions contribute to the enthalpic component (ΔH) of ΔG. These include strong electrostatic interactions and hydrogen bonding but also weaker van der Waals or hydrophobic interactions.

Although scoring functions are useful in virtual docking and are successful in sorting best binding ligands from poorer binders and providing useful preliminary information on the binding affinity of a ligand to its target, they have significant inaccuracy and should ideally be verified experimentally [87].

1.4 Thiamine monophosphate kinase

1.4.1 Thiamine pyrophosphate/ vitamin B1

Vitamin B1 or thiamine pyrophosphate (TPP) is a water-soluble co-factor for many enzymes in major biochemical pathways. Examples include α-ketoglutarate dehydrogenase, pyruvate dehydrogenase and transketolase. Pyruvate dehydrogenase and α-ketoglutarate dehydrogenase form part of glycolysis and the Krebs cycle, respectively, critical pathways for ATP production in living cells [64]. Transketolase forms part of the pentose phosphate pathway that produces NADPH and ribose-5-phosphate as raw materials in the synthesis of nucleotides, fatty acids and amino acids [88]. TPP therefore, is central to efficient production of many cellular macromolecules.

The role of TPP as a co-factor involves breaking and forming bonds between carbon and sulphur, nitrogen, hydrogen, oxygen and other carbon atoms with the reactive C2 carbon on the TPP thiazole group acting as a nucleophile. First the TPP C2 is deprotonated to form a carbanion, which attacks a carbonyl carbon on the substrate to form a TPP-substrate complex often releasing CO_2. The resulting molecule is covalently bound to the substrate. A shift of electrons results in the release of the aldehyde and regeneration of the carbanion [89-91].

De novo thiamine biosynthesis happens exclusively in bacteria, fungi, plants and lower eukaryotes, while mammals derive vitamin B1 from their diet [82, 88]. In bacteria, vitamin B1 is created by synthesising a thiazole (5-hydroxyethyl-4-methylthiazole phosphate) and a pyrimidine moiety (4-amino-5-hydroxymethyl-2-methyl pyrimidine pyrophosphate) (Fig. 1.3) and combining these to form thiamine monophosphate (TMP). The latter step is catalysed by ThiE, while ThiL then phosphorylates TMP to thiamine pyrophosphate (TPP) using ATP [88, 89, 92].
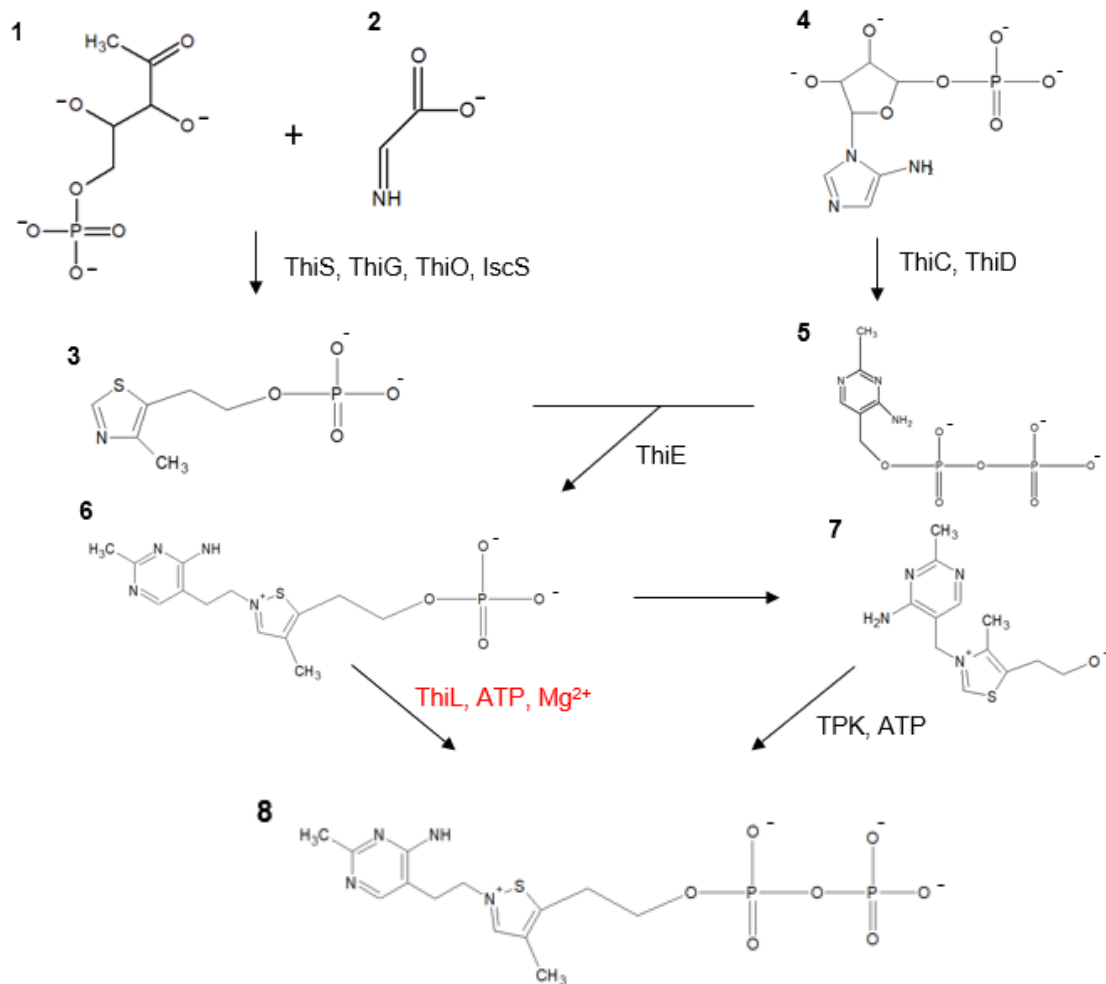
Figure 1.3: Bacterial thiamine biosynthesis. Thiamine monophosphate (TMP) is formed by combining a thiazole **(3)** and a pyrimidine moiety **(5)**. Deoxy-D-xylulose-5-phosphate **(1)**, dehydroglycine **(2)** (formed from glycine and tyrosine in *B. subtilis* and *E. coli* respectively) and, ThiS-COSH are combined through a condensation reaction by ThiG, to form thiazole phosphate **(3)**. The pyrimidine component of TMP, hydroxymethyl pyrimidine phosphate (HMP-P) **(5)** is formed from 5-aminoimidazole ribonucleotide **(4)** through rearrangement by ThiC and a phosphorylation by ThiD. Thiazole phosphate and HMP-P are combined by ThiE, to create TMP **(6)** followed by its phosphorylation by ThiL to the active co-factor thiamine pyrophosphate **(8)**.Adapted from [89, 93, 94].

In addition to synthesizing thiamine *de novo*, bacteria, archaea, fungi, and some eukaryotes have thiamine salvage pathways (Fig. 1.4) involving the directed uptake of extracellular intermediates or TMP to compensate for inefficiencies in their biosynthetic pathway [82, 95]. The absence of the TMP biosynthetic pathway in humans ensures that drugs targeting the pathway are specific to pathogens. This applies particularly to bacteria that lack thiamine salvage pathways, such as Mtb [82, 96, 97].
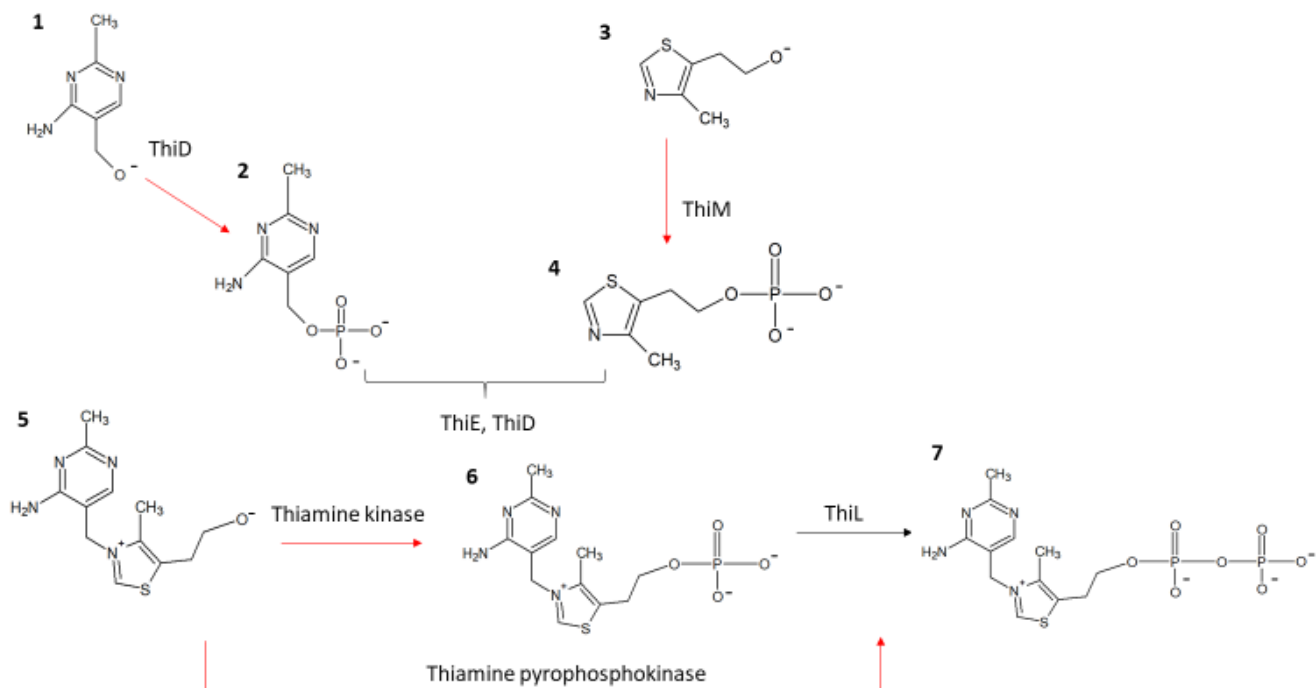
Figure 1.4: Thiamine salvage pathways in bacteria (*E.coli* and *B. subtilis*). Red arrows indicate salvage pathways where reaction intermediates are taken up from the extracellular medium. Thiamine kinase catalyses the conversion of free thiamine to TMP (*inter alia* in *E. coli*), while thiamine pyrophosphate kinase catalyses ATP dependent phosphorylation of free thiamine to TPP (e.g. in *B. subtilis*).

1.4.2 Thiamine monophosphate kinase

Thiamine monophosphate kinase (ThiL) is a ~36 kDa, dimeric enzyme that catalyses the last reaction in thiamine biosynthesis. Each monomer has both an ATP and a TMP binding site [90, 98, 99]. ThiL is a member of the nucleotide monophosphate kinases that catalyse the direct transfer of a γ-phosphate from ATP to a nucleoside/nucleotide monophosphate species (Fig. 1.3 reaction highlighted in red). ThiL requires divalent cations, typically Mg^{2+}, to catalyse the phosphorylation reaction [78].

1.4.3 Thiamine monophosphate kinase from *Mycobacterium tuberculosis*

Previous structural studies of thiamine monophosphate kinase from *Mycobacterium tuberculosis* (MtbThiL) yielded the crystal structure of apo-MtbThiL i.e. without any ligands (apo-structure) [100]. The tertiary structure of MtbThiL is similar to those of other ThiL enzymes. The A domain on each of the monomers has an ATP-binding fold in which a β-sheet is flanked by α-helices and

13

connecting loops. A variable region in the form of a loop and a short β-strand in ThiL proteins from bacterial species such as *Aquifex aeolicus* and *Acinetobacter baumanii* is distinctly longer in MtbThiL [90, 100].

Currently no crystal structure of MtbThiL with bound ligands is available. Structural analysis of MtbThiL-ligand complexes could clarify its catalytic mechanism and support attempts to develop an MtbThiL inhibitor. As the last enzyme in the thiamine biosynthetic pathway, ThiL is an attractive target for the design of anti-Mtb drugs.

1.5 Justification of study

Tuberculosis is a global threat to human health. The rapid development of resistance in strains of *Mycobacterium tuberculosis* has limited the effectiveness of most anti-TB drugs. There is thus an urgent need for new and repurposed drugs as well as the identification and study of new drug targets.

Thiamine is an important co-factor in various important biochemical processes. Thiamine deficiency may be linked to reduced intracellular growth of Mtb. Thiamine monophosphate kinase from *Mycobacterium tuberculosis* (MtbThiL) has not been studied as a potential drug target for anti-TB drugs. The first crystal structure of apo-MtbThiL was solved by Ntui 2015. The purpose of this study is to extend the work done on MtbThiL by solving its structure in complex with substrates and products. The MtbThiL crystal structure and kinetics will give greater insight into the reaction mechanism and its control. In addition, we hope to identify and study possible MtbThiL inhibitors, as a first step in developing possible drugs targeting MtbThiL.

1.6 Study aims and objectives

The overall aim of this study was to structurally and functionally characterise MtbThiL, to identify potential inhibitors and assess their inhibitory capacity.

This was achieved by addressing the following objectives:

1. To produce and purify MtbThiL

2. To obtain a high-resolution crystal structure of MtbThiL in complex with substrates and products and investigate its catalytic mechanism

3. Determine the kinetic parameters for MtbThiL

4. Identify and test the inhibitory capabilities of a potential MtbThiL inhibitor

2. Materials and Methods

2.1 Materials

2.1.1 General chemicals and reagents

All chemicals and reagents used in this work were of good laboratory standard.

Table 2.1: Kits, enzymes and, molecular weight standards used in this study

Use	Product	Supplier
Protein marker	Pierce Unstained Protein Molecular weight marker	Thermo Scientific (Waltham, MA, USA)
Substrates, products and inhibitors	Thiamine-monophosphate, Thiamine-pyrophosphate Oxythiamine Reduced nicotinamide adenine dinucleotide	Sigma Aldrich (St Louis, MO, USA)
Enzyme assay	Pyruvate kinase	Sigma Aldrich (St Louis, MO, USA)
Enzyme assay	Lactate dehydrogenase	Sigma Aldrich (St Louis, MO, USA)

2.1.2 Buffers, stock solutions and growth media

All buffers, stock solutions and growth media used in this study were sterilised by autoclaving, and/or filtering. All buffers used in chromatography experiments were also degassed prior to use. The pH of buffers was adjusted with concentrated solutions of HCl and NaOH.

Table 2.2: Growth media and protein production reagents

Use	Buffer/ Reagent	Preparation
Antibiotic	Kanamycin	1 mL 50 µg/mL kanamycin in 1 L LB media
Induction of gene expression	IPTG	100 µL 1 M IPTG in 1 L LB media
Liquid growth media	Luria-Bertani (LB) media	10 g/L tryptone, 5 g/L NaCl, 5 g/L yeast extract powder
Solid growth media	LB Agar plates	1.25 g/L tryptone, 1.25 g/L NaCl, 1.25 g/L nutrient agar and 50 µg/mL kanamycin

Table 2.3 Buffers and solutions

Name	Preparation
Resolving gel buffer	1.5 M Tris-HCl pH 8.8
Stacking gel buffer	0.5 M Tris-HCl pH 6.8
8 x SDS sample buffer	2.2 mL 0.5 M Tris-HCl pH 6.8, 4 mL 60% (v/v) glycerol, 16 mL 10% SDS, 200 µL β-mercapto-ethanol, 0.8 mL 0.1% bromophenol blue,20 mL dH$_2$O
Coomassie staining solution	1% (w/w) Coomassie R-250, 30% (v/v) ethanol, 10% (v/v) acetic acid, filtered using Whatman paper
De-staining solution	400 mL 40% (v/v) ethanol, 100 mL 10% (v/v) acetic acid
12% SDS-PAGE gel	Resolving gel: 7.5 mL 30 % (w/v) acrylamide, 5 mL 1.5 M Tris-HCl pH 8.8, 200 µL 10% (w/v) SDS, 50 µL 25 % (w/v) APS, 20 µL TEMED, 6.2 mL ddH$_2$O
	Stacking gel: 1.4 mL 30 % (w/v) acrylamide, 2.5 mL 0.5 M Tris-HCl pH 6.8, 25 µL 25 % (w/v) APS, 1.5 µL TEMED, 6.5 mL ddH$_2$O

Table 2.4: Protein electrophoresis buffers and reagents

Buffer/solution	Purpose	Preparation
Harvesting buffer	Ni-NTA affinity chromatography	20 mM Tris pH 8.0, 400 mM NaCl, 30 mM imidazole
Wash buffer		20 mM Tris pH 7.0, 35 mM imidazole, 250 mM NaCl
Elution buffer		20 mM Tris pH 7.9, 500 mM NaCl, 200 mM imidazole
1 x Running buffer	SDS-PAGE	25 mM Tris, 190 mM glycine, 0.1% SDS
Crystallisation buffer	Crystallisation	5 mM NaCl, 1 mM MgCl$_2$, and 1 mM DTT, 10 mM Tris pH 7.5
Imidazole-HCl buffer	Enzyme assay	3 M imidazole, 1.7 mL HCl pH 7.4
Isothermal titration calorimetry (ITC) buffer	ITC	20 mM Tris pH 7.4, 10 mM NaCl, 100 mM KCl, 20 mM MgCl$_2$

Table 2.5: List of equipment

Equipment	Model	Manufacturer
Affinity chromatography drip column	Mobicol 10 mL column, 35 µm filter pore size	Boca Scientific Inc. (Boca Raton Florida, USA)
Centrifuges	Sorvall RC6 Heraeus Megafuge 8R	ThermoScientific (Waltham, MA, USA)
	Sorvall 5417C	Eppendorf (Hamburg, Germany)
Centrifugal filter units	Amicon Ultra-15	Millipore Sigma (Burlington, MA, USA)
Electrophoresis chamber	Mini-PROTEAN	Bio-Rad (Hercules, CA, USA)
Electrophoresis power pack	Power Pac Basic	Bio-Rad (Hercules, CA, USA)
Nanodrop spectrophotometer	ND-1000	Nanodrop Technologies Inc. (Wilmington, DE, USA)
pH meter	Unspecified	Bio-Rad (Hercules, CA, USA)
Roller mixer	SRT6	Stuart (Staffordshire, UK)
Shaking incubator	Multitron II	INFORS (Bottmingen, Switzerland)
Sonicator	Qsonica	QSonica, (Newtown, CT, USA)
HiTrap Q HP anion exchange chromatography 5 mL column	Q HP 5 mL	GE Healthcare Life Sciences (Chicago, IL, USA)
Gel Imager	Molecular Imager GelDoc XR and UV transilluminator	Bio-Rad (Hercules, CA, USA)
Particle analyser	DelsaNano Submicron Particle size and Zeta Potential	Beckman Coulter (Brea, CA, USA)
Isothermal titration calorimetry	MicroCal iTC 200 Isothermal Titration Calorimeter	GE Healthcare Life Sciences (Chicago, IL, USA)

