# MAKERERE UNIVERSITY

**STRUCTURAL ANALYSIS OF THE INTERACTION OF embB PROTEIN MUTATIONS OF *Mycobacterium tuberculosis* WITH ETHAMBUTOL AND THEIR EFFECT ON THE INTERACTION**

**BY**

**ARINDA DEAN**
**21/U/06126/EVE**
**2100706126**
**BBLT III**

**A SPECIAL RESEARCH PROJECT REPORT SUBMITTED TO THE COLLEGE OF VETERINARY MEDICINE, ANIMAL RESOURCES AND BIOSECURITY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF A DEGREE OF BACHELOR OF BIOMEDICAL LABORATORY TECHNOLOGY OF MAKERERE UNIVERSITY**

**AUGUST, 2024**

## DECLARATION AND APPROVAL

I **Dean Arinda,** declare that this special project report is original and has never been submitted to any academic institution of higher learning for any academic award. Reference citations are included where work of other people has been used.

Signature: ..... Date: 9/08/2024

This research project report has been submitted under supervision, guidance and approval from my supervisor.

**Dr. Kato Charles Drago (PhD)**

Lecturer,
Department of Biomolecular Resources and Biolaboratory Sciences
School of Biotechnical, Biosecurity and Laboratory Sciences
College of Veterinary Medicine, Animal Resources and Biosecurity
Makerere University, P. O Box 7062, Kampala

Signature: ..... Date: 9/8/24

## DEDICATION

I dedicate this research project report to my beloved parents, Mr. Tumuhimbise Isaac and Mrs. Asasiirwe Jennipher Tumuhimbise who have endeavoured to support me through my academic journey.

## ACKNOWLEDGEMENTS

I would like to thank the almighty God for the life and knowledge granted to come up with this report. Special thanks go to my supervisor, Dr. Kato Charles Drago for accepting to supervise me in the first place. I also thank him for the guidance and assistance in writing this research report.

I also extend my sincere gratitude to my family for the financial support and prayers to excel in this course. Sincere appreciation to my friends and classmates for the encouragement throughout. May the almighty bless you all for me.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| Acp: | Acyl carrier protein |
| Aft: | Arabinofuranosyltransferases |
| AG: | Arabinogalactan |
| AraTs: | Arabinosyltransferases |
| CBMs: | Carbohydrate Binding molecules |
| DOTS: | Directly Observed Short Treatment Course |
| DPA: | Decaprenyl Phosphoryl-D-Arabinofuranose |
| EMB: | Ethambutol |
| ERDR: | Ethambutol Resistance Determining Region |
| HIV: | Human Immunodeficiency Virus |
| INH: | Isoniazid |
| LAM: | Lipoarabinomannan |
| MA: | Mycolic Acid |
| mAGP: | mycolyl- Arabinogalactan- Peptidoglycan complex |
| MDR: | Multi Drug Resistant |
| MTB: | Mycobacterium tuberculosis |
| MTBC: | Mycobacterium tuberculosis Complex |
| NTM: | Non tuberculosis mycobacterial disease |
| PDB: | Protein Data Bank |
| PG: | Peptidoglycan |
| PZA: | Pyrazinamide |
| RMP: | Rifampicin |
| RR: | Rifampicin Resistant |

| | |
|---|---|
| SDF: | Structural Data File |
| TB: | Tuberculosis |
| TDR: | Totally Drug Resistant |
| TM: | Trans Membrane |
| TMH: | Trans membrane helical |
| UniProtKB: | Universal Protein Knowledge Base |
| WHO: | World Health Organisation |
| XDR: | Extensively Drug Resistant |

## ABSTRACT

Tuberculosis (TB), caused by *Mycobacterium tuberculosis*, remains a global health challenge exacerbated by the emergence of drug-resistant strains. Ethambutol is a critical first-line anti-TB drug targeting the embB protein, an enzyme essential for mycobacterial cell wall biosynthesis. Mutations in the embB gene are associated with resistance to ethambutol, reducing treatment efficacy and complicating disease management.

This study aimed at investigating the interaction between the wild-type and specific mutant embB proteins and ethambutol and to investigate the effects of these specific embB mutations on this interaction.

Computational docking and structural interaction analysis of the wild-type embB protein with Protein Data Bank (PDB) ID: 7BVF and its common mutants (M306V, M306L, E378A) with ethambutol was performed using PyRx software and LigPlot respectively. Binding affinities were calculated and the interaction sites were identified and analysed using molecular docking techniques and 2D ligand interaction diagrams respectively.

The wild-type embB protein demonstrated a binding affinity of -4.8 kcal/mol with ethambutol, involving key residues such as Leu561, Thr564, Tyr502 and Asp498, forming a network of hydrogen bonds ranging from 2.87 Å to 3.26 Å. In contrast, embB mutants exhibited reduced binding affinities: -3.8 kcal/mol for M306V, -3.9 kcal/mol for M306L, and -4.7 kcal/mol for E378A. These mutations disrupted the hydrogen bond network, involving fewer residues and exhibiting weaker interactions with ethambutol.

The findings highlight the critical role of specific residues in the embB protein for maintaining high-affinity interactions with ethambutol. Mutations in these residues significantly reduce drug binding, contributing to resistance. This study underscores the importance of targeting these mutations for the development of novel therapeutic agents and suggests the need for enhanced diagnostic tools to detect embB mutations in clinical isolates for effective TB management.



**Keywords:** *Mycobacterium tuberculosis*, embB protein, ethambutol, drug resistance, protein-ligand interactions, structural analysis.

# CHAPTER ONE

# INTRODUCTION

## 1.1 Background

The *Mycobacterium tuberculosis* is a tubercle bacillus characterized by its slow growth, dormancy and intracellular pathogenesis. It is a Gram-positive bacterium and its genome comprises about 4.4 megabase pairs (Yang *et al*., 2019). The MTB is also an acid-fast organism which contains large amounts of mycolic acids within the cell wall. The *Mycobacterium tuberculosis* is the cause of a highly infectious chronic airborne disease, tuberculosis and accounts for considerable morbidity and mortality. In the year of 2021, the World Health Organization (WHO) estimated 10.6million cases and 1.4 million deaths of tuberculosis (WHO, 2022). Uganda is one of the world's thirty (30) high-burden countries for TB and TB/HIV co-infection. Each year, approximately 91,000 people in Uganda get sick of TB with 32% of them being HIV-infected (WHO, 2023).

The treatment of TB is strenuous and necessitates prolonged therapy, initially 2 months with rifampicin, isoniazid, pyrazinamide and ethambutol, and only rifampicin and isoniazid for 4 months (Vinod *et al*., 2020). Drug resistance in tuberculosis chemotherapy is fast becoming a health crisis on a global scale. The emergence of multidrug-resistant (MDR), extensively drug-resistant (XDR), and totally drug-resistant (TDR) strains of *Mycobacterium tuberculosis* (MTB) has been observed as a result of ineffective directly observed treatment short-course (DOTS) among a myriad of other factors (Mugumbate *et al.*, 2021). In 2021, there were an estimated 450,000 incident cases of MDR-TB by WHO and an estimated 191,000 deaths occurred due to MDR-TB (WHO, 2022). In Uganda, 12% of previously treated TB cases and 1.6% of new cases have MDR-TB and require specialized treatment and care (Batte *et al.*, 2021). Two out of every 100 people with TB have drug-resistant TB that is not cured by first-line drugs, while approximately 15% of TB cases in Uganda are children aged below 14 years (WHO, 2023). Treating DR-TB continues to be challenging due to the longer duration of treatment, lower cure rates and necessity of using more expensive drugs with higher toxicity which imposes a burden on the health care resources (He *et al*., 2021). Contrasting with other bacterial pathogens that have evolved to spread drug resistance in populations via horizontal gene transfer, drug resistance in MTB is mainly

due to mutations in chromosomal genes (Singh & Chibale, 2021). This genotypic drug resistance may develop due to insertions and deletions in genes coding for drug targets or converting enzymes (Phelan *et al*., 2016).

