Novel inhibitors for biotin biosynthesis pathway in *Mycobacterium tuberculosis*

by

Wanisa Salaemae, M.Sc.

A thesis submitted to the University of Adelaide, South Australia in fulfillment of the requirements for the degree of Doctor of Philosophy

School of Molecular and Biomedical Sciences
Discipline of Biochemistry
University of Adelaide
South Australia

February, 2015
# Table of Contents

**Abbreviations** .................................................................................................................. iv

**Abstract** .............................................................................................................................. vi

**Statement of originality** ...................................................................................................... ix

**Communications and presentations** .................................................................................. x

**Acknowledgements** ........................................................................................................... xi

## Chapter 1: General introduction ......................................................................................... 1

1.1 Tuberculosis and its pathogens ......................................................................................... 2

1.2 Diagnosis and current treatment for tuberculosis ............................................................. 3

1.3 Drug-resistant tuberculosis: Need for new anti-TB drugs ............................................... 4

1.4 Biotin ................................................................................................................................ 6

1.5 Biotin-dependent enzymes ............................................................................................... 6

1.6 Biotin biosynthetic pathway ............................................................................................. 8

1.7 Biotin biosynthesis as promising target for anti-TB drugs development ........................ 10

1.8 Structural biology of biotin biosynthetic enzymes .......................................................... 12

1.8.1 BioC-O-methyltransferase ......................................................................................... 12

1.8.2 BioH carboxylase ....................................................................................................... 13

1.8.3 7-Keto-8-aminopelargonic acid synthase ..................................................................... 16

1.8.4 7,8-Diaminopelargonic acid synthase ........................................................................ 18

1.8.5 Dethiobiotin synthetase ............................................................................................ 21

1.8.6 Biotin synthase .......................................................................................................... 22

1.9 Inhibitors of biotin biosynthetic enzymes ....................................................................... 27

1.10 Fragment-based drug discovery ...................................................................................... 32

1.11 Selection of enzyme target .............................................................................................. 33

1.12 Objectives ....................................................................................................................... 34

## Chapter 2: General materials and methods ....................................................................... 36

2.1 Materials ........................................................................................................................... 37

2.1.1 Chemical reagents ..................................................................................................... 37

2.1.2 General materials ..................................................................................................... 38

2.1.3 Antibodies ................................................................................................................ 38
2.1.4 Commercial kits ..........................................................38
2.1.5 Sequencing oligonucleotides ...........................................39
2.1.6 Bacterial strains .........................................................39
2.1.7 Bacterial media ..........................................................39
2.1.8 Plasmids .................................................................40
2.1.9 Buffers and solutions ..................................................40
2.1.10 Computer software ..................................................41
2.1.11 Web resources ..........................................................41
2.2 General methods ..........................................................42
2.2.1 Plasmid DNA extraction ...............................................42
2.2.2 Agarose gel electrophoresis .........................................42
2.2.3 Restriction digestion of DNA .......................................43
2.2.4 Ligation of DNA fragments .........................................43
2.2.5 Site-directed mutagenesis ..........................................43
2.2.6 Transformation of recombinant plasmids by heat-shocked method .......44
2.2.7 Glycerol stocks ..........................................................45
2.2.8 Sequencing ..............................................................45
2.2.9 Expression of His₆-tagged proteins ................................45
2.2.10 Purification of His₆-tagged proteins using Profinia™ ...............46
2.2.11 Refolding of recombinant protein by dialysis .......................47
2.2.12 Concentration of the proteins ......................................48
2.2.13 Determination of protein concentration ..........................48
2.2.14 SDS-PAGE electrophoresis and gel staining ......................49
2.2.15 Western blotting ......................................................49
2.2.16 Size exclusion chromatography - multi-angle light scattering analysis .......50
2.2.17 Liquid chromatography-electrospray ionization tandem mass spectrometry ....50
2.2.18 MALDI mass spectrometry .........................................51

Chapter 3: Assay development .................................................52
3.1 Introduction ..................................................................53
3.2 Specific methods ..........................................................54
  3.2.1 Cloning, expression and purification of DTBS enzymes ....................54
  3.2.2 Enzyme assay ..........................................................55
  3.2.3 ATP-binding assay ....................................................57
3.2.4 SPR-binding assay ................................................................. 58

3.3 Results and discussion ............................................................................. 59
  3.3.1 Purification of DTBS proteins ............................................................... 59
  3.3.2 Enzyme assay .......................................................................................... 64
  3.3.3 ATP-binding assay .................................................................................. 71
  3.3.4 Surface plasmon resonance (SPR)-binding assay ................................ 75

Chapter 4: Nucleotide triphosphate promiscuity in Mycobacterium tuberculosis .... 81
dethiobiotin synthetase
Statement of authorship .................................................................................. 82
Manuscript ........................................................................................................... 84

Chapter 5: Fragment-based screening of Mycobacterium tuberculosis ............... 114
dethiobiotin synthetase inhibitors
Statement of authorship .................................................................................. 115
Manuscript ........................................................................................................... 116

Chapter 6: Final discussion and future directions .............................................. 153
  6.1 Lead optimization ....................................................................................... 154
  6.2 NTP-dependent enzymes ......................................................................... 160
  6.3 Transcriptional regulation mechanism of biotin biosynthesis .................... 163

References .......................................................................................................... 167

Appendix: Biotin biosynthesis in Mycobacterium tuberculosis: Physiology, ....... 182
biochemistry and molecular intervention
Published review article .................................................................................... 183
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>Acetyl-CoA carboxylase</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine 5′-triphosphate</td>
</tr>
<tr>
<td>ADP</td>
<td>Adenosine 5′-diphosphate</td>
</tr>
<tr>
<td>Apo</td>
<td>Unliganded enzyme</td>
</tr>
<tr>
<td>BirA</td>
<td>Biotin inducible repressor A</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic local alignment search tool</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>BPL</td>
<td>Biotin protein ligase</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>CTP</td>
<td>Cytosine 5′-triphosphate</td>
</tr>
<tr>
<td>°C</td>
<td>Degrees Celsius</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiotheitol</td>
</tr>
<tr>
<td>EcDTBS</td>
<td><em>Escherichia coli</em> dethiobiotin synthetase</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and drug administration</td>
</tr>
<tr>
<td>GTP</td>
<td>Guanosine 5′-triphosphate</td>
</tr>
<tr>
<td>Holo</td>
<td>Ligand bound enzyme</td>
</tr>
<tr>
<td>hr</td>
<td>Hour</td>
</tr>
<tr>
<td>HTS</td>
<td>High throughput screening</td>
</tr>
<tr>
<td>$IC_{50}$</td>
<td>Inhibition concentration at 50% enzyme activity</td>
</tr>
<tr>
<td>ITP</td>
<td>Inosine 5′-triphosphate</td>
</tr>
<tr>
<td>kb</td>
<td>Kilobase pair</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilodalton</td>
</tr>
<tr>
<td>$k_{cat}$</td>
<td>Catalytic constant</td>
</tr>
<tr>
<td>$K_D$</td>
<td>Dissociation constant</td>
</tr>
<tr>
<td>$K_m$</td>
<td>Michaelis-Menten constant</td>
</tr>
<tr>
<td>$K_i$</td>
<td>Inhibition constant</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>LB</td>
<td>Luria broth</td>
</tr>
<tr>
<td>MW</td>
<td>Molecular weight</td>
</tr>
<tr>
<td>min</td>
<td>Minute, minutes</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>MIC</td>
<td>Minimal inhibitory concentration</td>
</tr>
<tr>
<td>MtDTBS</td>
<td>Mycobacterium tuberculosis dethiobiotin synthetase</td>
</tr>
<tr>
<td>MW</td>
<td>Molecular weight</td>
</tr>
<tr>
<td>MWCO</td>
<td>Molecular weight cut-off</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein data bank</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethylsulfonylfluoride</td>
</tr>
<tr>
<td>RMSD</td>
<td>Root mean square deviation</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>RU</td>
<td>Resonance units</td>
</tr>
<tr>
<td>s</td>
<td>Seconds</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>SPR</td>
<td>Surface plasmon resonance</td>
</tr>
<tr>
<td>TTP</td>
<td>Thymidine 5′-triphosphate</td>
</tr>
<tr>
<td>Tris</td>
<td>2-amino-2-hydroxymethylpropane-1,3-diol</td>
</tr>
<tr>
<td>UTP</td>
<td>Uridine 5′-triphosphate</td>
</tr>
<tr>
<td>$V_{\text{max}}$</td>
<td>Maximum velocity</td>
</tr>
</tbody>
</table>
Abstract

Tuberculosis (TB) ranks alongside HIV-AIDS and malaria as one of the leading causes of death by infectious disease worldwide. The 2015 Millennium Development Goals (MDGs) for reducing the mortality rate and the incidence of new patients to 50% of the 1990 incidences have nearly been reached. However, the prevalence of multidrug-resistant TB (MDR-TB) remains well off-track and needs to be addressed as a public health crisis. Within the past 40 years only one new anti-TB drug with a unique mode of action, namely bedaquiline (FDA approved in 2012), has become available for drug resistant strains of TB but even this has concerning side effects. Clearly, there is an urgent need for safer drugs that have no pre-existing resistance mechanism to terminate the global prevalence of drug resistance TB.

Biotin biosynthesis has been proposed as a promising druggable target for anti-TB drug discovery. Biotin is an essential metabolite required for growth of all living cells. Biotin is synthesized de novo in microorganism, plants, and fungi. The absence of this metabolic pathway in humans makes biotin biosynthesis attractive for antibiotic discovery. In particular, biotin biosynthesis plays an important metabolic role as the sole source of biotin in all stages of the tuberculi life cycle due to the lack of biotin transporter. Biotin is intimately associated with lipid synthesis where the products form key components of the cell membrane that is critical for bacterial survival. Therefore, enzymes involving in lipid synthesis and biotin biosynthesis have been designated as an excellent target for the development of new anti-TB drugs to combat drug resistant TB.
Dethiobiotin synthetase (DTBS) catalyzes the penultimate step of the biotin biosynthetic pathway. It was selected as a target for screening inhibitors for *M. tuberculosis* biotin biosynthesis in this study due to its essential role in the growth and virulence of tuberculi. X-ray crystal structures of *M. tuberculosis* DTBS (MtDTBS) reveals two preformed adjacent ligand-binding pockets that allowed DAPA and NTP substrates to bind independently, making both pockets attractive for drug discovery. Enabling technologies were developed for the characterization of DTBS enzymes, including *in silico* screening coupled with X-ray crystal structures and three new facile assays for identifying ligand binding in the NTP pocket, namely an enzyme assay, a competitive ATP-binding assay and surface plasmon resonance (SPR) analysis. Unexpectedly, MtDTBS was shown to have broad specificity for a variety of nucleotide triphosphates, although the enzyme had the highest affinity for CTP in competitive binding and SPR assays. For the first time, X-ray crystal structure of MtDTBS bound to a nucleotide triphosphate (CTP) has been reported, showing that the nucleoside base is stabilized in its pocket through hydrogen-bonding interactions with the protein backbone, rather than amino acid side chains. These novel findings for MtDTBS are in contrast to other DTBS orthologs, for example *Escherichia coli* DTBS (EcDTBS) preferentially binds ATP primarily through hydrogen-bonds between the adenine base and the carboxamide side chain of a key asparagine. Mutational analysis performed alongside *in silico* experiments revealed a gate-keeper role of asparagine at position 175 in *E. coli* DTBS that excludes binding of other nucleotide triphosphates. Due to the absence of the gate-keeper at an equivalent position, MtDTBS thus has the broad specificity to multiple types of nucleotide triphosphates.

The X-ray crystal structure of MtDTBS in complex with CTP was used in an *in silico*, fragment-based approach to drug discovery. Compounds identified by *in silico* docking were verified using an SPR binding assay and DTBS enzyme assay. Total six hits (namely CT6,
CT7, B1, B3, B7, and B9) were identified that were predicted to bind to the protein “hot spot” located in the phosphate-binding loop at the junction of the two ligand binding pockets. Lineweaver-Burk analysis revealed one compound, gemcitabine, was competitive against DAPA and ATP. The low molecular weight (< 300 Da), low chemical complexity, and good ligand efficiency (LE) (0.2-0.3 kcal/mol/heavy atom) of the hits make them attractive targets for chemical development into more drug-like leads. Interestingly, the anti-cancer gemcitabine CT6 clearly showed in vitro inhibitory activity against MtDTBS, suggesting an application of this existing drug as a new anti-TB agent. As proof of concept, the potential optimization of leads has been proposed via merging CT6 with DAPA carbamate in order to avoid potential toxicity that might cause through off-target of human NTP-dependent enzymes. Finally, the potential transcriptional regulation of M. tuberculosis biotin biosynthesis has been firstly proposed in order to understand the biotin metabolism of tuberculi and to combat TB effectively.

Thesis layout:
The thesis will be presented as a combination of conventional chapters and publication formats. Two manuscripts are presented in chapter 4 and 5 with their own referencing. A general introduction, general materials and method, and final discussion and future direction is also be included to link together all the research conducted during my candidate. A publishing agreement with all co-authors involved with the work is included.
Statement of Originality

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution to Wanisa Salaema and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University’s digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time

.................................................. ........................................

Wanisa Salaemae  Date
Communications and Presentations


Acknowledgement

I would like to express my profound gratitude to my supervisor, Grant Booker, for having a place for me to work in his lab. Grant, I still remember your first reply within a few hours after I sent email to express my interest of doing PhD in your lab. Your enthusiasm was encouraging and I am glad that I decided to come and completed my PhD here. Thank you for your expertise and knowledge about structure-based drug design and computational approaches that really fulfill my interest of drug discovery field.

I extend gratitude to my co-supervisor, Steven Polyak. I appreciate the amount of time and support that you gave me throughout my PhD. This thesis would not have been completed without your constant and untiring advice, guidance, expertise and assistance. Thank you for encouraging me to start this project from my questions. I used to think it is very difficult to develop new assays but finally, I can make it because of your great advice.

I was the recipient of the Royal Thai Government Scholarship, which funded my PhD studies, living expense and travel costs for attending an international conference in Boston. I truly appreciate the financial assistance.

I would like to thank Min Yap and Matthew Wilce from Monash University for solving X-ray crystal structure of MtDTBS in complex with CTP. Without this structure, my PhD project should not be completed. Thank you to Staphan Meding, James Eddes and Chris Crusaro from the Adelaide Proteomics Centre for your assistance with the mass spectrometry works. Thanks to Colleen Bindloss for your assistance with gel filtration and ASTRA. Thanks to the Wallace & Frobe Lab, the Whitelaw lab and the Brunning lab for being my go
to place when I need to borrow some chemicals. To everyone in the department, the different labs and the lovely people from CSU and other support staffs, thank you for the smiles and chats in the hallways.

Thanks to the Booker lab member for their support during my PhD. Especially Kate Wegener for helping me with all aspects of structural analysis. Thank you Al Azhar for being a good lab college, a wonderful brother who always listen and give advice whenever I had problem, and also being a great friend who I spent fun time for sports, parties, or either helping me move houses. Jiulia, thanks for your neurofens and of course chocolate therapy (you still need them until you finish PhD 😊). Thanks Ashleigh, Andrew, Louise and David for making the lab the positive place to work. Thanks previous lab members, Mike (special thanks for allowing me to hit your arm when I got bad assay results 😊), Sushil, Lungisa and Wei Wei for being so helpful. Thanks all my students for listening to my supervision and helping me many parts of my project.

I would like to thank the most important people in my life, Dad, Mom, and my brothers for their love and belief in me. Dad, now I become Dr Wanisa as I have promised you since I was 15 years old. Thank to Toori for constant cheering me up and trying to understand me when I needed particularly during thesis writing. Thank all of my friends, especially Dylan, Shee Chee, Cher-Lynn, Yassir, Silvia, Nisa, Wimie, P’Bee, P’Noks, P’Lukyee, Prae and Yulee, that have supported and gave me accompany during my stay in Australia. You make my serious life more interesting and enjoyable. I love you all and I look forward to see you again somewhere soon.
CHAPTER 1

GENERAL INTRODUCTION
1.1 Tuberculosis and its Pathogens

Tuberculosis (TB) has been a major health problem for many thousands of years (1-3). It took almost 200 years after the discovery of the disease in early 1689 to identify the causative pathogen, namely *Mycobacterium tuberculosis*, by Robert Koch who awarded a Nobel prize in physiology or medicine in 1905 for this discovery (4). In addition to *M. tuberculosis*, TB can also be caused by other mycobacterium species from the *Mycobacterium tuberculosis* complex such as *M. bovis*, *M. microti*, *M. africanum* (1,3). *M. tuberculosis* remains the principal pathogen for humans whilst patients infected with *M. africanum* are rarely found and the two other species commonly infect other animals (1,4).

Currently, TB is a global pandemic that ranks alongside HIV-AIDS and malaria as the leading cause of death by infectious disease with the highest incidence rates in South-East Asian, African and Western Pacific countries (5,6). In 1993 the WHO declared TB as a global health emergency and set the Millennium Development Goal (MDG) of reducing the prevalence and mortality rates to 50% of those observed in 1990 by the 2015 deadline (7). Although the rates of new TB cases and mortality have reduced over the last decade within reach of the 2015 target, the number of TB patients and the prevalence of drug resistant strains are out of control (6). Multidrug-resistant TB (MDR-TB) needs to be addressed as a public health crisis in order to achieve the ambitious MDG target of complete elimination of TB as a public health concern by 2050 (5,6)

Our understanding of the pathogenesis of TB is still insufficient. As far as we know, TB can affect anyone regardless of the status of their immune system, causing two different stages of disease; the active stage and asymptomatic latent stage (4). For asymptomatic subjects, the immune system confines the bacterial pathogen to survive in macrophages resulting in the latent stage of TB infection (4). Meanwhile, the active TB causes an airborne
disease that transmits easily through the air from a patient having active respiratory TB disease via coughing or sneezing to others nearby (8). The bacterial pathogen mainly infects the lungs causing symptoms associated with respiratory disease such as severe coughing with blood and pains in the chest and other parts of the body (8,9). Individuals with weakened immune systems, for example patients with a HIV co-infection, are more susceptible to infection with active TB. If left untreated, the mortality rate for HIV patients with TB is over 50% (10).

1.2 Diagnosis and Current Treatment for Tuberculosis

There are two common diagnostic tests for TB; the Mantoux tuberculin skin test and a TB blood test (8). The skin test is initially performed by injecting a small volume of tuberculin fluid into the skin, usually on the lower part of the arm. After 2 or 3 days, swelling at the tested skin area is monitored for signs of a positive TB infection. Notably, false positive results can occur in persons who have been previously immunized with the BCG vaccine (Bacille Calmette-Guèrin strain of M. bovis). In contrast, the TB blood test that detects TB antigens in the blood is not affected by the BCG vaccination. Positive results from these tests can diagnose the presence of the infectious agent but not specify the stage of the disease. Therefore, other tests such as a chest X-ray and a sputum culture are required to identify active TB disease (8).

Treatments are applied differently to patients harboring the active or latent forms of TB. Current guidelines for directly observed therapy (also known as DOT), a strategy developed by WHO for TB management (7), recommend a minimum of 6 months of TB treatment with first-line drugs (isoniazid, rifampicin, ethambutol and pyrazinamide) combined with second-line drugs (ie fluoroquinolones, aminoglycosides and cycloserine) if
needed. The standard short course is divided into the intensive phase (for 2 months) using a combination of isoniazid, rifampicin, ethambutol, and pyrazinamide and followed by the continuation phase (at least 4 months) using isoniazid and rifampicin administered daily, twice-, or thrice-weekly (11,12). For HIV-positive TB patients, anti-retroviral co-treatment must also be applied in this regimen. Available anti-retroviral agents that are currently used include nucleoside and non-nucleoside reverse transcriptase inhibitors, protease inhibitors, integrase inhibitors, chemokine receptor antagonists, and entry (fusion) inhibitors (11). Completion of the anti-TB treatment is usually required prior to commencement of the anti-retroviral therapies unless patients show poor clinical status (in late stage 3 or 4 or having CD4 count < 350/mm³) and need to start the anti-retroviral therapy sooner (13). Four treatment options are approved for treating latent TB infections (8,14). Principally, isoniazid can be administered for 6 or 9 months daily to HIV patients taking anti-retroviral therapy and children 2-11 years of age (14). Another effective treatment option is a 12-dose regimen of isoniazid and rifapentine taken once a week for 3 months under DOT. However, this regimen is not suitable for patients younger than 12 years old nor pregnant females due to the toxicity of rifapentine (14). The administration of rifampicin for 4 months is used for patients whom suffer from hepatotoxicity of isoniazid (15). Noteworthy, these regimens are recommended only for individuals who are at high risk of developing active TB. Whilst isoniazid prevents the development of active TB, only rifampicin and rifapentine are competent to combat the latent tuberculi (16,17).

1.3 Drug-Resistant Tuberculosis: Need for New Anti-TB Drugs

The situation of worldwide TB epidemic has been aggravated by several factors such as the rising epidemic of HIV-AIDS co-infection (18), decreasing efficacy of the BCG
vaccine (19), inadequate or inefficient administration of chemotherapies and non-compliance of treatment (4). These factors have lead to the rise of *M. tuberculosis* strains that are multidrug-resistant TB (MDR-TB; resistance to isoniazid and rifampicin) and extensively drug-resistant TB (XDR-TB; resistant to isoniazid, rifampicin, fluoroquinolones, and second-line injectable drugs) (20,21). As mentioned earlier, MDR-TB should be considered as a public health crisis. In 2012, MDR-TB was found in 3.6% of new TB cases and 20% of treated TB patients. Approximately, 10% of MDR-TB develop XDR-TB (6). Of immense concern is the rise of totally drug-resistant TB (TDR-TB) strains, discovered in Italy (in 2007), Iran (in 2009) and India (in 2012), that are resistant to anti-TB drugs (22-25). This new definition of resistance has not yet been formally recommended by WHO (26).

It is clear that existing anti-TB chemotherapies, that have been used in the clinic for the past 40 years, are losing their efficacy against the drug-resistant strains of *M. tuberculosis* (27). Indeed, resistance to the first-line drugs isoniazid and rifampicin makes treatment more complicated and expensive with longer courses, higher doses and associated toxicity. In late 2012, a novel anti-TB drug namely bedaquiline was approved by the FDA for MDR- and XDR-TB treatment (6). This compound appears to have significant associated toxicity as clinical trials showed 5-fold higher rates of deaths in the patients receiving bedaquiline compared to the control group (28,29). Therefore, there is a critical need for better treatment options. More precisely, identification of novel chemical classes affecting new drug targets that are not subject to pre-existing resistance mechanisms are now required for combating drug resistance.
1.4 Biotin

Biotin (also known as vitamin B7, vitamin H or coenzyme R) is a water-soluble vitamin that is required for growth and pathogenicity of *M. tuberculosis*. Biotin is an essential cofactor for several biotin-dependent enzymes, which is involved in membrane lipid synthesis, gluconeogenesis and amino acids metabolism (see section 1.5 for more detail) (30-34). Although biotin is essential for all living cells, only microorganisms, plants and some fungi synthesize biotin *de novo* (35). Human and other mammals lack this metabolic pathway, thereby relying on an uptake either from their diet or from intestinal biotin-producing bacteria (36). Hence, the enzymes that synthesize biotin represent promising drug targets for new antibiotics, and are the subject of this thesis.

1.5 Biotin-Dependent Enzymes

Generally, several carboxylase enzymes have been identified to require biotin as a cofactor such as acetyl-CoA carboxylase, geranyl-CoA carboxylase, 3-methylcrotonyl-CoA carboxylase, methylmalonyl-CoA decarboxylase, propionyl-CoA carboxylase, pyruvate carboxylase, oxaloacetate decarboxylase, and urea carboxylase (37-40). These enzymes are involved in the synthesis of fatty acids and polyketide, amino acids metabolism, gluconeogenesis, and urea utilization (38). Based on currently available literature and annotation of the genome, *M. tuberculosis* seems to have functional acetyl-CoA carboxylase, propionyl-CoA carboxylase and pyruvate carboxylase. Instead of acetyl-CoA carboxylase and propionyl-CoA carboxylase, acyl-CoA carboxylases was used as a term for enzymes catalyzing the carboxylation of acyl-CoA substrates, such as acetyl-CoA, propionyl-CoA and butyryl-CoA, to generate malonyl-CoA and methylmalonyl-CoA that will be utilized in fatty acid synthesis and polyketide synthesis pathways (34,41,42). The end products of these
pathways, namely mycolic acids, linear fatty acids and methyl-branched lipids, are major components of the lipid bi-layer in mycobacterial cell envelope (33,34,42). Due to an important function of cell membrane in bacterial defence, metabolic enzymes involved in the synthesis of mycolic acids are highlighted as promising targets for antibiotic drug discovery (43-45). Noteworthy, our understanding of the role of pyruvate carboxylase in M. tuberculosis is still limited. Only one study on pyruvate carboxylase from M. smegmatis suggested that the enzymes from M. smegmatis and M. tuberculosis are highly homologous in that the enzymes formed a homomultimer of ~121 kDa, required acyl-CoA as substrate and were inhibited by ADP and an excess of Mg\(^{2+}\), Co\(^{2+}\), and Mn\(^{2+}\) (46). Pyruvate carboxylase from M. tuberculosis was proposed to contain a putative PLP cofactor binding site and play an important role in gluconeogenesis via regulation of the oxaloacetate synthesis, suggesting the same function as pyruvate carboxylase in other microorganisms (46).

The biotin-dependent enzymes contain several highly conserved domains required for catalysis, namely biotin carboxylase, carboxyltransferase and biotin carboxyl carrier protein (38,47,48). Biotin is covalently attached to a conserved lysine residue of the biotin carboxyl carrier protein domain by the function of biotin protein ligase (BPL, encoded by birA) (49). The attachment of biotin or “biotinylation” is an ATP-dependent reaction. BPL is generally classified into three classes based on the nature of the N-terminal domain whilst the central domain for catalyzing biotinylation is highly conserved in all classes. Class I BPLs such as from M. tuberculosis is the short peptide chain compared to other two classes due to the lack of a DNA binding domain (49). Class II BPLs, such as from E. coli, B. subtilis and S. aureus, and Class III BPLs found in eukaryotes contain the DNA binding domain at their N-terminal region, making them bi-functional enzymes that acts as both a transcriptional repressor and a catalytic enzyme in biotinylation (50-53). Notably, Class III BPLs also have an extended N-
terminal region with unknown function. Due to its important function in biotinylation, BPL is also proposed as a promising target for anti-TB drug development (54).

1.6 Biotin Biosynthetic Pathway

The biotin biosynthetic scheme can be divided into two stages; 1) synthesis of the pimelate precursor and 2) the conserved metabolic pathway catalyzing the final four steps that yield biotin. Figure 1.1 shows the biosynthetic pathway and the chemical structures of the synthetic intermediates. The steps leading to the formation of a pimeloyl-thioester precursor (either linked to CoA or acyl carrier protein) are variable among biotin-producing organisms. The best understood pathways are in E. coli (ie BioC-BioH pathway) and B. subtilis (ie BioI-BioW pathway) (reviewed (55,56)). The presence of BioC and BioH homologues in mycobacteria genomes suggest that M. tuberculosis employs the same pathway as E. coli (30,57). Here, a 3-carbon malonyl-thioester 1 is firstly methylated by the BioC-O-methyltransferase to produce a methyl ester 2 that serves as the precursor for two iterations of alkyl chain elongation using the fatty acid synthetic pathway, producing a 7-carbon pimeloyl-ACP methyl ester 4. Alternatively, the pimeloyl-ACP methyl ester 4 can be synthesized from a pimelate monomethyl ester 3 by the activity of Acyl-ACP synthetase (58). The methyl group of the pimeloyl-ACP methyl ester 4 is then hydrolyzed by BioH carboxylesterase to generate the pimeloyl-ACP 5 (55,56,59). Pimeloyl-ACP, rather than pimeloyl-CoA, is believed to be the physiological intermediate required in the subsequent conserved metabolic pathway (56,58).
Figure 1.1 Biotin biosynthetic pathway in *M. tuberculosis*. The proposed precursor synthesis (dashed box) and the conserved metabolic pathway (black box) are shown. The modified chemical atoms were highlighted in red. Abbreviations: ACP, acyl carrier protein; AaaS, Acyl-ACP synthetase; AMTB, S-adenosyl-2-oxo-4-methylthiobutyric acid; DOA, 5'-deoxyadenosine; FAS, fatty acid biosynthesis; SAM, S-adenosyl-L-methionine; SAH, S-adenosylhomocysteine. The figure was adapted from Salaemae et al. 2011 (30).

Unlike the synthesis of the pimelate precursor, the final four reactions in the pathway that assemble the biocyclic rings of biotin are highly conserved amongst microorganisms and plants. The pimeloyl-ACP 5 is converted to biotin by the activities of 7-keto-8-aminopelargonic acid synthase (KAPAS), 7,8-diaminopelargonic acid synthase (DAPAS), dethiobiotin synthetase (DTBS) and biotin synthase (BS) that are encoded by *bioF*, *bioA*, *bioD*, and *bioB* respectively (30,60-62). Amongst biotin producing organisms,
Saccharomyces cerevisiae is the only species reported that employed only the last three steps of the conserved metabolic pathway due to the presence of a 7-keto-8-aminopelargonic acid (KAPA) transporter (63,64). Briefly, KAPAS converts pimeloyl-ACP 5 to KAPA 6 by using L-alanine as an amino donor and releasing ACP and CO₂ (65). At the antepenultimate step, DAPAS catalyzes the conversion of KAPA 6 to 7,8-diaminopelargonic acid (DAPA) 7 using S-adenosyl-L-methionine as an amino group donor (62). Next, DTBS catalyzes closure of the ureido ring of dethiobiotin (DTB) 8 from DAPA 7 using CO₂ and ATP with the release of ADP and inorganic phosphate (62). Finally, closure of the thiophane ring of biotin 9 by biotin synthase requires the insertion of a sulfur atom between the C6 carbon and the non-reactive methyl C9 of DTB 8 (66).

1.7 Biotin Biosynthesis As Promising Target for Anti-TB Drugs Development

De novo biotin biosynthesis is the sole source of biotin in M. tuberculosis due to the lack of a biotin transporter. Several biotin transporters have recently been identified and characterized in eubacteria and archae. BioY works in corporation with an ATP-dependent energy coupling system (BioMN) and has been well characterized in many bacterial species (67,68). Alternatively, the YigM protein has been reported in E. coli (69,70) after 42 years of numerous attempts to identify this transport system (71-79). A homologue of YigM, designated as MadN, has also been identified in certain Gram-negative bacteria (69). Noteworthy, genome annotation studies have failed to identify a BioY homologue in mycobacteria (67,80). Likewise, nucleotide sequence searching using tblastx algorithm could not recognize a homologue of YigM or MadN in M. tuberculosis genome. Furthermore inhibition of mycobacterial growth using two natural compounds, namely amiclenomycin and actithiazic acid, that targets biotin biosynthesis enzymes (see section 1.9 for more detail)
strongly suggest that mycobacteria do not have a high affinity biotin transport system (81-85). The inhibitory activity of amiclenomycin could not be compromised unless the growth media was supplemented with a high concentration of biotin (> 0.4 nM) (84,85). This is 10-fold higher than the concentration of biotin normally found in human plasma (86), implying that biotin can diffuse into mycobacterial cells at non-physiological concentrations but not through a high-affinity, active transporter.

Biotin biosynthesis is one of the pathways associated with membrane lipid synthesis that has been designated as promising targets for the development of new antibiotics. The products of lipid synthesis are key components of cell membrane that plays critical role in bacterial survival and defence (43-45). Targeting enzymes involved in membrane lipid synthesis therefore inhibits bacteria to grow. This approach is clinically validated by isoniazid that targets InhA in the fatty acid synthase II system (87,88). Importantly, screening of drugs effective against latent M. tuberculosis revealed that certain lipid biosynthesis enzymes can be targeted against tuberculi at this stage (89,90). These data validate biotin biosynthesis as an excellent target for the development of new anti-TB drugs that are able to combat both the active and latent stages of TB. Lastly, the absence of pathway homologue in human makes biotin biosynthesis more attractive (30,91).