Table 2.6: Computer programs used in this study

Program/Software	Purpose
ACD ChemSketch	Sketch and modify chemical structures
GraphPad Prism	Construction and analysis of Michaelis-Menten graphs
Phenix	Molecular replacement and structure refinement
Pymol	Modelling and docking data validation
Schrodinger (Maestro Suite, Glide)	Docking and scoring
WinCoot	Crystal structure refinement and analysis
Origin	Isothermal titration calorimetry data analysis and fitting
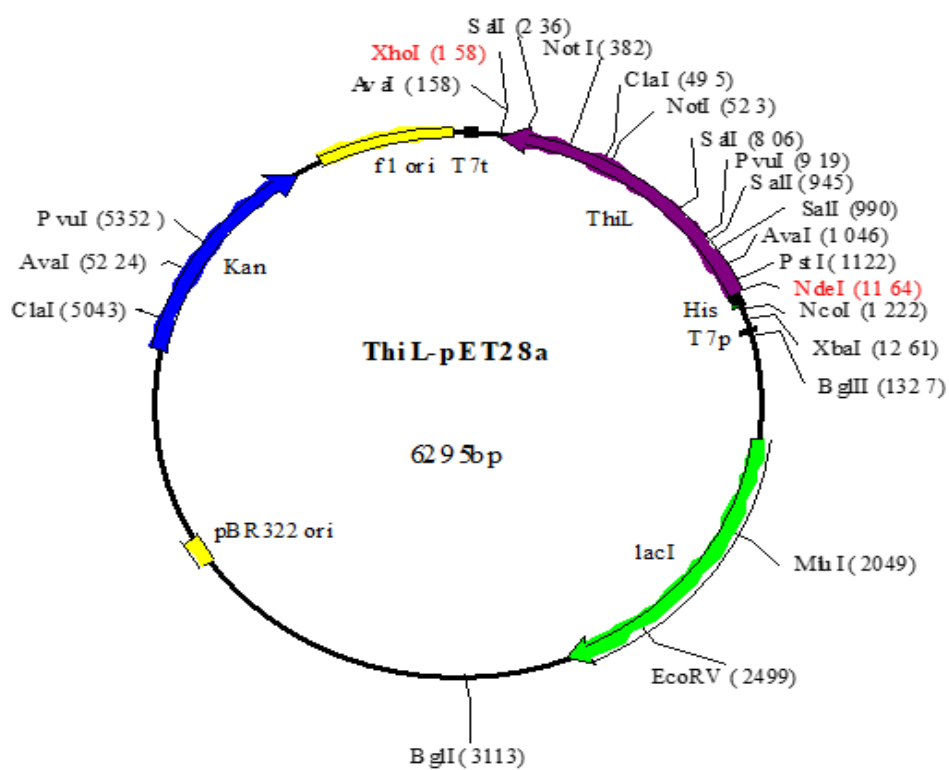
Figure 2.1: pET28a-*mtbthil* plasmid map. Provided by Dr Collin Kenyon, Council for Scientific and Industrial Research (CSIR) Biosciences.

2.2 Methods

2.2.1 Recombinant protein production, purification and, analysis

A. Large-scale, recombinant protein production

Cloning refers to the introduction of a DNA insert, usually containing a gene of interest into a plasmid vector resulting in a recombinant plasmid. Most plasmids contain an antibiotic resistance gene to ensure selectivity and growth of only transformed cells, an isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible *lac* or T7 promoter, and encode a purification tag that enables the rapid isolation of the target protein from the cell lysate [101, 102].

A pre-culture was prepared by inoculating 50 mL of LB media supplemented with 50 µL (50 µg/mL) kanamycin, with a scraping of glycerol stock of BL21-CodonPlus cells transformed with the pET28a-*mtbthil* plasmid. The culture was incubated overnight at 37 °C in a shaking incubator at 170 rpm. The 50 mL cell culture was inoculated into 1 L LB supplemented with 1 mL of 50 µg/mL kanamycin. The cells were grown in a 37 °C incubator with shaking at 190 rpm until an OD_{600} of between 0.4 and 0.6 was reached. A 1 mL "before induction" sample of the culture was taken before adding IPTG (final concentration 0.1 mM) to the culture to induce target gene over-expression. The temperature was lowered to 28 °C and the culture grown overnight with shaking at 200 rpm.

B. Cell harvesting and lysis

A 1 mL "after induction" sample was taken to quantify the amount of protein produced. The cells were harvested by centrifugation at 7000 x *g* for 15 min at 4 °C in a Sorvall Lynx 6000 centrifuge. The largely cell-free supernatant was discarded, and the cells resuspended in 20 mL of cold harvesting buffer. The resuspended cells were lysed by sonication at an amplitude of 45 m, for six cycles of 30 s sonication interspersed with 30 s breaks on a QSonica sonicator.

Ultrasonication ruptures cell walls and membranes through high frequency sound waves that create localized areas of high and low pressure with associated shearing forces [103]. Because the cell contents may be heated excessively by this procedure, the process is interrupted to cool the cells in a beaker of ice. Other mechanical method of releasing a target protein from bacterial cells under non-denaturing conditions include blenders, French presses or bead milling as well as biochemical cell rupturing.

The lysed cells were centrifuged at 37000 x g and 4 °C for 1 h in a Sorval Lynx centrifuge using a RC6 rotor to separate soluble from insoluble cellular components. A 20 μL sample of the supernatant was set aside to determine the amount of soluble protein. To confirm that the target protein was not insoluble produced, a fraction of the pellet was solubilised in 500 μL of 8 M urea and a 20 μL sample set aside.

C. Immobilised metal affinity chromatography

To allow the rapid purification of milligram amounts of protein, encoding genes are often cloned adjacent to tag-encoding sequences, to allow rapid isolation of the resulting protein through specific interactions of a column resin. Tags may vary in size from a few amino acids, such as His_6 or Streptavidin tags, to entire protein domains, such as glutathione-S-transferase or maltose binding protein. Immobilised metal affinity chromatography (IMAC) relies on the binding of a His_6 tag to transition metal ions such as Ni^{2+} or Co^{2+} immobilised on a hydrophilic resin. Non-target proteins will mostly have a much lower affinity for the metal ions and will be eluted. The target protein is eventually released and eluted by displacing the His_6 tag from the metal centres by high concentrations of imidazole, the functional group of histidine [104-106].

Here, Ni^{2+}-NTA beads (Protino) were equilibrated with three column volumes of harvesting buffer. The beads were then added to the soluble cellular fraction in a 50 mL Falcon tube. The

mixture was incubated with gentle agitation on a roller mixer for 3 h to allow the target protein to bind to the Ni-NTA resin.

The bead-lysate mixture was transferred to a Mobicol drip column and excess buffer containing unbound cell constituents was allowed to elute. A 20 μL 'flow-through' sample was collected to verify that target protein was not being lost. Non-target proteins with low affinity to the resin were removed by washing with 700 mL of wash buffer. The first two 15 mL samples were collected for analysis as wash fractions to ensure minimal loss of target protein. Elution buffer with 200 mM imidazole was then allowed to pass through the column, to displace and elute the bound target protein. The 'elution' fractions were collected in three 15 mL samples. A 20 μL sample of the beads was collected after the elution to ascertain the quantitative elution of the target protein. All samples were analysed using 12 % (w/v) sodium dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE) to confirm protein production and determine the amount of impurities. The elution fractions were concentrated in preparation for further purification.

D. Size exclusion chromatography

Size exclusion chromatography (SEC) separates molecules by molecular mass or, more accurately, their hydrodynamic radius. Smaller molecules are retained in the porous resin and eluted later since their movement is retarded while larger molecules are eluted earlier. SEC can also be used to determine the oligomeric status of a protein molecule since the different oligomerization states will have different retention volumes [107].

MtbThiL (64 mg) was purified by SEC on a Superdex 200 10/300 GL column. The column was equilibrated with crystallisation buffer and connected to an FPLC Äkta UV 900 system. A flow rate of 0.1 mL/min was maintained and the absorbance measured at 280 nm. Two millilitre fractions were collected and all fractions that demonstrated an absorbance higher than a threshold were

analysed by 12 % SDS-PAGE as outlined in section 2.2.3 G. Where non-target protein bands remained after SEC, ion exchange chromatography was used as an additional purification step.

E. General protein quantification

Protein concentrations can be determined spectrophotometrically by using assays such as the Bradford assay. In this study the Nanodrop ND-1000 spectrophotometer was used to indirectly estimate the amount of the protein after each purification step. The Nanodrop measures the absorbance at 280 nm, the wavelength at which aromatic residues especially tryptophan reach peak absorption. The concentration is determined using the Beer-Lambert equation ($A_\lambda = \varepsilon_\lambda\, l c$), where the A_λ is the absorbance of the protein at a specific wavelength; ε_λ is the extinction coefficient specific to a protein due to its amino acid composition; l is the path length; and c is the concentration of the protein. The absorbance of a blank (same buffer without protein) is subtracted from the raw absorbance before calculating the concentration.

F. Sodium dodecyl sulphate polyacrylamide gel electrophoresis

Protein production and purity were analysed by SDS-PAGE which separates proteins roughly by molecular weight using an electric field. The detergent SDS confers a net negative charge to proteins, while a reducing agent, such as DTT or β-mercaptoethanol disrupts disulphide bond formation. The negative charge promotes migration of the proteins to the anode during electrophoresis. Proteins move through the polyacrylamide matrix, with smaller proteins migrating faster towards the anode than larger ones [108].

Samples of 10 µL were collected from the supernatant and the pellet, after cell lysis and centrifugation to represent the soluble and insoluble fractions, respectively. Twenty microlitres of 8 x sample buffer was added to each sample, while 10 µL of sample buffer was added to 15 µL samples collected in section 2.2.1 C. Sample/buffer mixtures were heated at 95 °C for 2 min to denature the protein and loaded on the gel. For optimal visibility of the protein bands, 5 µL of

the soluble and solubilised insoluble fraction and 15 µL of all other samples were transferred to the SDS-PAGE wells. Five µL of molecular weight standard was applied to the first well to infer sizes of protein in sample lanes. A current of 40 mA and a potential of 160 V was then applied for 30 min in a Mini-PROTEAN electrophoresis chamber. The resulting gel was stained with Coomassie staining solution, heated in the microwave for 30 s. The gel was rinsed with water and agitated in de-stain solution until protein bands were visible.

2.2.2 Dynamic light scattering

Dynamic light scattering (DLS) provides information on the size distribution of particles in the nm range in a solution by determining their hydrodynamic radii, where the relationship between particle movement and particle size is described by the Stokes-Einstein equation [109, 110]. Monochromatic laser light is passed through the sample and the scattering of light at different angles relates to particle movement due to Brownian motion. Applied to samples of purified proteins, DLS data may be used to provide information protein oligomerisation state(s).

The oligomerisation state of MtbThiL in the crystallisation buffer was inferred from the size distribution determined using DLS. Before analysis, the MtbThiL solution was centrifuged at 12 000 x g and 4 °C for 5 min, to remove potentially denatured or aggregated protein. The protein was sterile filtered with a 0.02 µm filter. DLS measurements were performed in a Delsa Nano Submicron particle analyser with a 1 cm quartz cuvette at 4 °C and the resulting data analysed using the CONTIN method of Delsa Nano UI Software.

2.2.3 Protein crystallisation

To determine the 3D structures of proteins using X-ray crystallography (discussed in section 2.2.4), millions of copies of protein molecules must be arranged in a highly ordered manner, achieved through protein crystals. Proteins are crystallised by the controlled supersaturation of the protein in solution through sitting- and hanging-drop vapour diffusion. A drop of protein

sample and reagent is placed either on top (hanging drop) or next to (sitting drop), a reservoir of the reagent, the system is sealed. Since the drop with the protein sample has lower concentration of the reagent, water vapour diffuses out of it into the reservoir. Dehydration of the solution causes the protein to be supersaturated, and slightly insoluble resulting in the protein crystallising [111].

A. Screening for crystallisation conditions

To identify lead conditions that could be optimized for the crystallization of MtbThiL and its substrate/product/inhibitor complexes, a sitting-drop screen was set up manually using the Hampton Research PEG Rx 2 suite containing 96 solutions. Seventy microlitres of each solution were transferred to the well of a crystallisation plate as the reservoir. Of this, 1 μL was transferred to the crystallisation platform and mixed with 1 μL of protein at 20 mg/mL. The crystallisation plate was sealed by applying transparent, self-adhesive tape and incubated at 4 ˚C. This procedure was repeated for incubation at 7 and 18 °C. The plate was optically inspected at regular intervals to assess the progress and possible crystallization of the protein.

B. Crystallisation optimisation

Suitable lead crystallisation conditions from **A** above, were optimised by varying critical parameters to improve crystal size and morphology using the hanging drop vapour diffusion method in 24-well plates with 500 μL reservoir solutions. One microlitre of the reservoir solution was transferred to the centre of a coverslip and mixed with 1 μL of protein/protein-ligand complex. The experiment was sealed by inverting the coverslip over the well and allowed to equilibrate for 1 h at 4 ˚C, whereafter 0.8 μL of a solution of MtbThiL crystallization nuclei from a 10- to 10000-fold diluted solution was added. Crystallization nuclei were prepared by combining 48 μL reservoir solution with 2 μL from a previous successful crystallization experiment. The mixture is vortexed with a glass bead to break crystals into nuclei. A 1:10 dilution is repeated serially to ensure only a few nuclei are added to each crystallization drop.

C. Co-crystallisation and soaking

To demonstrate the binding mode of a ligand to a protein by X-ray crystallography, a crystal with a stoichiometric ratio of protein to ligand is required. This can be achieved by co-crystallisation or "soaking". For ligand/protein co-crystallisation, the ligand is added to the protein solution prior to setting up the crystallisation experiment. For soaking experiments, the protein crystal is incubated in a drop of mother liquor complemented with the ligand [112]. In both methods the ligand concentration can be significantly higher than that of the protein to ensure a high occupancy of the binding site.

Apo-MtbThiL was co-crystallised with concentrated solutions of each substrate, in ratios between 1:1 and 1:10 of protein to the substrate. Ten milligrams of MtbThiL was added to equal volumes of 20 mM $MgCl_2$ and substrate and filled up to a final volume of 28 µL with crystallisation buffer. This mixture was incubated for 1 h before being used in crystallisation experiments (see section 2.2.3 B).

Substrates for which co-crystallisation experiments had not yielded an occupied binding site were next investigated by soaking experiments. Apo-MtbThiL crystals were crystallised as described (2.2.3 B). Using 5 µm cryoloops (Hampton Research), MtbThiL crystals were transferred from the mother liquor on the coverslip to solutions of 20 mM ATP, ADP, TMP or TPP, incubated for 15 min, cryocooled in liquid nitrogen and stored in a Dewar for later diffraction experiments.

2.2.4 X-ray crystallography and molecular replacement

X-ray crystallography can reveal the atomic 3D structures of small molecules and macromolecules. When a beam of X-rays, each wave of photons with identical wavelength encounters a protein crystal some of its energy is absorbed by the electrons of the atoms that make up the protein. Some of the energy that was not absorbed is re-emitted or scattered. Where the re-emitted waves interfere constructively i.e. with identical phase and wavelength, the signal is amplified and appears as spots on the detector creating a diffraction pattern. The

relationship between the incident and emergent X-rays is described by Bragg's Law. The Fourier transform is a mathematical model that relates the intensity of the diffraction spots to the amplitude of each wave to construct an electron density map [113].The phases of the wave correspond to positions of high electron density and determine the positions of atoms in the unit cell, information on the phases of the emergent X-rays is lost during diffraction [114, 115].

Various techniques are used to estimate the unknown phase of the structure factors including molecular replacement. In molecular replacement, phases of a known structure model, are calculated and used as approximate phases of the unknown protein. Combining the experimentally determined structure factor amplitudes with these approximate phases allows an electron density distribution map for the unknown protein to be computed and used to improve its partial structure. X-ray diffraction data was collected remotely at the Diamond Light Source on a 32x20 μm beam with an exposure time of 0.015 s per image. A total of 3600 images were collected per crystal with a rotation angle of 0.05° per image. Diffraction data was processed using xia2 3dii.

The previously solved structure of MtbThiL was used as a model, with 100 % sequence similarity. Experimental phases were calculated using the Phaser single wavelength anomalous dispersion (SAD) pipeline in the CCP4 i2 platform [116] and the model built on WinCoot [117], the structure was refined using the phenix.refine program in the Phenix suite [118].

2.2.5 Structure-based inhibitor identification

Molecular docking programs optimize the ligand position to maximize protein-ligand interactions for the most energetically favourable position. Each orientation or pose is assigned an energy score based on the sum of electrostatic and van der Waals interaction as a measure of the structural complementarity between protein and ligand.

All inhibitor identification experiments were conducted in parallel with protein production, purification and crystallisation experiments. In addition, regions at both termini could not be

modelled in the previously solved structure of MtbThiL limiting the usefulness of this structure in docking experiments due to these regions possibly participating in ligand binding. Thus, a homolog of MtbThiL, thiamine monophosphate kinase from *Aquifex aeolicus* (AaThiL) was used in place of MtbThiL for docking. The amino acid sequence of MtbThiL and that of AaThiL were analysed for coverage and sequence similarity, using the basic local alignment search tool (BLAST). Although overall sequence identity between MtbThiL and AaThiL was relatively low (29 %), active site residues across ThiL proteins are conserved.

A. Compound library and protein structure preparation

Compounds were downloaded from open-access databases PubChem [119], ZINC15 [120] and ChEMBL [121], and selected on the basis of their structural and chemical similarities to TMP. One hundred compounds were selected and energy minimised 3D structures of compounds prepared using LigPrep on Schrödinger [122].

Thiamine monophosphate kinase from *Aquifex aeolicus* (AaThiL) PDB (Protein data bank) code 3C9T [90] was downloaded and prepared for docking using the Protein Preparation Wizard on Schrödinger [122]. AaThiL was prepared to correct for errors that may have occurred during molecular replacement. These include filling in missing hydrogen atoms and incomplete side chains and correcting the orientations of amino acid functional groups on the protein structure. A flexible receptor grid isolating the target binding region was created to prevent interaction with other residues outside of the binding site.