The Ethambutol (EMB) is a first line drug in tuberculosis treatment and plays an important role in the second line regimens of drug resistant TB (Shi *et al*., 2011). The EMB is bacteriostatic and acts against bacterium by targeting arabinan synthesis through inhibiting arabinosyltransferases encoded by the embCAB operon with three genes (embC, embA and embB) which affects the mycobacterial cell wall. The resistance to EMB has continuously been correlated with alterations in the embB gene, particularly in embB codon 306, referred to as EMB resistance determining region (ERDR) (B *et al*., 2013). Mutations at embB406 and embB497 have also been found to be associated with ethambutol resistance. Nonetheless, this association is unresolved because all these codons have also been found mutated in isolates susceptible to EMB. The advances in computational techniques and expansions in bioinformatics and chemoinformatics have brought alleviation in the study of mutations and provided a rapid drug susceptibility testing important in the detection and control of MDR/XDR TB. The aim of this study therefore was to use bioinformatics tools to understand molecular target mutations in presenting ethambutol resistance and also analyse the interactions of the mutated protein with ethambutol.

### 1.2 Problem statement

Tuberculosis (TB) is a global health crisis, with approximately 10.6million cases and 1.4 million deaths reported in 2021, according to the World Health Organization (WHO, 2022). Each year, approximately 91,000 people in Uganda get sick of TB with 32% of them being HIV-infected (WHO, 2023). The growing concern is in the emergence of the drug resistant TB strains including the resistance to ethambutol, with around 450,000 individuals diagnosed with multidrug resistant TB (MDR-TB) in 2021 with 191,000 deaths (WHO, 2022).

Mutations in the embB gene have been identified as a key contributor to ethambutol resistance (Cui *et al.,* 2014)**.** However, a comprehensive structural analysis of embB protein mutations and their interaction with ethambutol is currently lacking. Employing advanced computational structural biology techniques, such as molecular dynamics simulation and protein-ligand docking studies, this research project aimed to explain

the three-dimensional structure of embB protein mutations in the *M. tuberculosis* and investigate their impact on ethambutol binding which could contribute to the development of targeted drug design approaches and informatics-driven therapeutic strategies to alleviate the global burden of drug resistant TB.

## 1.3 Objectives

### 1.3.1 General objective

To investigate the interaction of embB protein mutations of *Mycobacterium tuberculosis* with ethambutol and their effects on the interaction.

### 1.3.2 Specific objectives

i. To describe the interaction between wild type embB protein with ethambutol.
ii. To investigate the effect of mutations in embB protein on ethambutol binding.

## 1.4 Research questions

1. What are the interactions between wild type embB protein with ethambutol?
2. What are the effects of mutations in the embB protein with ethambutol?

## 1.5 Justification and significance

Tuberculosis (TB) has been a fatal enemy to humanity for thousands of years. In spite of the introduction of chemotherapeutic drug treatments, TB today remains the leading cause of death by infectious diseases. The developing mutations among TB micro-organisms have led to increased resistance to anti-tuberculosis drugs including ethambutol, which has made treatment of TB rather strenuous. They have also led to high mortality rates among patients. Therefore, analysis of the interaction of the embB protein mutations with ethambutol and their effect on the interaction is crucial. This will help in the development of new and better interventions against tuberculosis, which will make TB management easier.

# CHAPTER TWO

# LITERATURE REVIEW

## 2.1 *Mycobacterium tuberculosis*

The *Mycobacterium tuberculosis* (MTB) is a bacterium transmitted through respiratory droplets. It is one of the most successful human pathogens with approximately 10 million cases and 1.45 million associated deaths per year. It is an obligate aerobe and acid-fast bacillus due to the high lipid content of the cell wall (Levinson, 2014). The MTB grows slowly. It is slender, straight or slightly curved rod with rounded ends in pairs or small clumps. The bacilli are non-motile, non-sporing and non-capsulated (Surinder, 2016).

### 2.1.1 Classification of *Mycobacterium tuberculosis*

The *Mycobacterium tuberculosis* belongs to the genus *Mycobacterium*, which is classified within the family Mycobacteriaceae. The genus *Mycobacterium* includes over 190 species of bacteria, many of which are pathogenic to humans, animals and plants (Surinder, 2016). The *Mycobacterium tuberculosis* complex (MTBC) consists of a group of closely related Mycobacterium species, including M. tuberculosis, M. bovis, M. africanum, M. microti, and M. caprae (Bayraktar *et al*., 2011). The most important to humans is *Mycobacterium tuberculosis.*

The *Mycobacterium tuberculosis* can be classified as follows; **Kingdom**: Bacteria **Phylum**: Actinobacteria **Order**: Actinomycetales **Family**: Mycobacteriaceae **Genus**: *Mycobacterium* **Species:** *Mycobacterium tuberculosis*

The MTB can also be classified based on the resistance profiles; Monoresistance: Resistance to one first-line anti-TB drug only. Polydrug resistance: Resistance to more than one first-line anti-TB drug, other than both isoniazid and rifampicin. Multidrug resistance (MDR): Resistance to at least both isoniazid and rifampicin. Rifampicin resistance (RR): Resistance to rifampicin detected using phenotypic or genotypic methods, with or without resistance to other anti-TB drugs. It includes any resistance to rifampicin, whether monoresistance, multidrug resistance, polydrug resistance, or extensive drug resistance. Extensive drug resistance (XDR): Resistance to any

fluoroquinolone, and at least one of three second-line injectable drugs (capreomycin, kanamycin, and amikacin), in addition to multidrug resistance (Seung *et al*., 2015).

### 2.2 Burden of tuberculosis

The World Health Organization (WHO), 2022 announced that TB remains a major global health problem causing deaths among millions of people every year. TB is the ninth leading cause of death worldwide and the leading cause from a single infectious agent (Diriba & Churiso, 2022). Most of the deaths are a result of coinfection with HIV. Globally, an estimated 10.6 million people fell ill with TB and 1.4 million died from the disease in 2021 (WHO, 2022). Sub-Saharan Africa (SSA) accounts for a quarter of the global TB cases of which 32% are co-infected with HIV (Baluku *et al*., 2022). Uganda is one of the countries in the world with a high prevalence of TB disease estimated at 1,675 cases per 59,751 people representing 2.8% (Samuel Mwesige *et al*., 2022).

As a result of spontaneous gene mutations of MTB, there is emergence of drug resistance in TB (Palomino & Martin, 2014). Non- compliance to the treatment regimen is signaled as the 1$^{st}$ cause to the emerging resistance to anti-TB drugs. According to (Salari *et al*., 2023), the global prevalence of multidrug-resistant, mono drug-resistant, isoniazid, and rifampicin tuberculosis are 11.6%, 11.8%, 15.7%, and 9.4%, respectively. In a study carried out in 2022 (Molla *et al*., 2022) in East Africa, the pooled prevalence of MDR-TB among newly diagnosed TB cases was 4%. Among newly diagnosed MDR-TB cases, the highest prevalence was reported to be 20% in Sudan and the overall prevalence of MDR-TB among previously treated TB cases was 21%. In Uganda, 12% of previously treated TB cases and 1.6% of new cases have MDR-TB (Batte *et al*., 2021).

### 2.3 The embB protein and its role

The arabinosyltransferase B (EmbB) belongs to a family of membrane-bound glycosyltransferases that build the lipidated polysaccharides of the mycobacterial cell envelope, and are targets of anti-tuberculosis drug ethambutol.