Validation of biotin biosynthesis as a druggable anti-TB target is further supported by genetic knockout studies. Firstly, deletion of MMAR_2770, a homologue of Rv1882c in M. tuberculosis that encodes a putative short chain dehydrogenase/reductase catalyzing synthesis of a biotin precursor, defected the growth of M. marinum on blood agar unless supplementing with high concentrations of biotin (>1 µM) (57). The mutant also failed to develop pathogenicity in murine macrophage and zebrafish models of disease (57). In separate studies, disruption of bioA in the biotin biosynthetic operon impaired survival of M. smegmatis during stationary phase on carbon-depleted media (92,93) while disruption of
*bioF* and *bioB* debilitated the bacterial growth during and post infection of murine macrophages (94). Likewise, genome-wide genetic screens revealed that *bioF*, *bioA*, and *bioB* are essential for virulence as shown by impeded growth of the deletion strains in murine macrophages (95). In addition, *bioF*<sup>−/−</sup> *M. tuberculosis* showed poor recovery from mouse lung and spleen in an infection challenge experiment (95). Together, these genetic studies validated biotin biosynthesis as a key metabolic process during growth, infection, and survival in the latent life cycle of mycobacteria.

### 1.8 Structural Biology of Biotin Biosynthetic Enzymes

#### 1.8.1 BioC-O-methyltransferase

BioC-O-methyltransferase [BioC; EC 2.1.1.197, encoded by *bioC*] is an S-adenosyl-L-methionine (SAM, also known as AdoMet)-dependent methyltransferase (58). The native form of BioC from *E. coli* and *Bacillus cereus* is monomeric with molecular mass of ~31 kDa (96). No crystal structure of BioC has yet been reported. The function of BioC is to generate a methyl ester 2 by transferring a methyl group from SAM to the free carboxyl group of a malonyl-thioester 1, which is either linked to CoA or ACP (Figure 1.1) (58,96). This methylation step is essential to neutralize the negative charge of the carboxyl group prior to interaction with the extremely hydrophobic active sites of the fatty acid synthases (96). Without BioC, the hydrophilic malonyl-thioester cannot enter fatty acid synthesis for assembling the pimelate moiety. Noteworthy, malonyl-ACP was shown to be preferred acceptor of methyl groups from SAM than malonyl-CoA (96). It is also an early precursor of the canonical fatty acid synthesis pathway (96). Therefore, the expression of BioC and its activity must be tightly controlled to avoid the depletion of the malonyl-ACP pool causing impaired cell growth (96).
1.8.2 BioH Carboxylesterase

As described above for BioC, hydrophobic molecules are often substrates for fatty acid synthetic enzymes (58,97). Thus the carboxyl group of the pimeloyl-ACP methyl ester 4 must be liberated after two cycles of carbon elongation in order to leave the fatty acid synthesis machinery. Indeed, the free carboxyl group is further required for biotin protein ligase (BPL), the enzyme responsible for protein biotinylation (as discussed in section 1.5) (98). BioH carboxylesterase [BioH; EC 3.1.1.85, encoded by bioH] plays an essential role in the cleavage of the methyl ester 4 producing a pimeloyl-ACP precursor 5 required for KAPAS in the subsequent conserved metabolic pathway (Figure 1.1) (58,59).

BioH is an active monomeric protein of ~28 kDa (99). It belongs to the hydrolase superfamily, a group of enzymes containing a classical Ser-His-Asp catalytic triad and a pentapeptide Gly-Xaa-Ser-Xaa-Gly motif (100-102). BioH is a carboxylesterase that can employ short acyl chain as substrates (100,103). Three crystal structures of BioH have been determined from *E. coli, Shigella flexneri* and *Salmonella enterica* (Table 1.1), showing a two-domain containing protein (Figure 1.2) (59,100). The α/β/α three layer large N-terminal domain, also known as the Rossman fold, consists of a twisted seven β-sheet in the middle and sandwiched by five α-helices at both ends. Meanwhile, the small C-terminal domain consists of four α-helices. The catalytic triad active site is located at the interface between these two domains (59,100). The overall structure of BioH is similar to bromoperoxidase [EC 1.11.1.10], aminopeptidase [EC 3.4.11.5], epoxide hydrolase [EC 3.3.2.3], haloalkane dehalogenase [EC3.8.1.5] and lyase [EC 4.2.1.39] (100). BioH captures the pimeloyl-ACP methyl ester substrate in association with an acyl carrier protein (Figure 1.2). The acyl carrier protein consists of four helices where its second helix (α2) interacts with the BioH small domain to facilitate the capture of the phosphopantetheine arm of the pimeloyl-ACP methyl ester substrate through Ser35 positioned at the N-terminus of the α2 helix (59). This acyl
carrier protein-dependent complex is also found in other enzymes such as a P450-BioI and a castor desaturase (59,104,105).

Although various ACP-bound methyl ester compounds, such as glutaryl (C5), adipyl (C6), suberyl (C8), and azelayl (C9), can interact with BioH, pimeloyl-ACP methyl ester 4 (C7) is the preferred substrate. Its length is efficiently organized in the enzyme hydrophobic cavity to reach the catalytic triad and, thus, is hydrolyzed faster than the unnatural substrates (59). Indeed, shorter carbon chains cannot reach across the hydrophobic cavity and reach the catalytic triad, whilst longer alkyl chains cause steric clashes within the cavity. This suggests that the length of the substrate is critical for catalysis. The reaction mechanism of BioH has also been studied in some detail. A recent study of the Baylis-Hillman reaction, which is the formation of a C-C bond from an aldehyde and an activated alkene, revealed that E. coli BioH has the ability to catalyze this reaction using a range of aldehydes and alkenes (106). The catalytic and binding mechanisms of this reaction in BioH have not yet been characterized. In addition, a few studies of an enantioselectivity, which refers to the enzyme’s capacity to produce an excess of one product over the other configuration, reported that BioH has the enantioselectivity that involves the chiral synthesis by hydrolyzing aryl prochiral diesters and alcohols (107-109). Indeed, the introduction of an aromatic side chain into the L83F/L86F BioH mutant strain showed a dramatic increase of the BioH enantioselectivity toward the S-product of dimethyl-3-phenylglutarate (96%, compared to 25% that of wildtype BioH) (107). Within the small and highly hydrophobic binding pocket of BioH, the protein “hot-spot” was identified, indicating besides the catalytic triad key residues such as Phe111, Lue86, Leu83, Tyr27 and Ile71 are essential for interaction with ligands, for example to a benzene ring of dimethyl 3-phenylglutarate (108). Together, the combined knowledge of the productive carbon length of the ligand that is required for binding, and the substrate promiscuity of BioH provides useful information for rational design of BioH inhibitors.
In certain BioC-containing bacteria, BioH is replaced by other enzymes, such as BioG, BioI, and BioK, to allow the growth of an *E. coli* ΔBioH biotin auxotrophic strain on biotin-free medium (80,98,110). Although, these enzymes share low sequence similarity to BioH, they belong to the same α,β-hydrolase family with associated esterase activity (98,110). A recent study of BioJ demonstrated that the ΔBioJ strain of *F. novicida* had attenuated growth in minimal medium. Importantly, the replication of this mutant strain is 5-folds lower than that of wildtype after infection in murine bone marrow-derived macrophages, suggesting the importance of BioH-like enzymes for bacterial virulence (110).

Figure 1.2 Overall Structure of BioH. The crystal BioH S82A from *S. flexneri* is shown in complex with pimeloyl-ACP methyl ester (shown in pink) substrate and an acyl carrier protein partner (shown in blue) (PDB 4ETW). The α/β/α large domain contains seven β-sheets that are sandwiched by five α-helices (shown in grey). The small domain consists of four α-helices (shown in yellow). The residues of the catalytic triad, namely Ser82, Asp207, His235 shown in green, are located at the interface between the two domains. The Ser82 was mutated to alanine.
1.8.3 7-Keto-8-Aminopelargonic Acid Synthase

7-keto-8-aminopelargonic acid synthase [KAPAS; EC 2.3.1.47, encoded by *bioF*], also known as 8-amino-7-oxononanoate synthase (AONS), catalyzes the first step of the conserved biotin biosynthetic pathway. It is classified as type I pyridoxal 5′-phosphate (PLP or vitamin B6)-dependent enzyme in aminotransferase subclass II (61,97). Unusually, there are two *bioF* genes in the genome of *M. tuberculosis*; *bioF* and a putative *bioF*2 that are translated to 386 and 771 amino acids products respectively (111). The function of an additional N-terminal part of *bioF*2 is still unclear although it contains the putative conserved acetyltransferase (GNAT) domain (Accession: pfam13480), similar to enzymes in the N-acyltransferase superfamily.

KAPAS converts a pimelate moiety to 7-keto-8-aminopelargonic acid (KAPA) requiring L-alanine as an amino donor and PLP as a cofactor. Unlike the homologues from other species that require only L-alanine, *M. tuberculosis* KAPAS can use both L-alanine and D-alanine as amino donor substrates (112). The mechanism of KAPAS is similar to other aminotransferase enzymes where the reaction proceeds through several intermediates, namely an analine-bound external aldimine complex, a quinonoid intermediate, a lysine-bound internal aldimine complex, and a 3-oxoacid aldimine complex. This leads to DAPA 7 production via the addition of two carbons (C8, C9) and a nitrogen (N8) atom onto the pimeloyl-ACP substrate 5 (Figure 1.1) (97).

Currently, the structure of KAPAS from *M. tuberculosis* has not been determined. However, available structures of the *E. coli* and *F. tularensis* orthologs determined in the apo- and holo-forms suggest that KAPAS is active as a homodimer (Table 1.1) (113-115). Each subunit assembles into three domains, 1) a large domain containing a seven-stranded β-sheet, 2) a small C-terminal domain consisting of a four-stranded β-sheet with α-helices and
3) a small N-terminal domain forming three-stranded β-sheet that extended the C-terminal domain (61,113). The C’4 atom of PLP cofactor covalently linked to the ε-amino group of a conserved Lys236 (numbering in *E. coli*) in the active site, positioning the PLP molecule at the interface between subunits (Figure 1.3) (113,115). Apart from Lys236, active site residues His133, Ser179, Asp204, His207, Thr233 are also important for making direct contact with PLP (115,116). These residues are highly conserved in all species the KAPAS enzymes as well as other enzymes in the α-oxoamine synthase family (116). A Mg$^{2+}$ ion is essential for coordination with the O3 and hydroxyl group of Ser179, the N’8 of the KAPA intermediate, and two water molecules. A conformational change within the C-terminus upon KAPA binding results in movement of Arg349 to align antiparallel with Arg21 from the N-terminal domain allowing these two arginine residues to H-bond with the carboxyl group of KAPA (115). In addition to H-bonding, the ligand is stabilized via hydrophobic interactions between the methylene chain of KAPA and Val79 from the subunit that interacts with PLP, and Tyr264 and Ile263 from the partner subunit. This molecular detail will provide powerful information for future structure guided drug design.
Figure 1.3 Crystal Structure of KAPAS. One subunit of the *E. coli* KAPAS homodimer is shown in complex with KAPA-PLP aldimine intermediate (shown in red connected to blue respectively) (PDB 1DJ9). The protein folds into three domains that are the N-terminal small domain, the large domain containing seven-stranded β-sheet, and the C-terminal small domain. The Mg$^{2+}$ ion is shown in green.

1.8.4 7,8-Diaminopelargonic Acid Synthase

7,8-diaminopelargonic acid synthase [DAPAS; EC 2.6.1.62], encoded by *bioA*, is also known as Adenosylmethionine-8-amino-7-oxononanoate aminotransferase. DAPAS is analogous to KAPAS as both belong to the same subclass the aminotransferase family (61). Briefly, DAPAS catalyzes the attachment of a nitrogen atom derived from SAM to KAPA 6 that becomes the N7 moiety of biotin in DAPA 7 (Figure 1.1). The aminotransferase reaction of DAPAS is similar to KAPAS in that it proceeds through two partial reactions that require a PLP cofactor and an amino group donor. In contrast to KAPAS that utilizes L-alanine, here
S-adenosyl-L-methionine serves as the amino group donor \((61,117,118)\). The first partial reaction catalyzes the sequential formation of intermediates as the aldime that formed between DAPAS and PLP interacts with SAM to generate subsequent quinonoid, ketamine, and then PMP \((118)\). Following the second partial reaction, the PMP interacts with KAPA substrate generating the intermediates in a sequential, reversible direction of the first partial reaction as producing ketamine, quinonoid, and then aldime intermediates before the release of the final DAPA product and recycled PLP \((112,118)\). Interestingly, although KAPA itself has an amino group, it cannot serve as amino donor in DAPAS reaction \((117)\).

The interaction of DAPAS in the first partial reaction with PLP and KAPA, instead of SAM, is unproductive. This can be explained by the crystal structure of DAPAS in complex with KAPA that shows the substrate N8 amino group is orientated away from C4’ carbon of the PLP cofactor, thus preventing transmination and formation of the internal aldime intermediate \((117)\). This suggests specificity for the amino acid donor and revealed the important interaction of the C4’ carbon of the PLP cofactor for transamination, thereby blocking this interaction. This finding paves the way to design the inhibitors for DAPAS based on chemical analogues of the reaction intermediates such as the aldime.

A number of DAPAS crystal structures have been reported from *M. tuberculosis*, *E. coli*, *Arabidopsis thaliana*, *Chromobacterium violaceum* and *bacillus subtilis* (Table 1.1) \((62,84,117-123)\). The crystal structure of *M. tuberculosis* DAPAS in complex with PLP cofactor and KAPA substrate (PDB 4CXQ) shows that an active form of DAPAS is a homodimer with a molecular mass of ~46 kDa per monomer (Figure 1.4) \((119)\). As with other aminotransferases, a conserved lysine (Lys283 in *M. tuberculosis* DAPA) in the active site is essential for catalysis by forming an interaction with PLP. Upon the binding of KAPA, a conformational change repositions Arg400 such that it can H-bond with the carboxyl group of KAPA \((62,117,119)\). Whilst there are a number of X-ray structures of the TB enzyme,
most of the mechanistic studies on DAPAS have been performed with the *E. coli* ortholog. The equivalent arginine is also important for making interactions to the carboxyl group of DAPA as shown by the 180-fold increase of $K_m$ for DAPA in the R391A mutant (numbering in *E. coli*) compared to wildtype (118). The amino group of KAPA forms tetrahedral H-bonding to the hydroxyl group of Tyr157, the carbonyl oxygen of Gly316 and the phosphate group of PLP with an additional hydrophobic interaction to the aromatic ring of Trp64 (119). Currently, there is no crystal structure of DAPAS bound to SAM available, however, SAM is proposed to bind the enzyme at the KAPA binding site. Based upon mutagenesis studies of *E. coli* DAPAS, the DAPAS harboring R253A mutant (at equivalent position of Arg262 in *M. tuberculosis*) increased $>3$-folds higher $K_m$ for SAM than wildtype, indicating this conserve residue is important for SAM binding (123). These key residues provide useful structural information required for ligand binding and can be used in rational drug design.

**Figure 1.4 Structure of *M. tuberculosis* DAPAS.** The homodimer of DAPAS formed by two subunits, chain A (grey) and chain B (blue). The enzyme was crystallized in complex with PLP cofactor and KAPA substrate where the binding pockets are shown as green and red meshes respectively (PDB 4CXQ).
1.8.5 Dethiobiobin Synthetase

Dethiobiobin synthetase [DTBS; EC 6.3.3.3] is encoded by bioD. DTBS catalyzes the penultimate step of biotin biosynthetic pathway to produce dethiobiobin (DTB) \( \text{DTB} \) from DAPA \( \text{DAPA} \) using ATP and CO\(_2\) (Figure 1.1) (61). The DTBS reaction proceeds through three discrete steps; 1) formation of N7-DAPA carbamate, 2) formation of carbamic phosphoric acid anhydride and 3) closure of the ureido ring of DTB with the release of inorganic phosphate and ADP (124). Indeed, following carbamic phosphoric acid anhydride formation, a tetrahedral intermediate is proposed to occur before proceeding to the last step and closure of the ureido ring with the release of inorganic phosphate. However, this mechanism has not yet been verified by crystallographic studies (124,125).

The structures of DTBS have been reported from \( M. \) tuberculosis, \( E. \) coli, \( Helicobacter pylori \) and \( F. \) tularensis either in apo- or in complex with ligands such as ATP or DAPA (Table 1.1) (62,124-129). The active DTBS is a homodimer with a molecular mass of ~46 kDa (Figure 1.5). Two active sites are placed at the interface between the two subunits in anti-parallel directions 25 Å apart (129). Each subunit folds into a single \( \alpha/\beta \) globular domain consisting of a seven-stranded parallel \( \beta \) sheet in the core surrounded by helices, similar to certain GTP-dependent enzymes such as adenylsuccinate synthase and p21ras (125,129). Each active site contains adjacent binding pockets for DAPA and ATP substrates. The DAPA pocket is formed by amino acids from the two subunits whereas the ATP pocket is located solely in one subunit. Four residues, including Lys37, Thr41 from the first subunit and Leu146, Asn147 form the partner subunit (numbering in \( M. \) tuberculosis), are required for DAPA binding (129). Mutagenesis studies performed upon \( E. \) coli DTBS reveal that the mutations in Lys37 reduced \( k_{cat} \) (< 0.9% of wildtype) and increased the \( K_m \) for DAPA (>100-folds higher than wildtype) (130). Moreover, the phosphate binding loop termed P-loop (Gly8-Xaa9-Xaa10-Thr11-Xaa12-Xaa13-Gly14-Lys15-Thr16, numbering in \( M. \) tuberculosis)
is crucial for hydrogen bonding with the α, β and γ phosphate groups of ATP. Binding of ATP to *E. coli* DTBS induces conformational changes of active site loop regions including the phosphate loop (125). In particular, the replacement of Thr11 results in a 24,000-fold increase in $K_m$ for ATP while Lys15 is a critical residue for both catalysis and ATP binding (0.01% of wildtype $k_{cat}$ and 1800-fold higher $K_m$ than wildtype respectively) (125). This data suggests key residues for enzyme-ligands interaction that is useful for structure-based drug design.

**Figure 1.5 Structure of *M. tuberculosis* DTBS.** The homodimer of DTBS is formed by two subunits; chain A (grey) and chain B (blue). Two active sites are located at the interface between subunits where each active site contains two adjacent binding pockets of DAPA (pink mesh) and ATP (tred mesh). The DAPA carbamate-bound structure was determined in *M. tuberculosis* DTBS (PDB 3FMF) whereas the interaction between ATP and the enzyme has not been clarified.

### 1.8.6 Biotin Synthase

Biotin synthase [BS; EC 2.8.1.6], encoded by *bioB*, catalyzes the final step of the biotin biosynthetic pathway. Here a sulfur atom is inserted between the methyl carbons (C6
and C9) of DTB 8, thus creating the thiophane ring (Figure 1.1). BS belongs to a radical SAM (or AdoMet) superfamily that uses S-adenosyl-L-methionine for radical generation. BS contains a conserved CysxxxCysxxCys sequence motif necessary for binding of the [4Fe-4S]2+/1+ cluster (131,132). Only one structure of BS has been resolved from E. coli in complex with both SAM and DTB (PDB 1R30) (Table 1.1) (Figure 1.6) (66). Each subunit of the homodimer folds into a triosephosphate isomerase type (α/β)8 barrel flanked with two helices at the N-terminus and a disarranged C-terminal region. Two iron clusters are present within the protein; [2Fe-2S] is located deep within a TIM barrel whilst a [4Fe-4S]2+/1+ is positioned at the C-terminal end of the barrel. These iron-sulfur clusters serve different functions in BS. The [4Fe-4S]2+/1+ cluster is involved in the extraction of a hydrogen atom from the methyl and methylene groups of DTB. The cluster is initially in the [4Fe-4S]1+ state by the electron transfer system containing NADPH, flavodoxin and ferredoxin (or ferredoxin reductase) (131,133). The reduced [4Fe-4S]1+ cluster catalyzes the reductive cleavage of SAM in order to generate a 5′′-deoxyadenosyl radical and methionine. This radical can then extract one proton from each of C6 and C9 of DTB substrate (131). As a result, this first half reaction requires two SAM equivalents per molecule of biotin formed. In fact, the dimer of BS has a single functional active site with a 2:1:1 stoichiometry of SAM:DTB:BS dimer binding (131,134). In the second half reaction, the [2Fe-2S] cluster has been proposed to close the thiophane ring of biotin by donating a sulfur atom (134,135). Thus, BS itself appears to act as a substrate in vitro rather than an enzyme by producing less than one molecule of biotin per molecule of BS protein (0.3-0.4 equiv/monomer) (136). In vivo, BS has an extremely modest catalysis rate of 10-60 turnovers per monomer, implying a cellular mechanism exists to repair the [2Fe-2S] cluster (137,138). A mitochondrial matrix protein, namely Isa2, has been proposed to play a role in the regeneration of the [2Fe-2S] cluster. BS
is inactive in the Isa mutant strain of Saccharomyces cerevisiae, leading to the failure of the cell growth in minimum media or DTB supplementing media unless biotin is supplied (137).

Similar to other SAM-dependent enzymes, three conserved cysteine residues in the CysxxxCysxxCys sequence motif contribute to ligand binding to BS. Indeed, the replacement of Cys53, Cys57 and Cys60 with alanine was shown to abolish the SAM reduction activity of E. coli BS in the [4Fe-4S]$^{2+/1+}$ cluster (139,140). In addition, Cys97, Cys128 and Cys188 that do not belong to the motif, and a conserved Arg260 are essential for binding of the [2Fe-2S] cluster (140,141). This implies the function of the homologous Cys85, Cys89, Cys92 in CxxxCxxC sequence motif, Cys128, Cys161, Cys220 and Arg290 in M. tuberculosis BS.

**Figure 1.6 Structure of BS.** The available structure of E. coli BS was determined in complex with SAM (green) and DTB (purple) (PDB 1R30). Each subunit, chain A (grey) and chain B (blue), of homodimer folds as a triosephosphate isomerase (TIM) type (α/β)$_8$ barrel with additional regions at the N- and C-terminal ends. BS contains one [4Fe-4S] and one [2Fe-2S] per monomer as highlighted in yellow.
Table 1.1 Protein structure database of biotin biosynthetic enzymes

Crystal structures of biotin biosynthetic enzymes have been recorded in protein database as shown in PDB number. These proteins were crystallized in apo- and holo-forms with different ligands.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Ligands</th>
<th>Organism</th>
<th>PDB</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BioH S82A</td>
<td>Pimeloyl-ACP methyl ester</td>
<td>S. flexneri</td>
<td>4ETW</td>
<td>Agarwal et al. 2012 (59)</td>
</tr>
<tr>
<td>BioH</td>
<td>3-hydroxy-propanoic acid</td>
<td>E. coli</td>
<td>1M33</td>
<td>Sanishvili et al. 2013 (100)</td>
</tr>
<tr>
<td>BioH</td>
<td>Cl-, Dihydroxyethyl ether</td>
<td>S. enterica</td>
<td>4NMW</td>
<td>Kim et al. (unpublished)</td>
</tr>
<tr>
<td>KAPAS</td>
<td>PLP, (SO_{4})^{2-}</td>
<td>E. coli</td>
<td>1DJE</td>
<td>Webster et al. 2000 (115)</td>
</tr>
<tr>
<td>KAPAS</td>
<td>PLP, KAPA, (SO_{4})^{2-}</td>
<td>E. coli</td>
<td>1DJ9</td>
<td>Webster et al. 2000 (115)</td>
</tr>
<tr>
<td>KAPAS</td>
<td>-</td>
<td>E. coli</td>
<td>1BS0</td>
<td>Alexeev et al. 1998 (113)</td>
</tr>
<tr>
<td>KAPAS</td>
<td>Trifluoroalanine</td>
<td>E. coli</td>
<td>2G6W</td>
<td>Alexeev et al. 2006 (114)</td>
</tr>
<tr>
<td>KAPAS</td>
<td>-</td>
<td>F. tularensis</td>
<td>4IW7</td>
<td>Newcomb et al. (unpublished)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>PLP, Ethylene glycol, (SO_{4})-1,3-benzothiazol-2-ylmethylamine</td>
<td>M. tuberculosis</td>
<td>4CXR</td>
<td>Dai et al. 2014 (119)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>PLP, KAPA, Dihydroxyethyl ether</td>
<td>E. coli</td>
<td>4CXQ</td>
<td>Dai et al. 2014 (119)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>2B9, Ethylene glycol</td>
<td>M. tuberculosis</td>
<td>4MQR</td>
<td>Dai et al. 2014 (119)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>Ethylene glycol, Cl-, 2B6, Imidazole</td>
<td>M. tuberculosis</td>
<td>4MQQ</td>
<td>Dai et al. 2014 (119)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>2B1, Ethylene glycol Dihydroxyethyl ether</td>
<td>M. tuberculosis</td>
<td>4MQP</td>
<td>Dai et al. 2014 (119)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>PL8, Dimethyl sulfoxide</td>
<td>M. tuberculosis</td>
<td>3TFU</td>
<td>Shi et al. 2011 (122)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>PLP, Dimethyl sulfoxide</td>
<td>M. tuberculosis</td>
<td>3TFT</td>
<td>Shi et al. 2011 (122)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>PLP, Adenosyl-ornitine</td>
<td>M. tuberculosis</td>
<td>3LV2</td>
<td>Dey et al. 2010 (62)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>PLP</td>
<td>M. tuberculosis</td>
<td>3BVO</td>
<td>Dey et al. 2010 (62)</td>
</tr>
<tr>
<td>DAPAS Y17F</td>
<td>Na^{+}</td>
<td>E. coli</td>
<td>1SOA</td>
<td>Sandmark et al. 2004 (123)</td>
</tr>
<tr>
<td>DAPAS Y144F</td>
<td>Na^{+}</td>
<td>E. coli</td>
<td>1SO9</td>
<td>Sandmark et al. 2004 (123)</td>
</tr>
<tr>
<td>DAPAS D147N</td>
<td>Na^{+}</td>
<td>E. coli</td>
<td>1SO8</td>
<td>Sandmark et al. 2004 (123)</td>
</tr>
<tr>
<td>DAPAS R253A</td>
<td>2-propanol, Na^{+}</td>
<td>E. coli</td>
<td>1SO7</td>
<td>Sandmark et al. 2004 (123)</td>
</tr>
<tr>
<td>DAPAS R253K</td>
<td>Na^{+}</td>
<td>E. coli</td>
<td>1SO6</td>
<td>Sandmark et al. 2004 (123)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>PLP, Trans-amniclenomycin, Na^{+}</td>
<td>E. coli</td>
<td>1MLZ</td>
<td>Sandmark et al. 2002 (84)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>PLP, Cis-amniclenomycin, Na^{+}</td>
<td>E. coli</td>
<td>1MLY</td>
<td>Sandmark et al. 2002 (84)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>PLP, K^{+}</td>
<td>E. coli</td>
<td>1QJ5</td>
<td>Kack et al. 1999 (117)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>PLP, KAPA, Na^{+}</td>
<td>E. coli</td>
<td>1QJ3</td>
<td>Kack et al. 1999 (117)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>PLP, Na^{+}</td>
<td>E. coli</td>
<td>1DTY</td>
<td>Alexeev et al. (unpublished)</td>
</tr>
<tr>
<td>DAPAS R391A</td>
<td>2-propanol, Na^{+}</td>
<td>E. coli</td>
<td>1MGV</td>
<td>Eliot et al. 2002 (118)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>PLP, KAPA</td>
<td>B. subtilis</td>
<td>3DU4</td>
<td>Dey et al. 2010 (62)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>-</td>
<td>B. subtilis</td>
<td>3DRD</td>
<td>Dey et al. 2010 (62)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>PLP</td>
<td>B. subtilis</td>
<td>3DOD</td>
<td>Dey et al. 2010 (62)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>Thiocyanate Ion, Na^{+}</td>
<td>C. violaceum</td>
<td>4A6U</td>
<td>Humble et al. 2012 (121)</td>
</tr>
<tr>
<td>DAPAS</td>
<td>PLP</td>
<td>C. violaceum</td>
<td>4A6T</td>
<td>Humble et al. 2012 (121)</td>
</tr>
<tr>
<td>DAPAS &amp; DTBS</td>
<td>PLP, DTB, L(+)-Tartaric acid</td>
<td>A. thaliana</td>
<td>4A0R</td>
<td>Cobessi et al. 2012 (120)</td>
</tr>
</tbody>
</table>

*Protein was co-crystallized with a protein partner namely Acyl carrier protein.
Table 1.1 (continue) Protein structure database of biotin biosynthetic enzymes

Crystal structures of biotin biosynthetic enzymes have been recorded in protein database as shown in PDB number. These proteins were crystallized in apo- and holo-forms with different ligands.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Ligands</th>
<th>Organism</th>
<th>PDB</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTBS</td>
<td>DTP, Mg^{2+}, (PO_4)^2-</td>
<td>M. tuberculosis</td>
<td>3FPA</td>
<td>Dey et al. 2010 (62)</td>
</tr>
<tr>
<td>DTBS</td>
<td>KAPA, (SO_4)^2-</td>
<td>M. tuberculosis</td>
<td>3FMI</td>
<td>Dey et al. 2010 (62)</td>
</tr>
<tr>
<td>DTBS</td>
<td>DAPA carbamate</td>
<td>M. tuberculosis</td>
<td>3FMF</td>
<td>Dey et al. 2010 (62)</td>
</tr>
<tr>
<td>DTBS</td>
<td>-</td>
<td>M. tuberculosis</td>
<td>3FGN</td>
<td>Dey et al. 2010 (62)</td>
</tr>
<tr>
<td>DTBS</td>
<td>-</td>
<td>E. coli</td>
<td>1DTS</td>
<td>Huang et al. 1994 (126)</td>
</tr>
<tr>
<td>DTBS</td>
<td>(SO_4)^2-</td>
<td>E. coli</td>
<td>1DBS</td>
<td>Alexeev et al. 1994 (127)</td>
</tr>
<tr>
<td>DTBS</td>
<td>DAPA carbamate</td>
<td>E. coli</td>
<td>1DAI</td>
<td>Huang et al. 1995 (125)</td>
</tr>
<tr>
<td>DTBS</td>
<td>ADP, DTP, Mg^{2+}</td>
<td>E. coli</td>
<td>1DAM</td>
<td>Kack et al. 1998 (124)</td>
</tr>
<tr>
<td>DTBS</td>
<td>ADP, DTP, Mg^{2+}, (PO_4)^2-</td>
<td>E. coli</td>
<td>1DAF</td>
<td>Huang et al. 1995 (125)</td>
</tr>
<tr>
<td>DTBS</td>
<td>DAPA carbamate, ADP, Ca^{2+}</td>
<td>E. coli</td>
<td>1DAF</td>
<td>Huang et al. 1995 (125)</td>
</tr>
<tr>
<td>DTBS</td>
<td>3-(1-aminoethyl)nonanedioic acid</td>
<td>E. coli</td>
<td>1DAE</td>
<td>Huang et al. 1995 (125)</td>
</tr>
<tr>
<td>DTBS</td>
<td>ADP</td>
<td>E. coli</td>
<td>1DBS</td>
<td>Huang et al. 1995 (125)</td>
</tr>
<tr>
<td>DTBS</td>
<td>-</td>
<td>E. coli</td>
<td>1BYI</td>
<td>Sandalova et al. 1999 (129)</td>
</tr>
<tr>
<td>DTBS</td>
<td>DAPA, ATP, Mg^{2+}</td>
<td>E. coli</td>
<td>1A82</td>
<td>Kack et al. 1998 (124)</td>
</tr>
<tr>
<td>DTBS</td>
<td>ADP, DTP, Mg^{2+}, (PO_4)^2-</td>
<td>E. coli</td>
<td>1DAM</td>
<td>Kack et al. 1998 (124)</td>
</tr>
<tr>
<td>DTBS</td>
<td>ADP, Mg^{2+}</td>
<td>E. coli</td>
<td>1BS1</td>
<td>Kack et al. 1998 (124)</td>
</tr>
<tr>
<td>DTBS</td>
<td>DAPA carbamate with Aluminum F</td>
<td>E. coli</td>
<td>1DAK</td>
<td>Kack et al. 1998 (124)</td>
</tr>
<tr>
<td>DTBS</td>
<td>Cl</td>
<td>H. pylori</td>
<td>2QMO</td>
<td>Porebski et al. 2012 (128)</td>
</tr>
<tr>
<td>DTBS</td>
<td>GDP, (PO_4)^2-, Mg^{2+}</td>
<td>H. pylori</td>
<td>3QY0</td>
<td>Porebski et al. 2012 (128)</td>
</tr>
<tr>
<td>DTBS</td>
<td>Ethylene glycol</td>
<td>H. pylori</td>
<td>3QXX</td>
<td>Porebski et al. 2012 (128)</td>
</tr>
<tr>
<td>DTBS</td>
<td>8-aminocaproic acid, GDP Mg^{2+}, (PO_4)^2-, Ethylene glycol</td>
<td>H. pylori</td>
<td>3QXS</td>
<td>Porebski et al. 2012 (128)</td>
</tr>
<tr>
<td>DTBS</td>
<td>Adenyl Imidodiphosphate Mg^{2+}, HNO_3, Ethylene glycol</td>
<td>H. pylori</td>
<td>3QJ</td>
<td>Porebski et al. 2012 (128)</td>
</tr>
<tr>
<td>DTBS</td>
<td>GTP, Mg^{2+}, HNO_3, Ethylene glycol</td>
<td>H. pylori</td>
<td>3QXJ</td>
<td>Porebski et al. 2012 (128)</td>
</tr>
<tr>
<td>DTBS</td>
<td>8-aminocaproic acid, ADP Mg^{2+}, (PO_4)^2-, Ethylene glycol</td>
<td>H. pylori</td>
<td>3QXH</td>
<td>Porebski et al. 2012 (128)</td>
</tr>
<tr>
<td>DTBS</td>
<td>ATP, Mg^{2+}, HNO_3, Ethylene glycol, Glycerin</td>
<td>H. pylori</td>
<td>3QXH</td>
<td>Porebski et al. 2012 (128)</td>
</tr>
<tr>
<td>DTBS</td>
<td>(PO_4)^2+, Mg^{2+}, HNO_3, Cl^-</td>
<td>8-aminocaproic acid</td>
<td>H. pylori</td>
<td>3MLE</td>
</tr>
<tr>
<td>DTBS</td>
<td>Acetate ion, Na^+</td>
<td>F. tularensis</td>
<td>3OF5</td>
<td>Brunzelle et al. (unpublished)</td>
</tr>
<tr>
<td>BS</td>
<td>SAM, DTB, Tris buffer</td>
<td>E. coli</td>
<td>1R30</td>
<td>Berkovitch et al. 2004 (66)</td>
</tr>
</tbody>
</table>