B. Docking

Docking was conducted in two phases using Glide of Schrödinger [123]. In both phases TMP was first docked and the results used as a reference point for selecting best binding compounds. In each phase, the 3D compound was docked and oriented in the TMP binding site to get the 'best pose' i.e. the orientation of the compound that forms the most interactions while also being the

most energetically favourable. The first docking phase was with Glide standard precision (GlideSP) which has less stringent docking parameters. This phase was used to select compounds that met the minimum docking score and binding energy determined by TMP GlideSP results. A second and stricter docking protocol, Glide extra precision (GlideXP) was used to refine the data resulting in only compounds that had binding energies that were better than that of TMP. From the one hundred candidate compounds ten were chosen as 'leads' based on their binding energies. The ADMET properties of compounds selected for downstream experiments were determined *in silico* using QikProp on Schrödinger.

C. Docking data validation

To confirm the results from the docking experiment, docked structures were analysed using Pymol and WinCoot [117, 124]. For all predicted interactions, interaction distances and angles were measured, analysed and compared to TMP docking data and AaThiL-TMP co-crystallised structures. Ten compounds had higher binding energies and Glide scores than TMP and were selected as potential MtbThiL inhibitors.

D. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is a label-free biophysical technique used to quantify the energetics of molecular interactions. During biochemical reactions, heat changes occur as a result of molecular interactions between the reacting molecules. Titrating one partner (ligand) into the other (sample), initiates the process of binding resulting in the absorption or release of heat. This heat is directly proportional to the amount of ligand that binds the sample and the unique binding enthalpy of the reaction. The heat released or absorbed is a direct measure of the change in enthalpy (ΔH), while change in entropy (ΔS), equilibrium dissociation constant (K_D) and the reaction stoichiometry (N) may be derived by assuming a type of reaction and fitting a

corresponding curve to the derived data. These parameters describe specific protein-protein and protein-ligand interactions [125].

The concentration of MtbThiL was determined using the Beer-Lambert equation [126] and the extinction coefficient of MtbThiL determined from its amino acid sequence. Purified MtbThiL was equilibrated in ITC buffer by dialysis. The protein was dialysed using a 10 kDa molecular weight limit Dialysis SnakeSkin tubing and kept overnight at 4 °C in 5 L of ITC buffer with constant agitation. Dialysed MtbThiL was then concentrated to 20 mg or higher. Stock solutions of TMP, ATP, and oxythiamine were prepared in ITC buffer to a concentration of 5 mM each. The ITC experiment was carried out on a MicroCal iTC 200 Titration Calorimetry system.

All experiments were carried out at 25 °C with a stirring speed of 1000 rpm and 21 injections of 1.2 µL each. Pre-injection delays (0.8 s) were used to determine a flat baseline for the experiment. Buffer-to-buffer control titrations were used to eliminate false heat changes due to heat of dilution that occurs when the sample and ligand are in different buffers. Curves were fitted using the one binding sites model on Origin 7 software.

2.2.6 Thiamine monophosphate kinase from *Mycobacterium tuberculosis*, Pyruvate kinase-Lactate dehydrogenase coupled enzyme assay

Enzyme kinetics aims to quantify the rate (V_{max}) and the Michaelis constant (K_M) of an enzyme catalysed reaction. Coupled enzyme reactions determine the activity of a specific enzyme by coupling it with another enzymatic reaction(s) that is easily detectable.

The pyruvate kinase-lactate dehydrogenase (PK-LDH) enzyme assay employs two reactions involved in glycolysis to determine the kinetics of MtbThiL. The end products of the MtbThiL catalysed reaction are TPP and ADP (Fig. 1.3, highlighted in red). The ADP produced in the MtbThiL catalysed reaction is coupled to the PK reaction where ADP is converted to ATP and pyruvate. The pyruvate is then coupled to the LDH reaction where its reaction with pre-defined

amounts of reduced nicotinamide adenine dinucleotide (NADH), produces lactate and oxidised nicotinamide adenine dinucleotide (NAD$^+$) (Fig. 2.2). The decay in NADH is monitored spectrophotometrically at 340 nm and the absorbance used to determine the rate of ADP production by the MtbThiL catalysed reaction.

$$\text{Phosphoenolpyruvate + ADP} \xrightarrow{\text{Pyruvate kinase, Mg}^{2+}} \text{Pyruvate + ATP}$$

$$\text{Pyruvate + NADH} \xrightarrow{\text{Lactate dehydrogenase}} \text{Lactate + NAD}^+$$
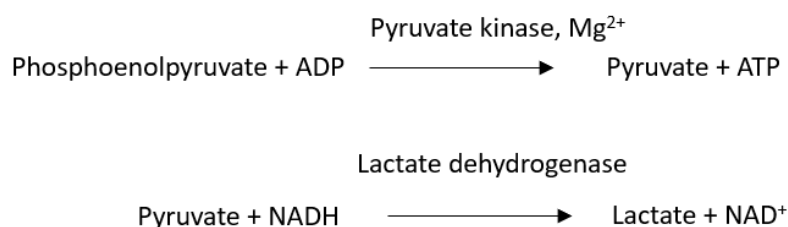
Figure 2.2: Coupled reporter reaction. Pyruvate kinase and lactate dehydrogenase catalysed reactions used to detect the amount of ADP released by the MtbThiL catalysed reaction.

A. MtbThiL reaction set-up

To determine the initial velocity (V_0), the reaction was set up at constant TMP concentration (10 mM). Different concentrations of ATP (1, 5, 10 and 15 mM) were added to 5 µM of MtbThiL, 50 mM Tris-HCl pH 7.8, 5 mM MgCl$_2$, 10 mM KCl and brought to a final volume of 300 µL with dH$_2$O. All reactions were incubated at 37 °C with shaking at 100 rpm in an Eppendorf Thermomixer 5436. A 50 µL sample of each reaction was collected at time intervals (0, 1, 3, 5 and 7 h) and the reaction stopped by heating at 95 °C. The reaction contents were centrifuged at 10 000 x g and the supernatant collected.

B. Reporter reaction set-up

One hundred millimolar imidazole-HCl buffer, 13 mM NADH and 25 mM phosphoenolpyruvate were added to the reaction mixture described in section 2.2.6 A. Pyruvate kinase (0.1 U) and lactate dehydrogenase (10 µL) were then added consecutively to the reaction and incubated at 30 °C for 10 min with shaking at 140 rpm on an Eppendorf Thermomixer 5436. The amount of ADP produced by the MtbThiL reaction was indirectly detected through the amount of NADH

remaining in solution and spectrophotometrically determined at 340 nm on a SpectraMax Paradigm Multi-Mode Detection Platform. An absorbance mono cartridge with an endpoint read type was used, the absorbance was detected at 25 °C. Each reading was repeated 5 times and the experiment was done in triplicate.

To test the inhibitory capabilities of the potential inhibitor identified in section 2.2.5 B (oxythiamine) equal amounts of TMP and the inhibitor (10 mM) were added in addition to the reaction mixture, the amount of ADP produced was determined as outlined above.

C. Kinetic data analysis

The change in absorbance was calculated by subtracting the blank measurement from the absorbance measurement of the reaction. Graphs of change in absorbance versus time were plotted for each ATP concentration and the initial velocity determined from the slope of each graph and the extinction coefficient of NADH. V_0 was calculated in units of ΔAbs_{340}/h. The Beer-Lambert equation and the extinction coefficient of NADH were used to convert the absorbance values into units of M/h. Each V_0 value was plotted against substrate concentration to construct the Michaelis-Menten graph on GraphPad Prism 3.0 [127] , following observations that MtbThiL was not following Michaelis-Menten kinetics. We used Microsoft Excel to plot the graphs. The Michaelis-Menten graph (initial velocity against substrate concentration) was plotted to determine kinetic parameters, K_M, V_{max}, and k_{cat}.

The inhibition reaction was set-up as previously described with concentrations of oxythiamine (10 mM) equal to those of TMP and treated as described above. Graphs of these reactions were also constructed on GraphPad Prism. To determine whether there was any statistically significant difference between the repeats, a one-way analysis of variance (ANOVA) was used to calculate the p-value at a 5 % confidence interval on R statistical software [128].

3. Results

3.1 Confirming thiamine monophosphate kinase gene in plasmid expression vector

The recombinant *mtbthil*-pET28a construct was provided by Dr Collin Kenyon (CSIR, Biosciences). We confirmed the presence of the *mtbthil* gene (H37Rv strain) in the *mtbthil*-pET28a construct by Sanger sequencing and restriction enzyme digestion in a previous study [100, 129, 130].

Restriction enzyme digestion showed the presence of the expected ~1.0 kbp *mtbthil* gene insert in the ~5.6 kbp pET28a plasmid. The insert had been codon-harmonised, for increased protein yield, prior to restriction enzyme digestion. Competent BL21-CodonPlus cells were transformed using the plasmid for protein production. Protein production had been optimised in our previous studies [100, 129]. Ideal target protein production conditions were identified at an induction temperature of 28 °C and an IPTG concentration of 0.1 mM.

3.2 Protein production and affinity purification of thiamine monophosphate kinase from *Mycobacterium tuberculosis*

MtbThiL was produced as an N-terminally His$_6$-tagged fusion protein in BL21-CodonPlus cells as outlined in section 2.2.1. Protein production and affinity purification was analysed using SDS-PAGE.
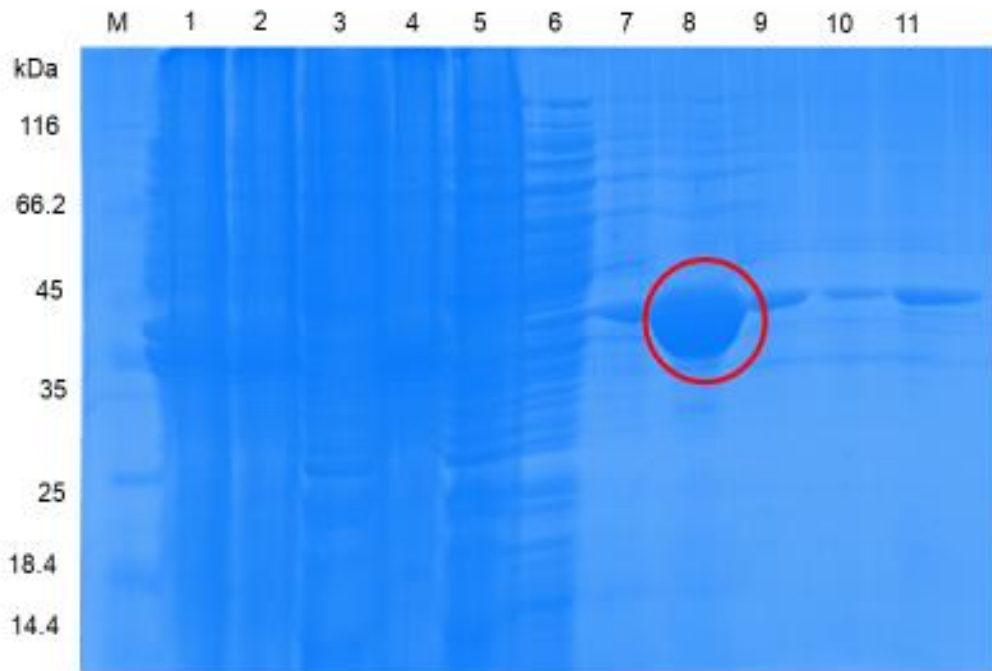
Figure 3.1: Analysing the production and purification of MtbThiL-His6 fusion protein. M: Molecular weight marker, **1**: Sample before induction, **2**: Sample after induction **3**: Soluble fraction, **4**: Insoluble fraction, **5**: Flow-through of non-binding proteins, **6**: Wash fraction 1, **7**: Wash fraction 2, **8**: Elution 1 (red circle), **9**: Elution 2, **10**: Elution 3, **11**: Beads after elution of target protein.

Analysing MtbThiL protein production by SDS-PAGE revealed the appearance of a ~36 kDa protein band after induction with 0.1 mM IPTG (lane 2). As the size corresponds to that expected for MtbThiL, this was taken to confirm the successful production of the target protein. Moreover, the fusion protein was observed in the soluble fraction (lane 3), indicating that MtbThiL was solubly produced in the cytoplasm.

Samples from an affinity chromatography experiment were analysed on the same gel, lanes 6 to 11. Both wash fractions in lanes 6 and 7 showed a large number of protein bands alongside a weaker band at ~36 kDa, indicating that although many contaminants were removed, some target protein is lost as well. Samples of the eluted target protein, lanes 8-10, showed a prominent band at ~36 kDa plus other weaker bands, indicating that soluble MtbThiL had previously bound to the Ni-NTA column and was now being released from the column by

competition with excess imidazole. The MtbThiL band is still visible on the beads after elution

sample, lane 11, indicating that a fraction of the target protein had not been eluted from the

beads. The protein remaining on the beads was eluted by increasing the amount of imidazole in

the elution buffer. Approximately 32 mg of partially pure MtbThiL was obtained from 1 L of

medium. This was deemed sufficient for downstream experiments.

To further purify the enriched MtbThiL sample from affinity chromatography and to exchange

buffers from elution to crystallisation buffer, we further purified the protein by size exclusion

chromatography using a Superdex 200 10/300 GL column equilibrated with 50 mL of crystallisation

buffer.
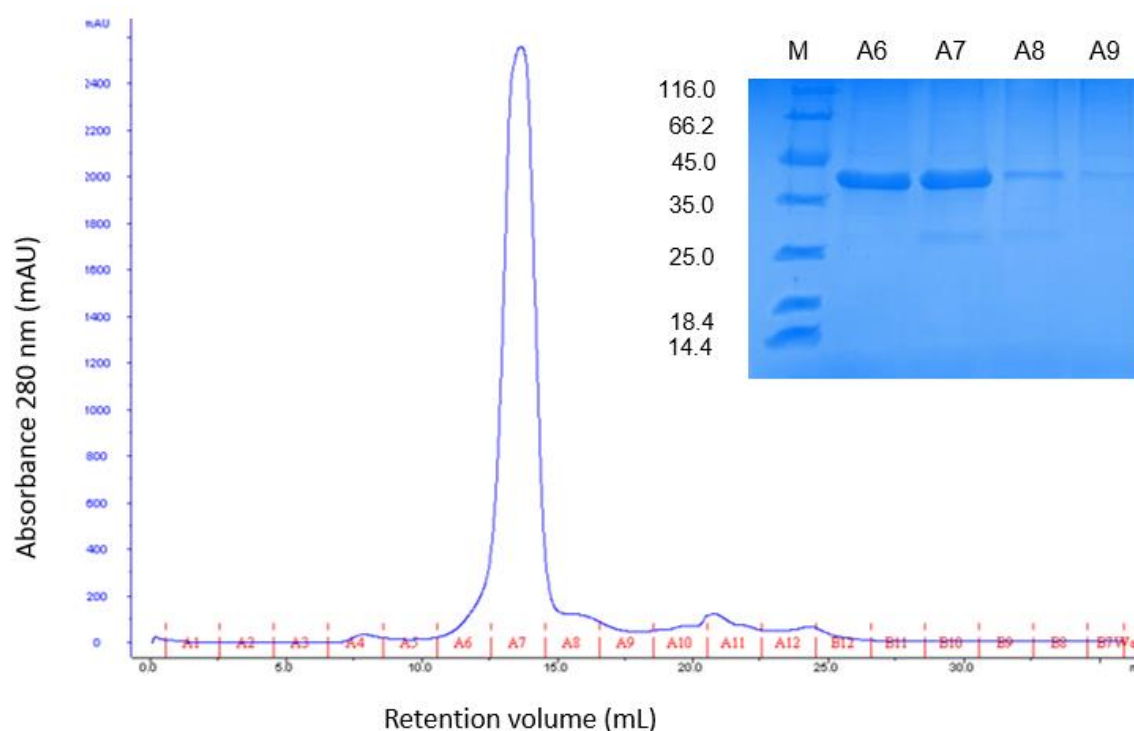
Figure 3.2: Purification of MtbThiL by size exclusion chromatography. A single absorbance peak was
observed at 280 nm between elution volumes 13 and 15 mL. Fractions corresponding to the peak were
analysed on SDS-PAGE (insert). Lane 1 (M): protein molecular weight marker. Fractions A6 to 8 from the
chromatogram correspond to lanes 2 to 4.

Size exclusion chromatography revealed a single absorbance peak with a retention volume of

~13.7 mL (Figure 3.2). Analysing the corresponding fractions by SDS-PAGE (gel insert, lanes 6-8)

showed prominent protein bands at ~36 kDa implying the purification of MtbThiL by SEC was

largely successful. Additional minor protein bands on the gel indicate possible degradation during storage.

3.3 Biophysical analysis through dynamic light scattering

To determine the homogeneity of particle sizes in the MtbThiL protein solution dynamic light scattering was used.
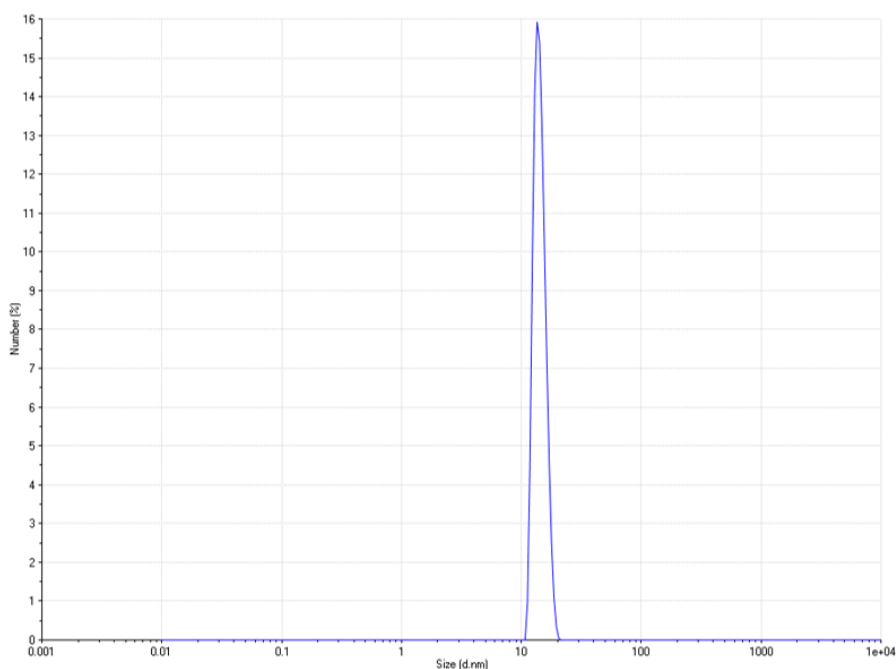
Figure 3.3: Dynamic light scattering size distribution graph for MtbThiL. The graph shows the percentage of differently sized particles in solution plotted against their hydrodynamic diameter in nm.