The cell envelope is crucial for growth and virulence of pathogenic mycobacteria like *M. tuberculosis* and is a major contributor to resistance against common antibiotics (Tan *et al*., 2020). Its main component is the mycolyl-arabinogalactan-peptidoglycan

(mAGP) complex, which consists of three highly unusual elements covalently linked together; a cross-linked network of peptidoglycan (PG), a branched heteropolysaccharide arabinogalactan (AG) and long chain mycolic acids (MA) (Fig. 1). Another major component is the lipidated heteropolysaccharide lipoarabinomannan (LAM), a species of phosphatidyl-myo-inositol derived glycolipids containing mannan and arabinan domains and is an important virulence factor playing a key role in host-pathogen interactions, and in modulating the host immune response during infection (L. Zhang, Zhang, *et al*., 2020). Of the enzymes involved in mycobacterial cell wall biosynthesis, arabinofuranosyltransferases are responsible for the addition of D-arabinofuranose sugar moieties to AG and LAM. These transmembrane (TM) enzymes utilize decaprenyl phosphoryl-D-arabinofuranose (DPA) to transfer an arabinofuranose unit to the growing lipidated polysaccharides of the cell envelope.

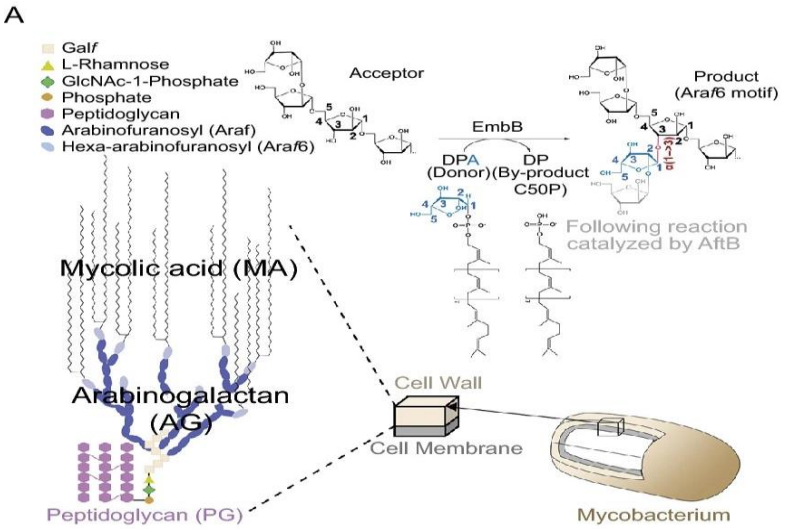



Figure 1: The *Mycobacterium* cell wall

The arabinosyltransferase B (EmbB) is a 117 kDa integral membrane enzyme involved in the α-(1→5)-linked extension of the AG arabinan chain (Tan *et al*., 2020.). Its gene belongs to an operon coding for two other homologous arabinosyltransferases EmbA and EmbC. `The operon was named due to the sensitivity of these gene products to ethambutol, a first-line antibiotic against tuberculosis and nontuberculous mycobacterial (NTM) disease (Tan *et al*., 2020.). Because the Emb proteins belong to the same family of AraTs, their amino acid sequences are highly similar (sharing ~40% identity). Both EmbA and EmbB have important roles in the formation of the α (1→3) linkage on the terminal hexaarabinofuranosyl motif of AG. Furthermore, these proteins

are suggested to function in a coordinating way by forming a heterodimer within cells, whose reaction product is further catalysed by AftB by forming the terminal β (1→2) linkage at the AG nonreducing end. The product of the reactions catalysed by EmbA and EmbB and by AftB serves as the mycolic acid attachment site for AG. EmbB and EmbC have mutations known to lead to ethambutol resistance, while a clear effect of the drug on EmbA activity has not been established (L. Zhang, Zhao, *et al*., 2020).

### 2.3.1 Structure of embB protein

The EmbB is a fifteen TMH spanning protein which catalyses the transfer of arabinose from the donor decaprenyl-phosphate-arabinose (DPA) to its arabinosyl acceptor. On the cytoplasmic surface of EmbB is an acyl-carrier protein (AcpM), which is suggested to be associated with the transferase activity (Bendre *et al.*, 2021). The EmbB appears as a monomer, consisting of 15 TM helices and two distinct periplasmic carbohydrate binding modules (CBMs). The first two TM helices are not found in other glycosyltransferase structures solved to date, and seem to serve to anchor the N-terminal CBM (N-CBM) to the membrane. The next 11 TM helices adopt a typical GT-C glycosyltransferase fold14, structurally similar to enzymes from various glycosyltransferase families. The last two TM helices are shared only with ArnT. Thereafter, the polypeptide chain exits the membrane in the periplasm to form the second C-terminal CBM (C-CBM). The C-CBM then loops back around to complete a β-sheet with the N-CBM, likely to secure the N-terminal domain in place (Tan *et al*., 2020).

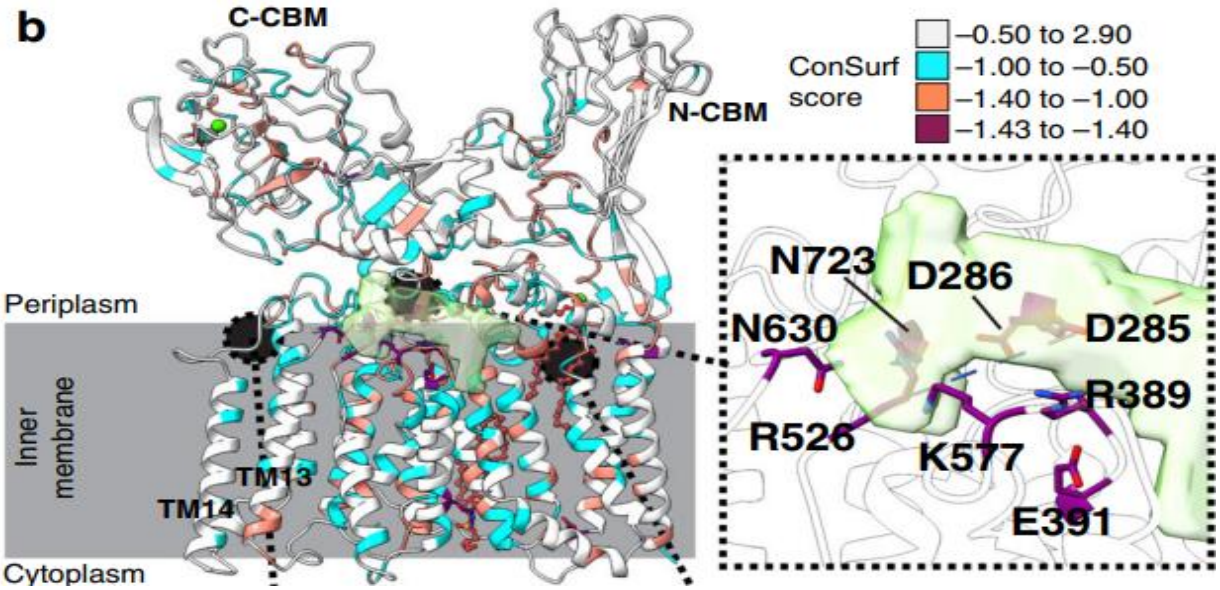


Figure 2: The structure of embB protein, rendered in cartoon and coloured based on ConSurf score for sequence conservation. The more negative the score, the more conserved the residue. The putative active site cavity, generated by the Voss Volume Voxelator server is colored in semi-transparent green. The insert shows the putative active site cavity with the strictly conserved residues labelled (Tan *et al*., 2020).