26
1.9 Inhibitors of Biotin Biosynthetic Enzymes

There have been a number of studies investigating inhibitors of biotin biosynthetic enzymes. Most studies have reported on the *in vitro* characterization of compounds as enzyme inhibitors, and a few considered the antibacterial properties. Like other SAM-dependent methyltransferases, BioC is inhibited by chemical analogues of SAM substrate (10 in Figure 1.7). Demethylated SAM, namely *S*-adenosylhomocysteine (SAH) 11, is the product of the methyl transfer catalyzed by BioC. SAH inhibits BioC in a concentration-dependent manner in which 1 µM of SAH reduced ~40% of the *E. coli* BioC activity (96). Sinefungin 12, a natural antibiotic isolated from *Streptomyces griseolus* has also been shown to inhibit BioC (96,142). Sinefungin is a steric and electrostatic mimic of SAM but has more potent inhibitory activity than SAH (142). It reduced ~60% of the enzyme activity at 0.1 µM and completely abolished activity at 1 µM (96). The inhibition of KAPAS has also been pursued using chemical analogues of either substrate, reaction intermediate or products. The L-alanine substrate (13) analogue, namely L-trifluoroalanine 14 (Figure 1.8) is a slow suicide inhibitor of *E. coli* KAPAS ($t_{1/2} \sim 20$ min) that covalently binds to the active site lysine and forms an irreversible 2-(pyridoximine phosphate) acetoxy protein adduct (114). D-alanine 15, which is the enantiomer of the native L-alanine substrate, was found to competitively inhibit the *E. coli* KAPAS with a $K_i$ of 0.59 mM (64). In contrast, D-alanine is not a competitive inhibitor for *M. tuberculosis* KAPAS as the enzyme can utilize both L- and D-configurations (112). Interestingly, the product of the D-alanine utilized reaction, D-KAPA 16, was found to inhibit KAPAS ($K_i = 115$ µM) (112). Several analogues of the aldimine intermediate 17 (Figure 1.8), such as (±)-8-amino-7-oxo-8-phosphononanoic acid 18, 4-carboxybutyl-(1-amino-1-carboxyethyl)-phosphonate 19 and 2-amino-3-hydroxy-2-methylnonadioic acid 20 are competitive inhibitors of *E. coli* KAPAS with respect to L-alanine with $K_i$ values of 7, 68 and 80 mM respectively (Figure 1.8) (64). Amiclenomycin 21 (Figure 1.9), another suicide
inhibitor isolated from *Streptomyces sp.*, showed inhibitory activity against DAPAS from *E. coli* and *M. tuberculosis* \( (K_i = 2 \, \mu M \text{ and } 12 \, \mu M \text{ respectively}) \) (82-84,143-146). The crystal structures of *E. coli* DAPAS in complex with amicleomycin suggest that the inhibition of the enzyme is stereoselective as the cis-isomer \( 21 \) (PDB 1MLY) is a potent inhibitor whilst the trans-isomer (PDB 1MLZ) causes a steric hindrance at the active site that results in a significantly less potent inhibitor (84). The design of amicleomycin analogs revealed that the cis-configuration, but not the amino acid moiety, is essential for inhibitory activity (85,145). In addition to KAPAS as mentioned above, D-KAPA \( 22 \) (or \( (R)\)-KAPA) also showed inhibitory activity against *M. tuberculosis* DAPAS with a \( K_i \) of 5.9 \( \mu M \) (112,147). An 8-amino-7-oxooctanoic acid (also known as Desmethyl-KAPA \( 23 \), an achiral analog of the KAPA substrate lacking the methyl group) was an inhibitor with a \( K_i \) of 4.2 \( \mu M \) (147). Recently, the aryl hydrazine \( 24 \) has also been identified from a fragment library as a competitive inhibitor of *M. tuberculosis* DAPAS with respect to SAM \( (K_i = 10.4 \, \mu M \text{ (119)}) \). The crystal structure (PDB 4MQP) revealed that the aryl hydrazine forms a reversible covalent adduct with the PLP cofactor bound to DAPAS. To inhibit the DTBS activity, a series of DAPA, DAPA carbamate, and ATP mimics were rationally designed. For the purpose of herbicide development, a total of 54 compounds, such as \( 25 \) (Figure 1.10), were synthesized and tested for inhibitory activity against *E. coli* DTBS, but none showed submillimolar inhibition constants (148). In a separate study using available X-ray data of ATP bound to *E. coli* DTBS, a pharmacophore was proposed and employed to design an inhibitor that targeted the ATP binding pocket. Consequently, 6-hydroxypyrimidin-4(3H)-one (6-HP4)) \( 26 \) was developed and shown to have a \( K_i \) of 11 mM (149). Finally, a number of BS inhibitors have also been reported. Actithiazic acid \( 27 \) (Figure 1.11) (also known as acidomycin) isolated from *Actinomyces virginiae* and *Streptomyces sp.* inhibited BS from *E. coli* \( (K_i = 0.45 \, \mu M \text{ and } M. tuberculosis \text{ (no record of } K_i, \text{ see the antibacterial activity that} \)
will be discussed later) (150-153). α-methyldethiobiotin 28 and α-methylbiotin 29, isolated from Streptomyces lydicus, were also shown to inhibit the E. coli BS activity (only \( K_i = 1.1 \mu M \) for α-methyldethiobiotin was reported) (150,154,155). It can inhibit the growth of E. coli and B. subtilis unless supplemented with biotin in growth media (150).

The antibacterial activity for some of the inhibitors that target DAPAS and BS have been investigated against several strains of mycobacteria. Amiclenomycin and actithiazic acid inhibited growth of M. smegmatis with MIC of 12.5 and 0.4 µg/ml respectively, but failed to reduce the bacterial burden in a murine model of infection (83,146,150,152,153). Meanwhile, the α-methyldethiobiotin and α-methylbiotin were found to effectively inhibit M. fortuitum, M. smegmatis, M. avium, M. phlei, and M. salmoniphilum with MIC values of 0.8-80 µg/ml and 12.5-200 µg/ml respectively (154). The published inhibitors have a low inhibition activity, with inhibition constants in the micromolar to millimolar level. This suggests that the design strategies that have been pursued have not been successful in the discovery of new anti-TB agents. New approaches are now required.

![Chemical structures of substrate and inhibitors of BioC. A, S-adenosylmethionine (SAM) substrate. B, S-adenosylhomocysteine (SAH) product of BioC reaction. C, sinefungin.](image)

Figure 1.7 Chemical structures of substrate and inhibitors of BioC. A, S-adenosylmethionine (SAM) substrate. B, S-adenosylhomocysteine (SAH) product of BioC reaction. C, sinefungin.
Figure 1.8 Chemical structures of KAPAS inhibitors. A, L-alanine. B, L-trifluoroalanine. C, D-alanine. D, D-KAPA. E, the aldimine intermediate and its analogues; F, (±)-8-amino-7-oxo-8-phosphononanoic acid, G, 4-carboxybutyl (1-amino-1-carboxyethyl) phosphate, and H, 2-amino-3-hydroxy-2-methylnonadiolic acid. Abbreviation: Pyr, pyrimidine ring.

Figure 1.9 Chemical structures of DAPAS inhibitors. A, cis-amiclenomycin. B, D-KAPA (or (R)-KAPA). C, desmethyl-KAPA. D, aryl hydrazine
Figure 1.10 Chemical structure of DTBS inhibitor. A, phosphonic acid is a phosphate-based mimic of DAPA carbamate. B, 6-hydroxypyrimidin-4(3H)-one (also known as 6-HP4) is a minimal adenine replacement ligand that showed inhibitory activity against *E. coli* DTBS.

Figure 1.11 Chemical structures of BS inhibitors. A, actithiazic acid. B, α-methyldebiotin. C, α-methylbiotin.
1.10 Fragment-Based Drug Discovery

Fragment-based drug discovery (FBDD) has become a powerful approach over the past decade for early stage hit discovery (156-159). This technique aims to identify small starting structures that can be optimized into drug like compounds. FBDD has enjoyed success in drug discovery with one FDA approved drug (Zelboraf, also known as Vemurafenib) approved for melanoma treatment and more than 10 compounds in clinical trials for treating leukemia, myeloma, coronary artery disease, chronic obstructive pulmonary disease, diabetes, and bacterial skin infections (160-162). Conceptually, FBDD is the discovery of small fragments that can bind to specific target sites. Fragments are defined by a “Rule of Three” with low molecular weight < 300 Da, less than 3 hydrogen bond donors and acceptors and cLogP < 3 (163). Once fragment hits had been identified, larger lead molecules with higher affinity can be created by growing or modifying the chemical structure of one fragment or by linking or merging two adjacent fragments (164-167). Fragments are preferable as starting points for hit to lead development than larger (MW ≤ 500) drug-like compounds obtained from conventional high throughput screenings (HTS). Firstly, the low molecular mass often results in fragments that have high ligand efficiency, where the binding affinity is calculated per heavy atom of the ligand (167-170). Secondly, fragments libraries can be chemically diverse such that they can probe chemical space more effectively than larger compounds (160,171,172). Lastly, fragments with less complexity can bind to various sites of the protein target resulting in screening a fragment library generates a high rate of hit identification (160).

While FBDD has been emphasized in several fields of drug discovery, its application in the field of antibiotic discovery is still underutilized. The previous large-scale failure of hit identification from natural compound libraries, that have limited permeability through the mycobacterial cell envelope, restricts success in target-based HTS approach (173). Instead of
moving forward to target-based FBDD after a decade of HTS, the anti-TB drug discovery field mostly reversed to phenotypic whole cell screens (173-175). However, a recent study of synthetic anti-TB drugs and prodrugs suggested that many of these drugs, such as isoniazid, ethionamide, para-aminosalicylic acid, and pyrazinamide, are considered as small fragments (MW ≤ 300) that are reactive inside the cell to manifest the antimicrobial activity (176). Indeed, being a small fragment with moderately lipophilic (clogP < 3) seems to be a positive feature for penetration through the complex mycobacterial cell envelope (176). There is only one recent report of an inhibitor of \textit{M. tuberculosis} DAPAS identified by FBDD, namely aryl hydrazine (as discussed previously in section 1.8) (119). These successes suggest that fragment-based screening should be re-introduced in current anti-TB drug discovery efforts. Therefore, FBDD for screening inhibitors targeting specifically to various enzymes in \textit{M. tuberculosis} biotin biosynthesis is a novel approach to new anti-TB agents.

1.11 Selection of Enzyme Target

Of the six enzymes required for biotin biosynthesis in \textit{M. tuberculosis}, DTBS appears to be the most promising target for anti-TB drug development, and is the focus of this study. Genetic knockout studies have demonstrated DTBS is an essential enzyme for the growth and virulence of \textit{M. tuberculosis} (177). Several structures of DTBS are available from many species such as \textit{M. tuberculosis}, \textit{E. coli}, \textit{H. pylori}, and \textit{F. tularensis}, showing two adjacent binding pockets located in the active sites of homodimer DTBS that are ideally required for hits identification in fragment based drug discovery programs (125,128). These crystal structures are powerful tools for \textit{in silico} screening of the hits and also provide useful molecular details of the enzyme-ligand interaction for further lead optimization.
Although genetic studies have also confirmed the importance of genes encoding KAPAS, DAPAS and BS in the conserved pathway (178), I argue that these enzymes are less attractive as drug targets. As mentioned earlier, *M. tuberculosis* has two *BioF* genes, *BioF* and *BioF₂*. In fact, the structure and function of *BioF₂* encoding product has not yet been studied. However, amino acid sequence alignments revealed the existence of the KAPAS catalytic domain at the C-terminus of *BioF₂*. Having two isoforms provides tuberculi with a chance to avoid drug activity through target-based mutagenesis. Meanwhile, BS has extremely modest enzyme properties as shown by low turnover rates *in vivo* (138) resulting in it functioning more like a substrate than an enzyme as described previously. Therefore, high doses of an inhibitor would be required to completely inhibit the last step of biotin biosynthesis. DAPAS was initially considered as a target for this study. However, the *M. tuberculosis* DAPAS failed to over-express as a soluble protein in the *E. coli* host, causing the difficulty of protein handling, particular for further enzyme and crystallography studies.

### 1.12 Objectives

The overall goal of this study was to discover novel inhibitors for DTBS from *M. tuberculosis*. Due to its crucial roles in the biotin biosynthetic pathway as described previously, DTBS catalyzing the penultimate step of the pathway has been selected as a target for the inhibitors screening. This study can be divided into three aims.

The first aim was to establish enabling technologies for DTBS that are more expedient than other published assays. Due to the lack of facile, homogenous assays suitable for high-throughput screening (HTS) applications, previous attempts to study DTBS have been restricted. Thus, three new assays were developed, namely a DTBS enzyme assay, a
competitive ATP-binding assay, and surface plasmon resonance (SPR) binding analysis (Chapter 3, Assay development).

The second aim was to characterize the biochemical and structural properties of *M. tuberculosis* DTBS (*MtDTBS*). In particular, the molecular details of nucleotide triphosphate (NTP) substrate binding are less well defined due to the lack of a NTP-bound X-ray crystal structure. Comparisons of the bacterial DTBS orthologs from *M. tuberculosis*, *E. coli*, *H. pylori*, and *F. tularensis* revealed considerable structural differences in nucleoside recognition at the C-terminal region of DTBS (128). While a conformational change of *E. coli* DTBS (*EcDTBS*) was observed following ATP binding (130), the homologous conformational change is not observed upon the binding of (PO₄)⁻ in *MtDTBS* (62). In order to explore the NTP binding pocket as a drug target, it is necessary to first obtain a better understanding of this site (described in Chapter 4).

The final aim of this study is to identify hits for *MtDTBS* using fragment-based screening in combination with virtual screening. Based on the novel finding in Chapter 4, CTP analogues were also introduced for screening inhibitors specific to *MtDTBS*. A total of 93,904 fragments from the ZINC database and 57 CTP analogues were applied for initial virtual hits identification. Binding of the hits to the DTBS was validated by SPR, DTBS enzyme assay and docking studies (Chapter 5).
CHAPTER 2

GENERAL MATERIALS AND METHODS
2.1 Materials

2.1.1 Chemical reagents

All chemicals and reagents were of analytical grade or of the highest purity available.

<table>
<thead>
<tr>
<th>Reagents</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 kb and 100 bp DNA ladders</td>
<td>New England Biolabs, MA, USA</td>
</tr>
<tr>
<td>Bradford protein reagent concentrate</td>
<td>Bio-Rad Laboratories Inc., CA, USA</td>
</tr>
<tr>
<td>T4 DNA ligase</td>
<td>New England Biolabs, MA, USA</td>
</tr>
<tr>
<td>VENT® DNA polymerase</td>
<td>New England Biolabs, MA, USA</td>
</tr>
<tr>
<td>Phusion® DNA polymerase</td>
<td>New England Biolabs, MA, USA</td>
</tr>
<tr>
<td>Pfu DNA polymerase</td>
<td>New England Biolabs, MA, USA</td>
</tr>
<tr>
<td>Phenylmethanesulfonyl fluoride (PMSF)</td>
<td>Sigma-Aldrich, MO, USA</td>
</tr>
<tr>
<td>Dithiothreitol (DTT)</td>
<td>Sigma-Aldrich, MO, USA</td>
</tr>
<tr>
<td>Isopropyl β-D-1-thiogalactopyranoside (IPTG)</td>
<td>BioVecra, PE, USA</td>
</tr>
<tr>
<td>MOPS/MES SDS running buffer</td>
<td>Invitrogen Life Technologies Inc., NY, USA</td>
</tr>
<tr>
<td>Restriction endonucleases</td>
<td>New England Biolabs, MA, USA</td>
</tr>
<tr>
<td>SeeBlue® Plus2 Prestained Protein Marker</td>
<td>Invitrogen Life Technologies Inc., NY, USA</td>
</tr>
<tr>
<td>GelRED™ Nucleic acid gel stain</td>
<td>BioVecra, PE, USA</td>
</tr>
<tr>
<td>BigDye (version 3) reaction mix</td>
<td>Sigma-Aldrich, MO, USA</td>
</tr>
<tr>
<td>Streptavidin from Streptomyces avidinii</td>
<td>Biotium Inc., CA, USA</td>
</tr>
<tr>
<td>7,8-diaminopelargonic acid (DAPA)</td>
<td>Toronto Research Chemicals, Inc., CANADA</td>
</tr>
<tr>
<td>DL-Desthiobiotin (DTB)</td>
<td>Sigma-Aldrich, MO, USA</td>
</tr>
<tr>
<td>Fluorescein-biotin</td>
<td>Anaspec Inc., CA, USA</td>
</tr>
<tr>
<td>Ampicillin</td>
<td>Sigma-Aldrich, MO, USA</td>
</tr>
<tr>
<td>Cytidine 5′-triphosphate disodium salt (CTP)</td>
<td>Sigma-Aldrich, MO, USA</td>
</tr>
<tr>
<td>Guanosine 5′-triphosphate sodium salt hydrate (GTP)</td>
<td>Sigma-Aldrich, MO, USA</td>
</tr>
<tr>
<td>Uridine 5′-triphosphate trisodium salt hydrate (UTP)</td>
<td>Sigma-Aldrich, MO, USA</td>
</tr>
<tr>
<td>Thymidine 5′-triphosphate sodium salt (TTP)</td>
<td>Sigma-Aldrich, MO, USA</td>
</tr>
<tr>
<td>Adenosine 5′-diphosphate sodium salt (ADP)</td>
<td>Sigma-Aldrich, MO, USA</td>
</tr>
<tr>
<td>Adenosine 5′-triphosphate sodium salt (ATP)</td>
<td>Sigma-Aldrich, MO, USA</td>
</tr>
<tr>
<td>Inosine 5′-triphosphate sodium salt (ITP)</td>
<td>Sigma-Aldrich, MO, USA</td>
</tr>
</tbody>
</table>
2.1.2 General materials

<table>
<thead>
<tr>
<th>Materials</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amicon® centrifugal filter devices (10000 MWCO)</td>
<td>Millipore, MA, USA</td>
</tr>
<tr>
<td>Minisart syringe filter 0.2 µm, 0.45 µm or 0.80 µm</td>
<td>Sartorius, Goettingen, Germany</td>
</tr>
<tr>
<td>PVDF membrane (Hybond P)</td>
<td>GE Healthcare, Buckinghamshire, England</td>
</tr>
<tr>
<td>CelluSepT1 MWCO 3500 dialysis tube</td>
<td>Adelab Scientific, Australia</td>
</tr>
<tr>
<td>Standard dialysis tubing closure, 35 mm width</td>
<td>Spectrum Laboratories, Inc., CA, USA</td>
</tr>
<tr>
<td>NuPage® 4-12% Bis-Tris polyacrylamide gels</td>
<td>Invitrogen, CA, USA</td>
</tr>
<tr>
<td>0.1 µm VVPP Durapore® membrane filter</td>
<td>Millipore, MA, USA</td>
</tr>
<tr>
<td>Mixed Bed Resin</td>
<td>Sigma-Aldrich, MO, USA</td>
</tr>
</tbody>
</table>

2.1.3 Antibodies

Antibodies were reconstituted and stored as the manufacturers’ instructions. Antibodies were used at the manufacturers’ recommended dilutions.

<table>
<thead>
<tr>
<th>Reagents</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>6xHis Monoclonal Antibody (Albumin Free)</td>
<td>Clontech Laboratories, Inc., CA, USA</td>
</tr>
<tr>
<td>Cy3® 3-conjugated AffiniPure Donkey Anti-mouse IgG</td>
<td>Jackson ImmunoResearch Laboratories, PA, USA</td>
</tr>
</tbody>
</table>

2.1.4 Commercial kits

<table>
<thead>
<tr>
<th>Kits</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>QIAprep Miniprep kit</td>
<td>QIAGEN, GmbH, Germany</td>
</tr>
<tr>
<td>QIAquick Gel Extraction kit</td>
<td>QIAGEN, GmbH, Germany</td>
</tr>
<tr>
<td>QIAquick PCR Purification kit</td>
<td>QIAGEN, GmbH, Germany</td>
</tr>
<tr>
<td>QuickChange™ Site-Directed Mutagenesis kit</td>
<td>Stratagene, CA, USA</td>
</tr>
</tbody>
</table>
2.1.5 Sequencing oligonucleotides

Oligonucleotides were purchased from Geneworks Pty Ltd., Hindmarsh, South Australia. All primers used were of sequencing grade.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence 5’ - 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>T7-26mer Forward</td>
<td>CGA AAT TAA TAC GAC TCA CTA TAG GG</td>
</tr>
<tr>
<td>T7-23mer Reverse</td>
<td>CAA GAA TTC TCA TGT TTG ACA GC</td>
</tr>
</tbody>
</table>

2.1.6 Bacterial strains

*E. coli DH5α*: *supE44ΔlacU169*(p80lacZΔM15) *hsdR*17 *recA1* *endA1* *gyrA96* *thi-1* *relA1* host for recombinant plasmids (New England, Biolabs, CA, USA).

*E. coli BL21(DE3)*: *fhuA2* [lon] *ompT* *gal* (*λ* DE3) [dcm] *ΔhsdS*

*λ* DE3 = *λ* *sBamH*Io *ΔEcoRI-B int* (*lacI::PlacUV5::T7* gene1) *i21 Δnin5* for expression of recombinant proteins (Stratagene, La Jolla, CA, USA).

2.1.7 Bacterial media

**Luria Broth (LB):** 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% NaCl, adjusted to pH 7.0 with 5N NaOH.

**LB agar:** LB supplemented with 1.5% (w/v) bacto-agar.

Ampicillin was added to liquid and solid media at 100 µg/mL for selection of bacteria bearing plasmid.
2.1.8 Plasmids

Plasmids obtained for use in this study:

pMA-His6MtBioD (GeneArt® gene synthesis, Life Technologies Australia Pty Ltd.) was synthesized using codon optimization in the purpose of expression in E. coli host.

pMA-His6EcBioD (GeneArt® gene synthesis, Life Technologies Australia Pty Ltd.) was synthesized according to nucleotide sequence obtained from NCBI database.

pET16b (Novagen): 5711 bp cloning and bacterial expression vector, used to generate fusions of a known protein with N-terminal-His$_6$ tag.

2.1.9 Buffers and solutions

Blocking solution (for Western Blotting): 1% BSA (w/v) in PBS containing 0.05% (v/v) Tween20.

SDS-PAGE Coomassie Blue staining solution: 0.2% (w/v) Coomassie brilliant blue, 10% (v/v) Methanol, 10% (v/v) Acetic acid.

SDS-PAGE destaining solution: 10% (v/v) Methanol, 10% (v/v) Acetic acid.

Loading buffer (for DNA) 6X: 0.5x TBE, 40% (v/v) glycerol, 1 mg/ml bromophenol blue.

PBS: 0.137 M NaCl, 2.7 mM KCl, 1.46 mM KH$_2$PO$_4$, 8.1 mM Na$_2$HPO$_4$ (pH 7.4).

PBS-Tween: PBS, 0.05% (v/v) Tween20.

SDS-PAGE sample buffer (for proteins) 5X: 0.25 M Tris (pH 6.8), 10% (w/v) SDS, 0.5% (w/v) bromophenol blue, 50% (w/v) glycerol.
**TBE:** 216 g Trizma base, 100 g boric acid, 18.6 g EDTA to 1 L with MilliQ water

**TBS:** 25 mM Tris pH 7.5, 150 mM NaCl.

**TE:** 10 mM Tris pH 7.5, 1 mM EDTA.

**Towbin western transfer buffer:** 20 mM Tris-HCl pH 7.5, 1.15 M Glycine, 20% Methanol, 0.1% (w/v) SDS.

**Storage buffer (for purified soluble proteins) 1X:** 25 mM Tris buffer (pH 7.5), 30 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol (DTT) and 5% (v/v) glycerol

**Denaturing buffer (for insoluble proteins) 1X:** 8 M Urea (pH 8.0), 100 mM Tris, 300 mM NaCl and 5 mM imidazole

### 2.1.10 Computer software

Data were analyzed using Graphpad Prism 6 (Graphpad Software, Inc., CA, USA). ApE-A plasmid Editor version 1.17 (by M. Wayne Davis) was used for sequence alignment. USCF Chimera version 1.8.1 (UCSF, CA, USA) was used for viewing and analysis of PDB files. AutoDock-Vina version 1.5.4 (UCSF, CA, USA) was used for *in silico* docking. ChemDraw® version 10.0 (PerkinElmer, Inc., MA, USA) was used for preparing chemical structures.

### 2.1.11 Web resources

NCBI (http://www.ncbi.nlm.nih.gov/) was used to access protein, nucleotide, and PubMed databases, the BLAST (Basic Local Alignment Search Tool) database and the
PubChem compound database. Multiple sequence alignments were performed using ClustalW (http://www.ebi.ac.uk/clustalW/). An Online SMILES Translator and Structure File Generator Software (http://cactus.nci.nih.gov/services/translate/) was used for generating 3D structures of compounds.

2.2 General methods

2.2.1 Plasmid DNA extraction

For purification of small (<10 µg) amounts of plasmid DNA, the QIAGEN QIAprep Miniprep kit was employed following the manufacturer’s instructions. For transformation and sequencing quality DNA, purified material from the kits was precipitated in 70% isopropanol, pelleted by centrifugation at 17 500 x g for 30 min and then washed in 70% (v/v) ethanol before resuspension in sterile MilliQ water. DNA was quantified by measuring absorbance at 260 nm using NANODrop 2000 Spectrophotometer (ThermoFisher Scientific Pty Ltd, Australia).

2.2.2 Agarose gel electrophoresis

Analysis of DNA and separation of DNA fragments was performed using agarose gel electrophoresis. Gel slabs were prepared by melting 0.7% (w/v) agarose in 1X TBE buffer. Prior to loading into wells, DNA samples were mixed with an appropriate volume of 6X DNA loading buffer. Samples were electrophoresed in 1X TBE buffer at 100 - 150 V and then stained in 1X GelRed™ Nucleic acid gel stain (Biotium Inc., CA, USA) for 30 min followed by destaining in distilled water briefly. DNA was visualized on a UV transilluminator and photographed using a Mitsubishi Video Processor.
2.2.3 Restriction digestion of DNA

A 1-5 µg of DNA was digested with 1-10 units of restriction enzyme in the appropriate NEB buffer for 1 hour at 37 °C. For cloning, DNA fragments were separated by agarose gel electrophoresis before purification from the excised gel slice using a QIAquick Gel Extraction kit.

2.2.4 Ligation of DNA fragments

Ligation reactions were carried out in a 10 µl reaction volume with an insert:vector molar ratio of 3:1 in 1X ligase buffer and 2 units of T4 DNA ligase for 1 hr at room temperature.

2.2.5 Site-directed mutagenesis

The QuickChange™ mutagenesis kit (Stratagene) was employed for site-directed mutagenesis according to the manufacturer’s recommendation. Complementary strands of synthetic oligonucleotide primers containing the desired mutation were extended by the activity of Pfu DNA polymerase during thermo cycling. The 50 µl reaction mixture was comprised of 1x reaction buffer, 50 ng DNA template, 1 mM dNTPs, 2.5 ng/µl each primer and 5 units of Pfu DNA polymerase. The reaction was carried out on a MJ Research PTC 2000 Thermal Cycler (GMI Inc., MN, USA) with the initial denaturation (94 °C) for 5 min followed by 16-20 cycles of denaturation (94 °C) for 1 min, annealing (50-60 °C) for 1 min, and extension (68 °C) for 9 min. Following the amplification, products were then treated with DpnI (target sequence: 5’-Gm6ATC-3’) to digest the methylated parental DNA template. The
digested vectors incorporating the desired mutation were then transformed into *E. coli* strain DH5α competent cells for amplification.

### 2.2.6 Transformation of recombinant plasmids by heat-shocked method

#### 2.2.6.1 Preparation of competent cells by the CaCl₂ method

A single colony of *E. coli* was inoculated into 2 ml of LB broth and incubated for overnight at 37 °C with rotation. The overnight culture was diluted 1:50 in LB broth and incubated at 37 °C with shaking until OD₆₀₀ 0.5-0.6. The culture was chilled on ice for 30 min prior to centrifugation at 3,200 x g for 10 min at 4 °C. The pellet was resuspended in 10 ml of ice-cold 0.1 M CaCl₂ and placed on ice for 30 min. After centrifugation at 3,200 x g for 10 min at 4 °C, the pellet was resuspended in 10 ml of ice-cold 0.1 M CaCl₂ and incubated on ice for 30 min before being pelleted as described above. The cell pellet was resuspended in 2 ml of ice-cold 0.1 M CaCl₂ for each 50 ml of original culture. Finally, glycerol was added to the cell suspension at 40% final concentration and stored in aliquots of the competent cells (100 μl) at -80 °C until needed.

#### 2.2.6.2 Transformation of plasmids into competent cells

An aliquot of 100 μl *E. coli* competent cells was gently mixed with 20 ng of DNA (or 5 μl of ligation mixture). Cells were placed on ice for 30 min, incubated at 42 °C for 90 sec then placed on ice for an additional 5 min. The transformed cells were mixed with 900 μl of LB broth, incubated at 37 °C for 45 min with rotation and then centrifuged at 2 656 x g for 1 min. The pellet was resuspend in 250 μl media and spread on LB agar plates containing ampicillin (100 μg/ml) and incubated overnight at 37 °C.
2.2.7 Glycerol stocks

For long-term storage of plasmids, an overnight culture of the *E. coli* harbouring the recombinant plasmid was prepared at 37 °C. An equal volumes of the overnight culture and 80% glycerol were mixed and stored at -80 °C.