A single peak was observed from the DLS size distribution graph, indicating that MtbThiL is monodisperse. This suggests that all particles in solution had uniform hydrodynamic diameters and thus the protein sample was pure. This also demonstrates that MtbThiL was stable in the crystallisation buffer i.e. it did not precipitate or aggregate nor were there any impurities, prior to crystallisation validating the results of our SEC experiment in section 3.2. The hydrodynamic diameter was ~11 nm corresponding to the approximate size of MtbThiL. The approximate size of MtbThiL was determined by its amino acid composition and the combination of attractive, repulsive and hydrophobic interactions that occur between the amino acids and with the solvent.

3.4 Protein crystallisation

Previous studies on ThiL proteins showed that they crystallise readily when polyethylene glycol (PEG) is used as the precipitant [98]. Thus, the Hamptons PEG II suite was used to screen for suitable MtbThiL crystallisation conditions. The sitting drop vapour diffusion method was used, and crystallisation set up in 96 well plates. The plates were incubated at 4, 7 and 18 °C while the reservoir and protein drop equilibrated, i.e. until the vapour pressure for the protein drop and the reservoir had equalized. Each plate was then analysed under the microscope for possible protein crystallisation.

Protein amounts between 2.5 and 12.5 mg were used for screening to determine the optimal amount of protein to crystallise. Crystals of varied shape and sizes were observed between three and four days without seeding. A single crystal from each well was fished and broken as a preliminary step to determine whether it was a salt or protein crystal. Positive screening conditions resulting in possible protein crystals were identified and optimised further using hanging drop vapour diffusion, by altering the concentration of each component in each crystallisation condition. The amount of glycerol was optimised in previous studies. The crystallisation experiment for MtbThiL was determined to be optimal at 4 °C and at MtbThiL amounts between 8 and 10 mg, protein amounts above this range resulted in protein precipitation and those below resulted in small needle shaped micro-crystals. Co-crystallisation was also optimised by adding protein and ligand(s) in different ratios. Ratios above 1: 1.3 protein to ligand resulted in protein precipitation and no crystal growth.

Table 3.1: Crystallisation screening results

Screening hit	Composition
1	0.1 M Tris-HCl pH 8.5, 25 % w/v PEG 400, 0.2 M LiSO$_4$, 1 % glycerol
2	30 % w/v PEG 4000, 1% glycerol
3	0.1 M MES 6.5, 20 mM MgCl$_2$, 10 % PEG 4000, 1% glycerol
4	0.1 M C$_2$H$_3$NaO$_2$ pH 4.6, 30% PEG 4000, 0.2 M (NH$_4$)$_2$S$_2$O$_8$, 1 % glycerol
5	0.1 M HEPES pH 7.5, 30 % w/v PEG 4000, 0.2 M CaCl$_2$, 1 % glycerol

(a) (b)
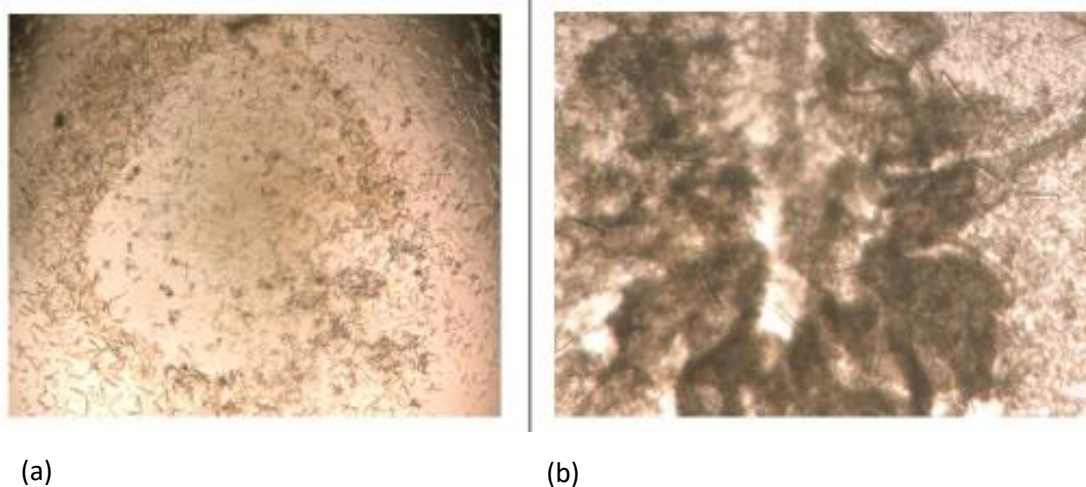
Figure 3.4: MtbThiL crystals obtained from screening hits and optimisation. (a) Micro-crystals obtained from screening condition 2. (b) MtbThiL hexagonal crystals obtained from optimised crystallisation condition 3.

Condition 5 yielded salt crystals and condition 2 yielded micro-crystals (Fig. 3.4 (a)). Condition 3 yielded rod-shaped hexagonal crystals that grew to ~25 µm after seeding and incubation for two days at 4 °C (Fig. 3.4 (b)). These crystals were fished and cryo-cooled in liquid nitrogen, the cryo-protectant (1 % glycerol) was included in the crystallisation condition to reduce handling of the crystals. Conditions 1 and 4 also yielded crystals but were not optimised.

3.5 X-ray crystallography

MtbThiL crystals were diffracted and data collected remotely at the Diamond Light Source (Table 3.2). All diffracted crystals were in the P6$_1$22 space group, the best crystal diffracted to a

resolution of 2.19 Å. The previously solved structure of MtbThiL [100] (resolution 2.8 Å, 100% sequence identity) was used as a model for molecular replacement on Phenix [118].The collected data (reflection file), the model and MtbThiL amino acid sequence files were used to generate phase data. The model and electron density map were used to build and refine the structure of MtbThiL on Phenix and WinCoot [117] (Table 3.3).

Table 3.2: Data collection statistics

Data collection statistics	
Beamline	MX20303-5
Wavelength range (Å)	0.916Å
Resolution range (Å)	51.3 - 2.19
Space group	$P6_122$
Unit cell parameters (a, b, c, α, β, γ)	60.4 Å,60.4 Å, 292.3 Å; 90°,90°,120°
Total reflections*	30244 (11503)
Unique reflections*	822 (1051)
Multiplicity	28.9
Completeness %*	100 (98.6)
Wilson B factor (Å^2)	47.7
R_{meas}*	0.05 (1.83)
I/sig(I)*	52.3 (1.2)
CC half*	0.5 (1)

*Values in parentheses represent the shell of highest resolution

Table 3.3: Refinement statistics

Refinement statistics	
Reflections used in refinement*	17162 (1642)
Reflections used for R-free*	834 (84)
R-work*	0.20 (0.25)
R-free*	0.27 (0.29)
Number of non-hydrogen atoms	2150
Protein residues	279
RMS (bonds lengths) (Å)	0.01
RMS (bond angles) (°)	0.98
Ramachandran favoured (%)	95.7
Ramachandran allowed (%)	1.0
Ramachandran outliers (%)	2.2
Clash score (%)	2.1
Average B-factor	15.7

*Values in parentheses represent the shell of highest resolution

The R-free and B-factor are measures of the quality of the atomic model obtained from the crystallographic data. The B-factor measures the certainty of atom positions in the crystal structure while the R-free value measures the consensus between the final model and the experimental data. High B-factors indicate uncertainty regarding atomic positions, while a high R-free value indicates the degree of difference between the model and the experimental data. Both values in our data are relatively low, an indication of a good atomic model.
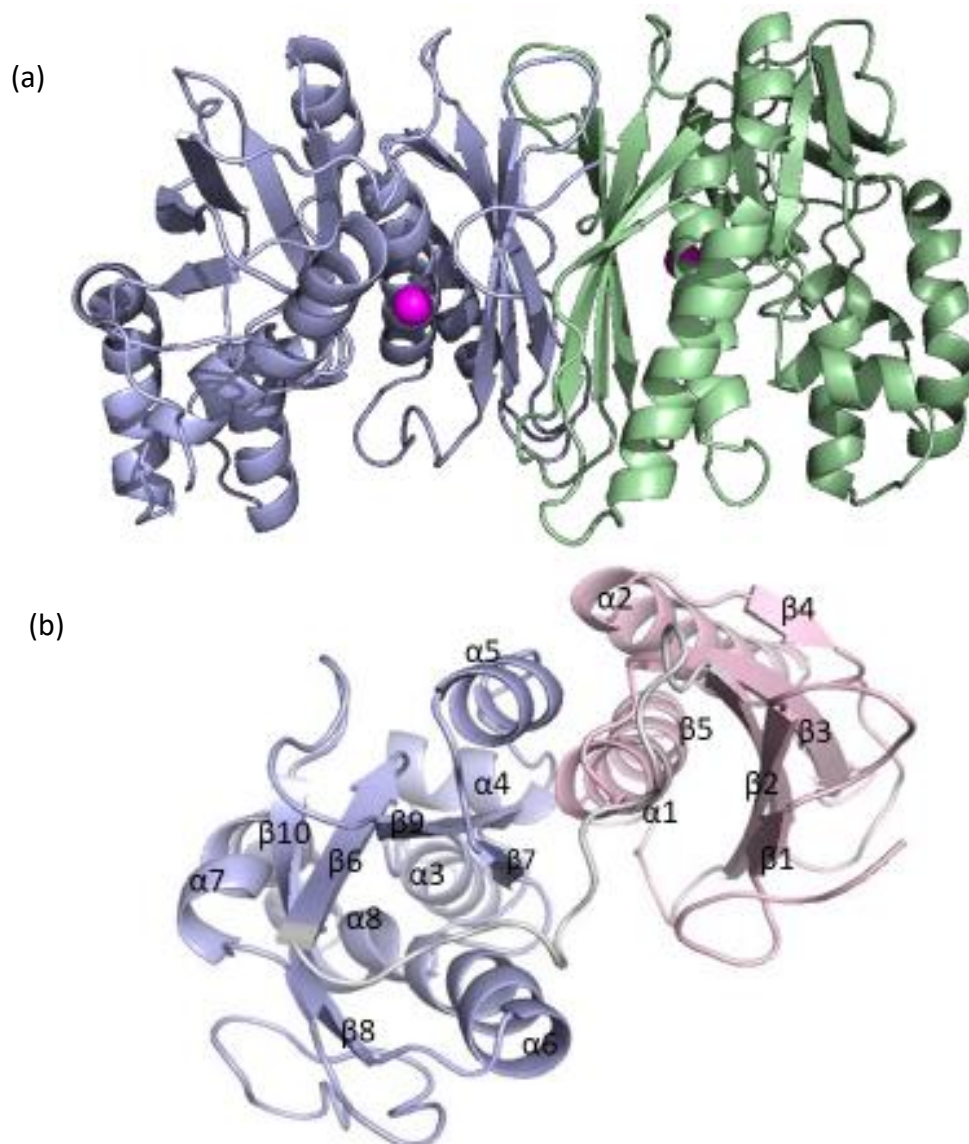
Figure 3.5: Crystal structure of MtbThiL homodimer. (a) MtbThiL occurs as a homodimer in solution and in the crystals. In the analysed crystals, a single monomer occupies the asymmetric unit, with the second monomer of the dimer related to the first by a crystallographic axis. Monomers are shown in blue and green. Magenta spheres represent Mg^{2+}. **(b)** The constituent domains of the MtbThiL monomer are shown in pink (A domain) and blue (B domain). They are linked by a loop shown in grey. Secondary structures are labelled for each domain.

MtbThiL is a homodimer with identical monomers related by crystallographic symmetry. The monomers interact through a central, eight stranded β-sandwich. The crystal structure comprises 276 of the possible 353 residues, with insufficient electron density to model the 59 N-terminal and 16 C-terminal residues. This occurs when regions are too flexible and disordered to assign electron density to. Each monomer has an A and B domain, the smaller N-terminal A domain is composed of 106 residues with an αβ fold. This domain is arranged in a βαβ motif, with predominant β strands of varying lengths. The C-terminal B domain consists of 127 residues with shorter secondary structures arranged in βαβ and αβα motifs. Peaks in electron density with octahedral arrangement of water molecules and co-ordination distances of ~2 Å were assigned to Mg^{2+} since it was the only metal ion with these properties in the crystallisation experiment and Mg^{2+} has been shown to bind to the active site of other ThiL proteins.

To compare the previously solved structure of MtbThiL and the one solved in this study, the two structures were superimposed and studied on Pymol.

Figure 3.6: Superimposed MtbThiL structures. The superimposed structure of MtbThiL solved in this study (blue and green) with the previously solved structure (grey). Differences between the structures 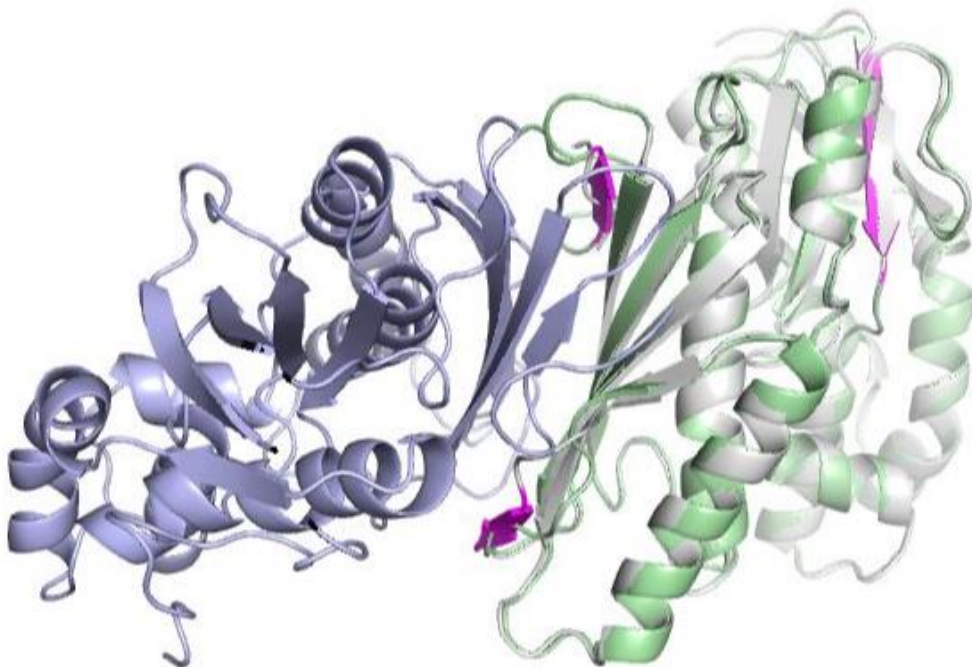are shown in magenta. The structure of MtbThiL solved in this study has an additional 22 residues that form a loop and a short β1-strand at the N-terminal. The previously solved structure has a β-strand between

residues 157-160 that form a loop in the structure solved in this study. The β6 strand in the B domain also appears longer in MtbThiL from this study. Image prepared on Pymol.

Compared to the previous crystal structure of MtbThiL, our structure includes an additional 22 residues at the N-terminus that form a loop and a short β-strand. This strand and loop also observed in other ThiL protein structures (magenta). An additional β-strand reported for the previous MtbThiL structure before β4 (A-domain), reflects as a loop in the current structure, in line with other ThiL structures. Similarly, β6 strand in the B domain (magenta) is longer in the current structure.

To compare the structure of MtbThiL with other ThiL proteins, MtbThiL was superimposed with the co-crystal structures of thiamine monophosphate kinase from *Acinetobacter baumanii* (AbThiL) (PDB code 5DD7) and thiamine monophosphate kinase from *Aquifex aeolicus* (AaThiL) (PDB code 3C9T).
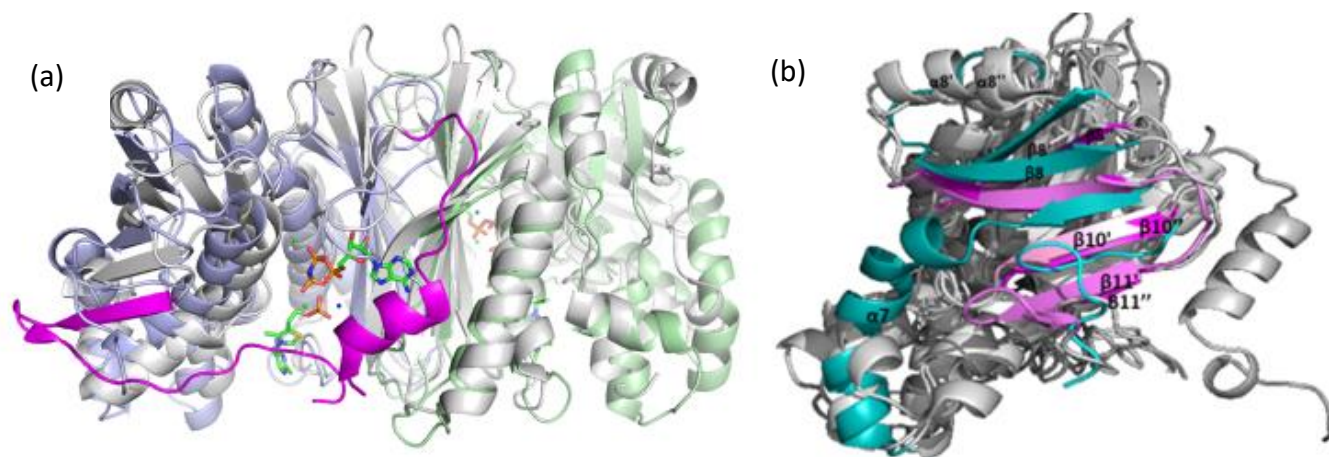
Figure 3.7: Superimposed ThiL structures. The superimposed structure of MtbThiL (blue and green) and AbThiL (PDB code 5DD7) (grey) left. The ANP-PNP and TMP bound onto the AbThiL structure show possible positions of the TMP and ATP binding sites on MtbThiL. Secondary structures shown in magenta represent the N- and C-terminal regions missing in our MtbThiL crystal structure. Also shown are co-factors, Mg^{2+} and K^+ near the substrates involved in metal co-ordination. Superimposing the B domains (right) revealed some differences between MtbThiL (cyan), AbThiL (light pink) and AaThiL (magenta). Secondary structure elements of AbThiL are indicated with a single apostrophe while those of AaThiL are indicated with two apostrophes.