### 2.3.2 Mutations in the embB protein

The embB gene mutations have the predominant role in EMB resistance, particularly at codons 306, 406, and 497, which are considered as hotspot resistance codons (Sun *et al*., 2018). Codon 306 was shown to be directly involved in EMB binding while codons 406 and 497 are not directly involved. Nevertheless, mutations at codon 497 cause conformational changes that affect codon 327, one of the EMB binding sites. Codon 406 mutations may also affect drug binding by causing protein conformation changes (Bwalya *et al*., 2022). Mutations in the residue M306 from EmbB have been identified in more than 68% of EMBR isolates (Y. Zhang & Yew, 2015). In a study carried out by Maladan et al., (2023), they identified M306L, M306V, D1024N, and E378A mutations using the mutation analysis performed using TBProfiler.

M306 mutations

Based on the structure of EmbB, M306 is directly involved in EMB binding, and mutations affecting this residue, or those residues that interact with it, including Y302 and E327, disturb the EMB binding affinity because of differences in the interaction between the protein active site and EMB (Rossini & Dias, 2023).

G406 mutations

The resistance mechanisms involved in mutations affecting G406 are proposed to lead to a steric hindrance and consequently cause conformational changes in EMB binding site, decreasing the ligand affinity (Y. Zhang & Yew, 2015).

## 2.4 Ethambutol

The Ethambutol (EMB), an arabinose analogue, is a bacteriostatic, antimycobacterial drug, which has been used for the treatment of TB since the mid-1960s (Vilch, 2020). The drug is routinely recommended for the intensive phase of TB therapy, as part of a four-drug regimen, including isoniazid (INH), rifampicin (RMP), and pyrazinamide

(PZA). The Ethambutol appears to target the cell wall of tubercle bacilli through interfering with arabinosyltransferases, encoded by the embCAB operon, comprised of three homologous genes, designated embC, embA, and embB (B *et al*., 2013).

Structurally, ethambutol is an ethylenediamine derivative that is ethane-1,2-diamine in which one hydrogen attached to each of the nitrogens is substituted by a 1-hydroxybutan-2-yl group (S,S-configuration) (from https://pubchem.ncbi.nlm.nih.gov/compound/Ethambutol).

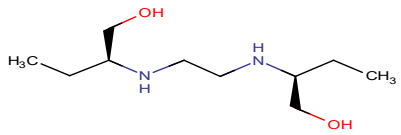



Figure 3: The 2D structure of ethambutol.

### 2.4.1 Mechanism of action of ethambutol

The EMB acts bacteriostatically by blocking the polymerization of arabinose subunits in the AG layer of the cell wall, which leads to loss of the MA layer (Schubert *et al*., 2017). Ethambutol diffuses into *Mycobacterium* cells. Once inside the cell, it inhibits the arabinosyltransferases (embA, embB, and embC), preventing formation of the cell wall components arabinogalactan and lipoarabinomannan, and preventing cell division. Decreased concentrations of arabinogalactan in the cell wall reduces the number of binding sites for mycolic acid, leading to the accumulation of mycolic acid, trehalose monomycolate, and trehalose dimycolate. Lipoarabinomannan is a component of a cell surface molecule involved in the interaction with host cells. Reduced levels of lipoarabinomannan may interfere with mycobacterial interaction with host cells (Goude *et al.*, 2009).

### 2.4.2 Resistance to ethambutol

Although ethambutol (EMB) plays a pivotal role in the chemotherapy of drug-resistant tuberculosis (TB), including multidrug-resistant tuberculosis (MDR-TB), there's growing resistance to EMB (Zhao *et al*., 2015). The EMB resistance mostly involves mutations in the embCAB operon, which encodes target proteins of EMB (Rossini &

Dias, 2023). Mutations in the residue M306 from EmbB have been identified in more than 68% of EMBR isolates (Rossini & Dias, 2023). Additionally, mutations in embC and the embC-embA intergenic space were also reported to EMB resistance (Cui *et al*., 2014). Mutations on the embC gene may lead to the production of mutant proteins with a lower EMB binding affinity due to substitutions near the protein binding site to EMB, while mutations in embC-embA intergenic space may have a strong impact on the mRNA expression of embA and embB genes (Cui *et al*., 2014; Zhang *et al*., 2020).

### 2.5 Structural analysis tools

The structural analysis tools are the different soft wares and databases in bioinformatics used for the analysis and prediction of the three-dimensional structure of biological macromolecules such as proteins, RNA, and DNA. They include among others, PyMol, Ligplot, PyRx, that I used specifically for this study.

#### 2.5.1 PyMol

PyMol is a user-sponsored molecular visualization system on an open-source foundation, maintained and distributed by Schrödinger (https://pymol.org/2/). It has been widely used for three-dimensional (3D) visualization of proteins, nucleic acids, small molecules, electron densities, surfaces, and trajectories. It is also capable of editing molecules, ray tracing, and making movies. This Python-based software, alongside many Python plugin tools, has been developed to enhance its utilities and facilitate the drug design in PyMol. There are various molecular modelling modules in PyMol which include those for visualization and analysis enhancement, protein–ligand modelling, molecular simulations, and drug screening (Yuan *et al*., 2017).

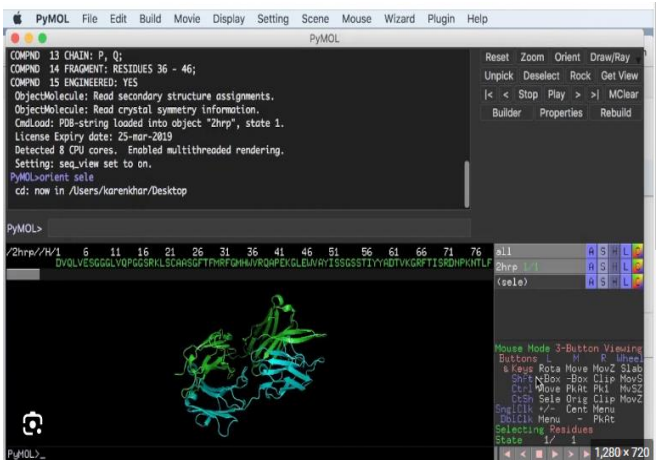



Figure 4: The PyMol interface

### 2.5.2 LIGPLOT

The LIGPLOT program automatically generates schematic 2-D representations of protein-ligand complexes from standard Protein Data Bank file input. The output is a colour, or black-and-white, PostScript file giving a simple and informative representation of the intermolecular interactions and their strengths, including hydrogen bonds, hydrophobic interactions and atom accessibilities. The program is completely general for any ligand and can also be used to show other types of interaction in proteins and nucleic acids. It was designed to facilitate the rapid inspection of many enzyme complexes, but has found many other applications (Wallace *et al*., 1995).

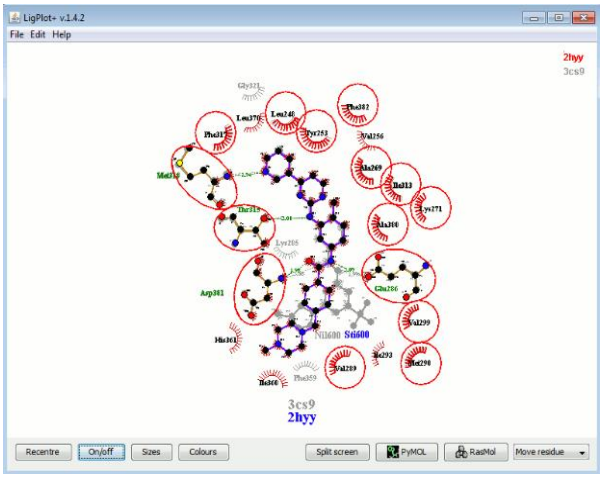



Figure 5: The LIGPLOT interface.