2.2.8 Sequencing

Plasmid DNA or PCR products were used as templates for DNA sequencing. A 20 µl reaction containing 4 µl DNA, 1 µl 100 ng of appropriate primer, and 1 µl BigDye™ (version 3) reaction mix (Perkin Elmer, Applied Biosystems, CA, USA), 4 µl 5X buffer and 10 µl autoclaved MilliQ water was prepared. The PCR profile consisted of 30 cycles of denaturation at 96 °C for 30 sec, annealing at 50 °C for 15 sec and extension at 60 °C for 4 min. After thermocycling, 80 µl of 75% (v/v) isopropanol was added to the PCR products, vortexed and left at room temperature for 30 min. Precipitated DNA was isolated by centrifugation at 17 500 x g for 20 min. The pellet was washed twice in 75% isopropanol followed by centrifugation for 5 min at 17 500 x g. The pellet was then dried in a 37 °C heating block prior to being submitted to the Gene Sequencing Service, SA Pathology (Adelaide, Australia).

2.2.9 Expression of His$_{6}$-tagged proteins

A single colony of *E. coli* strain BL21 (DE3) harbouring the recombinant plasmid was inoculated in 2 ml of LB broth containing 100 µg/ml ampicillin and grown overnight at 37 °C. The overnight culture was diluted 1:50 in the appropriate volume of LB broth containing 100 µg/ml ampicillin at 37 °C with shaking at 250 rpm until OD$_{600}$ 0.5-0.6. IPTG
was added to give the final concentration of 0.1 mM and incubated at either 25 °C for overnight or 37 °C for 3 hr depending on the protein being investigated. The cells can be kept as a pellet at -80 °C until needed. Cells were pelleted at 4 355 x g for 10 min and discarded the supernatant. Whole cell lysates were analyzed by SDS-PAGE to confirm protein expression.

2.2.10 Purification of His<sub>6</sub>-tagged proteins using Profinia<sup>TM</sup>

Purification of His<sub>6</sub>-tagged protein was performed using the Profinia<sup>TM</sup> Protein Purification System (BIO-RAD) as specified by the user manual. A 5 ml Bio-Scale<sup>TM</sup> Mini Profinity<sup>™</sup> IMAC cartridge was used together with a 50 ml Bio-Scale<sup>™</sup> Mini Bio-Gel® P-6 desalting cartridge (BIO-RAD). Two alternative purification protocols were employed depending upon whether the target protein was present in the soluble or insoluble fractions.

2.2.10.1 Purification of soluble His<sub>6</sub>-tagged proteins

Cell pellets were resuspended in Native IMAC lysis buffer or wash buffer 1 (300 mM KCl, 50 mM potassium phosphate buffer (pH 8.0), 5 mM imidazole) supplemented with 1 mM PMSF. Cells were lysed using a Microfluidics Cell Disruptor at 18 000 psi and debris removed by centrifugation at 48 384 x g for 30 min. His-tagged proteins were purified by immobilized nickel affinity chromatography using the Profinia® IMAC protocol following manufacturers instructions (BIO-RAD). The protein sample was loaded onto the column and first washed with wash buffer 1, followed by wash buffer 2 (300 mM KCl, 50 mM potassium phosphate buffer (pH 8.0), 40 mM imidazole). The purified protein was eluted in 50 mM potassium phosphate buffer (pH 8.0), 300 mM KCl and 250 mM imidazole and was subsequently buffer exchanged into storage buffer (25 mM Tris buffer (pH 7.5), 30 mM
NaCl, 1 mM EDTA, 1 mM dithiothreitol (DTT) and 5% (v/v) glycerol) using the desalting cartridge.

2.2.10.2 Purification of insoluble His<sub>6</sub>-tagged proteins

Lysed cells were centrifuged at 48,384 x g for 30 min to harvest insoluble inclusion bodies. The pellet was then resuspended in denaturing buffer (8 M Urea (pH 8.0), 100 mM Tris, 300 mM NaCl and 5 mM imidazole) that had been deionized using the Mixed Bed Resin (Sigma Aldrich®) following the manufacturers’ protocol. The protein solution was then centrifuged at 48,384 x g for 30 min to discard the cell debris. The urea solubilized fraction was collected and fractionated using the Profinia<sup>®</sup> Denaturing IMAC protocol following the manufacturer’s instruction (BIO-RAD). The proteins were eluted in 8 M Urea (pH 8.0), 100 mM Tris, 300 mM NaCl and 250 mM imidazole, then diluted to 0.2 mg/ml for refolding.

2.2.11 Refolding of recombinant protein by dialysis

A series of dialysis steps was performed in order to refold the insoluble protein. CelluSepT1 MWCO 3500 dialysis tubing (Adelab Scientific, Australia) was first prepared by soaking in MilliQ water at room temperature for 30 min to remove the preservative and then rinsed thoroughly in MilliQ water before use. The protein sample was diluted to the concentration of 0.2 mg/ml in 8 M Urea (pH 8.0), 100 mM Tris, 300 mM NaCl and 5 mM imidazole. Samples were dialyzed against four changes of the 100-fold (v/v) storage buffer (25 mM Tris buffer (pH 7.5), 30 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol (DTT) and 5% (v/v) glycerol) at 4 °C for 48 hr.
2.2.12 Concentration of the proteins

Concentration of protein solutions was performed using Amicon® Ultra-50 or -15 centrifugal filter devices (10000 MWCO) (Millipore, MA, USA) following manufacturer’s instruction manual. The columns were rinsed with MilliQ water and then equilibrated in storage buffer by centrifugation at 5000 x g at 4 °C for 30 min or until reaching a required retentate volume. The storage buffer was discarded prior adding protein sample into the spin-column. Likewise, protein was concentrated by centrifugation at 5000 x g at 4 °C until reaching a required retentate volume. For retentate recovery, the concentrate was collected using a pipette with 200 microliter tip to new pre-cold microcentrifuge tube. The protein was kept at -80 °C until needed. For storage of the Amicon® Ultra-50 or -15 centrifugal filter devices, the centrifuge tube was washed with distilled water to remove residual buffer components and kept in MilliQ water at 4 °C.

2.2.13 Determination of protein concentration

Protein concentration was assayed using the Bradford Reagent (Bio-Rad Laboratories Inc., CA, USA) assay method. A standard curve of bovine serum albumin (BSA) was generated from 0 to 1 mg/ml. A 10 µl of sample was mixed with 200 µl of Bradford reagent in a 96 well plate (Falcon). Absorbance at 620 nm wavelength was measured on a microplate reader (Molecular Devices, CA, USA). Standard curves were generated and linear regression used to calculate protein concentration using Graphpad Prism6.
2.2.14 SDS-PAGE electrophoresis and gel staining

Protein samples were diluted in Protein sample buffer to a final concentration of 1X. Samples were then boiled for 5 min, and centrifuged briefly to collect condensation from the top of the tube. The protein samples were fractioned on NuPage® 4-12% Bis-Tris polyacrylamide gels (Invitrogen) using 1X NuPAGE® MES running Buffer (Invitrogen) at 200 V for approximately 40 min or until the dye front reached the bottom of the gel. The protein bands were visualized using SDS-PAGE Coomassie Blue staining solution. The gel was first soaked in staining solution (0.2% Coomassie Blue, 50% ethanol and 10% acetic acid) at room temperature for 1 hr, before soaking in SDS-PAGE destaining solution (10% acetic acid and 5% methanol) overnight.

2.2.15 Western blotting

Proteins fractioned by PAGE were transferred onto a PVDF membrane using a semi-dry transfer unit (Hoefer SemiPhor, Amersham Pharmacia Biotech, CA, USA). Six sheets of Whatman filter paper and the PVDF membrane were pre-soaked in Towbin western transfer buffer prior to assembly of the 3:1:1:3 paper:gel:membrane:paper sandwich. Proteins were transferred for 1 hr at 80 mA per gel. The membrane was then blocked in 1% (w/v) skim milk blocking buffer for 1 hr at room temperature or at 4 °C overnight. The membrane was then washed three times with PBS-Tween before being probed with a 6xHis Monoclonal Antibody (diluted in 1:10000) for 1 hr at room temperature. The membrane was then washed three times in PBS-Tween before addition of a CyTM 3-conjugated AffiniPure Donkey Anti-mouse IgG (diluted in 1:5000) for 1 hr at room temperature. Finally the membrane was washed three times with PBS-Tween before being visualized using a Typhoon TRIO Variable Mode Imager.
2.2.16 Size exclusion chromatography – multi-angle light scattering analysis

Size exclusion chromatography – multi-angle light scattering (SEC-MALS) analysis was performed to estimate the molecular weight and stoichiometry of proteins. Purified protein samples at ≥ 1 mg/ml concentration were fractionated on a Superdex\textsuperscript{TM} 200 10/300 GL gel filtration column (GE Healthcare) connected in series with a miniDAWN TREOS light-scattering detector (Wyatt Technology) and an OPTILAB rEX interferometric refractometer detector (Wyatt Technology). Analytical size-exclusion chromatography was performed at 20° C using a mobile phase containing 25 mM Tris pH 7.5, 30 mM NaCl and 1 mM EDTA at a flow rate of 0.25 mL/min. Data analysis was performed using ASTRA V version 5.0.1.0 (Wyatt Technology) based on the manufacturer’s instruction. Bovine serum albumin (BSA) was employed as the reference sample during calibration.

2.2.17 Liquid chromatography-electrospray ionization tandem mass spectrometry

Liquid chromatography-electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) was carried to confirm the identity of proteins. LC-ESI-MS/MS was acquired using an 1100 series HPLC system (Agilent Technologies) coupled via an Advance CaptiveSpray source (Michrom Bioresources, Inc.) to an amaZon ETD mass spectrometer (Bruker Daltonics). Samples were initially run on a 4-12% gradient SDS-PAGE gel and stained with coomassie blue dye. Bands were excised from the gel manually and processed with a trypsin digestion protocol. Collision-induced dissociation (CID) spectra were acquired to select the two most abundant ionisable species in the 300-2000 m/z range at any point in the LC separation. The acquired spectra were subjected to peak detection and de-convolution using DataAnalysis version 4.0 SP4 Build 281 (Bruker Daltonics). Processed MS/MS spectra were exported to Mascot generic format (mgf) and submitted to Mascot version 2.3.02 for
identification by comparison to the amino acid sequences of known proteins and common contaminants (i.e. human keratins and porcine trypsin) from Uniprot.

2.2.18 MALDI mass spectrometry

MALDI mass spectrometry (MALDI-MS) was carried out to confirm the molecular weight in order to distinguish the origin of the proteins. MALDI-MS was acquired on an ultrafleXtreme MALDI-TOF/TOF mass spectrometer (Bruker Daltonics) operating in a linear positive mode under the control of flexControl software version 3.4 (Bruker Daltonik GmbH). Samples were desalted by spotting onto nitrocellulose filters 0.025 µm MF-membrane filters VSWP0500 (Millipore) and dialyzing in 800 ml of 0.1% TFA for 2 hr. One microliter of each sample was mixed with 1 µl of sinapinic acid saturated TA30 (30% ACN in 0.1% FA). Next, the sample matrix was prepared by spotting the 1 µl sample mix onto the dried matrix spots which were previously spotted with 2 µl of sinapinic acid saturated ethanol onto a ground steel target plate (Bruker Daltonics) and then air dried. Sample m/z range was set to 5,000 – 55,000 Da. 50,000 shots were collected for the external calibration and sample measurement. External calibration was performed using a mix of protein calibration standard I and II (Bruker Daltonics). Laser intensity and detector gain was manually adjusted for optimal resolution. The MS spectra obtained were analysed using flexAnalysis software version 3.3 (Bruker Daltonics) employing smoothing, background subtraction and peak detection algorithms. All mass spectrometry services were performed at the Adelaide Proteomics Center, The University of Adelaide.
CHAPTER 3

ASSAY DEVELOPMENT
3.1 Introduction

There are a number of published assays available for measuring DTBS activity. However, attempts to study DTBS are still restricted due to the lack of facile, homogenous assays suitable for high-throughput screening (HTS) applications. A spectrophotometric assay is the most commonly reported method used to monitor ADP product formation from DTBS activity through the disappearance of NADH in a coupled assay using pyruvate kinase and lactate dehydrogenase (130,179). Using an additional two enzymes in the assay is not optimal for HTS. An alternative radio-labeled carbon fixation assay has also been reported to detect the formation of acid-stable $[^{14}\text{C}]$DTB product in reaction supplemented with $[^{14}\text{C}]$CO$_2$ substrate (179). However, complicated procedures are required to remove excess $[^{14}\text{C}]$CO$_2$ prior to the $[^{14}\text{C}]$DTB product detection. Recently, a continuous fluorescent assay was developed to detect the coupled activity of two continuous enzymes in the biotin biosynthesis pathway, DAPAS and DTBS (as mentioned previously in Chapter 1.6). This relies on the displacement of a DTB-fluorescent probe from streptavidin by the DTB product (180). Although, it is possible to adapt this continuous fluorescent assay for detecting the DTBS activity, the fluorescent probe is not commercially available and must be synthesized via at least four synthetic steps (180,181).

To overcome these restrictions, the development of new facile assays was the first aim for this study. Three new assays were investigated, namely an enzyme assay, a competitive ATP-binding assay and Surface plasmon resonance (SPR) analysis. The DTBS enzyme and the ATP-binding assays were developed using fluorescence polarization (FP) based technology. FP is a light based technique that provides several advantages; 1. it is homogenous and does not require coupling enzymes, 2. high sensitivity, 3. fast and simple detection, 4. accurate quantitative measurement, 5. amenable to HTS, and 6. uses widely available instrumentation (182). Additionally, an alternative SPR binding method was also
pursued as it encompasses many advantages such as label-free detection and real-time monitoring (183). Together, these enabling technologies are potentially suitable for HTS and have utility in characterization of DTBS and in drug discovery efforts (as examined in chapter 4 and 5 respectively).

3.2 Specific Methods

3.2.1 Cloning, Expression and Purification of DTBS Enzymes

The *bioD* gene from *M. tuberculosis* strain H37Rv (Gene ID: 886338), containing an N-terminal hexa-histidine tag, was synthesized as a codon optimized gene for expression in *E. coli* and ligated into the pMA vector (GENEART® Gene Synthesis, Life Technologies Australia Pty Ltd.). Similarly, the *bioD* gene from *E. coli* strain K-12 (Gene ID: 945387) containing an N-terminal hexa-histidine tag was synthesized and inserted into the pMA vector (GENEART® Gene Synthesis, Life Technologies Australia Pty Ltd). The coding regions were then subcloned into *NcoI* and *HindIII* treated pET16b vector (Novagen) (See Chapter 2, General Materials and Methods). Both constructs were verified using DNA sequencing by the Gene Sequencing Service, SA Pathology (Adelaide, Australia).

DTBS proteins were recombinantly expressed in *E. coli* strain BL21(DE3) (Invitrogen). The overnight cultured cells were used to inoculate LB medium containing 0.1 mg ml⁻¹ ampicillin and grown at 37 °C until early log phase when protein expression was induced with 0.1 mM IPTG for 18 hr at 25 °C for *MtDTBS* (or for 3 hours at 37 °C for *EcDTBS*). Cells were harvested by centrifuging at 4,648 x g for 15 min and resuspended in 300 mM KCl, 50 mM potassium phosphate buffer (pH 8.0), 5 mM imidazole and 1 mM PMSF. Cells were lysed using a Microfluidics Cell Disruptor at 18,000 psi and debris
removed by centrifuging at 48,384 x g for 30 min. His-tagged DTBS proteins were purified by immobilized nickel affinity chromatography using the Profinia® IMAC protocol following manufacturers instructions (BIO-RAD). After elution in 50 mM potassium phosphate buffer (pH 8.0), 300 mM KCl and 250 mM imidazole, the proteins were subsequently buffer exchanged into storage buffer (25 mM Tris buffer (pH 7.5), 30 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol (DTT) and 5% (v/v) glycerol) using a 50 ml Bio-Scale™ Mini Bio-Gel® P-6 desalting cartridge (BIO-RAD). DTBS proteins were concentrated using Amicon Ultra-15 centrifugal filter devices (Millipore), followed by centrifugation at 15,700 x g for 15 min to discard protein aggregates. Protein concentrations were determined using a Bradford Protein Assay (BioRad) using bovine serum albumin (Sigma Aldrich®) as a standard. The purity of proteins was examined using SDS-PAGE gel electrophoresis. The identity of proteins was confirmed using western blot analysis, liquid chromatography-electron spray ionization tandem mass spectrophotometry (LC-ESI-MS/MS), and size exclusion chromatography – multi-angle light scattering (SEC-MALS) analysis (See Chapter 2, General Materials and Methods).

3.2.2 Enzyme Assay

The DTBS assay consists of three discrete steps; initiation of the DTBS reaction, termination of the reaction, and FP based detection. All experiments were performed in black 96 well microtiter plates (Costar®) preblocked with 1% casein dissolved in Tris-buffered saline. After blocking and drying, the plates were stored at 4 °C. The DTBS reactions were performed at 37 °C in FP buffer (50 mM Tris (pH 7.5), 150 mM NaCl, 5 mM MgCl₂, and 10 mM NaHCO₃ (62)) supplemented with saturating concentrations of 0.3 mM ATP and 0.1 mM DAPA (Toronto Research Chemicals Inc.). For determining the inhibition activity of
compounds, varying concentration of test compounds was added to the DTBS reaction mixture. The reactions were initiated with the addition of 0.25 – 3.0 µM of DTBS enzyme depending upon species or mutant enzyme employed. The reactions were terminated at 45 minutes by addition of 50 mM Tris (pH 7.5), 150 mM NaCl, 250 mM EDTA and then incubated for 15 min at 37 °C. The synthesized DTB product was detected by adding 50 mM Tris (pH 7.5), 150 mM NaCl buffer, and 20 nM streptavidin and incubating for 15 min at 37 °C, followed by adding 50 mM Tris (pH 7.5), 150 mM NaCl buffer containing 25 nM fluorescein biotin (Anaspec Inc.). Finally, the mixtures was incubated at 37 °C for 15 min before the fluorescence polarization value was measured by a PerkinElmer 2030 Plate Reader using polarization mode set to 485 nm excitation and 535 nm emission (184). FP is defined as the difference of the emission light intensity parallel ($I_\parallel$) and perpendicular ($I_\perp$) to the excitation light plane normalized by the total fluorescence emission intensity (Equation 1) (185,186). The FP is commonly represented as millipolarization (mP).

Equation 1:  
$$\text{FP} = \frac{(I_\parallel - I_\perp)}{(I_\parallel + I_\perp)}$$

The concentration of enzymatically generated DTB was calculated against a standard curve using commercially available DL-Desthiobiotin (Sigma Aldrich®). A standard curve of DTB concentration versus FP units was plotted using Equation 2; a linear regression binding model where $S$ is the slope of the standard curve, $Y$ is the Y intercept and $\Delta mP$ is the corrected FP value (ie subtract reading without streptavidin from the reading with streptavidin).

Equation 2:  
$$\log [\text{DTB (M)}] = - \frac{(\Delta mP + Y)}{S}$$
To calculate $V_{\text{max}}$ and $K_m$, the enzyme velocity measurements were fitted to the Michaelis–Menten equation against varying concentrations of substrate using GraphPad Prism 6 (GraphPad Software, Inc.).

The concentration of inhibitor that reduced DTBS activity by 50% ($IC_{50}$) was obtained by plotting the percentage of enzyme activity, correlated to 100% of enzyme activity in the control reaction, versus the log of compound concentrations. The data was fitted to a log (inhibitor) versus response – variable slope (four parameters) model in GraphPad Prism 6 (GraphPad Software, Inc.). The control reaction was performed without compounds in the assay. The inhibition constant ($K_i$) of the test compounds was calculated using Equation 3 (187), where $[S]$ is the substrate concentration, $K_m$ is the Michaelis-Menten constant, and $IC_{50}$ is the half maximal inhibitory concentration.

Equation 3: $K_i = IC_{50} / (1 + [S]/K_m)$

3.2.3 ATP-Binding Assay

The binding of a fluorescently-labeled ATP tracer, ATP-BODiPY (Life Technologies™), to DTBS was performed in black 96 well microtiter plates (Costar®) preblocked with 1% casein dissolved in Tris-buffered saline. The reactions were performed in binding buffer containing 100 µM DTBS enzyme, 50 mM Tris (pH 7.5), 150 mM NaCl, and 5 mM MgCl$_2$. Varying concentrations of MgNTPs were included and the reaction pre-equilibrated for 15 min before the addition of 25 nM ATP-BODiPY. The reaction was performed at 37 °C for 15 min and fluorescence polarization measured as described in 3.2.2.

The $IC_{50}$ values of tested compounds were calculated by plotting the corrected FP values ($\Delta mP = \text{subtract reading without DTB from the reading with DTBS}$) against log of
MgNTP concentration and fitting the data to a log (inhibitor) versus response – variable slope (four parameters) model in GraphPad Prism 6 (GraphPad Software, Inc.). The control reaction was performed in the absence of DTBS in binding buffer supplemented with 25 nM ATP-BODiPY and a constant concentration of ATP substrate and/or test compounds.

3.2.4 SPR-Binding Assay

The kinetics of the interaction between DTBS and compounds was analyzed using a BIAcore T100. DTBS enzymes were immobilized onto the surface of a CM5 sensor chip using amide coupling chemistry. After activating the surface with EDC and NHS, DTBS solution at a concentration of 0.2 mg ml\(^{-1}\) in 0.01 M NaOAc buffer (pH 5.2) was applied at a constant flow rate of 5 µl min\(^{-1}\) for 420 s. Approximately 10,000 response units (RU) of ligands were immobilized. Typically, no immobilized ligand (empty flow cell) and immobilized bovine serum albumin (BSA) were performed alongside immobilized DTBS in order to correct for bulk refractive index changes with buffer and distinguish non-specific binding events. Water-soluble compounds were diluted in running buffer (10 mM HEPES (pH 7.4), 150 mM NaCl and 0.005% (v/v) surfactant P20). For compounds that were poorly water soluble, samples were initially dissolved in DMSO and diluted in running buffer such that the final concentration of DMSO was 2.5%. To correct for variations in DMSO concentration during the preparation of these compounds, a solvent correction curve was included in the analysis by preparing a series of test solutions between 2% to 3.3% DMSO. Binding experiments were performed by injecting the analyte solutions into the instrument across the sensor surface of all flow cells at a flow rate of 30 µl min\(^{-1}\) with a contact time of 60 s followed by a dissociation time of 60 s (as a default setting of a wizard template). The time-dependent binding curves from all four flow cells were monitored simultaneously.
As analytes bound to the immobilized DTBS, the refractive index at the surface alters in proportion to the change in mass, resulting in a change of RU value. The binding capacity of the surface depends on the level and activity of immobilized ligand. The maximum binding capacity (Rmax) of the immobilized ligand was calculated using Equation 4 where MW is the molecular weight of the ligand and analyte, RL is the amount of immobilized ligand in RU, and Sm is the stoichiometry as defined by the number of binding sites on the ligand (188). The percent activity of the immobilized ligand that reflects the quality of binding was determined using Equation 5 (188).

Equation 4: \[ R_{\text{max}} = \frac{\text{MW}_{\text{analyte}}}{\text{MW}_{\text{ligand}}} \times RL \times Sm \]

Equation 5: \[ \% \text{ ligand activity} = \left( \frac{R_{\text{max,experiment}}}{R_{\text{max,theory}}} \right) \times 100 \]

The binding affinity (K_D) was determined by transforming the time-dependent binding curves into an affinity-steady state 1:1 model using BIAcore T100 evaluation software (GE Healthcare).

3.3 Results and Discussion

3.3.1 Purification of DTBS Proteins

DTBS proteins were recombinantly produced in *E. coli* using the pET-16b expression system containing a hexa-histidine tag at the N-terminus of the protein. This His_6-tag facilitated protein purification using immobilized nickel affinity chromatography. *EcDTBS* was expressed as a soluble protein at 37°C for 3 hr, whereas the optimal conditions for the *M. tuberculosis* equivalent was overnight at 25°C. SDS-PAGE analysis revealed a high purity (>95%) of the IMAC purified material, showing the band with the expected molecular mass of 23.5 kDa for *MtDTBS* when 50 mM potassium phosphate buffer (pH 8.0), 300 mM KCl and
40 mM imidazole was used in the second wash step to eliminate the contamination of other protein species (Figure 3.1A, Lane 9 and Figure 3.1B, Lane 1). Likewise, SDS-PAGE analysis of the IMAC purified EcDTBS revealed the protein was > 95% pure and migrated on SDS-PAGE gel at the expected molecular mass of 25.1 kDa (Figure 3.1B, Lane 2). Furthermore, these purified proteins were enzymatically active with specific activity of 250 - 300 µmol min\(^{-1}\) mg\(^{-1}\) (as will be discussed in 3.3.2, Enzyme assay). SEC-MALS analysis revealed that > 98% of the total proteins eluted at a molecular mass consistent with the expected homodimer (Figure 3.2A&B). Notably, the percentage of the eluted protein was calculated from the proportion of the area under the peaks on the chromatograms using ImageJ 1.4r (National Institute of Health, Maryland, USA). LC-ESI-MS/MS sequencing confirmed the existence of MtDTBS and EcDTBS in purified materials (Figure 3.3). Importantly, the endogenous \textit{E. coli} protein was not detected in the MtDTBS preparation implying that the recombinant protein did not form heterodimers with the host protein.
**Figure 3.1 SDS-PAGE analysis of DTBS enzymes.**

*A*, IMAC purification of *MtDTBS*; lane 1, non-induced crude extract; lane 2, crude extract induced with 0.1 mM IPTG; lane 3, soluble fraction; lane 4, flow through fraction; lane 5, wash 1 (5 mM imidazole); lane 6, wash 2 (20 mM imidazole); lane 7, purified *MtDTBS* after 20 mM imidazole wash; lane 8, wash 2 (40 mM imidazole); lane 9, purified *MtDTBS* after 40 mM imidazole wash.

*B*, The purified *MtDTBS* (lane 1) and *EcDTBS* (lane 2) migrated through the gel with a size consistent with their molecular mass of 23.5 kDa and 25 kDa, respectively.
Figure 3.2 SEC-MALS analysis of DTBS. The active forms of (A) MtDTBS and (B) EcDTBS were eluted with the calculated molecular mass of 47 and 46 kDa, representing the dimers. Its molecular weight was examined by multi-angle light scattering (Wyatt Technology Corporation) calibrated by (C) BSA as a reference standard representing in three forms ie monomer, dimer, and trimer.
**Figure 3.3 LC-ESI-MS/MS analysis of DTBS.** The various peptides obtained from a trypsin digestion show 47% of sequence coverage for *Mt*DTBS or 41% of that for *Ec*DTBS (as highlighted in yellow). LC-ESI-MS/MS was acquired using an 1100 series HPLC system (Agilent Technologies) coupled via an Advance CaptiveSpray source (Michrom Bioresources, Inc.) to an amaZon ETD mass spectrometer (Bruker Daltonics).
3.3.2 Enzyme Assay

The activity of DTBS was determined by quantitating the amount of synthesized dethiobiotin (DTB) product using FP based technology. Here, the large molecular mass of the fluorescently-labeled biotin:streptavidin complex impedes the rotation of the biotin tracer out of the plane of polarized light used for excitation. This provides a convenient way to measure the complex formation. DTB (chemical structure shown in Figure 1.1 and Figure 3.4B) binds reversibly to streptavidin (189) and can displace the binding of fluorescein biotin (Fl-biotin) (Anaspec Inc) (chemical structure shown in Figure 3.4C). Hence, high concentrations of DTB cause more displacement of the tracer thereby reducing the formation of the complex, leading to the lower FP value (Figure 3.5).

![Chemical structures of Biotin and its derivatives](image)

**Figure 3.4 Chemical structures of Biotin and its derivatives.** A, Biotin. B, Dethiobiotin (DTB). C, Fluorescein biotin (Fl-biotin). DTB is the precursor to biotin during biotin biosynthesis via the closure of its thiophane heterocycle. Fl-biotin is a fluorescently-labeled form of biotin commonly used in avidin/or streptavidin binding studies.
Figure 3.5 FP concept for enzyme assay. A, Without DTB, the Fl-biotin:streptavidin complex impedes the rotation of the small Fl-biotin (shown as the star biotin) out of the plane of polarized light resulting in a high FP value. B, The displacement of Fl-biotin by DTB reduces the formation of the complex, leading to lower FP values.

Development of the DTBS enzyme assay required the optimization of several parameters including the concentration of streptavidin and Fl-biotin tracer in the detection step, the concentration of enzyme, and the reaction time. The rotation of Fl-biotin out of polarized light reached the stable state when its concentration is $\geq 25$ nM, giving FP value of $\sim 50$ mP (Figure 3.6A). Meanwhile, 20 nM was the minimum concentration of streptavidin required for $> 90\%$ of the tracer to bind, resulting in the maximum FP value of $\sim 150$ mP (Figure 3.6B). Therefore, 25 nM Fl-biotin and 20 nM streptavidin were used in the detection step of all DTBS assays. The concentration of DTBS in the reaction varied depending on its species of origin such that DTB produced in the reaction was 0.5 - 20 µM. This enzymatically produced DTB was determined using a standard curve employing commercially available DTB (Figure 3.7). Indeed, the 1 - 20 µM DTB was aligned in a linear range of DTB concentrations that could be accurately detected. All DTBS reactions were performed for 45
min when the enzyme was working with steady state velocity and produced 20 µM DTB product (ie first 20% of the reaction), which was within the linear phase of the detection.

**Figure 3.6 Desired Concentration of FB and streptavidin correlated to the FP value.** *A*, Calibration curve for fluorescein biotin. *B*, Sigmoidal dose-response curve showing 25 nM Fl-biotin bound to streptavidin (0.032 nM – 1 µM). Optimal concentrations employed for subsequent experiments are shown with red arrows.

**Figure 3.7 The dynamic range of DTB detected by enzyme assay.** *A*, Sigmoidal dose-response curve using DL-Desthiobiotin, supplemented with 25 nM fluorescein biotin and 20 nM streptavidin in FP based detection. Insert, A linear range of DL-Desthiobiotin used as the standard curve for calculating the DTB product from enzyme assay. Experiments were performed in triplicate. The mean ± SEM are shown.
To further verify the accuracy of the DTBS assay, the effects of the concentration of DAPA substrate and DTBS enzyme on the FP based detection system were examined. Increasing the amount of DAPA (0.01 nM - 1 mM) in the assay revealed that DAPA did not displace the tracer from streptavidin (Figure 3.8A). Likewise, titrating increasing amounts of MtDTBS (0 – 2 µM) with Fl-biotin showed no binding of the tracer to the enzyme (Figure 3.8B). Together these data demonstrate the specificity of the detection system developed here.

![Graph A](image1)  ![Graph B](image2)

**Figure 3.8 Controls of FP based detection in DTBS enzyme assay.** *A*, DAPA did not displace the fluorescein biotin from streptavidin in the assay containing 25 nM fluorescein-biotin and 20 nM streptavidin. *B*, Fluorescein biotin did not bind to DTBS, resulting in readily bind to streptavidin in the assay.

The DTBS assay was employed to determine the kinetic parameters of DAPA, MgATP, and NaHCO₃ binding by varying the concentration of one substrate whilst maintaining two substrates at saturating concentrations. The $K_m$ values for DTBS from *M. tuberculosis* were determined to be $30.2 \pm 1.7 \mu M$ for DAPA, $30.2 \pm 1.7 \mu M$ for ATP, and $3.1 \pm 0.4 \text{ mM}$ for NaHCO₃ (Figure 3.9A, B, C). Noteworthy, the $K_m$ value of ATP obtained in this study is in good agreement with the published value (29 µM) using the coupled spectrophotometric assay (62). Meanwhile, the $K_m$ values for DTBS from *E. coli* were
measured as 15.2 ± 1.7 µM for DAPA, 10.5 ± 0.8 µM for ATP, and 0.6 ± 0.1 mM for NaHCO₃ (Figure 3.9D, E, F). Here, the $K_m$ for ATP in MtDTBS is 3-fold higher than that of EcDTBS. This was comparable to the 4-fold difference found in the previous study in which the published $K_m$ for ATP in EcDTBS was 7 µM (62,190). These reproducible kinetic data indicated the accuracy of determining the activity of DTBS using the enzyme assay developed in this study.