MtbThiL and AbThiL show overall sequence and structural homology with differences in both domains. AbThiL has an additional α-helix α1 and β-strand β0 in the N-terminal A domain. Corresponding residues are unresolved in MtbThiL. The β0-strand, α1 helix and the loop appear to enclose the TMP and ATP binding sites that are located on a single cleft in each domain (Fig.3.7 left shown in magenta). The C-terminal B domains are less homologous, with α7 (cyan helix) in MtbThiL composed of 18 residues while the corresponding helices in AbThiL and AaThiL are shorter with 6 residues each and the remaining residues in a loop. The α8' helix found in AbThiL corresponds to a loop in MtbThiL and a 2 residue α8''-helix in AaThiL. Strand β8 has opposite strand order when compared to AbThiL and AaThiL while strands β10 and β11 in both AbThiL and AaThiL are appear as loops in MtbThiL (Fig. 3.7 right).
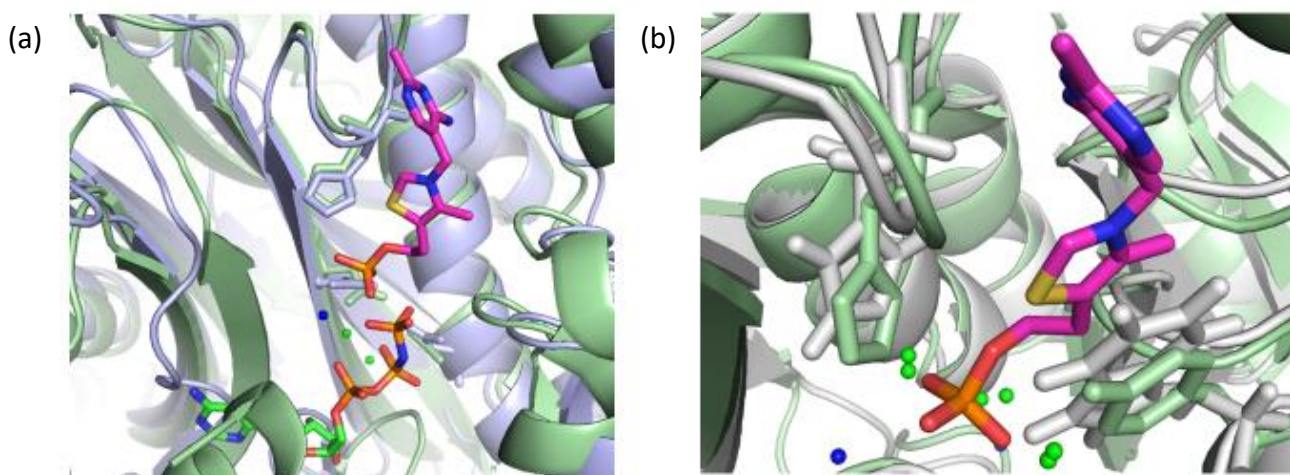
Figure 3.8: Close-up view of ThiL active site. (a) Image of superimposed MtbThiL (pale blue) and AbThiL (pale green) structures showing conserved residues. (b) AbThiL (pale green) and AaThiL (grey) superimposed structures showing C- and N-terminal regions that are ordered by TMP binding and conserved residues involved in interacting with TMP (shown in magenta in Fig. 3.7).

Conserved amino acids Asp 43, His 50, Asp 71 and, Asp 207 predicted to be involved in catalysis in studies done on AaThiL [90] occur at approximately the same positions and orientations in AbThiL and MtbThiL and interact with the same functional groups on AMP-PNP and TMP. The C-terminal loop that is absent in our crystal structure is ordered by TMP binding in AaThiL and

AbThiL. This suggests that the flexibility of this region in our structure could have been reduced by ligand binding.

3.6 Thiamine monophosphate kinase from *Mycobacterium tuberculosis* kinetics

To determine the kinetic parameters, K_M and V_{max}, and study oxythiamine inhibition on MtbThiL we used the pyruvate kinase-lactate dehydrogenase assay. The decrease in NADH concentration was recorded as absorbance at 340 nm and used to determine the initial velocity of the reaction at different substrate concentrations.
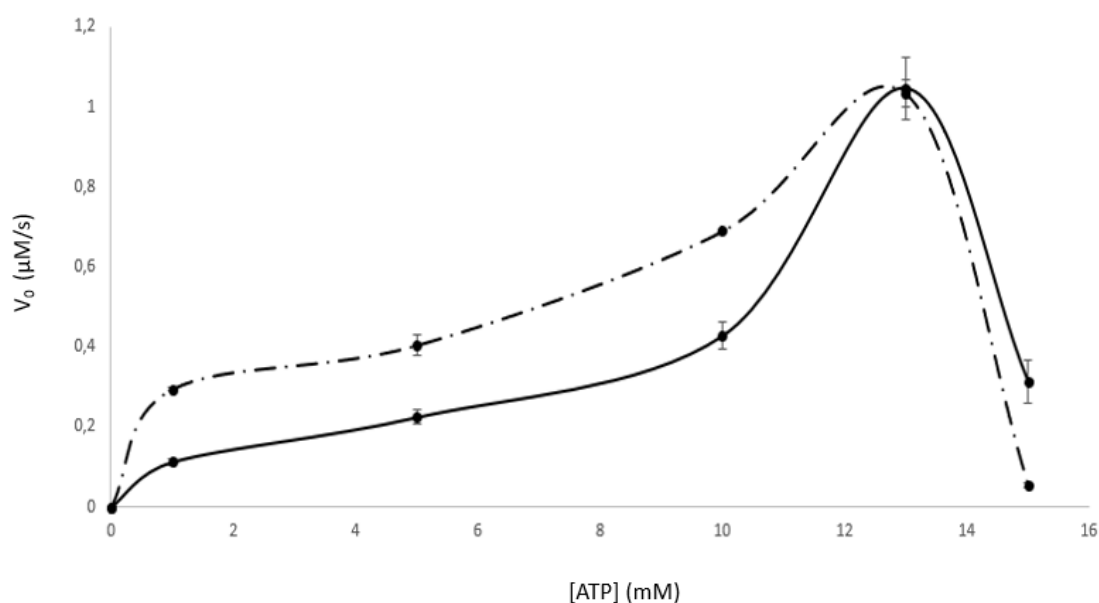
Figure 3.9: Graph describing MtbThiL kinetics. Graph of MtbThiL uninhibited reaction (solid line) and graph of MtbThiL inhibited reaction (dashed line).
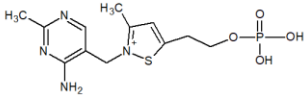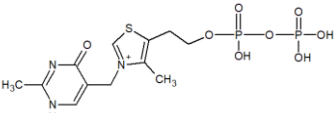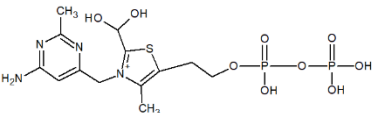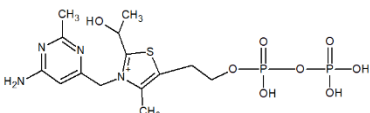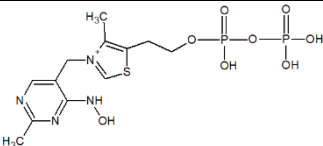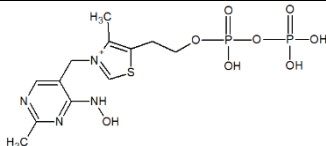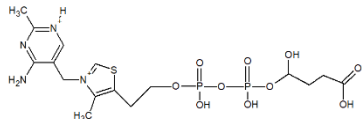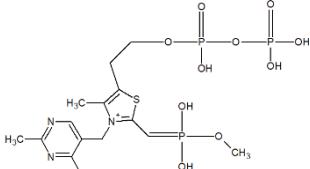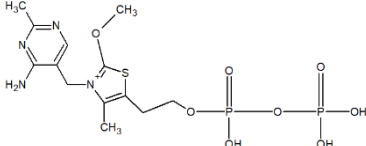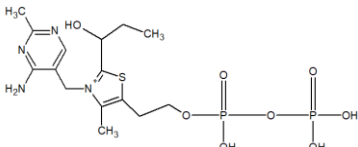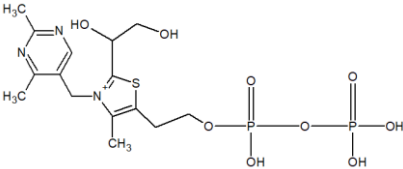
Both graphs indicate an increase in reaction velocity between 0 and 13 mM and a sharp decrease at 15 mM ATP. The graphs do not follow the classic Michaelis-Menten hyperbolic curve suggesting that the reaction cannot be described using Michaelis-Menten kinetics. The highest velocity reached for the uninhibited and inhibited reactions were 1.04 and 1.03 µM/s respectively. Since neither of the graphs plateau at a specific velocity, we cannot plot an asymptote to determine V_{max} or extrapolate to determine the K_M. From our kinetics data we

cannot conclude on the binding affinities of MtbThiL for TMP and oxythiamine. ANOVA statistical analysis showed no significant difference between the experiment and repeats at a 5 % confidence interval (ANOVA: p-value = 0.14 F-ratio value (variation between samples) = 2.06).

3.7 Structure-based inhibitor identification

To identify potential MtbThiL competitive inhibitors, one hundred compounds with a similar general structure to that of TMP were docked using Glide on Schrodinger's Maestro suite.

Table 3.4 Docking lead compounds, binding energies and docking scores

TMP
ΔG (kJ mol⁻¹):   -22.260
Docking score: -3.4850

Compound 1
ΔG (kJ mol⁻¹): -30.244
Docking score: -3.736

Compound 2
ΔG (kJ mol⁻¹):   -32.916
Docking score: -4.237

Compound 3
ΔG (kJ mol⁻¹): - 32.571
Docking score: -8.108

Compound 4
ΔG (kJ mol⁻¹): -34.053
Docking score: -4.283

Compound 5
ΔG (kJ mol⁻¹): -33.250
Docking score: -10.331

Compound 6
ΔG (kJ mol⁻¹): -131.847
Docking score: -8.123

Compound 7
ΔG (kJ mol⁻¹): -128.261
Docking score: -11.516

Compound 8
ΔG (kJ mol⁻¹): -126.440
Docking score: -6.366

Compound 9
ΔG (kJ mol⁻¹): -124.817
Docking score: -12.222

Compound 10
ΔG (kJ mol⁻¹): -121.5414
Docking score: -7.169

From our docking studies we identified ten compounds that bound to the TMP binding site with a binding energy that was higher than that of TMP. The pose prediction scores or docking scores of the best poses for each compound are also shown (Table 3.4). While other compounds had significantly higher binding energies in comparison to compound 1 (oxythiamine), they were not available for purchase requiring synthesis which was beyond the scope of this study. To validate the interactions predicted by Glide, the interacting or binding partners were studied on Pymol and WinCoot. As a control, TMP was docked onto the TMP binding site.
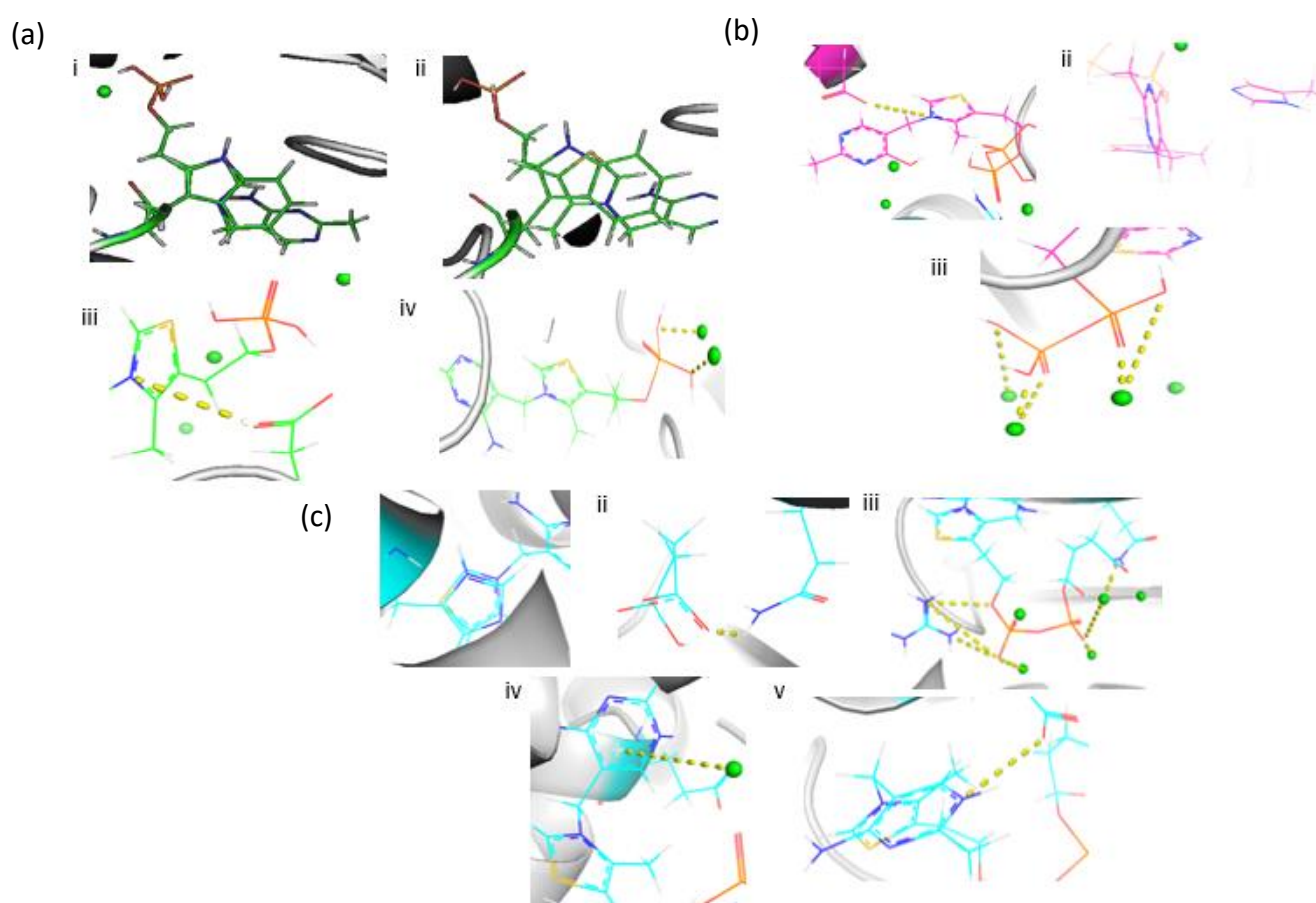
Figure 3.10: Validated docking data for interactions predicted by Glide. (a) TMP control docking showing (i) parallel π-π stacking interaction with pyrrole ring on the indole of Trp 303 (ii) staggered π-π stacking with the benzene ring on the indole of Trp 303 (iii) salt bridge interaction between N^+ and carboxyl oxygen of Glu 8 and (iv) salt bridge interactions between Mg^{2+} ions and hydroxyl oxygens of phosphate group on TMP. **(b)** oxythiamine docking showing (i) salt bridge interaction (ii) NH-π stacking interaction, (iii) four salt bridge interactions involving the phosphate groups of oxythiamine and the Mg^{2+} in the active site. **(c)** best binding compound 6 showing (i) parallel π-π stacking interaction, (ii) salt bridge interaction with side chain carboxyl group of Asn 119 (iii) two hydrogen bonds with Arg 142 and salt bridge interactions with Mg^{2+}, (iv) π-cation stack with Mg^{2+} (v) salt bridge interaction with carboxyl group of Glu 8.

The control docking experiment with TMP revealed mostly π- and salt bridge interactions. Validated TMP interactions (Fig. 3.10 (**a**)) include the following interactions between the following partners with bond distances and angles: pyrrole of Trp 303: thiazole parallel π-π stacking 3.7 Å, 0° dihedral angle (i), benzene of Trp 303: thiazole staggered π-π stacking 3.2 Å, 7.5° dihedral angle (ii), pyrrole Trp 303: N^+ of thiazole π-cation ~3.4 Å (iii), sp^2 hybridized carbonyl oxygen of Glu 260: N^+ salt bridge 3.4 Å 99 ° (iv), Mg^{2+} 309 and Mg^{2+} 311: β and (v) γ-phosphate groups metal co-ordination and a covalent bond 3.2 and 3.1 Å.

Docking analysis of oxythiamine showed fewer interactions with TMP binding site residues. Some of the predicted interactions could not be validated since the binding partners were too far apart for the interaction type to occur. The validated interactions predicted for oxythiamine binding (Fig. 3.10 (**b**)) are as follows: carboxyl oxygen of Glu8: N^+ on thiazole ring salt bridge (a), NH on imidazole of His 208:pyrimidine on oxythiamine π-NH stacking interaction, Mg^{2+} 9 and 10: carboxyl and hydroxyl oxygen on β phosphate group two salt bridges, Mg^{2+} 9 and 11: carboxyl and hydroxyl oxygen on α phosphate group two salt bridges.

Compound 6 formed multiple interactions with active site residues and Mg^{2+}, some of which occur in TMP binding. The validated interactions (Fig. 3.10 (**c**)) include imidazole on His 218: thiazole on compound 6 π-π stacking 3.8 Å (a), carboxyl oxygen of Glu 8: NH on pyrimidine hydrogen bond 3.5 Å (b), sp^3 hybridized amino group of Arg 142: sp^3 hybridized oxygen on backbone of compound 6 two hydrogen bonds 3.3 and 3.0 Å (c), salt bridges between Mg^{2+} 309, Mg^{2+} 311, Mg^{2+} 312 and carboxyl and hydroxyl oxygens on phosphate group 4.2, 3.8, 2.0, 4.2 Å (c), pyrimidine ring of compound 6: Mg^{2+} 311 π-cation interaction 3.5 Å (d), carboxyl oxygen on Glu 8: amino group on pyrimidine ring salt bridge interaction 3.5 Å.