### 2.5.3 PyRx

PyRx is a Virtual Screening software for Computational Drug Discovery that can be used to screen libraries of compounds against potential drug targets. PyRx enables Medicinal Chemists to run Virtual Screening from any platform and helps users in every step of this process - from data preparation to job submission and analysis of the results. While it is true that there is no magic button in the drug discovery process, PyRx includes docking wizard with an easy-to-use user interface which makes it a valuable tool for Computer-Aided Drug Design. PyRx also includes chemical spreadsheet-like functionality and powerful visualization engine that are essential for structure-based drug design https://pyrx.sourceforge.io/.

# CHAPTER THREE

# MATERIALS AND METHODS

## 3.1 Study design

This study was a computer-based with extended use of the internet. Various databases and software were utilized for data retrieval such as UniProtKB, Protein Data Bank (PDB) and PubChem, protein receptor cleaning using PyMol and molecular docking using PyRx at the College of Veterinary Medicine, Animal resources and Biosecurity (COVAB), Makerere University- Kampala, Uganda.

## 3.2 Inclusion and exclusion criteria

All reverse protein structures that did not back up the literature were excluded. This ensured focus on only the embB protein and ethambutol interactions only. The study included only embB proteins that have at least been detected to be involved in the binding with ethambutol.

## 3.3 Data collection

### 3.3.1 Protein retrieval and preparation

The three-dimensional structure of the wild type embB protein (ID: 7BVF) was retrieved from Protein Data Bank (PDB) through a link https://www.rcbs.org/PDB/home/home.do . Key words were fed into the search window of the page. The key words were embB and ethambutol to retrieve the required structure. Results were brought and the result of choice was chosen by clicking against it. The download tab in the new window was clicked on. A dialogue box appeared within which PDB format was clicked on as it is recognized by most tools for viewing the structures. The PDB format of the file was downloaded and stored in a file for future retrieval. On retrieval of the saved protein, it was prepared by cleaning it to remove water molecules, iron and any other attachments using PyMol software.

The mutated type embB FASTA sequences were downloaded from UniProtKB. The UniProtKB identification numbers of the mutated proteins were fed into the search window (i.e. D3XF88 for M306V, AOA024CD61 for E378A and Q84B90 for M306L). The FASTA sequences were then uploaded into SWISS-MODEL, a fully automated

protein structure homology-modelling server, for the molecular modelling of the three-dimensional structures of the mutated embB proteins. The best models of the mutated proteins were then chosen, downloaded in pdb format and stored on the computer for future retrieval. They were also viewed in PyMol and cleaned to prepare them for docking.

### 3.3.2 Drug retrieval and preparation

The ethambutol 2D Structural Data File (SDF) structure was retrieved from PubChem. The downloaded 2D SDF structure was converted to a 3D structure PDB format using Open Babel plugin in the PyRx software to prepare it for molecular docking with the embB proteins.

### 3.4 3D visualization of the wild type and mutated embB protein

The 3D structures of the wild type and mutated embB proteins were visualized using PyMol software. The PyMol window was accessed through double-clicking to open as a new window. In the window displayed, the file option was selected and left-clicked against. A dialogue box appeared and then open option was selected. A new display window appeared from which the directory of the required file was selected. The directory having the stored file was located and the PDB file clicked against. The open option was clicked on and the desired protein structures of concern were then loaded and viewed in the PyMol window. In the newly opened PyMol window, a dialogue box appeared after right-clicking in the window area. Various options then displayed, from this, the all option was left clicked on and then the show option selected. Another dialogue box appeared from which the cartoon option was selected. The command window within PyMol was accessed from which various commands were fed in to manipulate the viewed structure in PyMol to the desired formats and conformations. The structures under investigation were rotated and moved around in different orientations to view the structure for best interpretation.

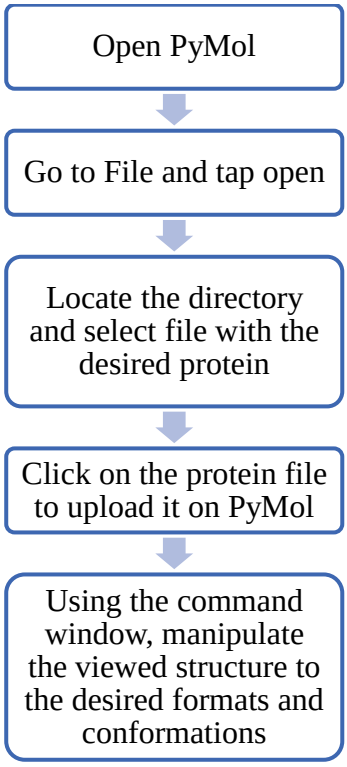



Figure 6: The schematic diagram showing steps to visualize the embB proteins in PyMol

### 3.5 Analysis of the structural interactions between wild type and mutated embB proteins with ethambutol

PyRx software was used for the analysis of the interactions between wild type and mutated embB proteins with ethambutol through molecular docking using the auto dock vina plugin in the software. The pdbqt format of the proteins was uploaded into PyRx and converted into AutoDock large macromolecule to prepare them for docking. Using the Open Babel plugin in the PyRx software, the 2D SDF file of the ligand ethambutol was uploaded into PyRx. The ligand was prepared via several options in the software to assemble it for docking. Using the auto dock plugin in the PyRx software, the two molecules were docked and the results were waited for, in like 20 minutes. The results were in form of binding affinity. The best model of the results was downloaded in PDB format and stored for further analysis using LigPlot and PyMol.

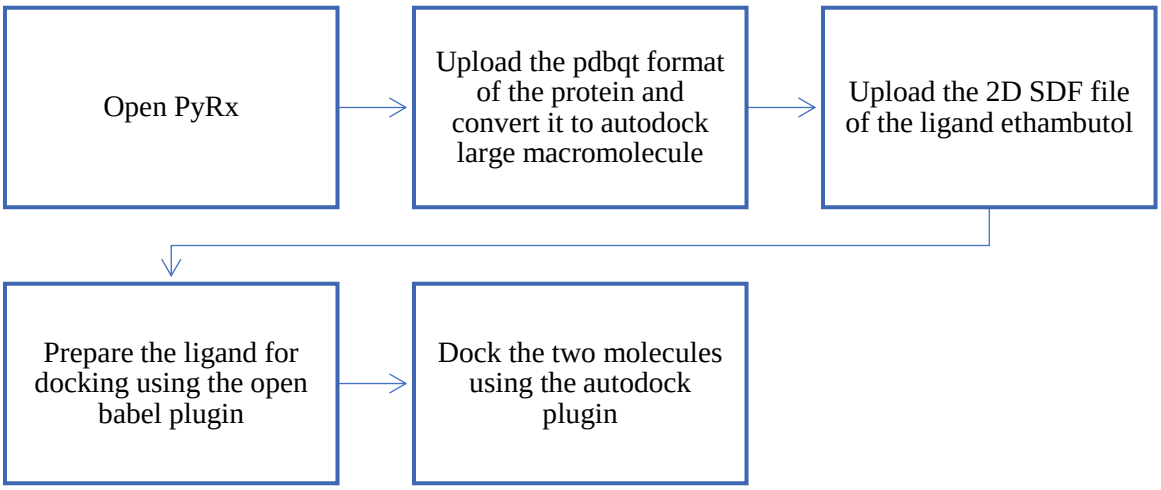



Figure 7: Steps for analysis of the structural interactions between the embB proteins and ethambutol

### 3.6 Protein- ligand interaction analysis

The protein-ligand structures were retrieved from the stored directory. Ligplot software was used to easily identify and view the complex. Ligplot was opened and complexed PDB file was retrieved from the storage and selected to open. A dialogue box was opened showing ligands within the complex, with which the ligand corresponding to drug of interest was chosen. This thus generated two-dimension representations of the protein to ligand complexes generated from the Protein Data Bank. This was achieved through unrolling the PDB format and thus flattened them onto a two-dimension page. Here the covalent and ionic interactions were observed. The obtained two dimensional images were stored to a desired file directory for future retrieval.