![Figure 3.9 Kinetic analysis of substrates for DTBS enzymes.](image)

The lines represent the nonlinear regression to the Michaelis-Menten equation. The $K_m$ values of MtDTBS (A, B, C) and EcDTBS (D, E, F) were obtained from three independent experiments.
The DTBS assay was adapted to investigate the inhibitory activity of certain compounds such as an adenosine diphosphate (ADP). In contrast to the ATP control, ADP was not a substrate for the DTBS enzymes as it did not produce DTB (Figure 3.10). To measure the inhibition activity of ADP, the enzyme assays were then performed in three discrete steps where a desired concentration of compounds was supplied to the DTBS reaction mixture obtaining a saturating concentration of DAPA and ATP (at 0.1 and 0.3 mM respectively). The result showed dose-responsive inhibition of MtDTBS by ADP with a $K_i$ value of 0.14 mM (Figure 3.11). This was in good agreement with a published study of an inhibition of DTBS activity by ADP using a radioactive-labeled carbon fixation assay (191). Indeed, ADP was found to be a competitive inhibitor for EcDTBS with respect to ATP ($K_i = 0.23$ mM). In summary, the kinetic analysis of ligand binding and characterization of the inhibitory activity of ADP demonstrated the suitability of this assay for further studies.

![Figure 3.10 Kinetic analysis of ADP versus ATP for MtDTBS.](image)

The lines represent the nonlinear regression to the Michaelis-Menten equation. MtDTBS obtained very low velocity in the enzyme reaction supplemented with ADP (shown in ■). In contrast, the $K_m$ value of ATP control (shown in •) was shown to be 33.4 µM. Likewise, ADP was not a substrate for EcDTBS (data not shown).
Figure 3.11 Inhibition of MtDTBS activity by ADP. The inhibition activity was calculated as the percentage of enzyme activity correlated to 100% of total enzyme activity without ADP in the reaction containing a saturating concentration of 0.1 mM DAPA, 10 mM NaHCO$_3$, and 0.3 mM ATP. Determined $IC_{50}$ value of ADP was 1.55 mM. $K_i$ value was calculated to be 0.14 mM with respect to ATP substrate.

Notably, there is a potential limitation of the enzyme assay for characterizing test compounds that can only be solubilized in dimethyl sulfoxide (DMSO) solvent as the presence of DMSO reduces the enzyme activity. DMSO at 2.5% (v/v) was found to be the highest concentration that could be included in the assay and still maintain greater than 90% MtDTBS activity (91.6% ± 4.7 versus 100%, p=0.26) (Figure 3.12). This may restrict the analysis of certain compounds, especially those with low affinity that requires high concentration in the assay. Beyond this limitation the new facile enzyme assay is a useful tool to study DTBS, particularly the characterization of DTBS mutants and screening of enzyme inhibitors. These topics will be discussed in Chapters 4 and 5, respectively.
Figure 3.12 Effect of DMSO to DTBS enzyme activity. The 2.5% is the highest concentration of DMSO that could maintain 91.6% ± 4.7 activity of MtDTBS in inhibition DTBS enzyme assay.

3.3.3 ATP-binding Assay

A competitive ATP-binding assay was developed to probe the binding of ligands in the nucleotide pocket of DTBS using FP based technology. Here, changes in FP were measured upon the binding of fluorescently-labeled ATP tracers to DTBS (Figure 3.13). The ATP tracers investigated here were ATP-BODiPY (Life Technologies™) and TNP-ATP (Life Technologies™). The chemical structures for these reagents are shown in Figure 3.14. ATP-BoDiPY at 25 nM and TNP-ATP at 0.04 mM were the minimum concentrations that reached the stable state of the rotation out of the plane of polarized light with FP value ~50 mP and ~300 mP respectively (Figure 3.15). The ATP-BoDiPY assay used 1,600 times less tracer than TNP-ATP to reach the stable state and, thus, was used further in all ATP-binding assays. DTBS enzyme at 0.1 mM was the optimal concentration providing maximum response for the detection of the DTBS:ATP-BODiPY complex (Figure 3.16). The binding of ATP-BODiPY to DTBS was inhibited by ATP with an $IC_{50}$ value of 349 ± 15 µM (Figure 3.17A). Likewise, ADP also inhibited tracer binding with similar potency ($IC_{50} = 249$ µM) (Figure 3.17B). These competition experiments indicated that the ATP tracer bound
specifically to the nucleotide binding site in DTBS. The binding of ATP tracer was also found to be independent of DAPA substrate. Adding increasing amounts of DAPA into the assay mixture did not alter formation of DTBS:ATP-BODiPY complex (Figure 3.18). Therefore, the ATP-binding assay was performed in the absence of the DAPA substrate.

The ATP-binding assay is complementary to the DTBS enzyme assay for determining the binding affinity of ligands and potential inhibitors described in Chapter 4 and 5. The selectivity for various NTPs was also investigated for the first time using this binding assay (Chapter 4).

Figure 3.13 FP concept for DTBS ATP-binding assay. A, Without a competitive inhibitor, the DTBS binds the ATP tracer, thereby impeding its rotation in solution (shown as the star ATP) out of the plane of polarized light resulting in a high FP value. B, The displacement of the ATP tracer by inhibitors reduces the formation of the complex, leading to a lower FP value.
Figure 3.14 Structure of ATP tracers. *A*, ATP-BODiPY is composed of a BODiPY FL fluorophore attached to the 2′ or 3′ position of the ribose ring via a linker (Life Technologies™). *B*, TNP-ATP is a modified ATP analog with an attachment of a trinitophenyl fluorophore to the 2′ and 3′ position of the ribose ring.

Figure 3.15 Desired concentration of ATP tracers. Calibration curve of (*A*) ATP-BODiPY or (*B*) TNP-ATP versus the FP value. The minimum concentration of ATP-BODiPY and TNP-ATP that showed a saturating level of fluorescence polarization was 25 nM and 0.04 mM, respectively.
Figure 3.16 Desired concentration of DTBS enzymes. Dose response of the formation of DTBS:ATP-BODiPY complex revealed that a 0.1 mM was the minimum concentration of MtDTBS showing the positive response in the complex formation with a desired concentration of ATP-BODiPY (at 25 nM).

Figure 3.17 Competitive binding of ATP and ADP. The bindings of (A) ATP substrate and (B) ADP to DTBS were detected by the assay using constant 25 nM ATP-BoDiPY and 0.1 mM MtDTBS with IC$_{50}$ values of 349 ± 15 µM (n=3) and 248 µM (n=1) respectively.
Figure 3.18 Effect of DAPA on ATP-BoDiPY binding to DTBS. The formation of DTBS:ATP-BoDiPY complex was not disrupted by increasing concentrations of DAPA in the binding reaction, implying that DTBS does not require DAPA for nucleotide binding.

3.3.4 Surface Plasmon Resonance (SPR)-Binding Assay

Surface plasmon resonance (SPR) has become a powerful tool for studying bimolecular binding kinetics in real time (as examined in Chapter 4 and 5). In this example, the SPR approach involves attaching proteins (referred as ligands) onto the surface of a sensor chip and then passing small molecules (referred as analytes) over the sensor surface, generating a response that is proportional to the bound mass at the sensor surface (192). There are several approaches for covalently attaching ligands to the sensor surface, however, an amine coupling method was used here. The carboxyl groups on the surface of the CM5 sensor chip were first activated with a mixture of EDC and NHS to give reactive succinimide esters. Ligands were then covalently linked with the dextran matrix through the reaction of amine groups with the esters on the activated surface (Figure 3.19) (192).
Figure 3.19 Amide coupling of DTBS to the CM5 sensor chip surface. The carboxyl dextran on the surface of the gold layer on sensor chip was activated with the mixture of EDC/NHS to give succinimide esters. The DTBS ligand was then immobilized randomly on the surface by the reaction occurred between its amide groups and the esters.

According to the binding experiments of known substrates for MtDTBS at a saturating concentration of 0.3 mM, DAPA showed a greater binding response (~50 RU) than ATP (~30 RU) (Figure 3.20A). The low response of ATP binding was reflected by the negative response of bulk refractive index change with ligand-solubilizing buffer different from the SPR running buffer, implying the underestimation of the binding response (Figure 3.20C). This consistently occurred in all ATP binding experiments may be caused by the need to adjust the pH to 7.0 during the preparation of MgATP resulting in a negative refraction index change. Since the maximum binding capacity (Rmax) of the immobilized ligand could be measured accurately in the range of 50 to 150 RU (188), the experimental Rmax obtained from the DAPA (MW = 261) binding was then used to assess the percent activity of the immobilized ligand. Typically ~10000 resonance units (RU) of MtDTBS (MW_{dimer} = 47000 and Sm_{dimer} = 2) was immobilized onto the sensor chips, leading to the theoretical Rmax of ~110. As a result, 45% of immobilized ligand on sensor chip was considered active.
To investigate the control analytes of SPR-binding assay, both DAPA and ATP were tested in binding experiments. Apart from the positive binding response to MtDTBS, DAPA also showed an unexpected response to a BSA control that was included for testing non-specific binding (Figure 3.20B). This could be explained as BSA is known to bind multiple ligands (193,194) and is therefore capable of binding small molecules like DAPA. In contrast, ATP bound only to MtDTBS and there was relatively low binding or no binding response to either the BSA control or the empty flow cell, thereby confirming the specificity of its binding. The protein was receptive to ATP binding for 4 days (Figure 3.21), suggesting that once immobilized, the enzyme is sufficiently stable to perform screening within this period of time. Notably, it is still unclear why the sensograms of ligand binding fell below the starting level of baseline after the dissociation in some binding experiments such as Day 1 in Figure 3.21. In addition, the binding affinity ($K_D$) of ATP to MtDTBS was calculated to be $335 \pm 96 \text{µM}$ using steady state analysis of binding (Figure 3.22). This was in good agreement with DTBS ATP-binding assay as the $IC_{50}$ value was $349 \pm 15 \text{µM}$, indicating a consistency of both binding assays developed in this study. All together, ATP was therefore used as a control analyte for all DTBS SPR-binding experiments and as an indicator for the stability of immobilized ligands.

SPR analysis is another important method for determining the binding affinity of ligands and potential inhibitors as examined in Chapter 4 and 5. The $K_D$ values obtained from SPR analysis well supports the binding data ($IC_{50}$ values) obtained from the competitive ATP-binding assay. Noteworthy, the detection system of the SPR technique is sensitive and it can detect the binding of DMSO-solubilized compounds. Instead of the competitive ATP-binding assay, SPR is therefore a better technique to detect the binding of weak inhibitors /or DMSO-solubilized compounds (as described in Chapter 5).
Figure 3.20 Surface plasmon resonance analysis. The sensograms represent the binding mode of DAPA (shown in red) and ATP (shown in green) to (A) MtDTBS, (B) BSA and (C) the empty flow cell. Notably, the sensograms shown the ligands binding to MtDTBS and BSA are illustrated by the subtraction of bulk refractive index changes with buffer (from the empty flow cell). The experiment was performed at a saturating concentration of analytes at 0.3 mM.
**Figure 3.21 Lifetime of immobilized MtDTBS.** The sensograms represent the binding mode of DAPA (shown in red) and ATP (shown in green) to MtDTBS from Day 1 to Day 4. The binding ability of MtDTBS to ATP was decreased over a period of time and was completely abolished on Day 4 while there was no effect to DAPA.
Figure 3.22 Dose-response SPR-binding analysis. A, Time-dependent binding curves of ATP correlated to dose-response. The concentrations of ATP are labeled in different colors as shown in box. B, Binding analysis based on affinity-steady state 1:1 model. The binding affinity ($K_D$) of ATP to $Mt$DTBS was calculated to be $335 \pm 96 \mu$M (n=3).
CHAPTER 4

NUCLEOTIDE TRIPHOSPHATE PROMISCUITY IN

MYCOBACTERIUM TUBERCULOSIS DETHIOBIOTIN
SYNTHETASE
# Statement of Authorship

<table>
<thead>
<tr>
<th>Title of Paper</th>
<th>Nucleotide triphosphate promiscuity in <em>Mycobacterium tuberculosis</em> deoxyribonucleotide synthetase.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publication Status</td>
<td>○ Published ○ Accepted for Publication ○ Submitted for Publication ○ Publication Style</td>
</tr>
</tbody>
</table>

## Author Contributions

By signing the Statement of Authorship, each author certifies that their stated contribution to the publication is accurate and that permission is granted for the publication to be included in the candidate's thesis.

<table>
<thead>
<tr>
<th>Name of Principal Author (Candidate)</th>
<th>Wanisa Salaemae</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contribution to the Paper</td>
<td>Performed DNA cloning, protein expression, protein purification, site-directed mutagenesis, assays development, SEC-MALS analysis, kinetic analysis, surface plasmon resonance experiments and competitive ATP-binding analysis. Primary role in preparation of the manuscript.</td>
</tr>
<tr>
<td>Signature</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>24/10/2014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name of Co-Author</th>
<th>Min Y. Yap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contribution to the Paper</td>
<td>Solved crystal structure of MtDTBS in complex with CTP.</td>
</tr>
<tr>
<td>Signature</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>24/10/2014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name of Co-Author</th>
<th>Kate L. Wegener</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contribution to the Paper</td>
<td>Provided direction in all aspects of structural analysis and interpretation of the results.</td>
</tr>
<tr>
<td>Signature</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>30/10/2014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name of Co-Author</th>
<th>Grant W. Booker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contribution to the Paper</td>
<td>Provided direction and discussion in all aspects of protein chemistry and biochemical assays.</td>
</tr>
<tr>
<td>Signature</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>30/10/2014</td>
</tr>
<tr>
<td>Name of Co-Author</td>
<td>Matthew C.J. Wilke</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Contribution to the Paper</td>
<td>Provided direction in all aspects of X-ray crystallography.</td>
</tr>
<tr>
<td>Signature</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>27/10/2014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name of Co-Author</th>
<th>Steven W. Polyrak</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contribution to the Paper</td>
<td>Provided project direction and corresponding author on manuscript.</td>
</tr>
<tr>
<td>Signature</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>27/10/2014</td>
</tr>
</tbody>
</table>
Nucleotide Triphosphate Promiscuity in *Mycobacterium tuberculosis* Dethiobiotin Synthetase*

Wanisa Salaemae‡, Min Y. Yap§, Kate L. Wegener‡, Grant W. Booker‡, Matthew C.J. Wilce§, and Steven W. Polyak†

†School of Molecular and Biomedical Science, The University of Adelaide, South Australia, Australia 5005

§Department of Biochemistry and Molecular Biology, School of Biomedical Science, Monash University, Victoria, Australia 3800

*Running title: Promiscuous NTP binding to mycobacterial Dethiobiotin synthetase

To whom correspondence may be addressed: Department of Biochemistry, School of Molecular and Biomedical Science, The University of Adelaide, South Australia, Australia 5005, Tel.: +61 8 8313 6042; E-mail: steven.polyak@adelaide.edu.au

Keywords: enzyme, biotin, Biotin biosynthesis; dethiobiotin synthetase; nucleotide triphosphate; *Mycobacterium tuberculosis*; *Escherichia coli*; crystal structure

**Background:** Biotin biosynthesis is a virulence factor for mycobacteria.

**Results:** Dethiobiotin synthetase (DTBS) from *Mycobacterium tuberculosis* has unexpected broad NTP specificity.

**Conclusion:** The absence of an asparagine at 169 results in the broad NTP specificity of DTBS in tuberculi.

**Significance:** The broad specificity provides tuberculi with the potential to utilize alternative nucleotides depending upon availability.

**ABSTRACT**

Dethiobiotin synthetase (DTBS§) plays a crucial role in biotin biosynthesis in microorganisms, fungi, and plants. Due to its importance in bacterial pathogenesis, and the absence of a human ortholog, DTBS is a promising target for the development of new antibacterials desperately needed to combat antibiotic resistance. Here we report the first X-ray structure of DTBS from *Mycobacterium tuberculosis* (MtDTBS) bound to a nucleotide triphosphate (CTP). The nucleoside base is stabilized in its pocket through hydrogen-bonding interactions with the protein backbone, rather than amino acid side chains. This resulted in the unexpected finding that MtDTBS could utilise ATP, CTP, GTP, ITP, TTP, or UTP with similar *K*<sub>m</sub> and *k*<sub>cat</sub> values, although the enzyme had the highest affinity for CTP in competitive binding and surface plasmon resonance assays. This is in contrast to other DTBS orthologs that preferentially bind ATP primarily through hydrogen-bonds between the purine base and the carboxamide side chain of a key asparagine. Mutational analysis performed alongside *in silico* experiments revealed a gate-keeper role for Asn175 in *Escherichia coli* DTBS that excludes binding of other nucleotide triphosphates. Here we provide evidence to show that MtDTBS has the broad specificity of nucleotide due to the absence of that gate-keeper.

*De novo* synthesis of biotin (aka vitamin H or B7) is an important metabolic activity for microorganisms, fungi and plants, which is not present in humans (1). In bacteria, biotin serves as an essential cofactor for biotin-dependent enzymes, namely pyruvate carboxylase (PC) and acyl-CoA carboxylase (ACC) (2-5). As the cell membrane provides a defense barrier against environmental toxins, host immune factors and antibiotic agents (6,7), the metabolic pathways related to the synthesis of membrane lipids have been suggested as promising targets for the development of new antibiotics. *De novo* biotin
Dethiobiotin synthetase (DTBS, EC 6.3.3.3, encoded by \textit{bioD}) controls the penultimate step of the biotin biosynthesis pathway, catalysing the closure of the ureido ring of dethiobiotin (DTB') from (7R, 8S)-7,8-diaminonoanoic acid (DAPA') using CO$_2$ and a nucleotide triphosphate (NTP). The reaction mechanism has been proposed to involve three discrete steps, which are (1) the formation of N7-carbamate, (2) the formation of the ceramic phosphoric acid anhydride, and (3) the closure of the ureido ring with the release of inorganic phosphate (13). The amino acid residues required for binding the DAPA substrate and subsequent catalysis have been well characterized, with X-ray crystal structures available for the enzymes from \textit{Escherichia coli}, \textit{M. tuberculosis}, \textit{Helicobacter pylori}, and \textit{Francisella tularensis} (14-17). However, the molecular details of NTP binding are less well defined, especially for \textit{M. tuberculosis} DTBS (\textit{Mt}DTBS) where there is currently no crystal structure of the enzyme in complex with NTP. From the available structural data, the NTP binding site is divided into two discrete subsites; the phosphate-binding region (also known as the P-loop or Walker A motif) and the nucleoside-binding pocket. In \textit{E. coli} DTBS (\textit{Ec}DTBS), a conformational change involving the P-loop (Gly8-X-X-X-X-Gly14-Lys15-Th16) was observed following phosphate binding (15). In contrast, a comparison of available \textit{Mt}DTBS structures revealed no homologous conformational change upon the binding of inorganic phosphate (14). Comparisons of the bacterial DTBS orthologs revealed considerable structural differences in nucleoside recognition at the C-terminal region of DTBS (17). For example, Asn175 in \textit{Ec}DTBS found within the [NQ]-[KR]-x-[DE] GTP specificity motif has been proposed to play a key role in the adenosine specificity observed for \textit{E. coli}, \textit{H. pylori} and \textit{F. tularensis}. However this amino acid is absent in \textit{Mt}DTBS, which instead contains a glycine at the equivalent position (Gly 169). As nucleotide analogs have proved a successful route to new drugs, there is the potential to likewise target the nucleotide pocket of DTBS for new antibiotics. Narrow spectrum therapeutics with anti-tuberculi activity would be especially welcome in the fight against tuberculosis. However, this requires a greater understanding of the nucleotide binding pockets of bacterial DTBS enzymes and the molecular basis of NTP binding.

Previous attempts to characterize DTBS have been restricted by the lack of facile assays. DTBS activity was commonly measured with a spectrophotometric assay by monitoring ADP production through the coupling of pyruvate kinase and lactate dehydrogenase (15,18). Using two additional enzymes in the reaction is less than desirable for high-throughput screening applications. An alternative radioactive-labeled carbon fixation assay has been reported for detecting the formation of the acid-stable [$^{14}$C]DTB product using [$^{14}$C]CO$_2$ (18). However, complicated procedures are required to remove excess [$^{14}$C]CO$_2$ from the reaction. Recently, a continuous fluorescent assay was developed to detect the coupled action of DAPAS and DTBS, based on the displacement of a dethiobiotin fluorescent probe bound to streptavidin by the DTB product (19). Although, it is possible to adapt this continuous fluorescent assay for detecting the DTBS activity alone, the fluorescent probe is not commercially available and must be synthesized via at least four steps (19,20).

In this study, we successfully developed two facile, homogeneous assays to study DTBS. These enabling technologies will be beneficial for future drug discovery efforts. Here the assays were employed to characterize DTBS from \textit{E. coli} and \textit{M. tuberculosis}. The specificity of NTP binding and utilization was investigated. Stark differences were observed between the two enzymes with \textit{Ec}DTBS having highest preference for ATP whereas \textit{Mt}DTBS showed equal potency for a broad panel of NTPs. A competitive ATP binding
assay together with surface plasmon resonance analysis showed that CTP bound to the mycobacterium ortholog with greatest affinity. For the first time, the structure of MdDTBS was resolved in complex with CTP. X-ray crystallography in combination with mutagenesis studies revealed a role for N175 in EcDTBS in discriminatory binding of NTPs.

**EXPERIMENTAL PROCEDURES**

**Plasmid Constructions** – The bioD genes encoding DTBS from *M. tuberculosis* strain H37Rv (Gene ID: 886338) and *E. coli* strain K-12 (Gene ID: 945387), both containing an N-terminal hexahistidine tag, were synthesized as codon optimized genes for expression in *E. coli* and ligated into the pMA vector (GENEART® Gene Synthesis, Life Technologies Australia Pty Ltd.). Site directed mutagenesis was performed using the QuickChange® site-directed PCR mutagenesis kit (Stratagene) with oligonucleotides synthesized at GeneWorks Pty Ltd (Adelaide, Australia). All oligonucleotides used in this study are listed in Supplementary Table S1. The MdDTBS-G169N mutation was constructed using oligonucleotides A1 and A2 whereas the two mutants of EcDTBS were constructed with oligonucleotides A3 and A4 for EcDTBS-N175G and A5 and A6 for EcDTBS-N175A. For recombinant expression in *E. coli*, the coding regions were subcloned into NcoI and HindIII treated pET16b vector (Novagen). All constructs were verified by DNA sequencing using the Gene Sequencing Service, SA Pathology (Adelaide, Australia).

**Expression and Purification** – DTBS proteins were recombinantly expressed in *E. coli* strain BL21(DE3) (Invitrogen). Cultured cells were used to inoculate LB medium containing 0.1 mg ml⁻¹ ampicillin and grown at 37 °C until early log phase when protein expression was induced with 0.1 mM IPTG. For EcDTBS and its mutants the cell cultures were induced at 37 °C for 3 h, whereas the MdDTBS and its mutant were grown for 18 hours at 25 °C. Cells were harvested by centrifuging at 4648 x g for 15 min and resuspended in 300 mM KCl, 50 mM potassium phosphate buffer (pH 8.0), 5 mM imidazole and 1 mM PMSF. Cells were lysed using a Microfluidics Cell Disruptor at 18 000 psi and debris removed by centrifuging at 48 384 x g for 30 min. His-tagged DTBS proteins were purified by immobilized nickel affinity chromatography (IMAC) using the Profinia® IMAC protocol following the manufacturers’ instructions (BIO-RAD). After elution in 50 mM potassium phosphate buffer (pH 8.0), 300 mM KCl and 250 mM imidazole, the proteins were subsequently buffer exchanged into storage buffer (25 mM Tris buffer (pH 7.5), 30 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol and 5% (v/v) glycerol) using a 50 ml Bio-Scale™ Mini Bio-Gel® P-6 desalting cartridge (BIO-RAD). For insoluble MdDTBS-G169N the inclusion bodies were solubilized in 8 M Urea (pH 8.0), 100 mM Tris, 300 mM NaCl and 5 mM imidazole. The urea buffer was first deionized using the Mixed Bed Resin (Sigma Aldrich®) following the manufacturers’ protocol. The solubilized sample was centrifuged at 48 384 x g for 30 min, before fractionation using the Profinia® Denaturing IMAC protocol following the manufacturers’ instructions (BIO-RAD). MdDTBS-G169N was eluted in 8 M Urea (pH 8.0), 100 mM Tris, 300 mM NaCl and 250 mM imidazole, then diluted to 0.2 mg ml⁻¹ for refolding. The protein was dialysed extensively against storage buffer for 48 hours by changing the buffer four times (1:100 (v/v)). DTBS proteins were concentrated using Amicon Ultra-15 centrifugal filter devices (Millipore), followed by centrifugation at 15 700 x g for 15 min to discard protein aggregates. Protein concentrations were determined using a Bradford Protein Assay (BioRad) using bovine serum albumin (Sigma Aldrich®) as a standard.

**Protein Methods** – Estimation of the molecular weight of active DTBS material was performed by size exclusion chromatography – multi-angle light scattering (SEC-MALS³) analysis. The instrument consisted of a Superdex 200 10/300 GL column (Amersham Biosciences) connected in series with a miniDAWN TREOS light-scattering detector (Wyatt Technology) and an OPTILAB rEX interferometric refractometer detector (Wyatt Technology). Analytical size-exclusion chromatography was performed at 20 °C using a mobile phase containing 25 mM Tris pH 7.5, 30 mM NaCl and 1 mM EDTA at a flow rate of 0.25 mL/min. Detector outputs were acquired using Astra software (Wyatt Technology).
The identities of the purified wildtype proteins were confirmed by enzyme assay and liquid chromatography-electrospray ionization tandem mass spectrometry, whilst the molecular weight of the mutants were confirmed using MALDI mass spectrometry, as described in Supporting Methods. All mass spectrometry services were performed at the Adelaide Proteomics Center, The University of Adelaide.

Preparation of Nucleotide triphosphates – Solutions of NTPs were prepared by dissolving solid in milliQ water and adjusting the pH to 7.0 with sodium hydroxide. The final concentration was determined by measuring absorbance at 260 nm, as described in Sambrook (21). For assays, the NTPs were first mixed with an equimolar concentration of MgCl₂ solution to form a MgNTP complex.

DTBS Enzyme Assay – The activity of DTBS was determined using fluorescence-polarization technology to quantify the amount of synthesized DTB that could displace fluorescein-biotin (Anaspec Inc) from streptavidin. All experiments were performed in black 96 well microtiter plates (Costar) preblocked with 1% casein dissolved in Tris-buffered saline. Assays were performed at 37 °C in FP buffer (50 mM Tris (pH 7.5), 150 mM NaCl, 5 mM MgCl₂, and 10 mM NaHCO₃ (or 50 mM NaHCO₃ for M₉DTBS-G169N)) supplemented with 0.3 mM ATP and 0.1 mM DAPA (Toronto Research Chemicals, Inc.). Assays were initiated with the addition of DTBS enzyme. The concentration of DTBS in the reaction (at 0.25 – 3.0 µM) varied depending on its species of origin or mutant employed. The reactions were terminated at 45 min by the addition of 50 mM Tris (pH 7.5), 150 mM NaCl, and 250 mM EDTA and then incubated for 15 min at 37 °C. The synthesized DTB product was detected by adding 20 nM streptavidin in 50 mM Tris (pH 7.5), 150 mM NaCl and incubating for 15 min at 37 °C before the addition of 25 nM fluorescein biotin in 50 mM Tris (pH 7.5), 150 mM NaCl. Finally, the reaction mixtures were incubated at 37°C for 15 min before fluorescence polarization (FP) was measured with a PerkinElmer 2030 Plate Reader using polarization mode set to 485 nm excitation and 535 nm emission (22). FP is defined as the difference of the emission light intensity parallel and perpendicular to the excitation light plane; normalized by the total fluorescence emission intensity (23,24). The concentration of enzymatically generated DTBS was calculated from a standard curve generated with each experiment (Supplementary Fig. S1) and using Equation 1; a linear regression binding model where S is the slope of the standard curve, Y is the Y intercept and ΔmP is the corrected FP value (ie mP (streptavidin) – mP (no streptavidin control)).

\[
\text{Log (DTB (M))} = \frac{-(\Delta mP + Y)}{S}
\]  

(1)

The data were fit to the Michaelis–Menten equation and the \(V_{\text{max}}\) and \(K_m\) were then calculated using GraphPad Prism 6 (GraphPad Software, Inc.).

DTBS ATP-binding Assay – The binding of a fluorescently-labeled ATP tracer, ATP-BODIPY (Life Technologies™), to DTBS was performed in binding buffer containing 0.1 mM DTBS, 50 mM Tris (pH 7.5), 150 mM NaCl, and 5 mM MgCl₂. Varying concentrations of MgNTPs were included and the reactions pre-equilibrated for 15 min before the addition of 25 nM ATP-BODIPY. The reaction was performed at 37 °C for 15 min and fluorescence polarization measured as described above. The \(IC_{50}\) was measured by plotting the corrected FP values (\(\Delta mP = mP(DTBS) - mP\) (no DTBS control)) against log of MgNTP concentration and fitting the data to a log (inhibitor) versus response – variable slope (four parameters) model in GraphPad Prism 6 (GraphPad Software, Inc.).

Surface Plasmon Resonance (SPR³) – The kinetics of the interaction between DTBS and NTP were analyzed using a BIACore T100. Enzymes were immobilized onto the surface of a CM5 sensor chip as previously described (25). Flow cell 1 of the sensor chip was used as reference to correct for bulk refractive index changes. Immobilized bovine serum albumin was also employed as control for non-specific binding. Varying concentrations of MgNTP were prepared in running buffer (10 mM HEPES (pH 7.4), 150 mM NaCl and 0.005% (v/v) surfactant P20) then injected across the sensor surface. The equilibrium binding constant (\(K_D\)) was determined using an affinity-steady state 1:1
model in BIACore T100 evaluation software (GE Healthcare).

X-ray Crystallography – MtDTBS was concentrated to 10 mg ml$^{-1}$, to which a final concentration of 1 mM CTP was added and incubated for 1 hour at room temperature. The complex was crystallized using the hanging drop vapour diffusion method at room temperature. Crystals grew from the following range of buffers: 0.8 - 1.2 M ammonium sulphate, 0.1 M Tris pH 7 - 8.5 with 10 % glycerol.

X-ray diffraction data was collected in-house using a Rigaku RUH2R rotating copper anode X-ray source equipped with Osmic confocal optics and an R-Axis IV detector. Crystals were flash frozen and the diffraction data collected with cryocooling using an Oxford Cryosystems 700 Series cryostream at 100 °K. The diffraction data was indexed, merged and integrated using XDS (26), which was also used to produce the final MTZ files required for structure refinement. Initial phases of the structure were determined with PHASER (27) using the coordinates of MtDTBS (PDB ID: 3FGN) as a search model. The model of MtDTBS in complex with CTP was built with cycles of manual modeling with COOT (28) and refinement with PHENIX (29).

RESULTS

Purification and biochemical characterization of M. tuberculosis and E. coli DTBS – Genes encoding the two bacterial DTBS enzymes where synthesized and cloned into the pET-16b plasmid for recombinant expression in E. coli. A hexahistidine tag was engineered onto the N-terminus of the proteins to facilitate purification using immobilized nickel affinity chromatography. EcDTBS expressed as a soluble protein at 37°C, whereas the optimal conditions for the M. tuberculosis equivalent were overnight at 25°C. SDS-PAGE analysis of the IMAC purified material revealed a high purity (>95%) of the IMAC purified material, and a single product corresponding to the expected molecular mass of 23.5 kDa for MtDTBS or 25.1 kDa for EcDTBS (Supplementary Fig. S2). Liquid chromatography-electrospray ionization tandem mass spectrometry analysis of the tryptic peptides confirmed the identification of the purified material, with 47% coverage for MtDTBS and 41% coverage for EcDTBS (Supplementary Table S2). SEC-MALS analysis revealed that >98% of the total proteins eluted at a molecular mass consistent with a homodimer (Supplementary Fig. S3A&B). Interestingly, MS/MS sequencing failed to detect endogenous E. coli protein in the MtDTBS preparation implying the recombinant mycobacterial protein did not form heterodimers with the host protein.