To experimentally determine the affinity of MtbThiL for TMP and oxythiamine we did ITC (refer to section 2.2.5 D). We did a buffer-to-buffer titration to ensure that the protein and ligand were in the same buffer thus eliminating heat changes not resulting from protein-ligand interaction.
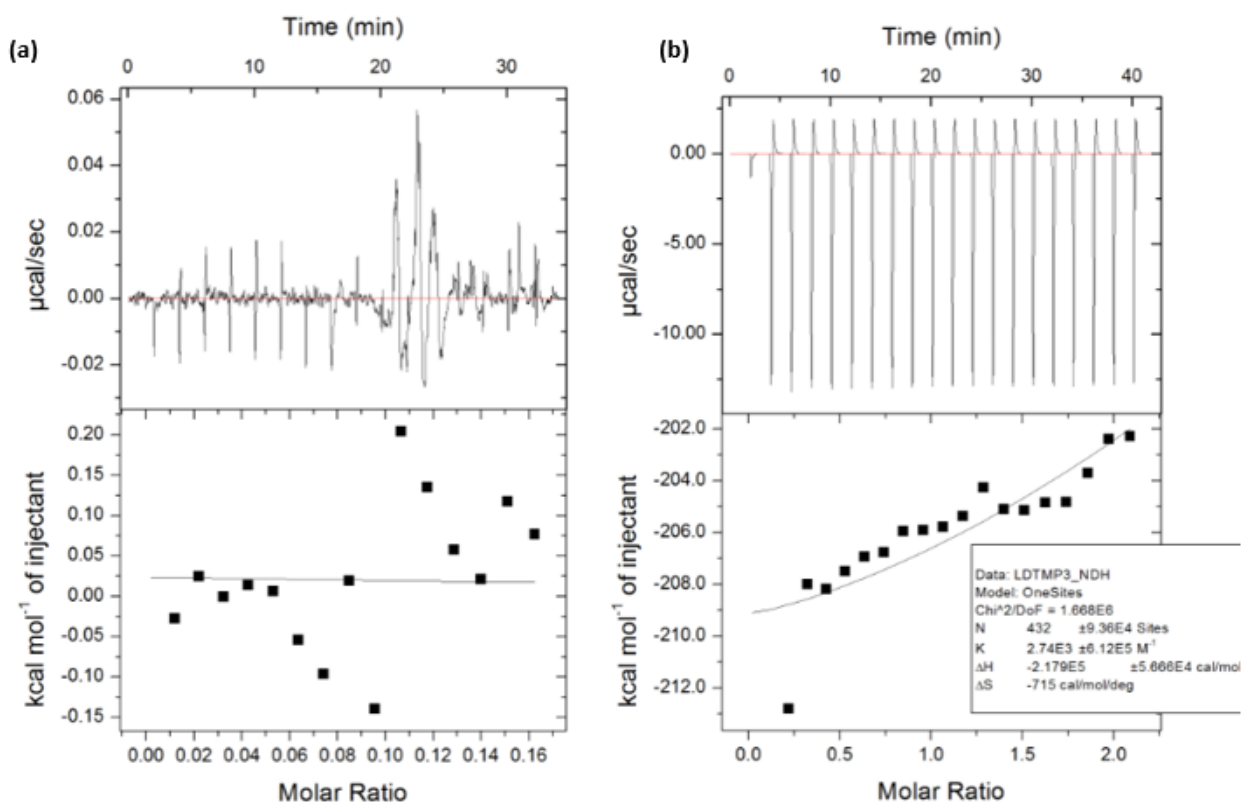
Figure 3.11: ITC control and concentration optimisation experiments. (a) The buffer to buffer control showing minimal heat changes per injection indicated by small deflections from the baseline. (b) Optimisation of the amounts of both interacting partners shows saturation after twenty-one injections in a 1:10 ratio of 20 μM protein to 200 μM ligand. All graphs were best fitted to a one binding site model equation.

The buffer-to-buffer control showed no significant heat changes indicating that the protein and ligands were in the same buffer ensuring that there would be no false signal resulting from heat of dilution (Fig. 3.11(a)). The reaction with 20 μM:200 μM showed saturation after twenty-one injections (Fig. 3.11 (b)). The concentrations were decreased to 10 μM:100 μM to reduce the number of injections before active sites were saturated with ligand. The reaction was therefore, optimised for a 1:10 ratio with 10:100 μM. To determine the binding affinities of MtbThiL for TMP and oxythiamine the optimised experiment was run for both ligands.
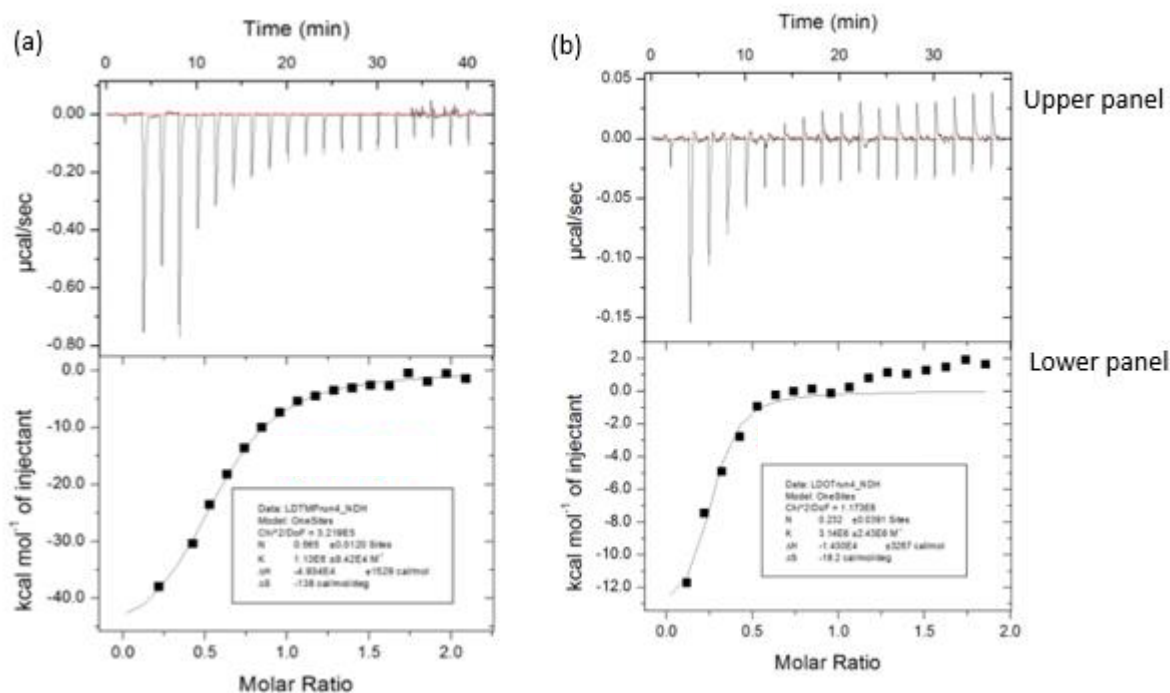
Figure 3.12: Calorimetric raw data and isotherms of ITC MtbThiL-oxythiamine (a) and MtbThiL-TMP interactions (b). The upper panel shows calorimetric raw data. Each spike represents each injection of the ligand where heat is released upon binding and the temperature of the sample cell brought back to the baseline. The energy required to bring the sample cell temperature back to the baseline is converted to power, measured in µcal/sec and plotted against time. The lower panel represents each injection heat standardised based on the amount of ligand injected and plotted as a function of molar ratio of ligand:protein. The characteristic sigmoid curve is fitted (lower panel). The slope of the curve represents the K_D, the stoichiometry is determined at the midpoint of the graph between complete and no binding and the enthalpy determined as the heat produced by strong ligand binding during the first injection (values shown in box on the lower panel).

Table 3.5: Integrated output data for TMP-MtbThiL and oxythiamine-MtbThiL titrations (label insert figure 3.12)

Data: Thiamine monophosphate Model: OneSites	
N	0.57 ± 0.01 Sites
K	$1.13^6 ± 9.42^4$ M^{-1}
ΔH	$-4.93^4 ± 1529$ cal/mol
ΔS	-138 cal/mol/deg
K_D	**0.9 µM**
Data: Oxythiamine Model: OneSites	
N	0.23 ± 0.04 Sites
K	$3.14^6 ±2.43^6$ M^{-1}
ΔH	$-1.43^4 ± 3267$ cal/mol
ΔS	-18.20 cal/mol/deg
K_D	**0.32 µM**

The downward peaks in (a) and (b) (Fig. 3.12) indicate that binding of TMP and oxythiamine to MtbThiL is exothermic. The first injection shows the largest heat change indicating strong ligand binding. Because the enzyme binding sites progressively get saturated with substrate, fewer interactions occur and the heat released decreases with each injection. The K_D value, reaction stoichiometry (N), change in enthalpy (ΔH), and change in entropy (ΔS) were computed on Origin software with one binding site model. The K_D for the MtbThiL-TMP and MtbThiL-oxythiamine interactions were 0.3 μM and 0.9 μM respectively. This demonstrates that MtbThiL has a higher affinity for TMP than oxythiamine.

The ΔG for each reaction was calculated using the equation ΔG = ΔH − TΔS. The ΔG for the TMP reaction was -8.2 x 10^3 kJ/mol and that of oxythiamine was computed to be -1.1 x 10^3 kJ/mol, these values indicate that both reactions were exothermic and spontaneous and that of TMP was more favourable indicated by the more negative value.

4. Discussion

This study sought to characterise the structure of thiamine monophosphate kinase from *Mycobacterium tuberculosis* and identify potential inhibitors. In this chapter, we discuss the main outcomes of this study.

4.1 Thiamine monophosphate kinase from *Mycobacterium tuberculosis* biophysical analysis

A protein crystal consists of countless copies of the unit cell, each with identical protein molecules that are translationally arranged. This uniformity in molecule type, size and arrangement ensures accurate and high-resolution diffraction revealing the correct positions of atoms. Dispersity is a measure of the size distribution of a protein sample. Polydisperse samples either indicate mixtures of distinct proteins or result from different aggregation states of a single protein. The latter state can impede the formation and growth of crystal nuclei preventing the formation of useful crystals. Our DLS data indicate a single peak in the size distribution of the sample (Fig. 3.3) confirming that only a single protein and more crucially a single oligomerization state of MtbThiL was present in the solution. This signifies a good starting point in creating good quality crystals for downstream analysis by X-ray diffraction [131].

4.2 Thiamine monophosphate kinase from *Mycobacterium tuberculosis* crystallisation

Structural characterisation of protein-ligand complexes often gives some insight into the mechanism of enzyme catalysed reactions. It provides some information on residues that may be important for catalysis and the order in which ligands and co-factors bind to the active site. To elucidate the reaction mechanism of MtbThiL we introduced ligands (TMP, TPP, ANP-PNP, ADP) into MtbThiL prior to crystal formation through co-crystallisation and onto already formed crystals through soaking. The co-crystallised and soaked crystals either did not yield a complex

or diffracted to too low a resolution to determine whether the ligands bound. Thus, our attempts to determine the reaction mechanism of MtbThiL were unsuccessful.

Published literature on phosphate transfer enzymes and the crystal structures of thiamine monophosphate kinases from *Aquifex aeolicus* (AaThiL) and *Acinetobacter baumanii* (AbThiL) suggest that these proteins require an additional ion to Mg^{2+} as a co-factor [132, 133]. We observed evidence of this in MtbThiL during our ITC and kinetics experiments, where the absence of a K^+ salt resulted in a lack of interaction and catalysis. Studies on phosphoryl transfer enzymes have shown that K^+ directly interact and orient phosphate groups on substrates for catalysis [134-136]. In addition, the crystal structures of AaThiL and AbThiL revealed multiple metal-ion binding sites located in the TMP and ATP binding sites. While all metal ions were identified as Mg^{2+} in the structure of AaThiL, the more recent study done on AbThiL showed that different metal ions were involved. They observed longer co-ordination distances more compatible with Ca^{2+}, Na^+ or K^+ [90, 98]. This indicates that K^+, Ca^{2+} or Na^+ might be important for ligand binding during co-crystallisation and soaking, a possibility we were unaware of during the crystallisation experiments.

Investigating the effect of additional ions on substrate binding through soaking and co-crystallisation would be next step in this study.

4.3 Analysis of thiamine monophosphate kinase from *Mycobacterium tuberculosis* crystal structure

The structure of MtbThiL was solved to a resolution of 2.19 Å, an improvement on the previously solved structure (2.8 Å). The first 59 N-terminal and last 16 C-terminal residues, could not be structurally characterised due to insufficient electron density. This phenomenon is not unique to this particular crystal structure and does not imply that the residues involved were proteolytically removed from the enzyme. Instead, it is a case of inherent disorder frequently observed for

termini, loops, or even entire domains in proteins. All structural techniques including biological nuclear magnetic resonance (NMR) spectroscopy, X-ray crystallography and cryo-electron microscopy provide an averaged view of a particular protein or macromolecules. Correspondingly, only conformationally conserved regions are described unambiguously. Regions of proteins not constrained by interactions, adopt many unique conformations with the result that no unique information is obtained in the form of electron density or constraints. Proteins generally retain their biological activity when crystallized and to some degree remain flexible. As outlined above, the flexibility results in uncertainty regarding the atomic positions of these regions [137-139]. In enzymes, loops near the active site often adopt an ordered conformation following ligand binding [138, 140]. Thus, some active-site loops are disordered in the structure of apo-MtbThiL, whereas the same loops are ordered and hence visible in AaThiL/substrate complexes. The disordered C-terminal loop on apo-AaThiL is visibly ordered and forms a "lid" that encloses the substrates in the co-crystal structures. Despite different attempts at generating crystals of MtbThiL with bound substrates, this was not achieved resulting in disordered loops and termini.

Amino acid sequences are highly conserved for ThiL proteins from different organisms (Fig. A1 Appendix) confirming that the catalytic mechanism of ThiL proteins is presumably similar. These include residues, Asp 43, Asp 71 and Asp 109 that are shown in both AaThiL and AbThiL to interact with phosphate groups on TMP and ATP through metal ion co-ordination facilitated by Mg^{2+}. The reaction mechanism of ThiL proteins is proposed to occur through direct transfer of the γ-phosphate from ATP to TMP forming TPP [90] [98] and might be similar in MtbThiL.

4.4 Kinetics of thiamine monophosphate kinase from *Mycobacterium tuberculosis*

The results of our kinetic studies suggest that MtbThiL follows atypical, non-Michaelis-Menten kinetics. Our velocity vs substrate concentration graphs show a steady increase in velocity at ATP

concentrations between 0 and 13 mM, followed by a sharp decline after 13 mM (Fig. 3.9). The kinetics of enzymes that do not follow Michaelis Menten are poorly described but often result from allostery or substrate inhibition.

Allosteric enzymes are often composed of identical or distinct multiple subunits [68, 141, 142]. Consequently, each subunit has its own substrate binding site. The ligands bind in sequential order, the binding of one substrate onto its binding site leads to conformational and electrostatic changes to the second binding site, either increasing or decreasing the enzymes affinity for the second ligand [143]. Conversely, if the ligands were to bind onto their binding sites simultaneously, the binding of one ligand would not affect binding of the other ligand.

Our crystal structure clearly shows that MtbThiL is a dimer with identical binding sites located on each monomer (Fig. 3.5). Based on this information, we could initially assume that the kinetics suggest allosteric binding. However, this is not enough to come to this conclusion since whether ATP and TMP bind to both active sites simultaneously or sequentially is unclear. The reaction mechanism of ThiL has been described in AaThiL. ATP binds first followed by TMP forming a ternary ThiL-ATP-TMP complex before product formation and release [90]. Although this shows the sequence of ligand binding to a single binding site it does not reveal whether substrate binding to both sites is simultaneous or sequential. In addition, our graphs do not support this theory since they do not show the sigmoidal curve that is typical of allosteric enzymes and indicative of co-operative substrate binding. Instead, the marked decrease in V_{max} points to possible substrate inhibition.

Substrate inhibition refers to decreased enzyme activity caused by high substrate concentrations. It occurs when the enzyme-substrate complex formed after binding both substrates breaks down at a progressively slow rate. This decreases the concentration of free enzyme available to catalyse the reaction, and consequently the rate of the reaction [144]. Substrate inhibition is common in

enzymes with multiple active sites where binding of the second substrate leads to a decrease in V_{max} and is described by graphs similar to ours [145]. Substrate inhibition was also observed for ThiL from *E. coli*, although the concentrations of ATP at which this occurred were lower [132, 146]. The same study also found that ThiL was inhibited by increasing ADP concentrations (product) suggesting regulation of the thiamine biosynthetic pathway. While substrate inhibition is often considered a non-physiological phenomenon brought about by experiments conducted *in vitro*, some studies have shown it to be a biological regulatory mechanism common to kinases [147, 148].

Our study could have included additional data points after 15 mM ATP. This would have allowed us to show a further or maintained decrease in velocity, definitively showing that MbThiL is inhibited by high concentrations of ATP.

4.5 Structure based inhibitor identification

Oxythiamine is a thiamine pyrophosphate (TPP) antagonist that inhibits the activity of TPP dependent enzymes [149]. It was identified, with nine other compounds as a potential ThiL inhibitor and its inhibitory capacity against MtbThiL was further investigated in this study. Oxythiamine was found to have higher K_M and K_D values compared to TMP in kinetic and ITC experiments. These values agree in implying that the affinity of MtbThiL for oxythiamine is lower than for TMP.

The structure of oxythiamine differs from that of TMP by the additional carbonyl group and replacement of the C1 carbon of the TMP pyrimidine ring by a secondary amine in oxythiamine. In the proposed binding mode of oxythiamine to ThiL, neither of the lone pairs on the carbonyl oxygen interact with active site residues or ions (Fig. 3.10 (2)). However, the Mg^{2+} in the active site interact, through salt bridge interactions, with the phosphate and hydroxyl groups on TMP

and oxythiamine respectively. This seems to confirm that Mg^{2+} has preference for phosphate groups as proposed in studies done on AaThiL and in studies on other nucleoside di- and triphosphates [90, 150, 151].

Protein-ligand interactions are often described using the thermodynamic properties associated with binding. The binding energy of a ligand is the sum of the enthalpy and entropy changes that occur during binding. Favourable, spontaneous reactions resulting in a negative ΔG correspond to a negative enthalpy change, a positive entropy change or both. [152, 153]. The more non-covalent interactions a ligand makes with a protein the more negative the enthalpy contribution which corresponds to a spontaneous reaction. Therefore, the more validated interactions and ligand conformation or pose affect ΔH and thus ΔG. Oxythiamine had fewer validated interactions in comparison to TMP, it is therefore not surprising that oxythiamine had higher binding energy and thus lower binding affinity as validated by our kinetics and ITC results.