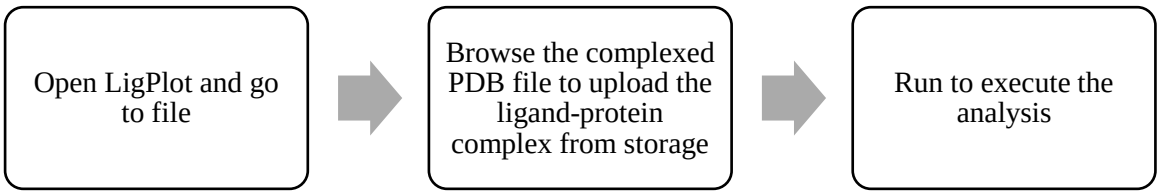



Figure 8: Steps for protein-ligand interaction analysis

### 3.7 Three-dimensional viewing of the complexes formed from Ligplot in three-dimensional format

Ligplot was opened from the desktop. The PyMol software plugin was added to Ligplot. The receptor and ligand complex were retrieved from storage directory by clicking on file, then opened PDB and PDB file located and selected. A two-dimensional image of complex was displayed by Ligplot. The PyMol button within Ligplot was used to open the software. PyMol was then used to analyse the amino acid residues in close proximity with the ligand. This was achieved through using the show command in PyMol. The amino acids per position were displayed in the view window. The desired amino acids in the window were selected and highlighted. These were right clicked against and a pop window appeared and then the colour option was selected and gave the residue of interest a desired colour.

# CHAPTER FOUR

# RESULTS

## 4.1 The interaction between wildtype embB protein with ethambutol

The study revealed a binding affinity of -4.8Kcal/mol of the wild type (7BVF) embB protein with ethambutol. Four residues (Leu561, Thr564, Tyr502 and Asp498) were involved in the interaction using hydrogen bonding of different lengths in angstroms (Å) as shown in table 1 below.

**Table 1: The residues involved in the interaction between wild type embB protein (7BVF) with ethambutol**

| Residues involved | Number of hydrogen bonds | Length (Å) |
|---|---|---|
| **Leu561** | 2 | 2.87, 3.12 |
| **Thr564** | 2 | 3.12, 3.10 |
| **Tyr502** | 1 | 2.97 |
| **Asp498** | 1 | 3.26 |

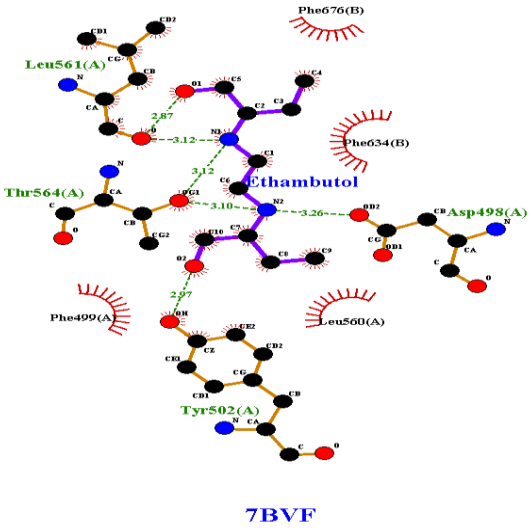


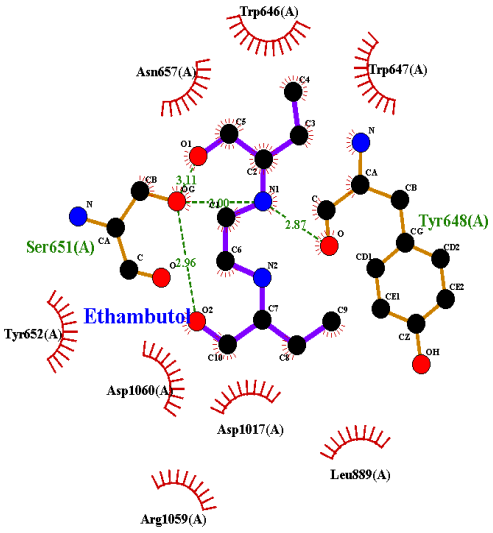

Figure 9: The Ligplot 2D representation of the interaction between 7BVF and ethambutol. Blue dots: Nitrogen atoms, Red dots: Oxygen atoms, **Black dots**: Carbon atoms, Light blue lines: Ligand bonds, Orange Lines: Non- ligand bonds, Green doted lines: Hydrogen bonds, Brick red comb- like shapes: Pocket atoms.

### 4.2 The interaction between the mutations in embB protein and ethambutol

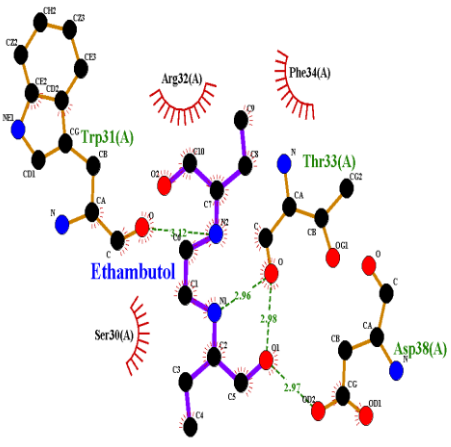
The study found a binding affinity of -3.8Kcal/mol in M306V, -3.9Kcal/mol in M306L and -4.7Kcal/mol in E378A mutations of embB protein. There was reduction in the number of hydrogen bonds involved in the interaction of ethambutol with the embB protein mutants; M306V (Gly33, Ser31, Arg27), M306L (Trp31, Thr33, Asp38) and E378A (Ser651 and Tyr648) with various lengths measured in angstroms (Å) as shown in table 2 below.

**Table 2: The interaction between the embB protein mutants and ethambutol**

| Mutation | Binding affinity (Kcal/mol) | Residues involved | Number of H bonds | Length (Å) |
|---|---|---|---|---|
| M306V | -3.8 | Gly33 | 1 | 3.19 |
| | | Ser31 | 4 | 2.92, 3.07, 2.95, 2.88 |
| | | Arg27 | 1 | 2.88 |
| M306L | -3.9 | Trp31 | 1 | 3.12 |
| | | Thr33 | 2 | 2.96, 2.98 |
| | | Asp38 | 1 | 2.97 |




E378A




M306L

**A** **B**

**C**

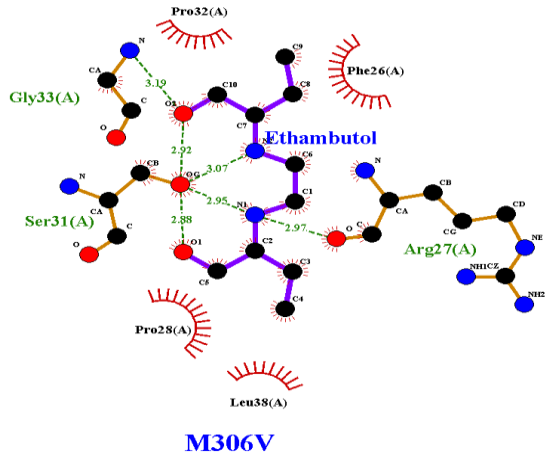



Figure 10: The LigPlot 2D representations of the interactions between the mutant proteins (A-E378A, B-M306L and C-M306V) and ethambutol. *Blue dots: Nitrogen atoms, Red dots: Oxygen atoms,* ***Black dots****: Carbon atoms, Light blue lines: Ligand bonds, Orange Lines: Non- ligand bonds, Green doted lines: Hydrogen bonds, Brick red comb- like shapes: Pocket atoms.*

# CHAPTER FIVE

# DISCUSSION

The embB gene of *M. tuberculosis* produces arabinosylindoylacetylinositol synthase. This enzyme plays a role in the biosynthesis of arabinan mycobacterial cell walls. Certain mutations in the embB gene can cause resistance to EMB.