Enzyme Activity and Assay Development - To characterize the purified enzymes two alternative assays were developed to measure enzyme activity and ATP binding. Both assays incorporated fluorescence polarization technology to quantify the difference in anisotropy of a fluorescent tracer between its free and protein-bound states (30-32). The enzyme assay exploits the observation that DTB, enzymatically produced by DTBS, effectively competes with commercially available fluorescein-biotin (M, 831 Da) to bind streptavidin (M, 67 000 Da). Notably, the tracer did not bind to DTBS (Supplementary Fig. S4A). Furthermore only DTB, but not the substrate DAPA, could displace the tracer from streptavidin (Supplementary Fig. S4B). The ATP binding assay exploits the large mass difference observed between free ATP-BODIPY (M, 933 Da) and the tracer in complex with DTBS (Mr 23 500 Da for MtDTBS; 25 100 Da for EcDTBS). The binding of ATP tracer to DTBS was found to be independent of DAPA substrate (Supplementary Fig. S4C). Both approaches provide convenient, homogenous assays using commercially available reagents that yield complementary data to quantify the mechanism of ligand binding. The apparent $K_M$ for DAPA, NaHCO$_3$ and ATP were obtained using steady-state kinetic analysis, varying the concentration of one substrate whilst maintaining the two others at saturating concentrations. The $K_M$ values for DAPA and NaHCO$_3$ were calculated to be 15.2 ± 1.7 µM and 0.6 ± 0.1 mM for EcDTBS and 30.2 ± 1.1 µM and 3.1 ± 0.4 mM for MtDTBS, respectively (Table 1). In this study, the $K_M$ for ATP for MtDTBS (30.2 ± 1.7 µM) was 3-fold higher than that of EcDTBS (10.5 ± 0.8 µM), but these values were in excellent agreement with published literature for the two enzymes (14,33).
The activity of the DTBS enzymes was initially measured in the presence of each of a panel of NTPs at 0.3 mM. Unexpectedly, MtDTBS could utilize any of ATP, CTP, GTP, ITP, TTP and UTP (Fig. 1). These observations were in stark contrast to the E. coli orthologue where the enzyme had greatest preference for ATP, a result consistent with published literature (33). Further kinetic analysis showed that MtDTBS utilized all nucleotides with similar $K_m$’s and $k_{cat}$’s while EcDTBS could use CTP and GTP albeit with more than 10-fold higher $K_m$ than ATP (Fig. 2 & Table 1). This nucleotide specificity was confirmed in the ATP-binding assay where all nucleotides could compete effectively against the ATP-BODIPY tracer for binding MtDTBS (Fig. 3). Here, CTP was observed to have the greatest binding affinity with a 6-fold lower $IC_{50}$ than ATP in MtDTBS (Table 2). Consistent with the enzyme assay, ATP was the preferred NTP in EcDTBS with >10-fold lower $IC_{50}$ values than CTP and GTP. SPR analysis was also employed to measure the bimolecular interactions between DTBS and nucleotide triphosphates. Again, this confirmed that both ATP and CTP bound to MtDTBS (Fig. 4A) with CTP contacting the highest affinity ($K_D$ CTP = 17.2 ± 1.7 μM; ATP = 335.1 ± 96.0 μM) (Supplementary Fig. S5A&B). In contrast, EcDTBS was highly selective for ATP ($K_D$ of 347.9 ± 75.1 μM) over CTP (Fig. 4B and Supplementary Fig. S5C). Injection of NTPs over either an empty flow cell or immobilized bovine serum albumin (BSA), performed concurrently with the immobilized DTBS, resulted in negligible response (Fig 4D, and data not shown). This confirmed the specificity of NTP binding to the DTBS enzymes.

X-ray Crystallography – To date there have been no X-ray structures available of MtDTBS in complex with a nucleotide triphosphate. To better understand the molecular basis of ligand binding, MtDTBS was crystallized in complex with CTP (Fig. 5). Structures were determined to 2.3 Å. The data collection, refinement, and geometric statistics are provided in Supplementary Table S3. Repeated attempts to crystallize the enzyme with ATP were unsuccessful. As expected, our structures were consistent with the α/β-fold observed in the published structures for E. coli, H. pylori and F. tularensis DTBS (17).

Superimposition of chain D of our structure against chain D of a previously reported apo-structure for M. tuberculosis (PDB # 3FGN (14)) had the lowest RMSD of 0.387 Å (for 224 Cα atoms), and the same fold of a seven-stranded parallel β-sheet surrounded by seven helices. Superimposition of the two dimeric forms of the crystallized structures (chain A and B) were also in excellent agreement with an RMSD of 0.403 Å (for 225 Cα atoms). These results indicated that there are no large conformational changes that accompany CTP binding. Whilst the enzyme crystallized as a dimer of dimers, the NTP-binding site is distant from the dimer interface and all the residues contacting the NTP are present within a single subunit. MtDTBS possesses the classical P-loop motif (Walker A motif) with Gly12, Val13, Gly14, Lys15 and Thr16 all forming hydrogen bonding interactions with the triphosphate moiety on CTP. A series of hydrophobic interactions involving C2 and the 2’ hydroxyl on the ribose with the side chain of Val 17 and the beta carbon of Ala 201 stabilize the sugar in its binding pocket (Supplementary Fig. S6). CTP adopted a bent conformation with rotation around the C-N bond between ribosyl and cytosinyl groups to allow the pyrimidine ring to bind in its appropriate pocket. Two hydrophobic interactions are observed between the pyrimidine ring and the protein, namely C5 with Ser170 and C6 with Val17 in the P-loop. The interaction is further stabilized through three hydrogen-bonding interactions with the pyrimidine ring. What is noteworthy is that all three of these bonds are with the protein backbone and not with amino acid side chains; O2 bonds with the nitrogen of Ala201 whilst N4 bonds with the backbone oxygens of Gly169 and Pro197.

The lack of side chain interactions contacting the nucleotide base observed in MtDTBS is in contrast to the E. coli orthologue, in which the carboxamide side chain on Asn175 forms hydrogen-bonding interactions with N6 and N7 on the purine ring of ATP (Fig. 6) (13). This side chain interaction is also present in H. pylori, which selectively binds ATP (17). It has previously been proposed that side chain interactions involving asparagine play a key role in nucleotide specificity (17,34), but this has not been tested experimentally. To uncover the function of this asparagine in NTP binding, an asparagine residue
was engineered into the *M. tuberculosis* enzyme at the structurally equivalent position to Asn175 in EcDTBS, thereby generating MtDTBS-G169N. Likewise, two muteins of the *E. coli* homologue were generated by substituting Asn175 with either glycine or alanine (producing EcDTBS-N175G and EcDTBS-N175A).

**Purification and biochemical characterization of mutants – EcDTBS-N175G and EcDTBS-N175A were expressed as soluble proteins in *E. coli* allowing their purification by IMAC. In contrast, MtDTBS-G169N formed inclusion bodies and, therefore, required refolding. MtDTBS-G169N was purified using IMAC under denaturing conditions and then refolded by a four step-wise dialysis procedure to restore enzyme activity. SEC-MALS analysis revealed that only 25% of the refolded MtDTBS-G169N material eluted as a homodimer, with the majority of the material forming inactive protein aggregates (Supplementary Fig. S3C). Following size exclusion chromatography the specific enzyme activity of active, dimeric MtDTBS-G169N (56.2 μM min⁻¹ mg⁻¹) was comparable to that measured for the wildtype MtDTBS (74.0 μM min⁻¹ mg⁻¹). To confirm the identity of the muteins, MALDI mass spectrometry was performed. The measured molecular mass corresponding to methionine removed DTBS [M+H]⁺ was in agreement with the predicted molecular mass (Table 3).

The mutant enzymes all showed *Kₐ* values for DAPA comparable to their wildtype counterparts (Table 1), suggesting that the introduced mutations had not compromised enzymatic activity nor the DAPA binding site. Whilst the two substitutions in the *E. coli* enzyme had no significant effect on the *Kₐ* for NaHCO₃, the G169N mutation in MtDTBS elevated the *Kₐ* for the carbonate 3-fold (WT vs G169N, *p*<0.05) (Table 1). The molecular explanation for this is unclear.

Substitution of N175 in EcDTBS with alanine or glycine had no effect on the *Kₐ* for ATP (EcDTBS-N175G vs WT, *p*=0.11; EcDTBS-N175A vs WT *p*=0.19). Whilst the *Kₐ* for CTP was not significantly different for EcDTBS-N175G, the introduction of alanine significantly lowered the *Kₐ* for CTP by 2.7-fold (EcDTBS-N175A vs WT, *p*<0.05) implying that hydrophobic interactions help stabilize the interaction between protein and NTP. This was supported by SPR analysis that also demonstrated that the binding of CTP was improved for EcDTBS-N175A (Fig. 4C). These kinetic data are in contrast to the MtDTBS-G169N mutant. Whilst the affinity for ATP was again unchanged (MtDTBS-G169N vs WT, *p* = 0.09), the asparagine substitution significantly elevated the *Kₐ* for both CTP and GTP by 1.8-fold and 3.8-fold respectively. Unfortunately, attempts to immobilize the MtDTBS-G169N for SPR analysis were unsuccessful. Together these data support a role for N175 in excluding other NTPs from the nucleoside-binding pocket, thus allowing selective binding of ATP. Without the ‘gate keeper’ asparagine, MtDTBS can accommodate a broad range of NTPs.

**DISCUSSION**

It has previously been proposed that DTBS is more similar to GTP-binding proteins than ATP-dependent enzymes based on its fold (35) and the presence of the [NO]-[KR]-x-[DE] GTP specificity motif in the *E. coli*, *H. pylori* and *F. tularensis* enzymes (containing NDVTP, NLKGN, and NCNDS motifs respectively) (17). The first asparagine in this motif (numbering N175 in EcDTBS) was suggested to be an evolutionary relic from a protein able to bind both adenosine and guanine nucleotides and provides nucleotide specificity (17,34). Curiously, in the current study, DTBS from *M. tuberculosis* was shown to be capable of binding ATP and GTP, as well as other nucleotides, with similar affinities despite the enzyme being devoid of the equivalent asparagine. Here, for the first time, we present the structure of MtDTBS in complex with a nucleotide (CTP) present. The structural data revealed that the hydrogen bonding interactions with the nucleoside occurred exclusively through the peptide backbone of the enzyme, and not via side chains. Indeed, with the absence of the Asn175 (equivalent position of G169 in MtDTBS) or the reduction of hydrophilic interaction at C5 of pyrimidine ring with S170 (MtDTBS-S170D, data not shown), the MtDTBS retained the ability to bind all nucleotides. Unlike EcDTBS, no large conformational changes were observed in MtDTBS upon the binding of nucleotide.
To further understand the molecular basis of NTP binding to DTBS, we superimposed the EcDTBS structure complexed with ATP (PDB # 1A82 (13)) and HpDTBS complexed with ATP or GDP (PDB # 3QXC, 3QXX (17)) onto the CTP bound MtDTBS structure. In the published X-ray structures ATP and GTP both adopt an extended conformation that would cause a significant steric clash with MtDTBS in the nucleotide-binding pocket. Compared to the more open pockets observed in EcDTBS and HpDTBS, MtDTBS has a more constrained pocket due to additional strands at the C-terminus of the enzyme. Possible flexibility at the loops formed by G169 – W171 and A195 – A201 would allow accommodation of the NTPs in the nucleotide pocket.

The biological significance of the selective binding reported here may be a function of the concentrations of NTPs found in different bacteria. For E. coli and Salmonella enterica, it has been reported that ATP is the most abundant NTP, with concentrations being at least 3 mM (36,37). This is >300 times above the $K_M$ for ATP measured in this study for EcDTBS suggesting that ATP is constantly at saturating concentrations in vivo. The levels of various NTPs can decrease as E. coli transitions from growth to stationary phase. Interestingly, of all the NTPs, CTP had the most constant concentration throughout the whole growth curve, with only minor fluctuations in its levels. In Mycobacteria there is an abundant supply of all NTPs during all stages of cell growth (38). Here nucleoside diphosphate kinase plays an important role in NTP synthesis by utilizing both purine and pyrimidine NDP as substrates to produce NTPs. Nucleoside diphosphate kinase can also form complexes with several proteins, and this has been shown to alter the specificity of substrate binding thereby directing the synthesis of specific NTPs as required during active cell growth. For example, M. tuberculosis cell wall protein A, a homologue of protein P$_{60}$ in Mycobacterium smegmatis, can induce the specificity of nucleoside diphosphate kinase towards CTP synthesis (39). Therefore, we propose that the broad specificity observed in this study for the MtDTBS provides the tuberculi with the potential to scavenge alternative NTPs depending upon availability. This ensures that the activity of this important enzyme, and the biotin biosynthetic pathway, is maintained.

Precisely why CTP binds with 6-fold higher affinity than ATP is still unknown. However, it does suggest that analogues of CTP may be attractive candidates for the development of MtDTBS inhibitors. CTP analogues have been clinically used in the treatment of HIV, hepatitis B virus (HBV) and cancer (40-44). For example, Lamivudine (β-D-(-)-2',3'-dideoxy-3’thiacytidine (-)3TC) is an anti-HIV drug that has been shown to be very potent (43). Effective doses of (-)3TC are well tolerated during therapy (45), and no signs of mitochondrial toxicity have been observed (44). These results suggest CTP analogues have promising drug-like characteristics. The enabling technologies developed here, namely the DTBS enzyme activity assay, SPR and X-ray crystal structures for in silico screening, will be of great utility for high throughput screening of CTP analogues as novel MtDTBS inhibitors. Targeting the DTBS enzyme holds great promise for treating drug-resistant strains of M. tuberculosis, since it is part of its essential biotin biosynthesis pathway, of which there are no known pre-existing resistance mechanisms.
REFERENCES


Promiscuous NTP Binding to Mycobacterial Dethiobiotin synthetase

Acknowledgements – The authors are grateful to Dr Stephan Meding, Mr James Eddes, and Mr Chris Cursaro of the Adelaide Proteomics Center at The University of Adelaide for help in mass spectrometry. We also acknowledge the computer resources of the Victorian Partnership for Advanced Computing and the Adelaide Protein Characterization Facility for access to the BIAcore. We also thank Rebecca Chao, Samuel Munday, and Kayla Downey for help in plasmid construction and preparation of the enzymes. The authors acknowledge that they have no conflict of interest to declare.

FOOTNOTES

*This work was supported by the National Health and Medical Research Council of Australia (applications APP1011806 & APP1068885) and the Centre of Molecular Pathology, University of Adelaide. WS was a recipient of the Royal Thai Government Scholarship. MCJW is an Australian National Health and Medical Research Council of Australia Senior Research Fellow.

2Current address: CSL Limited, 45 Poplar Road, Parkville, Victoria 3052, Australia

3Abbreviations: DAPA, (7R, 8S)-7,8-diaminonanoic acid; DTB, dethiobiotin; DTBS, dethiobiotin synthetase; IMAC, immobilized nickel affinity chromatography; SEC-MALS, size exclusion multi angle light scattering; SPR, surface plasmon resonance
FIGURE LEGENDS

Fig. 1. NTPs specificity. The specificity of MtDTBS (black bars) and EcDTBS (white bars) for different nucleotides was measured in the presence of 0.1 mM DAPA, 10 mM NaHCO₃, and 0.3 mM NTP. The specificity is presented as the percentage of enzyme activity normalized to 100% of enzyme activity using ATP.

Fig. 2. Enzyme analysis. The $K_m$ of nucleotide substrates for both MtDTBS (A – C) and EcDTBS (D – F) were determined by varying the concentrations of ATP, CTP and GTP in the reaction containing 0.1 mM DAPA and 10 mM NaHCO₃.

Fig. 3. NTP Binding to DTBS. Competitive binding assays using a fluorescent ATP tracer and various NTPs to (A) MtDTBS and (B) EcDTBS. NTPs employed were ATP (●), CTP (■), GTP (▲), ITP (★), TTP (◆), and UTP (●).

Fig. 4. Surface plasmon resonance analysis. The sensorgrams represent the binding of ATP (green lines), CTP (blue lines) to (A) MtDTBS, (B) EcDTBS and (C) EcDTBS-N175A. (D) BSA was included as a negative control to demonstrate specific binding of nucleotide to DTBS. HEPES running buffer alone (red line) was also included. The experiment was performed at a single concentration of nucleotide (0.3 mM).

Fig. 5. Crystal structure of MtDTBS in complex with CTP. A, Cartoon diagram of the MtDTBS dimer in complex with CTP. Subunit A is colored in gray and subunit B is colored in green. CTP is shown in pink stick representation. B, The hydrogen bond network formed between amino acid residues of MtDTBS and CTP is shown. Hydrogen bonds are indicated as dashed lines.

Fig. 6. Superimposition of MtDTBS and EcDTBS. The glycine at position Gly169 (grey), in MtDTBS, is located at the equivalent position of an Asn175 in EcDTBS, as shown in blue. While two hydrogen bonds (shown as dashed lines) formed between the side chain of Asn175 in EcDTBS and ATP (PDB # 1A82 (13), shown in cyan), these interactions are missing in MtDTBS.

TABLES

TABLE 1. Kinetic analysis of substrates.

TABLE 2. Nucleotide binding to MtDTBS and EcDTBS using DTBS ATP-binding assay

TABLE 3. Molecular mass of the wildtype and mutant DTBS measured by MALDI-MS.
Figure 1

![Graph showing enzyme activity for different nucleotides.](image1)

Figure 2

A. ![Graph showing velocity vs. ATP concentration.](image2)

B. ![Graph showing velocity vs. CTP concentration.](image3)

C. ![Graph showing velocity vs. GTP concentration.](image4)

D. ![Graph showing velocity vs. ATP concentration.](image5)

E. ![Graph showing velocity vs. CTP concentration.](image6)

F. ![Graph showing velocity vs. GTP concentration.](image7)
Figure 3

A. % Binding vs. Log (NTP concentration (M))

B. % Binding vs. Log (NTP concentration (M))
Figure 4
Figure 5

A.

B.

Figure 6
### TABLE 1 Kinetic analysis of substrates

<table>
<thead>
<tr>
<th>Enzymes</th>
<th>Substrates</th>
<th>$K_m$ (µM)</th>
<th>$k_{cat}$ (s$^{-1}$) (x 10$^{-3}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>MtDTBS</em></td>
<td>DAPA</td>
<td>30.2 ± 1.1</td>
<td>13.3 ± 0.7</td>
</tr>
<tr>
<td></td>
<td>NaHCO$_3$</td>
<td>3074.3 ± 352.4</td>
<td>9.9 ± 0.6</td>
</tr>
<tr>
<td></td>
<td>ATP</td>
<td>30.2 ± 1.7</td>
<td>11.7 ± 0.2</td>
</tr>
<tr>
<td></td>
<td>CTP</td>
<td>25.2 ± 1.7</td>
<td>15.2 ± 0.3</td>
</tr>
<tr>
<td></td>
<td>GTP</td>
<td>26.3 ± 0.5</td>
<td>11.1 ± 0.1</td>
</tr>
<tr>
<td></td>
<td>ITP</td>
<td>17.7 ± 1.7</td>
<td>10.2 ± 1.2</td>
</tr>
<tr>
<td></td>
<td>TTP</td>
<td>23.2 ± 1.4</td>
<td>11.8 ± 0.8</td>
</tr>
<tr>
<td></td>
<td>UTP</td>
<td>26.1 ± 0.5</td>
<td>11.4 ± 0.8</td>
</tr>
<tr>
<td><em>EcDTBS</em></td>
<td>DAPA</td>
<td>15.2 ± 1.7</td>
<td>49.0 ± 1.0</td>
</tr>
<tr>
<td></td>
<td>NaHCO$_3$</td>
<td>599.8 ± 72.3</td>
<td>38.4 ± 3.4</td>
</tr>
<tr>
<td></td>
<td>ATP</td>
<td>10.5 ± 0.8</td>
<td>38.6 ± 2.5</td>
</tr>
<tr>
<td></td>
<td>CTP</td>
<td>1089.8 ± 156.8</td>
<td>28.1 ± 0.8</td>
</tr>
<tr>
<td></td>
<td>GTP</td>
<td>1561.3 ± 299.2</td>
<td>29.0 ± 0.9</td>
</tr>
<tr>
<td></td>
<td>ITP</td>
<td>&gt; 0.3 mM</td>
<td>N/D</td>
</tr>
<tr>
<td></td>
<td>TTP</td>
<td>&gt; 0.3 mM</td>
<td>N/D</td>
</tr>
<tr>
<td></td>
<td>UTP</td>
<td>&gt; 0.3 mM</td>
<td>N/D</td>
</tr>
<tr>
<td><em>MtDTBS-G169N</em></td>
<td>DAPA</td>
<td>15.1 ± 5.7</td>
<td>0.9 ± 0.3</td>
</tr>
<tr>
<td></td>
<td>NaHCO$_3$</td>
<td>10385.3 ± 2248.0</td>
<td>1.7 ± 0.4</td>
</tr>
<tr>
<td></td>
<td>ATP</td>
<td>21.5 ± 3.4</td>
<td>1.3 ± 0.2</td>
</tr>
<tr>
<td></td>
<td>CTP</td>
<td>45.1 ± 4.9</td>
<td>1.0 ± 0.1</td>
</tr>
<tr>
<td></td>
<td>GTP</td>
<td>100.7 ± 12.8</td>
<td>1.2 ± 0.1</td>
</tr>
<tr>
<td><em>EcDTBS-N175G</em></td>
<td>DAPA</td>
<td>18.0 ± 2.3</td>
<td>23.2 ± 3.2</td>
</tr>
<tr>
<td></td>
<td>NaHCO$_3$</td>
<td>383.6 ± 64.0</td>
<td>15.4 ± 1.0</td>
</tr>
<tr>
<td></td>
<td>ATP</td>
<td>17.3 ± 2.6</td>
<td>23.6 ± 4.6</td>
</tr>
<tr>
<td></td>
<td>CTP</td>
<td>869.0 ± 55.9</td>
<td>19.6 ± 2.5</td>
</tr>
<tr>
<td></td>
<td>GTP</td>
<td>1386.8 ± 236.5</td>
<td>16.3 ± 1.6</td>
</tr>
<tr>
<td><em>EcDTBS-N175A</em></td>
<td>DAPA</td>
<td>18.8 ± 0.2</td>
<td>0.9 ± 0.3</td>
</tr>
<tr>
<td></td>
<td>NaHCO$_3$</td>
<td>433.0 ± 128.4</td>
<td>16.2 ± 1.9</td>
</tr>
<tr>
<td></td>
<td>ATP</td>
<td>12.3 ± 0.9</td>
<td>18.1 ± 3.3</td>
</tr>
<tr>
<td></td>
<td>CTP</td>
<td>390.9 ± 97.3</td>
<td>19.1 ± 1.6</td>
</tr>
<tr>
<td></td>
<td>GTP</td>
<td>1096.9 ± 121.3</td>
<td>18.3 ± 2.3</td>
</tr>
</tbody>
</table>

N/D represents not determined. The $K_m$ values were calculated from three independent experiments and reported as mean ± SEM.
### TABLE 2 Nucleotide binding to *MtDTBS* and *EcDTBS* using DTBS ATP-binding assay

<table>
<thead>
<tr>
<th>Enzymes</th>
<th>Substrates</th>
<th>$IC_{50}$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>MtDTBS</em></td>
<td>ATP</td>
<td>348.8 ± 15.3</td>
</tr>
<tr>
<td></td>
<td>CTP</td>
<td>54.4 ± 4.3</td>
</tr>
<tr>
<td></td>
<td>GTP</td>
<td>596.6 ± 51.3</td>
</tr>
<tr>
<td></td>
<td>ITP</td>
<td>520.3 ± 60.0</td>
</tr>
<tr>
<td></td>
<td>TTP</td>
<td>622.3 ± 19.5</td>
</tr>
<tr>
<td></td>
<td>UTP</td>
<td>636.4 ± 18.0</td>
</tr>
<tr>
<td><em>EcDTBS</em></td>
<td>ATP</td>
<td>436.9 ± 39.1</td>
</tr>
<tr>
<td></td>
<td>CTP</td>
<td>5752.6 ± 852.5</td>
</tr>
<tr>
<td></td>
<td>GTP</td>
<td>&gt; 2000</td>
</tr>
</tbody>
</table>

$IC_{50}$ values were calculated from three independent experiments and reported as mean ± SEM.

### TABLE 3 Molecular mass of the wildtype and mutant DTBS measured by MALDI-MS

<table>
<thead>
<tr>
<th>DTBS enzymes</th>
<th>Methionine removed mass (Da)</th>
<th>Calculated mass</th>
<th>Measured [M+H]$^+$mass$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>MtDTBS</em></td>
<td></td>
<td>23318.5</td>
<td>23319.5</td>
</tr>
<tr>
<td><em>MtDTBS</em>-G169N</td>
<td></td>
<td>23375.3</td>
<td>23376.3</td>
</tr>
<tr>
<td><em>EcDTBS</em></td>
<td></td>
<td>25003.1</td>
<td>249996.1</td>
</tr>
<tr>
<td><em>EcDTBS</em>-N175G</td>
<td></td>
<td>24960.1</td>
<td>24954.1</td>
</tr>
<tr>
<td><em>EcDTBS</em>-N175A</td>
<td></td>
<td>24946.0</td>
<td>24938.7</td>
</tr>
</tbody>
</table>

$^a$The measured mass considered the initiator methionine had been post-translationally removed.
Nucleotide triphosphate promiscuity in *Mycobacterium tuberculosis* dethiobiotin synthetase

SUPPLEMENTARY INFORMATION

Wanisa Salaemae, Min Y. Yap, Kate L. Wegener, Grant W. Booker, Matthew C.J. Wilce and Steven W. Polyak
SUPPLEMENTARY FIGURES

Fig S1. Representative calibration curve of DTB concentration.

The concentration of DTB was determined from a standard curve generated by incubating 25 nM fluorescein-biotin and 20 nM streptavidin with varying amounts of DTB. Each DTB concentration was assayed in triplicate, and the data points represented on the graph represent the mean ± SEM.
Fig S2. SDS-PAGE analysis of recombinant *Mt*DTBS and *Ec*DTBS.

SDS-PAGE analysis of IMAC purified *Mt*DTBS (23.5 kDa, lane 1) and *Ec*DTBS (25.1 kDa, lane 2). The migration of molecular mass standards is shown on the left. The samples were estimated to be > 95% pure. Both proteins migrated through the gel with a size consistent with their molecular mass.

$IC_{50} = 348.80$ (n= 3)
Fig S3. SEC-MALS analysis of *MtDTBS*, *EcdTBS*, and *MtDTBS-G169N*.

Proteins were fractionated on Superdex 200 10/300 GL gel filtration column connected in series with a miniDAWN TREOS light-scattering detector (Wyatt Technology) and an OPTILAB rEX interferometric refractometer detector (Wyatt Technology). Analytical size-exclusion chromatography was performed at 20°C using a mobile phase containing 25 mM Tris pH 7.5, 30 mM NaCl and 1 mM EDTA at a flow rate of 0.25 mL/min. BSA was employed as the reference sample during calibration. Chromatograms of the refractive index are shown for (A) *MtDTBS* and (B) *EcdTBS*. Both samples eluted with molecular masses of 47 and 46 kDa respectively, consistent with the proteins being dimers. (C) In contrast, only 25% of refolded *MtDTBS-G169N* eluted in the dimer peak, with most of the protein observed in a high molecular mass aggregate. The percentage of the eluted protein was calculated from the proportion of the area under the curves on the chromatograms using ImageJ 1.46r (National Institute of Health, Maryland, USA).
Fig S4. Controls of FP detection in DTBS enzyme assay.

A, Fluorescein-biotin does not bind to DTBS. B, DAPA did not displace the streptavidin:fluorescein-biotin complex in the assay containing 25 nM fluorescein-biotin and 20 nM streptavidin. Likewise, C, the DTBS:ATP-BODiPY complex was not disrupted by increasing concentrations of DAPA in the binding reaction implying that NTP binding is independent of DAPA. Each data point is the mean of three experiments ± SEM.
Fig S5. Surface plasmon resonance analysis.

Isotherms showing binding of (A) ATP and (B) CTP to MtDTBS and (C) ATP to EcDTBS. Data were analyzed using an affinity-steady state 1:1 model. The time-dependent binding curves of ligands correlated to dose-response are also shown (insert). The binding constant ($K_D$) of ATP and CTP to MtDTBS was calculated to be 335.1 ± 96.0 µM ($n = 3$) and 17.2 ± 1.7 µM ($n = 3$). The binding constant for ATP binding to EcDTBS was calculated to be 347.9 ± 75.1 µM ($n = 3$). No response was observed using CTP and EcDTBS (data not shown).
Fig S6. Mode of CTP binding within the MrDTBS active site.

Residues interacting with the substrate CTP in the DTBS active site are shown in this figure, prepared using Ligplot. The protein residues are colored brown, and the ligand CTP is colored purple. Hydrogen bonds are shown in green dotted lines. Hydrophobic interactions are shown in brick red sunbursts. Atoms are colored by type (Carbon, black; Oxygen, red; Nitrogen, blue).
SUPPLEMENTARY TABLES

TABLE S1. Oligonucleotides employed for site-directed mutagenesis

<table>
<thead>
<tr>
<th>ID</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>5’GCCGGTCTGATTATAATAGCTGAGCCCTGATCCGC</td>
</tr>
<tr>
<td>A2</td>
<td>5’GCAGATACGGCGACCTATTAATAACCAGACCCGC</td>
</tr>
<tr>
<td>A3</td>
<td>5’CTGCGGGGTTGCTGCTGCGGCGATGTATACGCTCGGGGAAACG</td>
</tr>
<tr>
<td>A4</td>
<td>5’CTGTTTCCCGAGGCGTAACATCGCCGCACCCACACCACCCGCAG</td>
</tr>
<tr>
<td>A5</td>
<td>5’CTGCGGGGTTGGCGCGGCGATGTATACGCTCGGGGAAACG</td>
</tr>
<tr>
<td>A6</td>
<td>5’CGTTTCCCGAGGCGTAACATCGCGGCGGACCCACACCACCCGCAG</td>
</tr>
</tbody>
</table>

TABLE S2. Mascot LC-ESI-MS/MS search results

<table>
<thead>
<tr>
<th>Samples</th>
<th>Matched Protein</th>
<th>IDa / Total queriesb</th>
<th>Sequence coverage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MtDTBS</td>
<td>His6-MtDTBS</td>
<td>158 / 593</td>
<td>47</td>
</tr>
<tr>
<td>EcDTBS</td>
<td>ATP-dependent EcDTBS (trIB1X7A8IB1X7A8_ECODH)</td>
<td>221 / 586</td>
<td>41</td>
</tr>
</tbody>
</table>

aThe number of queries that were matched to a single protein, bThe number of queries resulting from LC-ESI-MS/MS run
### TABLE S3. Data Collection, Refinement, and Geometry Statistics of MtDTBS in complex with CTP

<table>
<thead>
<tr>
<th></th>
<th>CTP-MtDTBS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.54 (Cu anode)</td>
</tr>
<tr>
<td>Space group</td>
<td>P 2_12_1</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
</tr>
<tr>
<td>(a, b, c) (Å)</td>
<td>54.7, 104.6, 152.7</td>
</tr>
<tr>
<td>(a, b, g) (°)</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>37.8 – 2.34° (2.42-</td>
</tr>
<tr>
<td>(R_{merge})</td>
<td>2.34)</td>
</tr>
<tr>
<td>CC(_{1/2})*</td>
<td>0.15(0.9)</td>
</tr>
<tr>
<td>(I/\sigma I)</td>
<td>0.99 (0.65)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>6.5 (1.9)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>99.3 (98.2)</td>
</tr>
<tr>
<td></td>
<td>5.3 (5.3)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>35 – 2.25</td>
</tr>
<tr>
<td>No. of reflections</td>
<td>79064</td>
</tr>
<tr>
<td>(R_{work} / R_{free})</td>
<td>0.19/0.25</td>
</tr>
<tr>
<td>No. of atoms</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>1565</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>10</td>
</tr>
<tr>
<td>Water</td>
<td>412</td>
</tr>
<tr>
<td>Average B Factors (Å(^2))</td>
<td>63.7</td>
</tr>
<tr>
<td>Protein</td>
<td>63.7</td>
</tr>
<tr>
<td>Solvent</td>
<td>68.2</td>
</tr>
<tr>
<td>Ligands</td>
<td>65.1</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.02</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.3</td>
</tr>
<tr>
<td><strong>Ramachandran Plot</strong></td>
<td></td>
</tr>
<tr>
<td>% most favored regions</td>
<td>99.3 %</td>
</tr>
</tbody>
</table>

* High-resolution limit determined using CC\(_{1/2}\) (1)
SUPPLEMENTARY EXPERIMENTAL PROCEDURES

Surface Plasmon Resonance

The kinetics of the interaction between DTBS and NTPs was analysed using a BIAcore T100. DTBS enzymes were immobilized onto the surface of a CM5 sensor chip using amide coupling chemistry. After activating the surface with EDC and NHS, DTBS solution at concentration of 0.2 mg ml\(^{-1}\) in 0.01 M NaOAc buffer pH 5.2) was applied at a constant flow rate of 5 µl min\(^{-1}\) for 420 s. Approximately 10,000 response units (RU) of ligands were immobilized. Typically, no immobilized ligand (empty flow cell) and immobilized bovine serum albumin (BSA) were performed alongside immobilized DTBS in order to correct for bulk refractive index changes with buffer and distinguish non-specific binding events, respectively. Water-soluble NTPs were diluted in running buffer (10 mM HEPES (pH 7.4), 150 mM NaCl and 0.005% (v/v) surfactant P20). Binding experiments were performed by injecting the analyte solutions across the sensor surface of all flow cells at a flow rate of 30 µl min\(^{-1}\) with a contact time of 60 sec followed by a dissociation time of 60 sec. The time-dependent binding curves from all four flow cells were monitored simultaneously.