The entropy contribution to free binding energy of protein-ligand binding is a sum of the water molecules displaced from the binding site, conformational changes of the ligand and protein and loss of translational and rotational freedom upon complex formation [93]. Compound 6 resulted in the lowest binding energy possibly resulting from its more validated interactions resulting in a more negative enthalpy contribution. In addition, the conformation and molecular weights of TMP, oxythiamine and compound 6 are very similar. This suggests that they occupy a similar volume in the active site, consequently displacing a similar number of water molecules. The entropy contribution of the three compounds is thus similar resulting in compound 6 having an overall lower binding free energy since it has the highest enthalpy contribution. Compound 6 is therefore, a good candidate for ThiL inhibition and may be tested in future studies for MtbThiL inhibition.

In humans, free thiamine (unphosphorylated) is absorbed and pyrophosphorylated into the active co-factor form, TPP by thiamine pyrophosphokinase (TPK) [154]. In addition, some bacteria have thiamine salvage pathways where they convert free thiamine into TPP using the same enzyme (Fig. 1.4) [155]. Although the crystal structures of ThiL and TPK are vastly different, they share amino acid similarity of catalytically important residues (Fig. A1 Appendix). Their reactions also proceed in similar fashion, through the direct transfer of a phosphate group(s) from ATP. ThiL and TPK cleave after the β- and α-phosphate groups respectively. These similarities may lead to unspecific binding of a ThiL inhibitor on TPK which could adversely affect normal metabolism in humans. This is an important point to consider in the design of anti-TB drugs targeting proteins of the thiamine biosynthetic pathway [94].

5. Conclusion and future work

In this study thiamine monophosphate kinase from *Mycobacterium tuberculosis* was successfully produced, purified and crystallised. Co-crystallisation was unsuccessful possibly due to the absence of a co-factor or activator that might play a pivotal and yet unexplored role in substrate binding and catalysis in MtbThiL. The structure of the apo-MtbThiL was solved to a high resolution and added structural features that were not present in earlier studies. MtbThiL is structurally similar to other ThiL proteins and has conserved residues that are important for catalysis and substrate binding. This suggests that ThiL proteins catalyse the reaction in a similar manner. To unequivocally determine the mechanism of the reaction catalysed by MtbThiL, we need to solve the MtbThiL-TMP/TPP/ADP/AMP-PNP co-crystal structures. Residues that have been identified to play a central role in catalysis could also be substituted using site-directed mutagenesis to confirm their importance and further describe the reaction mechanism of MtbThiL catalysis.

We screened for potential MtbThiL inhibitors and found ten hits that could be the starting point towards the discovery of an anti-TB drug with MtbThiL as the target. The compound we tested showed no inhibitory capability as well as poor binding. Testing the remaining nine compounds for their inhibitory abilities would be the next step in this regard.

6. Appendix

```
Mtuberculosis   MGSSHHHHHHSSGLVPRGSHMTTKDH---SLATESPTLQQLGEFAVIDRLVRGRR-QPA-      55
Aaeolicus       MGSHHHHHHDITSLYKKAGSAAAVLEENLYFQGSFTMRLKELGEFGLIDLIKKTLE----      56
Abaumanii       ------------------------------MAHHHHHHMAEFSIIDQYFNRQ--SHP      25
Ecoli           ----------------------------------------MACGEFSLIARYFDRVRSSRL      21
Styphimurium    ----------------------------------------MACGEFSLIARYFDRVRSSRL      21
Hsapiens        ------------------------------------------------------------       0


Mtuberculosis   TVLLGPGDDAALVSAGD-GRTVVSTDMLVQDSHFRLDWSTPQDVGRKAIAQNAADIEAMG     114
Aaeolicus       --SKVIGDDTAPVEYC-SKKLLLTTDVLNEGVHFLRSY-IPEAVGWKAISVNVSDVIANG     112
Abaumanii       DVALGIGDDSALITPPPNQQLVICADTLVAGRHFPLET-SPHAIGWKSVAVNLSDIAAMG      84
Ecoli           DVELGIGDDCALLNIPEKQTLAISTDTLVAGNHFLHDI-DPADLAYKALAVNLSDLAAMG      80
Styphimurium    DVETGIGDDCALLNIPEKQTLAISTDTLVAGNHFLPDI-DPADLAYKALAVNLSDLAAMG      80
Hsapiens        ------------------------------------MHHHHHHSSGRENLYFQG      18
                                                              :  : .  ::     *


Mtuberculosis   ARATAFVVGFGAPAE-----------------------------TPAAQASALVDGMWEE     145
Aaeolicus       GLPKWALISLNLPED-----------------------------LEVSYVERFYIGVKRA     143
Abaumanii       AKPHSILLAISLPQ------------------------------VDHEWLEGFSQGIYDC     114
Ecoli           ADPAWLTLALTLPD------------------------------VDEAWLESFSDSLFDL     110
Styphimurium    ADPAWLTLALTLPE------------------------------VDEPWLEAFSDSLFAL     110
Hsapiens        GNLKYCLVILNQPLDNYFRHLWNKALLRACADGGANRLYDITEGERESFLPEFINGDFDS      78
                .          : :  *                                      :  .


Mtuberculosis   A--------GRIGAGIVGGDLVSCRQWVVSVTAIGDLDGRAPVLRSGAKAGSVLAVVGEL     197
Aaeolicus       C--------EFYKCEVVGGNISKSEKIGISVFLVG--ETERFVGRDGARLGDSVFVSGTL     193
Abaumanii       C--------NQFGVALIGGDTTQGPHLTITVTAMGWIETGKAVLRSGAKVGDYVCVSGQI     166
Ecoli           L--------NYYDMQLIGGDTTRGPL-SMTLGIHGFVPMGRALTRSGAKPGDWIYVTGTP     161
Styphimurium    L--------NYYDMQLIGGDTTRGPL-SMTLGIHGYIPAGRALKRSGAKPGDWIYVTGTP     161
Hsapiens        IRPEVREYYATKGCELISTPDQDHTDFTKCLKM-----LQKKIEEKDLKVDVIV----TL     129
                          ::.              :              :  ...: .  :


Mtuberculosis   GRSAAGYALWCNGIEDFAE-----LRRRHLVPQPPYGHGAAAAAVG---AQAMIDVSDGL     249
Aaeolicus       GDSRAGLELLLMEKEEYEP-FELALIQRHLRPTARIDYVKHIQK-Y---ANASMDISDGL     248
Abaumanii       GDAAYGLQHLGHSL-----------QQRLDYPTPRCKLGEELKG-L----ASSMIDVSDGL     211
Ecoli           GDSAAGLAILQNRLQVADAKDADYLIKRHLRPSPRILQGGQALRD-L---ANSAIDLSDGL     217
Styphimurium    GDSAAGLAVLQNRLQVSEETDAHYLIQRHLRPTPRILHGGQALRD-I---ASAAIDLSDGL     217
Hsapiens        GGLAGRFDQI-------MASVNTLFQATHITPFPIIIIQEESLIYLLQPGKHRLHVDTGM     182
                *                          *                 .. :.:. *:


Mtuberculosis   LADLRHIAEASGVRIDLSAAALAADRDALTAAATALGTDPWPWVLSGGEDHALVACFVGP     309
Aaeolicus       VADANHLAQRSGVKIEILSEKLPLSN-ELKMYCEKYGKNPIEYALFGGEDYQL--LFTHP     305
Abaumanii       AQDLGHILKASKVGARLILEKLPVDP--VLQQIE--EQQRWQYALAGGDDYEL--CFTIT     265
Ecoli           ISDLGHIVKASDCGARIDLALLPFSD-ALSRHVE--PEQALRWALSGGEDYEL--CFTVP     272
Styphimurium    ISDLGHIVKASGCGARVDVDALPKSD-AMMRHVD--DGQALRWALSGGEDYEL--CFTVP     272
Hsapiens        EGDWCGL-------------IPVGQ-----PCMQVTTTGLKWNLTNDV-----------     212
                 *    :                          : *  ..


Mtuberculosis   V-------------PAGWRTIGRVLDGP--ARVLVDGE----EW-TGYAGWQSFGEPDNQ     349
Aaeolicus       KERWNPFL--------DMTEIGRVEEG---EGVFVDGK----KVEPKGWKHF------     342
Abaumanii       PQNYEKL--LQKQLDVKITMIGQIVEQ-TKLTFEHLGS----DYPLQIHGYQHFA-----     313
Ecoli           ELNRGALDVALGHLGVPFTCIGQMTADIEGLCFICDGE----PVTFDWKGYDHFATP---     325
Styphimurium    ELNRGALDVAIGQLGVPFTCIGQMSADIEGLSFVRDGM----PVTFDWKGYDHFATP---     325
Hsapiens        -LAFGTLV-------------STSNTYDGSGVVTVETDHPLLWTMAIKS----------     247
                                                .              .


Mtuberculosis   GSLG       353
Aaeolicus       ----       342
Abaumanii       ----       313
Ecoli           ----       325
Styphimurium    ----       325
Hsapiens        ----       247
```

Figure A1: Amino acid sequence alignment of ThiL proteins and human thiamine pyrophosphate kinase. Conserved amino acids in the ThiL structures are shown in red and highlighted in yellow. Amino acids that occur in ThiL proteins and thiamine pyrophosphate kinase are shown in red. An Asterix indicates amino acids that are conserved in all ThiL proteins and TPK. Amino acids shown in purple are conserved in at least more than three of the species. The amino acid sequence alignment was done using Pairwise sequence alignment from EMBL.

References

1.   Nguyen, L., *Antibiotic resistance mechanisms in M. tuberculosis: an update.* Archives of toxicology, 2016. **90**(7): p. 1585-1604.

2.   WHO, *Global Tuberculosis Report*. 2019.

3.   Guirado, E., et al., *Characterization of host and microbial determinants in individuals with latent tuberculosis infection using a human granuloma model.* MBio, 2015. **6**(1): p. e02537-14.

4.   Sulis, G., et al., *Tuberculosis: epidemiology and control.* Mediterr J Hematol Infect Dis, 2014. **6**(1): p. e2014070.

5.   Small PM, P.M., *Tuberculosis Diagnosis — Time for a Game Change.* The New England Journal of Medicine, 2010. **9**(11): p. 1005-15.

6.   Menzies, N.A., et al., *Population health impact and cost-effectiveness of tuberculosis diagnosis with Xpert MTB/RIF: a dynamic simulation and economic evaluation.* PLoS medicine, 2012. **9**(11): p. e1001347.

7.   Abubakar, I., et al., *Systematic review and meta-analysis of the current evidence on the duration of protection by bacillus Calmette-Guérin vaccination against tuberculosis.* Health technology assessment (Winchester, England), 2013. **17**(37): p. 1.

8.   Gengenbacher, M. and S.H. Kaufmann, *Mycobacterium tuberculosis: success through dormancy.* FEMS microbiology reviews, 2012. **36**(3): p. 514-532.

9.   Santosuosso, M., et al., *Intranasal boosting with an adenovirus-vectored vaccine markedly enhances protection by parenteral Mycobacterium bovis BCG immunization against pulmonary tuberculosis.* Infection and immunity, 2006. **74**(8): p. 4634-4643.

10.   Vilchèze, C., et al., *Transfer of a point mutation in Mycobacterium tuberculosis inhA resolves the target of isoniazid.* Nature medicine, 2006. **12**(9): p. 1027.

11.   Goude, R., et al., *The Arabinosyltransferase EmbC Is Inhibited by Ethambutol in Mycobacterium tuberculosis.* Antimicrobial Agents and Chemotherapy, 2009. **53**(10): p. 4138-4146.

12.   Shi, W., et al., *Pyrazinamide Inhibits Trans-Translation in Mycobacterium tuberculosis.* Science, 2011. **333**(6049): p. 1630.

13.   Shi, W., et al., *Aspartate decarboxylase (PanD) as a new target of pyrazinamide in Mycobacterium tuberculosis.* Emerging Microbes and Infections, 2014. **3**(1): p. 1-8.

14.   Organization, G.W.H. *WHO consolidated guidelines on drug-resistant tuberculosis treatment*. 2019   [cited   2019   25   April   2019];   Available   from:

https://apps.who.int/iris/bitstream/handle/10665/311389/9789241550529-eng.pdf?ua=1.

15. Palomino, J. and A. Martin, *Drug resistance mechanisms in Mycobacterium tuberculosis.* Antibiotics, 2014. **3**(3): p. 317-340.

16. Migliori, G.B., et al., *125 years after Robert Koch's discovery of the tubercle bacillus: the new XDR-TB threat. Is "science" enough to tackle the epidemic?* European Respiratory Journal, 2007. **29**(3): p. 423-427.

17. Asante-Poku, A., et al., *Molecular epidemiology of Mycobacterium africanum in Ghana.* BMC infectious diseases, 2016. **16**(1): p. 385.

18. Delogu, G., M. Sali, and G. Fadda, *The Biology of Mycobacterium tuberculosis infection.* Mediterranean journal of hematology and infectious diseases, 2013. **5**(1).

19. Dou, H.-Y., et al., *Lineage-specific SNPs for genotyping of Mycobacterium tuberculosis clinical isolates.* Scientific reports, 2017. **7**(1): p. 1425.

20. Barkan, D., et al., *Mycobacterium tuberculosis lacking all mycolic acid cyclopropanation is viable but highly attenuated and hyperinflammatory in mice.* Infection and immunity, 2012. **80**(6): p. 1958-1968.

21. Glickman, M.S., J.S. Cox, and W.R. Jacobs Jr, *A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of Mycobacterium tuberculosis.* Molecular cell, 2000. **5**(4): p. 717-727.

22. Vijay, S., et al., *Influence of stress and antibiotic resistance on cell-length distribution in Mycobacterium tuberculosis clinical isolates.* Frontiers in microbiology, 2017. **8**: p. 2296.

23. Bozzano, F., F. Marras, and A. De Maria, *Immunology of tuberculosis.* Mediterranean journal of hematology and infectious diseases, 2014. **6**(1): p. e2014027-e2014027.

24. Sala, A., P. Bordes, and P. Genevaux, *Multiple Toxin-Antitoxin Systems in Mycobacterium tuberculosis.* Toxins, 2014. **6**(3): p. 1002.

25. Nair, V.R., et al., *Microfold Cells Actively Translocate Mycobacterium tuberculosis to Initiate Infection.* Cell reports, 2016. **16**(5): p. 1253-1258.

26. Russell, D.G., C.E. Barry, 3rd, and J.L. Flynn, *Tuberculosis: what we don't know can, and does, hurt us.* Science (New York, N.Y.), 2010. **328**(5980): p. 852-856.

27. Teitelbaum, R., et al., *The M Cell as a Portal of Entry to the Lung for the Bacterial Pathogen Mycobacterium tuberculosis.* Immunity, 1999. **10**(6): p. 641-650.

28. Sunil Regmi, B.A., *Tubercular Prostatitis: A rare case of Genitourinary Tuberculosis.* Journal of Nobel Medical College, 2018. **7** (12): p. 65-67.

29.   Ahmad, S., *Pathogenesis, Immunology, and Diagnosis of Latent Mycobacterium tuberculosis Infection.* Clinical and Developmental Immunology, 2011. **2011**: p. 17.

30.   Mohareer, K., S. Asalla, and S. Banerjee, *Cell death at the cross roads of host-pathogen interaction in Mycobacterium tuberculosis infection.* Tuberculosis, 2018. **113**: p. 99-121.

31.   Shiloh, M.U., *Mechanisms of mycobacterial transmission: how does Mycobacterium tuberculosis enter and escape from the human host.* Future microbiology, 2016. **11**(12): p. 1503-1506.

32.   Kruh, N.A., et al., *Portrait of a Pathogen: the Mycobacterium tuberculosis proteome in vivo.* PLoS one, 2010. **5**(11): p. e13938.

33.   Smith, I., *Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence.* Clinical microbiology reviews, 2003. **16**(3): p. 463-496.

34.   Cadena, A.M., J.L. Flynn, and S.M. Fortune, *The importance of first impressions: early events in Mycobacterium tuberculosis infection influence outcome.* MBio, 2016. **7**(2): p. e00342-16.

35.   Polena, H., et al., *Mycobacterium tuberculosis exploits the formation of new blood vessels for its dissemination.* Scientific reports, 2016. **6**: p. 33162.

36.   Volkman, H.E., et al., *Tuberculous granuloma formation is enhanced by a mycobacterium virulence determinant.* PLoS biology, 2004. **2**(11): p. e367.

37.   Ramakrishnan, L., *Revisiting the role of the granuloma in tuberculosis.* Nature Reviews Immunology, 2012. **12**(5): p. 352.

38.   Davis J. Muse, R.L., *The Role of the Granuloma in Expansion and Dissemination of Early Tuberculous Infection.* Cell 2009. **136**(1): p. 37-49.

39.   Cliff, J.M., et al., *The human immune response to tuberculosis and its treatment: a view from the blood.* Immunological reviews, 2015. **264**(1): p. 88-102.

40.   Ernst, J.D., *The immunological life cycle of tuberculosis.* Nature Reviews Immunology, 2012. **12**(8): p. 581.

41.   WHO. *Tuberculosis.* 2018 [cited 2018 25 September 2018].

42.   Organization, W.H., *Global antimicrobial resistance surveillance system (GLASS) report: early implementation 2017-2018.* 2018.

43.   Munita, J.M. and C.A. Arias, *Mechanisms of antibiotic resistance.* Microbiology spectrum, 2016. **4**(2).

44.   Oliveira, P.H., et al., *The chromosomal organization of horizontal gene transfer in bacteria.* Nature communications, 2017. **8**(1): p. 841.

45. Dookie, N., et al., *Evolution of drug resistance in Mycobacterium tuberculosis: a review on the molecular determinants of resistance and implications for personalized care.* The Journal of antimicrobial chemotherapy, 2018. **73**(5): p. 1138-1151.

46. Richardson, L.A., *Understanding and overcoming antibiotic resistance.* PLoS biology, 2017. **15**(8): p. e2003775.

47. Alekshun, M.N. and S.B. Levy, *Molecular Mechanisms of Antibacterial Multidrug Resistance.* Cell, 2007. **128**(6): p. 1037-1050.

48. Magiorakos, A.-P., et al., *Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance.* Clinical Microbiology and Infection, 2012. **18**(3): p. 268-281.

49. Udwadia, Z.F., et al., *Totally Drug-Resistant Tuberculosis in India.* Clinical Infectious Diseases, 2012. **54**(4): p. 579-581.