The study found a binding affinity of -4.8Kcal/mol of ethambutol to the wild- type embB protein (7BVF), indicating a moderate interaction strength. This was consistent with a study carried out by Panja *et al.*, (2019). This affinity suggests that ethambutol has a fairly stable interaction with the embB protein, which is critical for its function as a therapeutic agent against *Mycobacterium tuberculosis*. Additionally, four residues (Leu561, Thr564, Tyr502 and Asp498) were identified as significant contributors to the interaction between the wild-type embB protein and ethambutol through hydrogen bonding.

These residues are likely situated within the binding pocket of the embB protein, providing a scaffold for the interaction with ethambutol. The hydrogen bonds further suggest a specific interaction where ethambutol is positioned within the active site of embB, facilitating its inhibitory action. The hydrogen bond length typically ranges from 2.7Å to 3.3Å. Shorter hydrogen bonds (closer to 2.7Å) are generally stronger than longer hydrogen bonds (closer to 3.3Å) (Yang *et al*., 2022). The 2.87Å and 2.97Å bond lengths at Leu561 and Tyr502 respectively signify stronger attractions contributing to a more stable protein-ligand complex. This increased stability was crucial for the efficacy of the wild- type embB protein to bind to ethambutol.

The study revealed a binding affinity of -3.8Kcal/mol, -3.9Kcal/mol and -4.7Kcal/mol in M306V, M306L and E378A mutations of embB protein respectively. These binding affinities were lower than the one of the wild type embB protein, suggesting a reduced interaction between ethambutol and the mutant proteins. The findings of this study were slightly different from a study by Maladan *et al*., (2023), which found out that in the M306L mutant, molecular docking results indicated an enhanced binding affinity with ethambutol.

The difference in the findings could be due to the fact that some mutations can alter the shape and electrostatic environment of the binding site, leading to a better fit for the

ligand. The mutations in this study, altered the binding pocket which likely weakened the interaction. Compared to the wild type, the mutants showed fewer hydrogen bonds as shown in table 2 above (page 18), indicating that mutations disrupted the optimal binding configuration of the ethambutol. This disruption could potentially lead to a decrease in ethambutol's efficacy contributing to drug resistance.

# CHAPTER SIX

# CONCLUSION AND RECOMMENDATIONS

## 6.1 Conclusion

The study revealed that the wild-type embB protein exhibited a stable interaction with ethambutol with a binding affinity of -4.8 kcal/mol, facilitated by specific hydrogen bonds with key residues such as Leu561 and Tyr502. Conversely, mutations such as M306V, M306L, and E378A significantly altered the binding pocket, reducing the number and strength of hydrogen bonds and subsequently decreasing the binding affinity to -3.8, -3.9 and -4.7kcal/mol respectively.

## 6.2 Recommendations

- Identifying and targeting alternative binding sites on embB or other proteins involved in cell wall biosynthesis that are less prone to resistance mutations.
- Testing combinations of ethambutol with other antibiotics to identify synergistic effects that could enhance overall treatment efficacy in resistant strains.
- Conducting detailed mapping of how different mutations affect the binding site geometry and dynamics aiding in the design of more effective inhibitors.
- Further modelling should be carried out to understand protein mutants and binding ability.

## REFERENCES


B, Z. B., Napiórkowska, A., Bielecki, J., T, E. A., Zwolska, Z., & Jagielski, T. (2013). *Mutations in the embB Gene and Their Association with Ethambutol Resistance in Multidrug-Resistant Mycobacterium tuberculosis Clinical Isolates from Poland*. *2013*.

Baluku, J. B., Nanyonjo, R., Ayo, J., Obwalatum, J. E., Nakaweesi, J., Senyimba, C., Lukoye, D., Lubwama, J., Ward, J., & Mukasa, B. (2022). Trends of notification rates and treatment outcomes of tuberculosis cases with and without HIV co-infection in eight rural districts of Uganda (2015 – 2019). *BMC Public Health*, *22*(1), 1–10. https://doi.org/10.1186/s12889-022-13111-1

Batte, C., Namusobya, M. S., Kirabo, R., Mukisa, J., Adakun, S., & Katamba, A. (2021). Prevalence and factors associated with non-adherence to multi-drug resistant tuberculosis (MDR-TB) treatment at Mulago National Referral Hospital, Kampala, Uganda. *African Health Sciences*, *21*(1), 238–247. https://doi.org/10.4314/ahs.v21i1.31

Bayraktar, B., Bulut, E., Bariş, A. B., Toksoy, B., Dalgic, N., Celikkan, C., & Sevgi, D. (2011). Species distribution of the Mycobacterium tuberculosis complex in clinical isolates from 2007 to 2010 in Turkey: A prospective study. *Journal of Clinical Microbiology*, *49*(11), 3837–3841. https://doi.org/10.1128/JCM.01172-11

Bendre, A. D., Peters, P. J., & Kumar, J. (2021). Recent Insights into the Structure and Function of Mycobacterial Membrane Proteins Facilitated by Cryo-EM. *Journal of Membrane Biology*, *254*(3), 321–341. https://doi.org/10.1007/s00232-021-00179-w

Bwalya, P., Solo, E. S., Chizimu, J. Y., Shrestha, D., Mbulo, G., Thapa, J., Nakajima, C., & Suzuki, Y. (2022). Characterization of embB mutations involved in ethambutol resistance in multi-drug resistant Mycobacterium tuberculosis isolates in Zambia. *Tuberculosis*, *133*, 1–17. https://doi.org/10.1016/j.tube.2022.102184

Cui, Z., Li, Y., Cheng, S., Yang, H., Lu, J., Hu, Z., & Ge, B. (2014). Mutations in the embC-embA intergenic region contribute to mycobacterium tuberculosis resistance to ethambutol. *Antimicrobial Agents and Chemotherapy*, *58*(11), 6837–

6843. https://doi.org/10.1128/AAC.03285-14

Diriba, K., & Churiso, G. (2022). The prevalence of Mycobacterium tuberculosis using Gene Xpert among tuberculosis suspected patients in Gedeo Zone, Southern Ethiopia. *European Journal of Medical Research*, *27*(1), 1–8. https://doi.org/10.1186/s40001-022-00650-x

Goude, R., Amin, A. G., Chatterjee, D., & Parish, T. (2009). The arabinosyltransferase EmbC is inhibited by ethambutol in Mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy*, *53*(10), 4138–4146. https://doi.org/10.1128/AAC.00162-09

He, W., Liu, C., Liu, D., Ma, A., Song, Y., He, P., Bao, J., Li, Y., Zhao, B., Fan, J., Cheng, Q., & Zhao, Y. (2021). Journal of Global Antimicrobial Resistance Prevalence of Mycobacterium tuberculosis resistant to b e daquiline and delamanid in China. *Journal of Global Antimicrobial Resistance*, *26*, 241–248. https://doi.org/10.1016/j.jgar.2021.06.007

Maladan, Y., Safari, D., & Parikesit, A. A. (2023). Structural dynamics insights into the M306L, M306V, and D1024N mutations in Mycobacterium tuberculosis inducing resistance to ethambutol. *Genomics and Informatics*, *21*(3), 1–11. https://doi.org/10.5808/gi.23019

Molla, K. A., Reta, M. A., & Ayene, Y. Y. (2022). Prevalence of multidrug-resistant tuberculosis in East Africa: A systematic review and meta-analysis. *PLoS ONE*, *17*(6 June), 1–13. https://doi.org/10.1371/journal.pone.0270272