As the analyte bound to the immobilized DTBS, the refractive index at the surface alters in proportion to the change in mass, resulting in change of RU value. The binding capacity of the surface depends on the level and activity of immobilized ligand. The maximum binding capacity (\(R_{\text{MAX}}\)) of the immobilized ligand was calculated using Equation 1 where MW is the molecular weight of the ligand and analyte, RL is the amount of immobilized ligand in RU, and \(S_M\) is the stoichiometry as defined by the number of binding sites on the ligand (1). The percent activity of the immobilized ligand that reflects the quality of binding was determined using Equation 2 (1).

\[
\text{Equation 1: } R_{\text{MAX}} = \left(\frac{\text{MW}_{\text{analyte}}}{\text{MW}_{\text{ligand}}}\right) \times \text{RL} \times S_M
\]

\[
\text{Equation 2: } \% \text{ ligand activity} = \left(\frac{R_{\text{max}}_{\text{experiment}}}{R_{\text{max}}_{\text{theory}}}\right) \times 100
\]

The binding affinity (\(K_D\)) was determined by transforming the time-dependent binding curves into an affinity-steady state 1:1 model using BIAcore T100 evaluation software (GE Healthcare).

Mass Spectrometry

LC-ESI-MS/MS was acquired using an 1100 series HPLC system (Agilent Technologies) coupled via an Advance CaptiveSpray source (Michrom Bioresources, Inc.) to an amaZon ETD mass spectrometer (Bruker Daltonics). Samples were initially run on a 4-12% gradient SDS-PAGE gel and stained with coomassie blue dye. Bands were excised from the gel manually and processed with a trypsin digestion protocol (2). Collision-induced dissociation (CID) spectra were acquired to select the two most abundant ionisable species in the 300-2000 m/z range at any point in the LC separation. The acquired spectra were subjected to peak detection and deconvolution using DataAnalysis version 4.0 SP4 Build 281 (Bruker Daltonics). Processed MS/MS spectra were exported to Mascot generic format (mgf) and submitted to Mascot version 2.3.02 for identification by comparison to the amino acid sequences of His\(_6\)-MtDTBS, His\(_6\)-
EcDTBS, MtDTBS, EcDTBS and common contaminants (i.e. human keratins and porcine trypsin) from Uniprot.

MALDI-MS was acquired on an ultrafleXtreme MALDI-TOF/TOF mass spectrometer (Bruker Daltonics) operating in a linear positive mode under the control of flexControl software version 3.4 (Bruker Daltonik GmbH). Samples were desalted by spotting onto nitrocellulose filters 0.025 µm MF-membrane filters VSWP0500 (Millipore) and dialyzing in 800 ml of 0.1% TFA for 2 hr. One microliter of each sample was mixed with 1 µl of sinapinic acid saturated TA30 (30% ACN in 0.1% FA). Next, the sample matrix was prepared by spotting the 1 µl sample mix onto the dried matrix spots which were previously spotted with 2 µl of sinapinic acid saturated ethanol onto a ground steel target plate (Bruker Daltonics) and then air dried. Sample m/z range was set to 5000 – 55000 Da. 50000 shots were collected for the external calibration and sample measurement. External calibration was performed using a mix of protein calibration standard I and II (Bruker Daltonics). Laser intensity and detector gain was manually adjusted for optimal resolution. The MS spectra obtained were analysed using flexAnalysis software version 3.3 (Bruker Daltonics) employing smoothing, background subtraction and peak detection algorithms.

SUPPLEMENTARY REFERENCES

CHAPTER 5

FRAGMENT BASED SCREENING OF

*MYCOBACTERIUM TUBERCULOSIS* DETHIOBIOTIN SYNTHETASE INHIBITORS
Statement of Authorship

Title of Paper: Fragment-based screening of Mycobacterium tuberculosis dethiobiotin synthetase inhibitors

Publication Status:
- Published
- Accepted for Publication
- Submitted for Publication
- Publication Style

Publication Details: N/D

Author Contributions

By signing the Statement of Authorship, each author certifies that their stated contribution to the publication is accurate and that permission is granted for the publication to be included in the candidate’s thesis.

Name of Principal Author (Candidate): Warrisa Salameh

Contribution to the Paper:
Performed DNA cloning, protein expression, protein purification, compound screening using in silico docking, enzyme assay and surface plasmon resonance analysis, and mode of binding analysis. Primary role in preparation of the manuscript.

Signature
Date: 30/10/2014

Name of Co-Author: Kate L. Wegener

Contribution to the Paper:
Provided discussion in all aspects of structural analysis and interpretation of the results.

Signature
Date: 30/10/2014

Name of Co-Author: Grant W. Booker

Contribution to the Paper:
Provided direction and discussion in all aspects of in silico docking, structural analysis and biochemical assays.

Signature
Date: 30/10/2014

Name of Co-Author: Stephen W. Polyak

Contribution to the Paper:
Provided overall project direction and supervision of SPR analysis. Corresponding author for the manuscript.

Signature
Date: 30/10/2014
Fragment-Based Screening of *Mycobacterium tuberculosis* Dethiobiotin Synthetase Inhibitors

(Short title: Fragment inhibitors of Mycobacterial Dethiobiotin Synthetase)

Wanisa Salaemae¹, Kate L. Wegener¹, Grant W. Booker¹, and Steven W. Polyak¹*

¹Department of Biochemistry, School of Molecular and Biomedical Science, The University of Adelaide, South Australia, Australia

*Corresponding author

E-mail: steven.polyak@adelaide.edu.au (SP)
Abstract

Dethiobiotin synthetase (DTBS), the penultimate enzyme in biotin biosynthesis, is proposed as a promising new target for anti-TB drug development due to its essential role in tuberculi virulence and pathogenesis. The X-ray crystal structure of the enzyme from *Mycobacterium tuberculosis* (MtDTBS) in complex with CTP was employed for virtual screening of two compound libraries, 93,904 fragments and 57 CTP analogues. Compounds were ranked based on their predicted binding affinity and the top-ranked commercially available hits were purchased for further analysis. Surface plasmon resonance (SPR) and DTBS *in vitro* enzyme assays were performed in parallel to assess the potency of binding and inhibition, respectively. A total of six hits, including the known anti-cancer drug gemcitabine, showed encouraging ligand efficiency (0.2 - 0.3 kcal/mol/heavy atom). *In silico* docking studies predicted binding of hits in the phosphate-binding loop between the binding pockets for DAPA and NTP. This was confirmed by Lineweaver-Burk analysis that showed competitive binding of gemcitabine against both DAPA and MgATP. The identified “hot-spot” for inhibitor binding, and the variety of chemical structures predicted here, provide valuable starting points for the optimization of new lead compounds for anti-TB development.

Keywords

Biotin biosynthesis, Dethiobiotin synthetase, Fragment based drug discovery, *Mycobacterium tuberculosis*
Introduction

Tuberculosis (TB) remains a major health challenge causing considerable morbidity and mortality worldwide (1,2). According to the 2013 WHO report, each year 5.7 million new cases will be diagnosed and 1.3 million people will die (3). Currently, the 2015 targets set as part of the Millennium Development Goals for the reduction of TB mortality and incidence to 50% of the 1990 levels have nearly been reached (3). However, efforts to combat multidrug-resistant TB (MDR-TB; resistance to first-line drugs i.e. rifampin and isoniazid) are still well off-track (3). The incidence of MDR-TB remains high, accounting for 3.6% of newly diagnosed TB patients and 20% of existing TB patients (3). In addition, approximately 10% of patients with MDR-TB develop extensively drug-resistant TB (XDR-TB; MDR plus resistance to second-line drugs i.e. fluoroquinolone and aminoglycoside) (3). Bedaquiline, approved by the FDA in 2012, is the only anti-TB drug with a unique mode of action to become available for MDR- and XDR-TB treatment in the past 40 years, but even this drug has concerning side effects (4,5). Clearly, there is a critical need for safer antibiotics that are not subject to pre-existing resistance mechanisms to end the global drug resistant TB epidemic.

Since sequencing of the *Mycobacterium tuberculosis* H37Rv genome was completed in 1998 (6), many potential drug targets for the development of new anti-TB drugs have been proposed. These include enzymes associated with the synthesis of membrane lipids; a metabolic pathway that is a validated drug target via the efficacy of the anti-TB drug isoniazid (7). Biotin biosynthesis is one such potential pathway due to its essential role in membrane lipid synthesis (8-10). Biotin serves as an essential cofactor for biotin-dependent acyl-CoA carboxylases (ACCs) that play a key role in the production of mycolic acids and multimethyl-branched fatty
acids present in the cell envelope (11,12). Although all living cells require biotin, only microorganisms, plants and a few fungi can synthesize biotin de novo. Humans and other mammals lack this capability thus making the pathway an attractive antibiotic target (11). For *M. tuberculosis*, de novo synthesis is the sole source of biotin since tuberculi apparently lack a high affinity transporter for scavenging exogenous biotin (11,13-15). In addition, genetic knockout studies have revealed that biotin biosynthesis is a key metabolic process in all stages of the tuberculi life cycle through active growth, infection, and survival in the latent phase (16-18). This is complementary to previous findings for other lipid biosynthetic enzymes that serve as effective targets for treating latent *M. tuberculosis* (19,20). Together these data suggest that the biotin biosynthetic pathway is a worthwhile target for the development of new anti-TB drugs that are able to combat TB.

Dethiobiotin synthetase (DTBS; EC 6.3.3.3, encoded by *bioD*), the penultimate enzyme in the biotin biosynthetic pathway, is highly conserved amongst microorganisms (21). DTBS catalyzes the production of dethiobiotin (DTB) from (7R, 8S)-7,8-diaminonanoic acid (DAPA), CO2 and nucleotide triphosphate (NTP) in the presence of Mg2+ (11,21). The reaction mechanism consists of three steps; formation of N7-DAPA carbamate, formation of an intermediate mixed carbamic phosphoric acid anhydride by transfer of the γ-phosphoryl group of NTP to a carbamate oxygen, and finally closure of the ureido ring on biotin with the release of inorganic phosphate (proceeding through a tetrahedral intermediate) (22). Previous attempts to design inhibitors for DTBS from *Escherichia coli* based on an understanding of this reaction mechanism, and generating analogues of either substrates or reaction intermediates, showed disappointing activity with inhibition constants in the millimolar range (23,24). This suggested that alternative strategies for the identification of the inhibitors are required. Noteworthy, our
previous study on the binding of an ATP analogue, namely ATP-BoDiPY, to *M. tuberculosis* DTBS in the presence of varying concentrations of DAPA revealed that nucleotide bound to the pocket in *M. tuberculosis* DTBS independently from the adjacent binding of DAPA (Manuscript (25)). X-ray crystal structures of MtDTBS in its apo-form and in complex with either DAPA carbamate (PDB 3FMF) or CTP (PDB 4WOP) confirmed the presence of distinct preformed ligand-binding pockets for DAPA and NTP respectively ((26), Manuscript (25)). In contrast, a conformational change at the P-loop was observed upon the binding of ATP in the *E. coli* ortholog. Indeed, the structural comparison of DTBS from *M. tuberculosis* with orthologs from *E. coli, Helicobacter pylori*, and *Francisella tularensis* demonstrates that the main difference between them is at the C-terminal region, resulting in differently shaped nucleotide binding pockets and conformational changes resulting from ligand binding (27,28). Preformed binding pockets are a unique feature of MtDTBS, making the DAPA and NTP pockets attractive targets for anti-TB drug discovery using *in silico* docking.

In this study, two alternative approaches were applied in this study to identify DTBS inhibitors. The first of this, fragment based drug discovery (FBDD) has become a powerful tool for discovering novel inhibitors with low molecular weight (MW < 300 Da) and low chemical complexity. In FBDD, it is desirable to identify hits that have good ligand efficiency (LE), which is the calculated binding affinity per heavy atom of the ligand (29-31). An acceptable value of LE for drug candidates is ≥0.3 kcal/mol/heavy atom (31,32). These small structures can then be elaborated through chemical optimization into larger, more potent drug-like leads. In addition to FBDD, our previous studies revealed that CTP binds with higher affinity to *M. tuberculosis* DTBS than ATP (Manuscript (25)). This strong preference for CTP led us to screen CTP analogues as potential inhibitors for the mycobacterial enzyme. The structure of MtDTBS in
complex with CTP (PDB 4WOP) (Manuscript (25)) was employed in virtual screening of the two compound libraries. Hits were evaluated using surface plasmon resonance (SPR) binding analysis and a DTBS inhibition assay. Using this strategy, four fragments and two CTP analogues were identified with promising LE (0.2 – 0.3 kcal/mol/heavy atom) including the known anti-cancer drug gemcitabine (33). These hits provide a promising foundation for further chemical development of potent lead compounds with the potential to combat TB through the inhibition of biotin biosynthesis.

**Materials and Methods**

**Construction of Computational Compound Libraries**

A fragment library was constructed from the ZINC database website version 11 (UCSF). The “In Stock” database was filtered for compounds complicit with “rule of three” criteria (34) yielding a library of 93,904 compounds with logP ≤ 2.5, MW ≤ 250 Da, H-bond donors and acceptors ≤ 3, tPSA (Polar Surface Area) ≤ 60 Å² and number of rotatable bonds ≤ 3,. A separate CTP analogue library consisting of 57 compounds was constructed using a structural search tool on the Sigma-Aldrich website (www.sigmaaldrich.com) with the criteria of obtaining >50% structural similarity to CTP and a molecular weight of 150 to 350 Daltons. To generate the 3D co-ordinates for each compound, the canonical SMILES string of each compound, obtained from PubChem compound database (NCBI), were translated into files using an Online SMILES Translator and Structure File Generator software (http://cactus.nci.nih.gov/services/translate/) (CNI/NCDD Group).
In Silico Docking Experiment

The structure of DTBS from \textit{M. tuberculosis} (MtDTBS) in complex with CTP (PDB 4WOP) (Manuscript (25)), was employed as the template for \textit{in silico} screening. Docking experiments were carried out using AutoDock – Vina version 1.5.4 (UCSF) (35). AutoDockTools was employed to prepare the co-ordinates of template and ligands as PDBQT files from PDB file and to define the boundaries of the target space covering the binding pocket of the nucleotide substrate. The predicted ligand binding posses of ligands were visualized using UCSF Chimera 1.8.1 (36). Hits were chosen for further consideration were based on their low binding energy, predicted binding mode in the pocket and their commercial availability.

Preparation and Assay of DTBS Enzymes

The cloning, expression and purification of recombinant DTBS from both \textit{M. tuberculosis} and \textit{E. coli} have been previously described (Manuscript (25)). The activity of DTBS was determined using a fluorescence-polarization based DTBS enzyme assay according to the published procedures (Manuscript ((25)).

To measure the inhibitory activity of compounds, DTBS activity was measured with varying concentrations of compound. Here, the assay mixture contained saturating concentrations of DAPA and MgATP (at 0.1 and 0.3 mM respectively). The concentration of inhibitor that reduced DTBS activity by 50\% (\textit{IC$_{50}$}) was calculated by plotting the percentage of enzyme activity as a percentage of control activity versus the log of compound concentration. The data were fitted to a log [inhibitor] vs. response – variable slope (four parameters) model.
GraphPad Prism 6 (GraphPad Software, Inc.). The control reaction was performed in the absence of test compounds. The inhibition constant \((K_i)\) of the test compounds was calculated using Equation 1 (37), where \([S]\) is the substrate concentration, \(K_m\) is the Michaelis-Menten constant, and \(IC_{50}\) is the half maximal inhibitory concentration.

\[
K_i = \frac{IC_{50}}{(1 + [S]/K_m)}
\]

**Surface Plasmon Resonance Analysis**

SPR analysis was performed using a BIAcore T100 instrument. Proteins were diluted to 0.2 mg ml\(^{-1}\) in 0.01 M NaOAc buffer (pH 5.2) and immobilized to the surface of a CM5 sensor chip using a constant flow rate of 5 µl min\(^{-1}\) and contact time of 420 s in order to reach approximately 10,000 final response units (RU). Flow cell 1 was left blank (ie no protein) to correct for bulk refractive index changes, flow cells 2 and 3 were used to immobilized \(Mt\)DTBS and \(Ec\)DTBS respectively. For water-soluble compounds, binding experiments were performed in running buffer (10 mM HEPES (pH 7.4), 150 mM NaCl and 0.005% (v/v) surfactant P20). For less water-soluble compounds, 2.5% dimethyl sulfoxide (DMSO) was added to the binding buffer to alleviate solubility problems. To correct for small dilution errors during the preparation of compounds, a solvent correction was performed using standards containing 2% to 3.3% (v/v) DMSO. MgATP or MgCTP, at concentrations of 0.3 mM, were passed over the sensor surfaces as controls. Hits were initially identified by SPR using a single concentration of compound (1 mM). Hits were then subjected to further binding analysis using varying concentrations in order to determine the binding affinity \((K_D)\). The \(K_D\) was determined using a steady-state kinetic model with BIAcore T100 evaluation software (GE Healthcare). Ligand efficiency (LE) of hits was
calculated using Equation 2 (31,32), where R is the ideal gas constant (1.987 x 10^-3 kcal/K/mole), T is the temperature in Kelvin (K), ln(K_D) is the natural log of the binding affinity (K_D), and HA is number of heavy (non-hydrogen) atoms.

Equation 2

LE = ((- RT x ln(K_D)) / HA

Mechanism of Inhibition

The enzyme activity was monitored using the DTBS enzyme assay with varying concentrations of both test compound and substrate. The data were fitted to the Lineweaver-Burk plot in Equation 3 where \( v \) is the reaction velocity, \( K_m \) is the Michaelis-Menten constant, \( V_{max} \) is the maximum reaction velocity, and \([S]\) is the substrate concentration (38). The plot of \( 1/v \) against \( 1/[S] \) was generated using linear regression analysis in GraphPad Prism 6 (GraphPad Software, Inc.).

Equation 3

\[
\frac{1}{v} = \left( \frac{K_m}{V_{max}} \right)(1/[S]) + \frac{1}{V_{max}}
\]

Results and Discussion

Hit Identification

Dethiobiotin synthetase (DTBS) is a highly promising anti-tuberculi target due to its essential metabolic role in bacteria, the absence of a human homologue, its well-characterized reaction mechanism, and the availability of structural information. The recently determined X-
ray crystal structure of MtDTBS in complex with CTP (PDB 4WOP) (Manuscript (25)) has the complete secondary structure and represents the stereometric changes upon the ligand binding, which provides a high possibility to predict the binding of ligands comparably to the actual binding. This structure was employed in in silico docking experiments, screening both a large chemical fragment library, as well as a smaller library of CTP analogues. Fragment based drug discovery offers the ability to screen a high proportion of chemical space, while the CTP analogues were included to exploit our recent finding that, while MtDTBS can utilize a broad range of nucleotide triphosphate substrates, it has a clear preference for binding to CTP (Manuscript (25)). Together, these two approaches were pursued with the goal of identifying tractable starting structures for future medicinal chemistry programs. Both MtDTBS and EcDTBS were used to address species selectivity of inhibitors. A four-stage screening cascade was devised to identify hits that bound DTBS (Figure 1). Hits were initially identified using virtual screening approaches (Stage 1), before being evaluated experimentally by two methods, namely an SPR binding assay and a DTBS in vitro enzyme assay. Here, only MtDTBS was investigated. These assays were performed in parallel using a single concentration of compound (Stage 2). Positive hits were then retested for dose-responsiveness (Stage 3). To consider specificity, both MtDTBS and EcDTBS were used at this stage. Finally, the mechanism of action was addressed using Lineweaver-Burk analysis and in silico docking (Stage 4).

A total of 93,904 ZINC fragments and 57 CTP analogues were screened against the MtDTBS structure (PDB 4WOP) (Manuscript (25)). Here, the target space for docking (x = 18, y = 14, z = 14 Å) was defined to cover two discrete subsites of the NTP binding pocket: the phosphate-binding region (also known as the P-loop) and the nucleoside-binding site (Figure 2A&B). AutoDock – Vina (35) was employed to calculate the predicted binding energy of each
compound. The docking software considered nine distinct binding poses for each molecule and ranked these according to their predicted affinity. As expected, the predicted mode of binding for CTP with the lowest binding energy (-9.0 kcal/mol) was in excellent agreement with the X-ray crystal data (Figure 2C). Furthermore, the predicted affinity of CTP was higher than that of ATP (-8.5 kcal/mol). This observed preference for CTP was in agreement with our recent finding using surface plasmon resonance (SPR) analysis and a competitive ATP-binding assay (Manuscript (25)). Hence, CTP was incorporated in virtual screening experiments as a positive control. The compounds from the virtual libraries were ranked according to their predicted binding energy. Of the nine predicted binding poses, the one with the lowest predicted binding energy was used for ranking and compared to the positive control. Those compounds that were predicted to have lower binding energy than the CTP control were selected for further analysis. A total of twenty-six commercially available compounds (15 ZINC fragments and 11 CTP analogues (Supplemental Table S1 and S2)) were then purchased for further testing.

**Experimental Hit Validation**

The twenty-six virtual hits were tested for binding and inhibitory activity using a single concentration with the compounds segregated into two separate classes depending upon their solubility in water. The SPR assays were performed using test compounds at 1 mM (Supplemental Fig. S1). Seven DMSO-solubilized fragments (namely B1, B3, B7, B9, B10, B13, and B14) yielded ≥ 2-fold increase in response units above the MgCTP and MgATP controls. Four water-soluble CTP analogues (namely CT3, CT6, CT7, and CT12) were considered to yield positive interaction to MfDTBS (ie ≥ 5 RU above the baseline), although their binding responses
were lower than that of the MgCTP and MgATP controls. In parallel to the SPR analysis, the twenty-six hits were also tested using the DTBS \textit{in vitro} enzyme assay (Supplemental Fig. S2). The $IC_{50}$ values were determined in an assay using saturating concentrations of DAPA, NaHCO$_3$, and MgATP substrates (ie 3x, 3x, and 10x $K_m$ respectively). Water soluble compounds were tested with this assay at a concentration of 15 mM, showing two water-soluble CTP analogues, namely CT6, and CT7, completely abolished the \textit{MtDTBS} activity or inhibited $> 70\%$ of the \textit{EcDTBS} (Supplemental Fig. S2A). The remaining compounds were tested in 1% DMSO, at the maximum compound concentration attainable (Supplemental Fig. S2B). The concentrations of DMSO up to 2.5% were shown to be tolerated by the enzyme without compromising activity (Supplemental Fig. S3). Maximal compound concentration ranged from 0.18 – 2.39 mM for the different fragments, and therefore the absence of any significant inhibition may be the result of inadequate concentrations. Therefore, hits from the group were selected for further investigation, based on the SPR binding data.

Encouraged by the initial results, the potency of binding was investigated for eleven compounds using dose-response curves (namely B1, B3, B7, B9, B10, B13, B14, CT3, CT6, CT7 and CT12). Dissociation equilibration constants ($K_D$) could be determined for six of the eleven compounds, using SPR with a steady state binding model, and measuring dose responsive binding. B1, B3, B7, B9, CT6 and CT7 bound to \textit{MtDTBS} with $K_D$ of 0.6 - 3.1 mM and \textit{EcDTBS} with $K_D$ of 0.4 – 4.0 mM (Table 1 and Supplemental Fig. S4). However, given their relatively small molecular mass, the compounds displayed encouraging LE in a range of 0.2 - 0.3 kcal/mol/heavy atom (Table 1). The $IC_{50}$ values were determined by a DTBS \textit{in vitro} enzyme assay for CT6 and CT7, showing the inhibitory activity with low millimolar level (Table 2). These are in excellent agreement with $K_D$ values obtained from SPR analysis. $K_i$ values were then
calculated assuming competitive binding with the ligands DAPA and MgATP, as predicted by the docking studies. The $K_i$ values for CT6 and CT7 for MgATP were calculated to be 0.7 ± 0.1 mM and 0.7 ± 0.1 mM, and 1.7 ± 0.2 mM and 1.6 ± 0.2 mM for DAPA. Thus, six confirmed binding hits (B1, B3, B7, B9, CT6, CT7) were obtained from this screening campaign with LEs in the range currently acceptable for hits ≥ 0.3 kcal/mol/heavy atom (31,32). No significant differences in $K_D$ and $IC_{50}$ values were found between $Mt$DTBS vs $Ec$DTBS ($p > 0.05$) (Table 1, 2 and Figure 3).

**Binding Mode of Hits**

To define the binding mode of the hits, *in silico* docking experiments were performed against the $Mt$DTBS structure (PDB 4WOP) (Manuscript (25)) using an enlarged volume of the target space (x = 22, y = 18, z = 14 Å) encompassing both DAPA and NTP binding pockets. Interestingly, all of the hits docked at the junction between the two binding pockets (Figure 4). This finding implied that the phosphate-binding loop Gly12-Thr16 (25,26) and the N’7 DAPA-binding region formed by Lys37-Thr41 and Gly109-Gly111 (26), at the junction of DAPA and NTP binding pockets, is a “hot-spot” for the majority of the hits binding to $Mt$DTBS. To verify this, the binding modes of CT6 and CT7 were determined experimentally by Lineweaver-Burk analysis. Here the activity of $Mt$DTBS was assayed with varying concentrations of inhibitor as well as either DAPA or MgATP substrates (Figure 5). For CT6, the compound showed competitive inhibition against both MgATP and DAPA as the $K_m$ for the ligands were affected by increasing concentrations of inhibitor. In contrast, mixed inhibition was observed for CT7,
with the compound affecting both $K_m$ and $V_{max}$. Together, the *in silico* analysis and kinetic data imply that the inhibitors occupy both the ATP and DAPA pockets, most likely at the P-loop.

**Conclusions**

Herein, we report the discovery of six *Mt*DTBS-binding compounds (CT6, CT7, B1, B3, B7, and B9), two of which have been shown to be competitive inhibitors of *Mt*DTBS with encouraging ligand efficiency (0.2 – 0.3 kcal/mol/heavy atom). *In silico* docking and Lineweaver-Burk analysis revealed the binding of these hits at the junction of the two binding pockets located at the phosphate-binding loop. This region is considered a protein “hot-spot” where ligands frequently bind (30). The identification of hot-spot area provides useful structural information for prediction favorable functional groups of drug-like leads while maintaining good LEs (30). The variable chemical groups of hits are good starting materials for further designing potent inhibitors. Interestingly, these hits are also beneficial for potential development of DTBS inhibitors with a broad spectrum of activity.

Of the six hits, CT6 seems to be the most promising starting point for the development of potent inhibitors against *Mt*DTBS. CT6, also known as gemcitabine, is one of drugs available for the treatment of solid tumors (33). The mechanistic action of the phosphorylated form of this CTP analogue (dFcCTP) has a well characterized role in inhibition of DNA synthesis (33). Interestingly, due to slow intracellular elimination, an effective plasma gemcitabine concentration can be maintained for a prolonged period of time, resulting in a higher anti-tumor activity compared to other anti-tumor drugs (33,39,40). In addition, this compound is taken up
into the cell via human nucleoside transporters (hNTs), and is largely excreted into the urine (40). These validate an excellent drug-like action of CT6 especially in terms of transporting into and excreting out of cells. Considering the fact that gemcitabine is a small molecule (MW < 300 Da) and binds at the junction of two binding pockets, its chemical structure will be useful starting point for medicinal chemistry programs to develop a series of more potent inhibitors. A previous study on the design and synthesis of inhibitors targeting EcDTBS (24), revealed that the carboxylic acid at C’1 end of DAPA carbamate is essential for forming a strong H-bonding network to the enzyme. It could not be replaced with alcohol, aldehyde, ester or lactone moiety that is consistent with the observation in the available crystal structures (24,41). Therefore, the possibility of derivatising DAPA at its amino end, by incorporation of gemcitabine or similar chemical groups, may increase the overall binding. Derivatives of gemcitabine, with increased potency for DTBS while engineered out the anti-tumor activity, may provide new insights into the treatment of tuberculosis.

Acknowledgement

The authors are grateful to Karen Druart for constructing the ZINC fragment library and Rebecca Chao for assistance generating the CTP analogue library.

References

5. FDA. (2012) Drugs for the treatment of tuberculosis, including drug resistant tuberculosis. Washington, DC


37. Cheng, Y., and Prusoff, W. (1973) Relationship between the inhibition constant (Ki) and the concentration of inhibitor which causes 50 percent inhibition (I50) of an enzymatic reaction. Biochem. Pharmacol. 22, 3099-3108


**Figure Legends**

**Fig. 1 Four-stage screening cascade for MtDTBS inhibitors.** The cascade consists of *in silico* hit identification, and two stages of experimental validation involving SPR binding and kinetic analysis. Finally the mode of binding was addressed by further docking studies and Lineweaver-Burk analysis.

**Fig. 2 In silico docking.** A, The X-ray crystal structure of dimeric MtDTBS (PDB 4WOP) is shown with CTP (pink) bound to one subunit and the grid box employed for *in silico* docking.
studies (shown in blue, green, red box) displayed on the partnering subunit. B, Close up view of the CTP binding site with the cytosine pocket shown in pink mesh and the phosphate binding region in orange mesh. C, A comparison of the predicted binding pose of CTP (blue) and crystallized CTP (pink) in MtDTBS.

**Fig. 3 Inhibition activity of CT6 and CT7 against M. tuberculosis and E. coli DTBS.** The selectivity of (A) CT6 and (B) CT7 was determined against MtDTBS (●) and EcDTBS (■). The inhibitory activity was calculated as the percentage of enzyme activity correlated to 100% of total enzyme activity without compound in the reaction containing a saturating concentration of DAPA, NaHCO₃, and MgATP substrates. Determined IC₅₀ values of CT6 were 7.3 ± 0.9 mM to MtDTBS and 8.9 ± 0.7 mM to EcDTBS (p = 0.23) while the IC₅₀ values of CT7 were calculated to be 7.0 ± 0.9 mM to MtDTBS and 12.0 ± 1.1 mM to EcDTBS (p = 0.06). Experiments were performed in triplicate. The mean ± SEM are shown.

**Fig. 4 Predicted binding site of validated hits to MtDTBS.** The poses with the lowest binding energy are shown for (A) CT6, (B) CT7, (C) B1, (D) B3, (E) B7, and (F) B9 in MtDTBS. DAPA carbamate (shown in red) (PDB 3FMA) was superimposed onto the CTP bound MtDTBS structure (PDB 4WOP). The docking studies predict that the inhibitors bind in the phosphate binding loop juxtaposed the DAPA and NTP binding pockets.