50. Klopper, M., et al., *Emergence and spread of extensively and totally drug-resistant tuberculosis, South Africa.* Emerging infectious diseases, 2013. **19**(3): p. 449.

51. Velayati, A.A., P. Farnia, and M.R. Masjedi, *The totally drug resistant tuberculosis (TDR-TB).* International journal of clinical and experimental medicine, 2013. **6**(4): p. 307-309.

52. Malik, B. and S. Bhattacharyya, *Antibiotic drug-resistance as a complex system driven by socio-economic growth and antibiotic misuse.* Scientific reports, 2019. **9**(1): p. 9788.

53. Kohanski, M.A., D.J. Dwyer, and J.J. Collins, *How antibiotics kill bacteria: from targets to networks.* Nature Reviews Microbiology, 2010. **8**(6): p. 423.

54. Marrakchi, H., M.-A. Lanéelle, and M. Daffé, *Mycolic Acids: Structures, Biosynthesis, and Beyond.* Chemistry & Biology, 2014. **21**(1): p. 67-85.

55. Baker, J.J. and R.B. Abramovitch, *Genetic and metabolic regulation of Mycobacterium tuberculosis acid growth arrest.* Scientific reports, 2018. **8**(1): p. 4168.

56. Garton, N.J., et al., *Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum.* PLoS medicine, 2008. **5**(4): p. e75-e75.

57. Lewis, K., *Multidrug tolerance of biofilms and persister cells*, in *Bacterial Biofilms*. 2008, Springer. p. 107-131.

58. Ford, C.B., et al., *Mycobacterium tuberculosis mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis.* Nature genetics, 2013. **45**(7): p. 784.

59. Wilson, M., et al., *Exploring drug-induced alterations in gene expression in Mycobacterium tuberculosis by microarray hybridization.* Proceedings of the National Academy of Sciences, 1999. **96**(22): p. 12833-12838.

60. Zhang, Y., *Genetics of drug resistance in Mycobacterium tuberculosis*. Second edition ed. Molecular genetics of Mycobacteria. 2000.

61. Li, G., et al., *Efflux pump gene expression in multidrug-resistant Mycobacterium tuberculosis clinical isolates.* PLoS One, 2015. **10**(2): p. e0119013.

62. Palmer, T. and P.L. Bonner, *15 - Investigation of Enzymes in Biological Preparations*, in *Enzymes (Second Edition)*, T. Palmer and P.L. Bonner, Editors. 2011, Woodhead Publishing. p. 274-292.

63. Voet, D., J.G. Voet, and C.W. Pratt, *Fundamentals of biochemistry: life at the molecular level*. 2013.

64. Campbell, M.K. and S.O. Farrell, *Biochemistry . Belmont, CA: Brooks/Cole Cengage Learning*. 2009, ISBN 978-0-8400-6858-3.

65. Cooper, G. and R. Hausman, *The cell: a molecular approach. Sinauer Associates.* Sunderland, MA, 2000.

66. Khan, M. and F. Khan, *Principles of Enzyme Technology*. 2015: PHI Learning Pvt. Ltd.

67. Broderick, J.B., *Coenzymes and cofactors.* Encyclopedia of life sciences (eLS). John Wiley & Sons Ltd, Chichester. http://www. els. net. doi, 2001. **10**.

68. Berg, J.M., et al., *Biochemistry. 5th edit*. 2002, WH Freeman and Company New York.

69. Kilpin, K.J. and P.J. Dyson, *Enzyme inhibition by metal complexes: concepts, strategies and applications.* Chemical Science, 2013. **4**(4): p. 1410-1419.

70. Lopina, O.D., *Enzyme Inhibitors and Activators*. 2016: IntechOpen.

71. Guo, X., Y. Chen, and C.T. Seto, *Rational design of novel irreversible inhibitors for human arginase.* Bioorganic & medicinal chemistry, 2018.

72. Bearne, S.L., *Illustrating enzyme inhibition using Gibbs energy profiles.* Journal of chemical education, 2012. **89**(6): p. 732-737.

73. Yang, X., et al., *Classification of difference between inhibition constants of an inhibitor to facilitate identifying the inhibition type.* Journal of Enzyme Inhibition and Medicinal Chemistry, 2013. **28**(1): p. 205-213.

74. Martin, J., K. Anamika, and N. Srinivasan, *Classification of protein kinases on the basis of both kinase and non-kinase regions.* PloS one, 2010. **5**(9): p. e12460-e12460.

75.   Ubersax, J.A. and J.E. Ferrell Jr, *Mechanisms of specificity in protein phosphorylation.* Nature Reviews Molecular Cell Biology, 2007. **8**: p. 530.

76.   Kenyon, C.P., et al., *Conserved phosphoryl transfer mechanisms within kinase families and the role of the C8 proton of ATP in the activation of phosphoryl transfer.* BMC research notes, 2012. **5**(1): p. 131.

77.   Manning, G., et al., *The protein kinase complement of the human genome.* Science, 2002. **298**(5600): p. 1912-1934.

78.   Verheijen, J.C., D.J. Richard, and A. Zask, *Chapter 6 Non-Protein Kinases as Therapeutic Targets*, in *Kinase Drug Discovery*. 2012, The Royal Society of Chemistry. p. 161-217.

79.   Li, C., et al., *X-ray crystal structure of aminoimidazole ribonucleotide synthetase (PurM), from the Escherichia coli purine biosynthetic pathway at 2.5 Å resolution.* Structure, 1999. **7**(9): p. 1155-1166.

80.   Bhullar, K.S., et al., *Kinase-targeted cancer therapies: progress, challenges and future directions.* Molecular cancer, 2018. **17**(1): p. 48.

81.   Meng, X.-Y., et al., *Molecular docking: a powerful approach for structure-based drug discovery.* Current computer-aided drug design, 2011. **7**(2): p. 146-157.

82.   Du, X., et al., *Insights into protein–ligand interactions: mechanisms, models, and methods.* International journal of molecular sciences, 2016. **17**(2): p. 144.

83.   Malhotra, S., et al., *Structure-guided, target-based drug discovery–exploiting genome information from HIV to mycobacterial infections.* Postepy biochemii, 2016. **62**(3): p. 262-272.

84.   Liu, J., et al., *Enhance the performance of current scoring functions with the aid of 3D protein-ligand interaction fingerprints.* BMC bioinformatics, 2017. **18**(1): p. 343.

85.   Lounnas, V., et al., *Current progress in structure-based rational drug design marks a new mindset in drug discovery.* Computational and structural biotechnology journal, 2013. **5**(6): p. e201302011.

86.   Rietman, E.A., et al., *Thermodynamic measures of cancer: Gibbs free energy and entropy of protein–protein interactions.* Journal of biological physics, 2016. **42**(3): p. 339-350.

87.   Wójcikowski, M., P.J. Ballester, and P. Siedlecki, *Performance of machine-learning scoring functions in structure-based virtual screening.* Scientific Reports, 2017. **7**: p. 46710.

88.   Drebes, J., et al., *Structure of ThiM from Vitamin B1 biosynthetic pathway of Staphylococcus aureus–insights into a novel pro-drug approach addressing MRSA infections.* Scientific reports, 2016. **6**: p. 22871.

89.   Begley, T.P., et al., *Thiamin biosynthesis in prokaryotes.* Archives of microbiology, 1999. **171**(5): p. 293-300.

90.   McCulloch, K.M., et al., *Structural studies of thiamin monophosphate kinase in complex with substrates and products.* Biochemistry, 2008. **47**(12): p. 3810-3821.

91.   Campbell, M.K. and S.O. Farrell, *Biochemistry.* 2012, Belmont, CA: Brooks/Cole, Cengage Learning.

92.   Drebes, J., et al., *Purification, crystallization and preliminary X-ray diffraction analysis of ThiM from Staphylococcus aureus.* Acta Crystallographica Section F: Structural Biology and Crystallization Communications, 2011. **67**(4): p. 479-481.

93.   Du, X., et al., *Insights into Protein–Ligand Interactions: Mechanisms, Models, and Methods.* International Journal of Molecular Sciences, 2016. **17**(2).

94.   Settembre, E., T.P. Begley, and S.E. Ealick, *Structural biology of enzymes of the thiamin biosynthesis pathway.* Current Opinion in Structural Biology, 2003. **13**(6): p. 739-747.

95.   Jenkins, A.H., et al., *A new thiamin salvage pathway.* Nature Chemical Biology, 2007. **3**: p. 492.

96.   Tyagi, G., et al., *Role of Vitamins B, C, and D in the fight against tuberculosis.* Int J Mycobacteriol, 2017. **6**(4): p. 328-332.

97.   Costliow, Z.A. and P.H. Degnan, *Thiamine acquisition strategies impact metabolism and competition in the gut microbe Bacteroides thetaiotaomicron.* MSystems, 2017. **2**(5): p. e00116-17.

98.   Sullivan, A.H., et al., *Crystal structures of thiamine monophosphate kinase from Acinetobacter baumannii in complex with substrates and products.* Scientific reports, 2019. **9**(1): p. 4392-4392.

99.   Jurgenson, C.T., T.P. Begley, and S.E. Ealick, *The structural and biochemical foundations of thiamin biosynthesis.* Annual review of biochemistry, 2009. **78**: p. 569-603.

100.   Ntui, C.M., *Structural and functional analysis of thiamine and homoserine kinases from Mycobacterium tuberculosis*, in *Biochemistry*. 2016, University of Pretoria. p. 45.

101.   Nallamsetty, S. and D.S. Waugh, *A generic protocol for the expression and purification of recombinant proteins in Escherichia coli using a combinatorial His6-maltose binding protein fusion tag.* Nature Protocols, 2007. **2**: p. 383.

102.   Sivashanmugam, A., et al., *Practical protocols for production of very high yields of recombinant proteins using Escherichia coli.* Protein science : a publication of the Protein Society, 2009. **18**(5): p. 936-948.

103. Saranya, N., et al., *Cells Disruption by Ultrasonication.* BioNanoScience, 2014. **4**(4): p. 335-337.

104. Porath, J., et al., *Metal chelate affinity chromatography, a new approach to protein fractionation.* Nature, 1975. **258**(5536): p. 598-599.

105. Smith, M.C., et al., *Chelating peptide-immobilized metal ion affinity chromatography. A new concept in affinity chromatography for recombinant proteins.* Journal of Biological Chemistry, 1988. **263**(15): p. 7211-7215.

106. Bornhorst, J.A. and J.J. Falke, *Purification of proteins using polyhistidine affinity tags.* Methods in enzymology, 2000. **326**: p. 245-254.

107. Mori, S. and H.G. Barth, *Size exclusion chromatography.* 2013: Springer Science & Business Media.

108. Manns, J.M., *SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) of Proteins.* Current Protocols in Microbiology, 2011. **22**(1): p. A.3M.1-A.3M.13.

109. Hassan, P.A., S. Rana, and G. Verma, *Making Sense of Brownian Motion: Colloid Characterization by Dynamic Light Scattering.* Langmuir, 2015. **31**(1): p. 3-12.

110. Lorber, B., et al., *Protein analysis by dynamic light scattering: Methods and techniques for students.* Biochemistry and Molecular Biology Education, 2012. **40**(6): p. 372-382.

111. McPherson, A. and J.A. Gavira, *Introduction to protein crystallization.* Acta crystallographica. Section F, Structural biology communications, 2013. **70**(Pt 1): p. 2-20.

112. Müller, I., *Guidelines for the successful generation of protein–ligand complex crystals.* Acta Crystallographica Section D: Structural Biology, 2017. **73**(2): p. 79-92.

113. Drenth, J., *Principles of protein X-ray crystallography.* 2007: Springer Science & Business Media.

114. Engh, R.A., *X-Ray Crystallography, Basic Principles*, in *Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine.* 2006, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 2026-2029.

115. Cowtan, K., *Phase Problem in X-ray Crystallography, and Its Solution.* e LS, 2001.

116. McCoy, A.J., et al., *Phaser crystallographic software.* Journal of applied crystallography, 2007. **40**(4): p. 658-674.

117. Emsley, P., et al., *Features and development of Coot.* Acta Crystallographica Section D: Biological Crystallography, 2010. **66**(4): p. 486-501.

118. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallographica Section D: Biological Crystallography, 2010. **66**(2): p. 213-221.

119. Kim, S., et al., *PubChem 2019 update: improved access to chemical data.* Nucleic acids research, 2018. **47**(D1): p. D1102-D1109.

120. Irwin, J.J. and B.K. Shoichet, *ZINC--a free database of commercially available compounds for virtual screening.* Journal of chemical information and modeling, 2005. **45**(1): p. 177-182.

121. Gaulton, A., et al., *The ChEMBL database in 2017.* Nucleic acids research, 2017. **45**(D1): p. D945-D954.

122. Sastry, G.M., et al., *Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments.* Journal of computer-aided molecular design, 2013. **27**(3): p. 221-234.

123. Friesner, R.A., et al., *Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein− ligand complexes.* Journal of medicinal chemistry, 2006. **49**(21): p. 6177-6196.

124. Pymol Molecular Graphics System Version 2.0 Schrödinger, L. [cited 2017 June 2017].

125. Leavitt, S. and E. Freire, *Direct measurement of protein binding energetics by isothermal titration calorimetry.* Current Opinion in Structural Biology, 2001. **11**(5): p. 560-566.

126. Swinehart, D., *The beer-lambert law.* Journal of chemical education, 1962. **39**(7): p. 333.

127. *GraphPad Software Prism version 3.00*. San Diego California USA.

128. Team, R.C., *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.* 2013.

129. Dlamini, L.S., *Structural and functional analysis of thiamine monophosphate kinase from Mycobacterium tuberculosis*, in *Department of Biochemistry* 2016, University of Pretoria: Pretoria. p. 40-41.f

130. Ntui, C.M., *Structural and functional analysis of thiaminephosphate and homoserine kinases from Mycobacterium tuberculosis*, in *Biochemistry*. 2016, University of Pretoria.

131. Roberts, S.M. and G.J. Davies, *The crystallization and structural analysis of cellulases (and other glycoside hydrolases): strategies and tactics*, in *Methods in enzymology*. 2012, Elsevier. p. 141-168.

132. Evans, H.J., *Effect of Potassium & Other Univalent Cations on Activity of Pyruvate Kinase in Pisum sativum.* Plant physiology, 1963. **38**(4): p. 397-402.

133. Larsen, T.M., et al., *Structure of Rabbit Muscle Pyruvate Kinase Complexed with Mn2+, K+, and Pyruvate.* Biochemistry, 1994. **33**(20): p. 6301-6309.

134. Vašák, M. and J. Schnabl, *Sodium and potassium ions in proteins and enzyme catalysis*, in *The Alkali Metal Ions: Their Role for Life*. 2016, Springer. p. 259-290.

135. Shibata, N., et al., *A new mode of B12 binding and the direct participation of a potassium ion in enzyme catalysis: X-ray structure of diol dehydratase.* Structure, 1999. **7**(8): p. 997-1008.

136. Suelter, C.H., *Enzymes Activated by Monovalent Cations.* Science, 1970. **168**(3933): p. 789-795.

137. Yuan, Z., J. Zhao, and Z.-X. Wang, *Flexibility analysis of enzyme active sites by crystallographic temperature factors.* Protein Engineering, Design and Selection, 2003. **16**(2): p. 109-114.

138. Rupp, B., *Biomolecular crystallography: principles, practice, and application to structural biology*. 2009: Garland Science.

139. Kiefer, J.R., et al., *Visualizing DNA replication in a catalytically active Bacillus DNA polymerase crystal.* Nature, 1998. **391**(6664): p. 304.

140. Fetrow, J.S., *Omega loops: nonregular secondary structures significant in protein function and stability.* The FASEB Journal, 1995. **9**(9): p. 708-717.

141. Muñoz, M.E. and E. Ponce, *Pyruvate kinase: current status of regulatory and functional properties.* Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology, 2003. **135**(2): p. 197-218.

142. DeLuca, D.C. and J. Lyndal York, *Enzymes*, in *Encyclopedia of Genetics*, S. Brenner and J.H. Miller, Editors. 2001, Academic Press: New York. p. 625-626.

143. Laskowski, R.A., F. Gerick, and J.M. Thornton, *The structural basis of allosteric regulation in proteins.* FEBS letters, 2009. **583**(11): p. 1692-1698.

144. Yoshino, M. and K. Murakami, *Analysis of the substrate inhibition of complete and partial types.* SpringerPlus, 2015. **4**(1): p. 292.

145. Atkins, W.M., *Non-Michaelis-Menten kinetics in cytochrome P450-catalyzed reactions.* Annu. Rev. Pharmacol. Toxicol., 2005. **45**: p. 291-310.

146. Nishino, H., *Biogenesis of Cocarboxylase in Escherichia coli.* The Journal of Biochemistry, 1972. **72**(5): p. 1093-1100.

147. Reed, M.C., A. Lieb, and H.F. Nijhout, *The biological significance of substrate inhibition: A mechanism with diverse functions.* BioEssays, 2010. **32**(5): p. 422-429.

148. Uyeda, K., E. Furuya, and L.J. Luby, *The effect of natural and synthetic D-fructose 2, 6-bisphosphate on the regulatory kinetic properties of liver and muscle phosphofructokinases.* Journal of Biological Chemistry, 1981. **256**(16): p. 8394-8399.

149. Tylicki, A., et al., *Modification of thiamine pyrophosphate dependent enzyme activity by oxythiamine in Saccharomyces cerevisiae cells.* Canadian journal of microbiology, 2005. **51**(10): p. 833-839.

150. Cowan, J., *Structural and catalytic chemistry of magnesium-dependent enzymes.* Biometals, 2002. **15**(3): p. 225-235.

151. Berg JM, T.J., Stryer L. , *Biochemistry*. 5th ed. 2002: New York: W H Freeman.

152. Dunitz, J.D., *Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions.* Chemistry & Biology, 1995. **2**(11): p. 709-712.

153. Olsson, T.S.G., et al., *Extent of enthalpy-entropy compensation in protein-ligand interactions.* Protein science : a publication of the Protein Society, 2011. **20**(9): p. 1607-1618.

154. Liu, J.-Y., D.E. Timm, and T.D. Hurley, *Pyrithiamine as a substrate for thiamine pyrophosphokinase.* Journal of Biological Chemistry, 2006. **281**(10): p. 6601-6607.

155. Melnick, J., et al., *Identification of the two missing bacterial genes involved in thiamine salvage: thiamine pyrophosphokinase and thiamine kinase.* Journal of bacteriology, 2004. **186**(11): p. 3660-3662.