Mugumbate, G., Nyathi, B., Zindoga, A., & Munyuki, G. (2021). Application of Computational Methods in Understanding Mutations in Mycobacterium tuberculosis Drug Resistance. *Frontiers in Molecular Biosciences*, *8*(September), 1–11. https://doi.org/10.3389/fmolb.2021.643849

Palomino, J. C., & Martin, A. (2014). Drug resistance mechanisms in Mycobacterium tuberculosis. *Antibiotics*, *3*(3), 317–340. https://doi.org/10.3390/antibiotics3030317

Panja, A. S., Sarkar, A., Biswas, R., Bandyopadhyay, B., & Bandopadhyay, R. (2019). Modification of drug-binding proteins associated with the efflux pump in MDR-

MTB in course of evolution: an unraveled clue based on in silico approach. *Journal of Antibiotics*, *72*(5), 282–290. https://doi.org/10.1038/s41429-019-0146-3

Phelan, J., Coll, F., Mcnerney, R., Ascher, D. B., Pires, D. E. V, Furnham, N., Coeck, N., Hill-cawthorne, G. A., Nair, M. B., Mallard, K., Ramsay, A., Campino, S., & Hibberd, M. L. (2016). Mycobacterium tuberculosis whole genome sequencing and protein structure modelling provides insights into anti-tuberculosis drug resistance. *BMC Medicine*, 1–13. https://doi.org/10.1186/s12916-016-0575-9

Rossini, N. de O., & Dias, M. V. B. (2023). Mutations and insights into the molecular mechanisms of resistance of Mycobacterium tuberculosis to first-line drugs. *Genetics and Molecular Biology*, *46*(1), 1–17. https://doi.org/10.1590/1678-4685-GMB-2022-0261

Salari, N., Kanjoori, A. H., Hosseinian-Far, A., Hasheminezhad, R., Mansouri, K., & Mohammadi, M. (2023). Global prevalence of drug-resistant tuberculosis: a systematic review and meta-analysis. *Infectious Diseases of Poverty*, *12*(1), 1–12. https://doi.org/10.1186/s40249-023-01107-x

Samuel Mwesige, Annet Nankwanga, Florence Tushabe, Ivan Kasamba, & Ruth Kateeba. (2022). Prevalence of pulmonary tuberculosis and the associated clinical symptoms in Western Uganda. *International Journal of Life Science Research Archive*, *3*(1), 155–162. https://doi.org/10.53771/ijlsra.2022.3.1.0098

Schubert, K., Sieger, B., Meyer, F., Giacomelli, G., Böhm, K., Rieblinger, A., Lindenthal, L., Sachs, N., & Wanner, G. (2017). *crossm The Antituberculosis Drug Ethambutol Selectively Blocks Apical Growth in CMN Group Bacteria*. *8*(1).

Seung, K. J., Keshavjee, S., & Rich, M. L. (2015). *Drug-Resistant Tuberculosis*. 1–20.

Singh, V., & Chibale, K. (2021). Strategies to Combat Multi-Drug Resistance in Tuberculosis. *Accounts of Chemical Research*, *54*(10), 2361–2376. https://doi.org/10.1021/acs.accounts.0c00878

Sun, Q., Xiao, T. Y., Liu, H. C., Zhao, X. Q., Liu, Z. G., Li, Y. N., Zeng, H., Zhao, L. L., & Wan, K. L. (2018). Mutations within embCAB are associated with variable level of ethambutol resistance in Mycobacterium tuberculosis isolates from China.

*Antimicrobial Agents and Chemotherapy*, *62*(1). https://doi.org/10.1128/AAC.01279-17

Tan, Y. Z., Rodrigues, J., Keener, J. E., Zheng, R. B., Brunton, R., Kloss, B., Giacometti, S. I., Rosário, A. L., Zhang, L., Niederweis, M., Clarke, O. B., Lowary, T. L., Marty, M. T., Archer, M., Potter, C. S., Carragher, B., & Mancia, F. (n.d.). Cryo-EM structure of arabinosyltransferase EmbB from Mycobacterium smegmatis. *Nature Communications*, *2020*, 1–10. https://doi.org/10.1038/s41467-020-17202-8

Vilch, C. (2020). *applied sciences Mycobacterial Cell Wall : A Source of Successful Targets for Old and New Drugs*.

Wallace, A. C., Laskowski, R. A., & Thornton, J. M. (1995). Ligplot: A program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering, Design and Selection*, *8*(2), 127–134. https://doi.org/10.1093/protein/8.2.127

Yang, Y., Chen, Q., Ruan, S., Ao, J., & Liao, S. G. (2022). Insights into the Inhibitory Mechanism of Viniferifuran on Xanthine Oxidase by Multiple Spectroscopic Techniques and Molecular Docking. *Molecules*, *27*(22). https://doi.org/10.3390/molecules27227730

Yuan, S., Chan, H. C. S., & Hu, Z. (2017). Using PyMOL as a platform for computational drug design. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *7*(2), 1–10. https://doi.org/10.1002/wcms.1298

Zhang, L., Zhang, L., Zhao, Y., Gao, Y., Wu, L., Gao, R., Zhang, Q., Wang, Y., Wu, C., Wu, F., Gurcha, S. S., Veerapen, N., Batt, S. M., Zhao, W., Qin, L., Yang, X., Wang, M., Zhu, Y., Zhang, B., … Rao, Z. (2020). *Structures of cell wall arabinosyltransferases with the anti-tuberculosis drug ethambutol*. *9102*, 1–15.

Zhang, L., Zhao, Y., Gao, Y., Wu, L., Gao, R., Zhang, Q., Wang, Y., Wu, C., Wu, F., Gurcha, S. S., Veerapen, N., Batt, S. M., Zhao, W., Qin, L., Yang, X., Wang, M., Zhu, Y., Zhang, B., Bi, L., … Rao, Z. (2020). Structures of cell wall arabinosyltransferases with the anti-tuberculosis drug ethambutol. *Science*, *368*(6496), 1211–1219. https://doi.org/10.1126/science.aba9102

Zhang, Y., & Yew, W. W. (2015). Mechanisms of drug resistance in Mycobacterium

tuberculosis: Update 2015 [Mecanismos de farmacorresistencia en Mycobacterium tuberculosis: Actualización 2015]. *International Journal of Tuberculosis and Lung Disease*, *19*(11), 1276–1289.

Zhao, L. L., Sun, Q., Liu, H. C., Wu, X. C., Xiao, T. Y., Zhao, X. Q., Li, G. L., Jiang, Y., Zeng, C. Y., & Wan, K. L. (2015). Analysis of embCAB mutations associated with ethambutol resistance in multidrug-resistant mycobacterium tuberculosis isolates from China. *Antimicrobial Agents and Chemotherapy*, *59*(4), 2045–2050. https://doi.org/10.1128/AAC.04933-14

# APPENDICES

## Appendix I: Amino acid three and one letter code conversion table

| Amino acid | Three letter code | One letter code |
|---|---|---|
| Alanine | Ala | A |
| Arginine | Arg | R |
| Asparagine | Asn | N |
| Aspartic acid or aspartate | Asp | D |
| Cysteine | Cys | C |
| Glutamic acid | Glu | E |
| Glutamine | Gln | Q |
| Glycine | Gly | G |
| Histidine | His | H |
| Isoleucine | Ile | I |
| Leucine | Leu | L |
| Lysine | Lys | K |
| Methionine | Met | M |
| Phenylalanine | Phe | F |
| Proline | Pro | P |
| Serine | Ser | S |
| Threonine | Thr | T |
| Tryptophan | Trp | W |
| Tyrosine | Tyr | Y |
| Valine | Val | V |