**Fig. 5 Mechanism of inhibition for CT6 and CT7.** Lineweaver-Burk analysis was performed by measuring MtDTBS activity with varying concentrations of both inhibitor and ligands, either MgATP (panels A and C) or DAPA (panels B and D). The concentration of inhibitor included in the assays were 0 mM (blue), 7.5 mM (red) and 10 mM (black).
Tables

Table 1 Binding analysis of Hits by SPR

Table 2 Hits validation by DTBS inhibition enzyme assay

Supporting Information

Table S1 Virtual hits from ZINC fragment library identified by in silico docking

Table S2 Virtual hits from CTP analogue library identified by in silico docking

Fig. S1 Initial hits validation by SPR analysis. A single concentration (at 1 mM) of fragments and CTP analogues was tested for binding to MtDTBS. MgATP and MgCTP at 0.3 mM concentration were used as positive controls. A and B, seven fragments namely B1, B3, B7, B9, B10, 7535142, and 7255880 showed the response (>40 RU) higher than the controls. C, Four hits from CTP analogs namely CT3, CT6, CT7, CT12 (shown in green, blue, pink and red) showed the response ≥5 RU although it is less than that of the controls. These compounds had been considered to bind MtDTBS.

Fig. S2 Initial hits validation by DTBS in vitro enzyme assay. The inhibitory activity of (A) water-soluble compounds (at 15 mM) and (B) DMSO-solubilized compounds (at < 3 mM) were tested against MtDTBS (shown in black bar). The inhibitory activity was presented through the percentage of the remaining enzyme activity correlated to 100% of total enzyme activity without compound in the reaction (shown as no CT) containing a saturating concentration of DAPA,
NaHCO₃, and MgATP substrates. C, only CT6, and CT7 could abolish the MtDTBS activity and were further tested against EcDTBS (shown in white bar), showing >70% inhibition to EcDTBS. The inhibitory values of CT6, and CT7 were obtained from three independent experiments.

**Fig. S3 Effect of DMSO concentration to DTBS enzyme activity.** The 2.5% is the highest concentration of DMSO which maintains >90% activity of MtDTBS. This concentration of DMSO had been used in both SPR analysis and inhibition DTBS enzyme assay.

**Fig. S4 SPR binding analysis.** The binding isotherms of CT6, CT7, B1, B3, B7 and B9 to MtDTBS were analyzed using an affinity-steady state 1:1 model. The binding constant (Kᵰ) as reported in Table 1 was calculated from three independent SPR experiments.
Stage 1: virtual screening
93,904 fragments
57 CTP analogues
In silico docking
MTBS

Stage 2: 1st validation - single concentration
15 fragments
11 CTP analogues
Surface plasmon resonance
Inhibition DTBS enzyme assay
MTBS

Stage 3: 2nd validation - dose response
7 fragments
4 CTP analogues
Surface plasmon resonance
Inhibition DTBS enzyme assay
MTBS
EcDTBS

Stage 4: Characterization of binding mode
4 fragments
2 CTP analogues
In silico docking
Linweaver-Burk analysis
MTBS

Hits (CT6, CT7, B1, B3, B7, B9)
Figure 2

A

B

C
Figure 3

A

% DTBS activity

0 50 100 150

-3.5 -3.0 -2.5 -2.0 -1.5 -1.0

Log (CT6 (M))

B

% DTBS activity

0 50 100 150

-3.5 -3.0 -2.5 -2.0 -1.5 -1.0

Log (CT7 (M))
Figure 4
Figure 5

A  MgATP
CT6

B  DAPA

C  MgATP
CT7

D  MgATP
Table 1 Binding analysis of Hits by SPR

<table>
<thead>
<tr>
<th>ID</th>
<th>Structure</th>
<th>$K_D$ (mM)</th>
<th>LE (kcal/mol/heavy atom)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MtDTBS</td>
<td>EcDTBS</td>
</tr>
<tr>
<td>B1</td>
<td><img src="image1" alt="Structure" /></td>
<td>1.2 ± 0.2</td>
<td>0.8 ± 0.1</td>
</tr>
<tr>
<td>B3</td>
<td><img src="image2" alt="Structure" /></td>
<td>0.6 ± 0.1</td>
<td>0.6 ± 0.0</td>
</tr>
<tr>
<td>B7</td>
<td><img src="image3" alt="Structure" /></td>
<td>0.7 ± 0.0</td>
<td>0.5 ± 0.1</td>
</tr>
<tr>
<td>B9</td>
<td><img src="image4" alt="Structure" /></td>
<td>0.6 ± 0.0</td>
<td>0.4 ± 0.2</td>
</tr>
<tr>
<td>CT6</td>
<td><img src="image5" alt="Structure" /></td>
<td>3.1 ± 1.2</td>
<td>4.0 ± 0.9</td>
</tr>
<tr>
<td>CT7</td>
<td><img src="image6" alt="Structure" /></td>
<td>2.3 ± 0.0</td>
<td>2.6 ± 0.8</td>
</tr>
</tbody>
</table>

All $K_D$ values were obtained from three independent experiments. B represents the compound from ZINC fragment library while CT represents the compound from CTP analogue library. LE (Ligand efficiency) was calculated using the mean of $K_D$ values. N/D represents not determined. Experiments were performed in triplicate. The mean ± SEM are shown.
**Table 2** Hits validation by DTBS inhibition enzyme assay

<table>
<thead>
<tr>
<th>ID</th>
<th>Structure</th>
<th>( IC_{50} ) (mM)</th>
<th>MtDTBS</th>
<th>EcDTBS</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT6</td>
<td><img src="structure1.png" alt="Structure" /></td>
<td>7.3 ± 0.9</td>
<td>8.9 ± 0.7</td>
<td></td>
</tr>
<tr>
<td>CT7</td>
<td><img src="structure2.png" alt="Structure" /></td>
<td>7.0 ± 0.9</td>
<td>11.0 ± 1.1</td>
<td></td>
</tr>
</tbody>
</table>

All values were obtained from three independent experiments. CT represents the compound from CTP analogue library.
 Supporting Information 

**Table S1** Virtual hits from ZINC fragment library identified by *in silico* docking

<table>
<thead>
<tr>
<th>Compound</th>
<th>Structure</th>
<th>ZINC ID</th>
<th>Affinity (kcal/mol)</th>
<th>MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td><img src="image1" alt="Structure B1" /></td>
<td>65409233</td>
<td>-8.1</td>
<td>239.3</td>
</tr>
<tr>
<td>B2</td>
<td><img src="image2" alt="Structure B2" /></td>
<td>00487559</td>
<td>-9.2</td>
<td>234.3</td>
</tr>
<tr>
<td>B3</td>
<td><img src="image3" alt="Structure B3" /></td>
<td>08771427</td>
<td>-8.8</td>
<td>230.2</td>
</tr>
<tr>
<td>B4</td>
<td><img src="image4" alt="Structure B4" /></td>
<td>08738503</td>
<td>-8.8</td>
<td>226.3</td>
</tr>
<tr>
<td>B5</td>
<td><img src="image5" alt="Structure B5" /></td>
<td>00262731</td>
<td>-8.0</td>
<td>239.3</td>
</tr>
<tr>
<td>B6</td>
<td><img src="image6" alt="Structure B6" /></td>
<td>00519029</td>
<td>-8.2</td>
<td>234.3</td>
</tr>
<tr>
<td>B7</td>
<td><img src="image7" alt="Structure B7" /></td>
<td>12956616</td>
<td>-8.3</td>
<td>242.2</td>
</tr>
</tbody>
</table>
Table S1 (continue) Virtual hits from ZINC fragment library identified by *in silico* docking

<table>
<thead>
<tr>
<th>Compound</th>
<th>Structure</th>
<th>ZINC ID</th>
<th>Affinity (kcal/mol)</th>
<th>MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>B8</td>
<td><img src="image1.png" alt="Structure" /></td>
<td>00083454</td>
<td>-8.8</td>
<td>238.3</td>
</tr>
<tr>
<td>B9</td>
<td><img src="image2.png" alt="Structure" /></td>
<td>04114260</td>
<td>-8.1</td>
<td>230.3</td>
</tr>
<tr>
<td>B10</td>
<td><img src="image3.png" alt="Structure" /></td>
<td>04114249</td>
<td>-8.1</td>
<td>244.3</td>
</tr>
<tr>
<td>B11</td>
<td><img src="image4.png" alt="Structure" /></td>
<td>00054281</td>
<td>-8.0</td>
<td>242.3</td>
</tr>
<tr>
<td>B12</td>
<td><img src="image5.png" alt="Structure" /></td>
<td>00400596</td>
<td>-8.1</td>
<td>244.3</td>
</tr>
<tr>
<td>B13</td>
<td><img src="image6.png" alt="Structure" /></td>
<td>04958092</td>
<td>-7.6</td>
<td>202.2</td>
</tr>
<tr>
<td>B14</td>
<td><img src="image7.png" alt="Structure" /></td>
<td>20149317</td>
<td>-8.9</td>
<td>244.3</td>
</tr>
</tbody>
</table>
Table S1 (continue) Virtual hits from ZINC fragment library identified by *in silico* docking

<table>
<thead>
<tr>
<th>Compound</th>
<th>Structure</th>
<th>ZINC ID</th>
<th>Affinity (kcal/mol)</th>
<th>MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>B15</td>
<td><img src="image" alt="Structure" /></td>
<td>23382588</td>
<td>-8.8</td>
<td>232.2</td>
</tr>
</tbody>
</table>
Table S2 Virtual hits from CTP analogue library identified by *in silico* docking

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Structure/Name</th>
<th>% Similarity to CTP</th>
<th>Affinity (kcal/mol)</th>
<th>MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT1</td>
<td>Lamivudine</td>
<td>56</td>
<td>-6.1</td>
<td>229.3</td>
</tr>
<tr>
<td>CT2</td>
<td>3'-Azido-3'-deoxythymidine</td>
<td>57</td>
<td>-7.5</td>
<td>267.2</td>
</tr>
<tr>
<td>CT3</td>
<td>5-Fluorocytidine</td>
<td>71</td>
<td>-7.5</td>
<td>261.2</td>
</tr>
<tr>
<td>CT4</td>
<td>Zubularine</td>
<td>74</td>
<td>-6.5</td>
<td>228.2</td>
</tr>
<tr>
<td>CT5</td>
<td>2'-C-Methylcytidine</td>
<td>78</td>
<td>-7.2</td>
<td>257.2</td>
</tr>
<tr>
<td>CT6</td>
<td>Gemcitabine hydrochloride</td>
<td>68</td>
<td>-7.3</td>
<td>299.7</td>
</tr>
<tr>
<td>CT7*</td>
<td>2'-Deoxycytidine hydrochloride</td>
<td>75</td>
<td>-7.0</td>
<td>263.7</td>
</tr>
</tbody>
</table>

*ZINC18286013
### Table S2 (continue) Hits from CTP analogue library identified by *in silico* docking

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Structure/Name</th>
<th>% Similarity to CTP</th>
<th>Affinity (kcal/mol)</th>
<th>MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT8</td>
<td><img src="image" alt="CT8 Structure" /> Cytidine 3',5'-cyclic monophosphate</td>
<td>76</td>
<td>-8.0</td>
<td>327.2</td>
</tr>
<tr>
<td>CT9</td>
<td><img src="image" alt="CT9 Structure" /> Cytidine 2':3'-cyclic monophosphate</td>
<td>73</td>
<td>-8.1</td>
<td>327.2</td>
</tr>
<tr>
<td>CT10</td>
<td><img src="image" alt="CT10 Structure" /> 2',3'-dideoxy-3'-fluorouridine</td>
<td>53</td>
<td>-6.7</td>
<td>230.2</td>
</tr>
<tr>
<td>CT12</td>
<td><img src="image" alt="CT12 Structure" /> Ancitabine hydrochloride</td>
<td>52</td>
<td>-7.1</td>
<td>261.7</td>
</tr>
</tbody>
</table>
Fig. S1 Initial hits validation by SPR analysis. A single concentration (at 1 mM) of fragments and CTP analogues was tested for binding to MtDTBS. MgATP and MgCTP at 0.3 mM concentration were used as positive controls. A and B, seven fragments namely B1, B3, B7, B9, B10, B13, and B14 showed the response higher than the controls. C, Four hits from CTP analogues namely CT3, CT6, CT7, CT12 (shown in green, blue, pink and red) showed the response ≥ 5 RU although it is less than that of the controls. These compounds had been considered to bind MtDTBS.
Fig. S2 Initial hits validation by DTBS in vitro enzyme assay. The inhibitory activity of (A) water-soluble compounds (at 15 mM) and (B) DMSO-solubilized compounds (at 3 mM, 1% final DMSO concentration) were tested against MtDTBS (shown in black bar). The inhibitory activity was presented through the percentage of the remaining enzyme activity correlated to 100% of total enzyme activity without compound in the reaction (shown as no CT) containing a saturating concentration of DAPA, NaHCO$_3$, and MgATP substrates. C, CT6, and CT7 abolished MtDTBS activity and were further tested against EcDTBS (shown in white bar), showing $>70\%$ inhibition to EcDTBS. The inhibitory values of CT6, and CT7 were obtained from three independent experiments.
Fig. S3 Effect of DMSO concentration to DTBS enzyme activity. The 2.5% is the highest concentration of DMSO which maintains > 90% activity of MtDTBS. This concentration of DMSO had been used in both SPR analysis and inhibition DTBS enzyme assay.
**Fig. S4 SPR binding analysis.** The binding isotherms of (A) CT6, (B) CT7, (C) B1, (D) B3, (E) B7 and (F) B9 to MtDTBS were analyzed using an affinity-steady state 1:1 model. The binding constant ($K_D$) as reported in Table 1 was calculated from three independent SPR experiments.
CHAPTER 6

FINAL DISCUSSION AND FUTURE DIRECTIONS
6.1 Lead Optimization

From the screening of MtDTBS inhibitors in Chapter 5, six hits were identified to bind the protein “hot spot” located in the phosphate-binding loop at the junction of the two ligand binding pockets. The hits have good ligand efficiency (0.2-0.3 kcal/mol/heavy atom) and are potentially suitable for further chemical optimization. In particular, the anti-cancer agent CT6 (gemcitabine) clearly showed *in vitro* inhibitory activity against the enzyme, suggesting derivatives of this drug could be applied for new anti-TB agents. Overall, these findings fulfill the goal of this study, which was to discover novel inhibitors of DTBS from *M. tuberculosis*.

Since the predicted binding sites of the hits are located at the junction of the two ligand binding pockets, there are two possibilities to generate larger molecules with expected higher binding affinity; either by growing the hits toward the DAPA or NTP binding sites. However, I argue that growing hits toward the NTP binding site is less attractive due to an increasing chance of toxicity of the leads by targeting numerous NTP-dependent enzymes including human orthologs (see section 6.2 for more detail). In contrast, DAPA is highly specific for DTBS for which there is no equivalent human homologue. The DAPA binding pocket is distinctive with a deeper pocket compared to the shallower and more open shape of the NTP binding pocket. Therefore, I propose to improve the *in vitro* potency of the hits toward the DAPA binding site. Previous studies that designed inhibitors targeting DTBS from *E. coli* revealed that the carboxylic acids at both C’1 end and at N’7 of DAPA carbamate are essential for forming H-bonds to the protein backbone, and these could not be replaced with alcohol, aldehyde, ester or lactone moieties (148). These bonding interactions are also observed in MtDTBS (PDB ID: 3FMA) where the H-bonding network stabilizes the
ligand-protein interaction (Figure 6.1) (125,148). Meanwhile, the center of the pocket, formed by the P71-A73, G111-E116 and T123-D126 loops, does not provide any specific hydrogen-bonding interactions to the ligand as most of these amino acid side chains face out of the pocket. As proof of concept, the first inhibitor that can be considered for chemical synthesis is CT6-1 (Figure 6.2). Here DAPA carbamate is joined onto the CT6 structure to yield a bi-substrate inhibitor. The chemical properties of CT6-1 were calculated using an online calculation of molecular properties and bioactivity score (http://www.molinspiration.com/cgi-bin/properties) (Molinspiration Cheminformatics, Slovak Republic), showing MW = 400.4 Da, clogP = 0.57, H-bond donors = 4, H-bond acceptors = 9, rotatable bonds = 7, and tPSA (Polar Surface Area) = 136.6 Å². Therefore, CT6-1 still abides by Lipinski’s drug like criteria of MW ≤ 500 Da, logP < 5, H-bond donors < 5, and H-bond acceptors < 10 (195). The moderate lipophilicity (clogP < 3) is also favorable for anti-TB drug development and will aid drug penetration through the mycobacterial cell envelope (176).

To understand the binding mode of the proposed compound, in silico docking is the most convenient tool that can be applied during the early stage drug design. The 3D structure of CT6-1 was generated by translation of the canonical SMILES that was determined using a JME Drawing Applet (www.chem.ucalgary.ca/courses/351/jme/applet.html) (Molinspiration Cheminformatics, Slovak Republic) into a PDB file using an Online SMILES Translator and Structure File Generator software (http://cactus.nci.nih.gov/services/translate/) (CNI/NCDD Group). The docking experiment was performed as described previously in Chapter 5 and revealed that the H-bonding network at the C’1 end of CT6-1 is consistent with that observed in the crystal structure of DTBS in complex with DAPA carbamate, with the exception that the O’3 atom of CT6-1 forms a H-bond to a side chain of Lys37 rather than the peptide backbone of Thr41 (Figure 6.3). The
alkyl chain of CT6-1 provides multiple hydrophobic interactions with the hydrophobic center of the DAPA binding pocket (Figure 6.4). These additional interactions provide CT6-1 with a more favorable predicted binding affinity (-8.8 kcal/mol) compared to the original parent compound CT6 (-7.4 kcal/mol). Furthermore, chemical optimization of CT6-1 by the introduction of various functional groups at positions R1, R2, R3 and R4 (see Figure 6.2) can generate extra H-bond interactions to the side chain of amino acids located at the protein binding hot spot such as Lys15, Thr16, Lys37, Thr41, Asp49 and Glu108 (Figure 6.3). This will generate the potent inhibitors for MtDTBS. However, for anti-TB drug development, it is necessary to test the toxicity of the compound that might occur due to its binding to off-targets.

Figure 6.1 H-bonding network in MtDTBS. DAPA carbamate intermediate (shown in red) from the MtDTBS structure (PDB ID: 3FMA) was superimposed to our CTP bound MtDTBS structure. The
binding site of CT6 (shown in yellow) was obtained from docking experiment, showing to bind at the hot spot. The protein formed multiple interactions to DAPA carbamate. In contrast, only one H-bond is predicted to form between O’2 of CT6 and the side chain of Lys37.

Figure 6.2 Structure of proposed CT6-1 lead. The lead optimization is performed here by combination of CT6 (shown in black) and DAPA carbamate (shown in blue). The red circle R1, R2, R3 and R4 are proposed for further optimization in order to provide extra interactions to the enzyme.
Figure 6.3 Predicted mode of binding for CT6-1. The binding of proposed compound CT6-1 lead (shown in purple) was obtained from docking experiment, indicating that the compound formed H-bond network at the carboxyl end (C’1) with the backbone of the enzyme same as those observed in DAPA carbamate. Additionally, the O’1 atom of CT6-1 is predicted to form a H-bond to a side chain of Lys37.
Figure 6.4 Ligplot of predicted binding of CT6-1 within the MtDTBS active site.

Predicted residues of MtDTBS (colored in brown) interacting with CT6-1 (colored in purple) are shown in this figure. Hydrogen bonds are shown in green dotted lines. Hydrophobic interactions are...
shown in brick red sunbursts. Atoms are colored by type (Carbon, black; Oxygen, red; Nitrogen, blue; Fluorine, green).

6.2 NTP-Dependent Enzymes

There are many enzymes in the human proteome that require a nucleotide triphosphate (NTP), mainly ATP or GTP. They can be classified into different EC (Enzyme Commission) groups based upon their biological functions (Table 6.1). Currently, the similarity of NTP binding pockets from all NTP-dependent enzymes has not been extensively investigated. However, the feasible binding of NTP to numerous enzymes has raised concerns about specificity and toxicity of the hits identified in this study. As discussed in Chapter 4, MtDTBS can utilize a broad selection of NTPs as a substrate where three phosphate groups of the NTP provide multiple electrostatic interactions with the enzyme via the phosphate-binding loop (Gly-X-X-X-X-Gly-Lys-Ser/Thr, also known as Walker A motif). Likewise, the interactions at the P-loop are found in other NTP-dependent enzymes (196-198), implying the possibility of the hits identified in this study to bind off-target at the P-loop of other NTP-dependent enzymes.

A number of human proteins have revealed a structural similarity to DTBS. From structural comparisons of DTBS from E. coli structure available in the Protein Data Bank (PDB), the monomers of three proteins showed similarity in three-dimensional folds as a single globular α/β domain, namely nitrogenase iron protein [EC 1.18.6.1], H-ras p21 [EC 3.6.5.2] and adenylosuccinate synthase [EC 6.3.4.4] (Table 6.1, highlighted in red) (125-127,199). Indeed, many of the helices and strands are superimposable with root mean square (rms) deviations of the Ca atoms at 2.3 - 2.8 Å (127). Using the P-loop sequence from DTBS
as the search query, concerned sequences were also identified in three additional ATP-dependent enzymes, namely cobyricin acid a,c-diamide synthase, cobyric acid synthase and glutamine synthetase (Table 6.1, highlighted in blue) (127,200). Noteworthy, only H-ras p21, adenylosuccinate synthase and glutamine synthetase are essential enzymes in humans. These should be considered for potential off target effects during the optimization of DTBS inhibitors (201-207). H-ras p21 is a Homo sapien enzyme belonging to the GTPase family in Ras subfamily that plays an important role in cell growth regulation, as its GTP bound active form functions as an activator of a Raf kinase in the MAPK/ERK pathway (206,208). Mutations in H-ras interfere with GTP binding/or hydrolysis leading to a constitutively active protein, thus causing uncontrollable cell growth in lung and thyroid cancer and leukemia (209-213). Therefore, H-ras has been identified as a therapeutic target for anti-cancer treatment (214,215). Adenylosuccinate synthase catalyzes the GTP-dependent conversion of inositol monophosphate and aspartic acid to GDP, Pi and N(6)-(1,2-dicarboxyethyl)-AMP in purine biosynthesis (207). A deficiency of adenylosuccinate synthase causes an elevated pool of inositol monophosphate leading to the excessive purine excretion that is linked to hyperuricemia and gout (216). Glutamine synthetase plays an important role in the modulation of glutamate turnover and detoxification of ammonia by catalyzing the conversion of glutamate to glutamine (205). Several reports have proposed glutamine synthetase inhibitors for medical applications, such as anti-hyperammonemic brain therapy, anti-cancer treatment and anti-TB agents (217-220). While the disruption of adenylosuccinate synthase activity is clearly clinically significant, it is still unknown whether the inhibition of H-ras and glutamine synthetase activity can be safely tolerated. To avoid potential toxicity, any DTBS inhibitors should be therefore necessarily optimized based on the selectivity over these three closely related NTP-dependent human enzymes.
<table>
<thead>
<tr>
<th>EC (Enzyme Commission) number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC1</td>
<td>Catalyse oxidation/reduction reaction by transferring Nitrogenase iron protein [EC1.18.6.1]</td>
</tr>
<tr>
<td>EC2</td>
<td>Transfer a functional group from one substance to another</td>
</tr>
<tr>
<td>EC3</td>
<td>Catalyse the hydrolysis of various bonds</td>
</tr>
<tr>
<td>EC4</td>
<td>Non-hydrolytic addition or removal of various bonds</td>
</tr>
<tr>
<td>EC5</td>
<td>Catalyse changes within one molecule</td>
</tr>
<tr>
<td>EC6</td>
<td>Join two molecules with concomitant hydrolysis of the diphosphate bond in ATP or a similar triphosphate</td>
</tr>
</tbody>
</table>

**Table 6.1 NTP-dependent enzymes.**

Enzymes requiring a nucleotide triphosphate as a cofactor are highlighted in blue. Enzymes with sequence conservation are highlighted in blue. Enzymes with structural similarity to DTBS are highlighted in red. The data is present in the online database, namely IntEnz: integrated relational Enzyme database (www.ebi.ac.uk/intenz/index.jsp (EMBL-EBI)). Enzymes with structural homology with the P-loop of DTBS (EC6.3.3.3), ADP-dependent short/medium-chain-acyl-CoA hydrolase [EC3.1.2.18/19], ADP-dependent short/medium-chain-acyl-CoA hydrolase [EC3.1.2.18/19], ATP citrate lyase [EC 4.1.3.8], ATP citrate lyase [EC 4.1.3.8], CTP synthetase [EC 4.6.1.8], H-Ras p21 [EC3.6.5.2], phenylalanine racemase [EC5.1.1.11], and isocitrate dehydrogenase with one molecule [EC1.1.1.29] are highlighted in red while enzymes having sequence conservation are highlighted in blue. Enzymes with structural similarity and similar P-loop sequence to DTBS were highlighted in red while enzymes only obtain similar P-loop sequence to DTBS were highlighted in blue.
6.3 Transcriptional Regulation Mechanism of Biotin Biosynthesis

Whilst biotin biosynthesis has been studied in *M. tuberculosis*, its transcriptional regulation is unknown. Since the level of biotin is critical for bacterial survival, particularly during infection, the inhibition of biotin biosynthesis might provoke a bacterial adaptation mechanism. One possible mechanism may be to overexpress biotin biosynthetic enzymes to overcome the inhibitory activity of compounds. Therefore, it is necessary to understand the transcriptional regulation of biotin biosynthesis in order to combat TB effectively.

Our understanding of the regulation of biotin biosynthesis is derived from studies primarily in *E. coli* and *Bacillus subtilis*. Here biotin protein ligase (also known as BPL or BirA in *E. coli*) is a key enzyme in this regulatory system (221-225). BirA belongs to class II BPL, a bi-functional protein that acts as both a transcriptional repressor of biotin biosynthesis operon and a catalytic enzyme in the biotinylation of biotin dependent enzymes (as described in chapter 1, section 1.5) (226-228). The N-terminal domain of BirA recognizes a specific palindromic DNA sequence of the bio operator, *bioO*, located upstream of the biotin biosynthesis operon (229). Biotin biosynthesis is a costly metabolic pathway requiring 20 ATP equivalents per one produced biotin molecule in *E. coli* (230) and, thus, it is tightly regulated. Instead of *de novo* synthesis, microorganisms can alternatively uptake biotin from exogenous sources via the activity of a biotin transporter. This is also regulated by BirA via the recognition of sequence located upstream of the biotin transporter gene, *bioY*. However, this is not the case in *M. tuberculosis*. As described in chapter 1 (section 1.7), biotin biosynthesis is the sole source of biotin due to the lack of biotin transporter, and it is a key metabolic process in all stages of tuberculi life cycle (30). Instead of Class II BPL, *M. tuberculosis* expresses a Class I BPL that lacks an N-terminal DNA-binding domain required for the transcriptional regulatory activity (49,231,232). This raises the question of whether or not biotin biosynthesis needs to be regulated in *M. tuberculosis*, and if so, how is this
Apart from BirA, there are two other regulatory systems that have been recently characterized, namely BioR and BioQ belonging to GntR and TetR transcription factor family respectively. These systems were identified in microorganisms that have Class I BPL. BioR functions as a repressor of the biotin operon \((\text{bioBFDAZ})\) in \(\alpha\)-proteobacteria such as \textit{Agrobacterium tumefaciens} (233). It binds to the \(\text{bioO}\) at a 10 bp inverted repeat (TTATC-TATAA) sequence (230). This BioR recognition sequence was also found upstream of the \(\text{bioY}\), and the \(\text{bioR}\) gene itself indicating an autoregulation (230). Likewise, BioQ can regulate the \(\text{bioO}, \text{bioY}\) and autoregulation of \(\text{bioQ}\) in actinobacteria such as \textit{Corynebacterium glutamicum} and \textit{Mycobacterium smegmatis} by recognition a TGAAC\(\ldots\)TTTAC consensus sequence (234,235). Interestingly, the identification of the BioQ consensus sequence in \textit{M. smegmatis}, \textit{M. sp JLS} and \textit{M. vanbaalenii}, implied that \textit{M. tuberculosis} might also conduct the regulation of biotin biosynthesis through a similar BioQ regulatory system, rather than BioR. Therefore, a search of BioQ consensus binding-sequence was performed manually within 500 bp upstream of the start codon of \(\text{bioF2, bioA, bioF1, bioD and bioB}\). A similar BioQ consensus sequence with one nucleotide (T) insertion was found at the upstream of \(\text{bioB}\) in an overlapping of an open reading frame (indicated as +1) and the -10 recognition site of RNA polymerase (Fig 6.5). Both the -10 and -35 recognition sites required by bacterial RNA polymerase, were predicted using an online BPROM software (http://linux1.softberry.com/berry.phtml) (Softberry, Inc) (236). The potential binding of BioQ at the predicted recognition sequence in \textit{M. tuberculosis} can inhibit the binding of RNA polymerase and, thus, repress the transcription of \(\text{bioB}\). Since biotin synthase encoded by \(\text{bioB}\) catalyzes the last step of biotin biosynthetic pathway, the repression of \(\text{bioB}\) transcription is ideally a good target to regulate the entire pathway. Notably, the existence of BioQ encoding sequence in \textit{M. tuberculosis} is still unclear. A
search using a tblastx algorithm and a well-characterized BioQ from Corynebacterium
glutamicum as a sequence query failed to identify the BioQ ortholog in the translated
nucleotide database of M. tuberculosis genome. However, based on the discovery of
predicted BioQ recognition sequence here, I strongly recommend that BioQ is a good
candidate that needs to be studied further to understand the regulatory system of biotin
biosynthesis in M. tuberculosis. Noteworthy, BioQ was suggested to act as a repressor of
biotin biosynthesis as the ΔBioQ strain of M. smegmatis significantly increased the
expression of bio operons (234). This implies that mycobacteria might overcome the low
level of biotin by stopping the manipulation of BioQ, resulting in unlimited expression of
biotin biosynthetic enzymes to synthesize more biotin. Consequently, higher concentrations
of inhibitors will be needed to completely abolish the biotin biosynthesis. Combination
therapy that targets both biotin biosynthesis and its transcriptional regulation system is,
therefore, an attractive strategy to combat TB.

Overall from this study, biotin biosynthesis is an excellent drug target and a
combination therapy with other pathways potentially gives the strengths of TB treatment in
terms of low dose and short course requirement. When taken in combination for example
with any clinical anti-TB drugs, the DTBS inhibitors might cause synergistic interactions that
can improve efficacy and also increase the clinical life of drugs. Lastly, as there is no pre-
existing resistant mechanism caused by targeting biotin biosynthesis, the introduction of any
successful developed drugs targeting enzymes in this pathway in TB treatment should be able
to combat the drug resistance.
Figure 6.5 Predicted site of BioQ recognition sequence in \textit{M. tuberculosis} genome. The predicted BioQ recognition sequence is shown in red letters, containing one inserted nucleotide (T, black boxed). The -10 and -35 are the predicted recognition sites of RNA polymerase. The codon sequence of \textit{bioB} gene is highlighted in grey while the predicted start transcriptional site (or an open reading frame (ORF)) is indicated at +1.


11. Organization, N. A. C. (2007) Antiretroviral therapy guideline for HIV-infected adults and adolescents including post-exposure prophylaxis (Section A7). New Delhi, India


CoA carboxylase beta subunit of Streptomyces coelicolor. *Biochemistry* **49**, 7367-7376


152. Dhyse, F., and Hertz, R. (1958) The effects of Actithiazic acid on egg white-induced biotin deficiency and upon the microbial formation of biotin vitamers in the rat. *Arch. Biochem. Biophys.* **74**, *7-16*


162. FDA. (2011) FY 2011 innovative drug approvals. FDA
186. Cheng, Y., and Prusoff, W. (1973) Relationship between the inhibition constant (Ki) and the concentration of inhibitor which causes 50 percent inhibition (I50) of an enzymatic reaction. *Biochem. Pharmacol.* **22**, 3099-3108


APPENDIX

NOTE:
This publication is included on pages 183 - 187 in the print copy of the thesis held in the University of Adelaide Library.

It is also available online to authorised users at:

[http://dx.doi.org/10.1007/s13238-011-1100-8](http://dx.doi.org/10.1007/s13238-011-1100-8)